LECTURE-21

Interactomics-Techniques to Study Protein-Protein Interactions

Handout

PREAMBLE

Proteins are the most important molecules of all biological entities. They are involved in every aspect of the cell structure and functioning – as backbones for support, cell signalling, hormones, enzymes, mediators of the immune system, and as key players in almost every metabolic pathway of the cell. Despite their wide reach in all the cellular processes, the factor that is arguably more than the protein itself in the cell functioning is the interaction between two proteins. Without the correct interaction between the right protein pairs, no biological function can take place. Conversely, a single incorrect interaction, or aberrant interaction could prove to be disastrous for the cell and possibly for the whole organism. This is one of the most fundamental characteristics in biology, be it for any metabolic pathway in heart, brain, immune system or even at any evolutionary level – right from simple bacteria to plants to humans. Thus, study of protein-protein interactions is one of the most pertinent topics for the understanding of biological functions. Another imperative reason why protein interactions are important to study is because they can cause major difference from normal to disease conditions. The excess or absence of a single metabolite might set off a chain of downstream processes, which eventually lead to an abnormal, diseased state. Thus, for achieving the crucial goal of understanding the disease, it is necessary to understand cell-

signaling pathways involved in it, and how the proteins interact under normal and abnormal conditions.

OUTLINE OF LECTURE

I. Interatomics and its significance

II. Methods for studying protein-protein Interactions

a. Traditional Approaches

i. Yeast Two-Hybrid systems

ii. Affinity chromatography

iii. Immunoprecipitation

b. High-throughput Approaches

Protein Microarrays

III. Applications

IV. Advantages and challenges

V. Conclusions

Box-1. Terminology

Interactome: Network of all the interactions in a cell. Interactomics comprises the study of interactions and their consequences between various proteins as well as other cellular components. The network of all such interactions, known as the 'interactome', aims to provide a better understanding of genome and proteome functions.

Yeast Two-Hybrid system: This is a novel molecular biology technique that is used for screening and discovery of protein-protein or protein-DNA interactions. In Y2H two proteins of our interest are tagged to a BD (Binding Domain) and an AD (Activation Domain) of the yeast transcription factor. Indirect interaction between the BD and AD also leads to transcription of the reported gene, thus proving interaction between the two proteins of interest.

Immunoprecipitation: Immunoprecipitation or tandem affinity purification (TAP) is a technique that is used to purify protein complexes and study their protein-protein interactions. Depending on the protein that needs to be purified, different tags can be attached to the bait protein.

Protein Microarrays: These are miniaturized arrays normally made of glass, polyacrylamide gel pads or microwells, onto which small quantities of many proteins are simultaneously immobilized.

I. INTERACTOMICS AND ITS SIGNIFICANCE

Interactomics is the study of a network of interactions. Interactome consists of all the interactions among biological pathways and the associated molecules. Interactomics utilizes bioinformatics approach along with the experimental data. The necessity for the usage of bioinformatics can be attributed to the immense amount of information obtained for each network – the key players, their upstream/downstream interactors, the type of interaction etc. But in all the interactions the most active players are the proteins – almost all major steps in all the pathways are mediated by proteins in the form of enzymes, hormones, receptors, metabolites etc. (Fig. 1). Therefore, to understand the mechanism of cellular molecular processes, studying the protein-protein interaction becomes most essential.

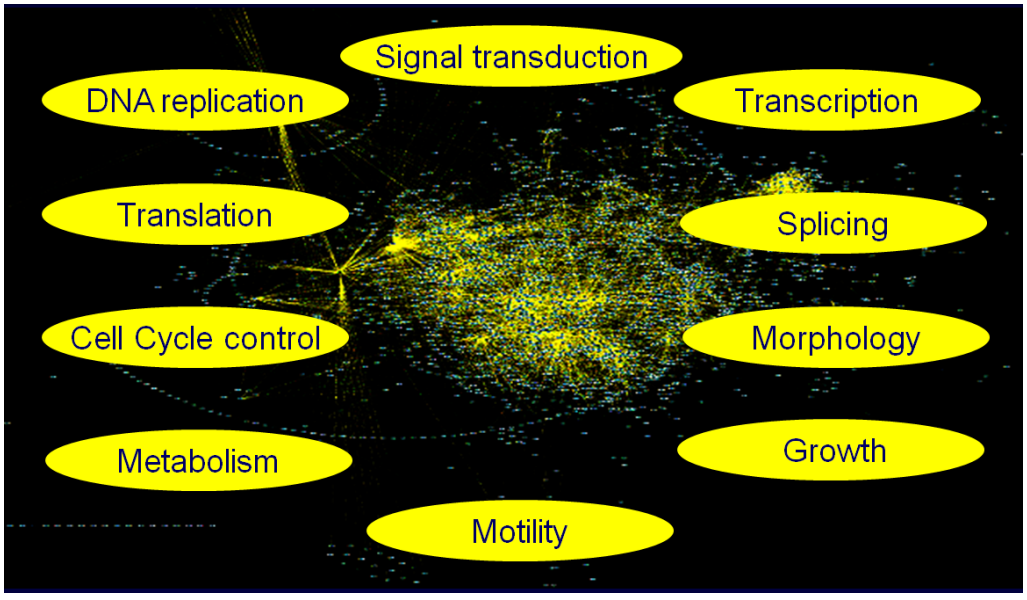

Fig 1. Protein-protein interactions: the potential cellular processes, which may get affected due to improper protein-protein interactions. There are different ways in which

these interactions take place. These interactions can be strong or weak, or they may be transient or permanent and may finally result in formation of homo-oligomers or hetero-oligomers. Any alteration in these interactions could lead to formation of new, incorrect oligomers or render a key factor inactive, or a change in the kinetics of a certain reaction.

Each interactome is highly controlled and tightly regulated. Any imbalance can hamper normal conditions and result in a disease-like state. Thus, studying the protein interactions is important as they interact with a wide variety of biomolecules such as lipids, nucleic acids, small drug inhibitors, etc. Proteins also interact with one another to form macromolecular complexes that regulate signal transduction & gene regulation. Study of these networks could also help in understanding the function of uncharacterized proteins, and also to find out new roles for characterized proteins. Also, new networks of protein interactions can be found out as the same protein may play a role in different pathways. Mechanisms to regulate protein activity can also be discovered, and this comprehensive knowledge will help us get an accurate picture of the real biological interactions. This understanding will further help us to manipulate the networks, thereby opening new avenues to study the diseases.

Illustration 1: Interactomics and its significance

Interaction studies of proteins with various biomolecules help in deciphering and understanding the functions of various proteins in the complex network of cellular pathways. Proteins interact with other biomolecules such as nucleic acids, lipids, hormones etc. to perform a multitude of functions in living organisms such as signal transduction, growth & regulation, and metabolism, to mention a few.

Protein interactions with other biomolecules can be of several different types. They may be weak or strong, obligate or non-obligate, transient or permanent. The physical basis for these interactions include electrostatic, hydrophobic, steric interactions, hydrogen bonds etc.

II. METHODS TO STUDY PROTEIN-PROTEIN INTERACTIONS

a. Traditional Approaches

i. Yeast Two-Hybrid System

The Yeast Two-Hybrid system was developed by Fields and Song in 1989. The Yeast Two-Hybrid system helps in detection of protein-protein interaction by utilizing a transcription factor involving DNA-BD (binding domain), and an AD (activator domain). The main principle which enables this system to be used for studying protein- protein interaction is that the transcription factor has a modular nature and hence the transcription activation can occur even when the two domains are weakly bound i.e. indirectly connected to each other. In Yeast Two-Hybrid system, the yeast transcription factor is divided into two separate parts, the DNA-BD and the AD. The two proteins of interest whose interaction is desired to study are bound to either domains. This is accomplished by generating genetically engineered plasmids, which express a fusion protein of BD and the desired protein, which is known as the Bait. A second plasmid is engineered which expresses a fusion protein of the AD with the other protein to be studied, which is termed as the Prey.

The Bait-BD and Prey-AD are jointly expressed in a yeast cell. If the bait and prey interact, they will bind and bring the BD and the AD in close proximity, which results in the activation of transcription of the reporter gene (Fig 2). If the bait and the prey do not interact, the BD and AD will not interact and the reporter gene transcription will not take place. Thus, one can find out whether two proteins interact with each other. This approach can also be used to screen a library of proteins.

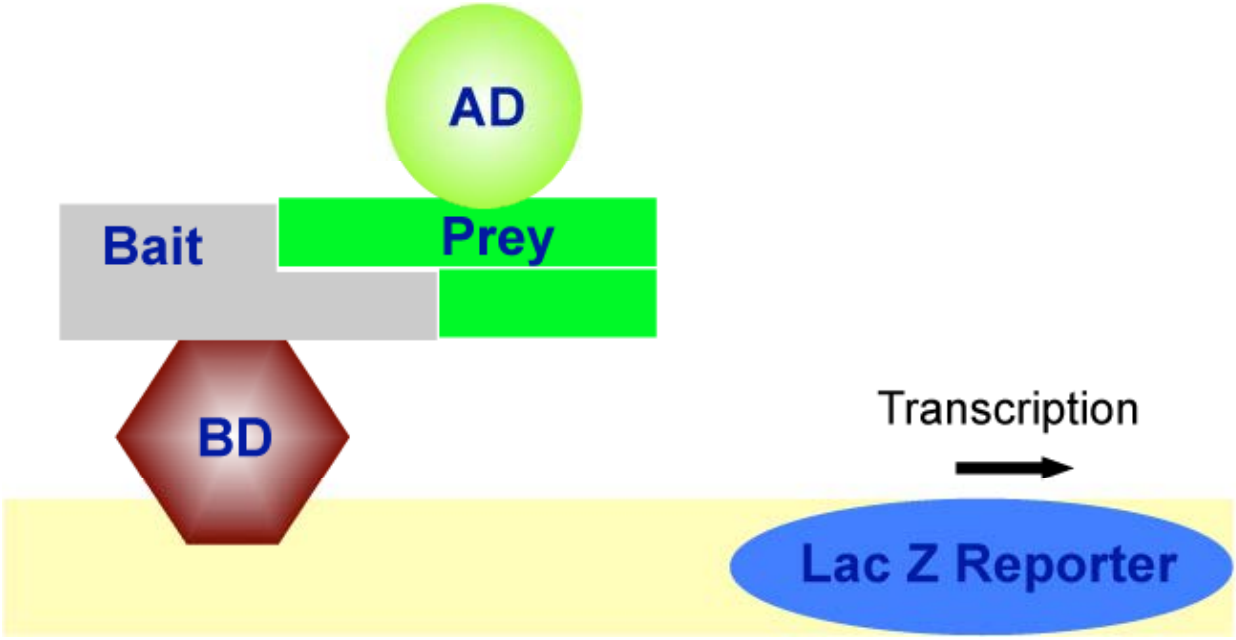

Fig 2. Yeast Two-Hybrid system: Interaction between AD and BD leads to transcription of reporter genes.

Illustration 2: Interactomics and its significance

Binding of transcriptional activator protein composed of binding domain and activation domain to the promoter region is essential for expression of the corresponding reporter gene located downstream of the promoter. The binding domain is fused with the bait protein, while the activation domain is fused with the prey protein. Binding of either one of the fusion proteins to the promoter is not sufficient to bring about transcription of the gene.

The two-hybrid screening protocol uses this interaction as the basis for screening for protein interactions. When the bait protein bound with the binding domain interacts with the prey protein fused with the activation domain, there will be expression of the reporter gene, which can easily be detected. lacZ is a commonly used reporter gene whose protein product β-galactosidase cleaves the substrate X-gal resulting in a blue color.

ii. Affinity Chromatography

One of the most robust methods of detecting and confirming protein-protein interactions is the purification of the actual multi-protein complex by the use of affinity chromatography. For affinity purification, plasmids are generated which tag the protein of interest, resulting in a fusion product of the protein with the tag, either at the N-terminus or the C-terminus. Such plasmids are grown in bacterial cells, resulting in production of the tagged-protein. The protein is then purified using an antibody specific for that tag such as His (histidine), FLAG, GST (glutathione S-transferase) and Halo. After purification, the target protein and its interacting partners get isolated from the sample.

iii. Immunoprecipitation

In immunoprecipitation protein is pulled down from the lysate using an antibody against the known protein. In a co-immunoprecipitation the proteins that are bound to the known protein i.e., the binding partners of the protein or the other members in a protein complex also get precipitated along with the known protein. Thus, it is possible to find out new, unknown interactors of the protein of our interest. Another technique used for protein-protein interactions purification is TAP (Tandem Affinity Purification) technique. Here, a fusion product with TAP tag, consisting of calmodulin binding peptide (CBP) from the N-terminal, followed by tobacco etch virus protease (TEV protease) cleavage site and Protein A is synthesized. The TAP tag binds strongly to IgG and thus the protein of interest along with its other interactors can be purified from the sample.

After precipitation of complexes from the lysate, they are subjected to PAGE (polyacrylamide gel electrophoresis), which separates all the members of multiprotein complex (Fig 3). Each band of protein separated in the gel is excised. After further treatment of the excised gel pieces, the proteins are analysed using Mass Spectrometry, which helps in identifying the unknown proteins that form the complex.

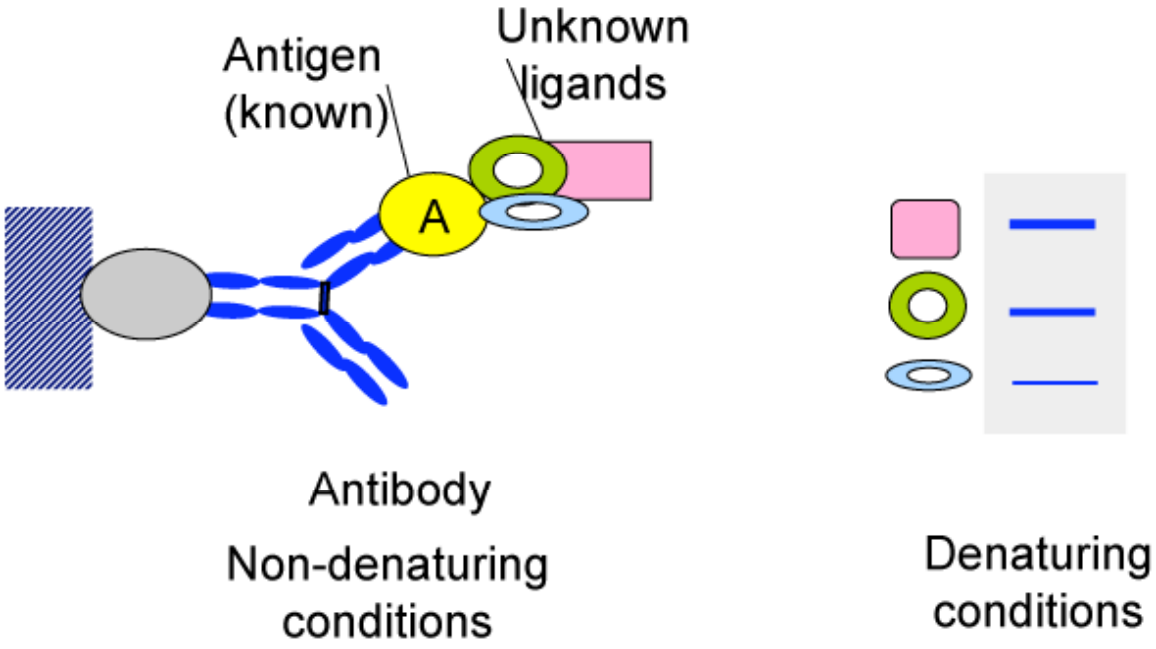

Fig. 3 Affinity purification of proteins: In this illustration a bait protein is tagged with affinity tags. Upon passage of this fusion protein through the affinity column, it binds to the affinity column. The elution of this fusion is done by adding another solution, which has higher affinity for the column. Upon collection of the eluted protein sample, the protein separation is performed using SDS-PAGE and further validation and characterization of the protein is done using MS-based techniques.

Illustration 3: Interactomics and its significance

The protein of interest is fused with a TAP tag, which contains a calmodulin binding peptide, a TEV cleavage site and protein A. Depending upon the proteins to be studied, this tag can be modified. The tag is then bound to a column through affinity interactions between the protein A and IgG. The protein mixtures whose interactions with the bait protein are to be studied, is then added. Some of the proteins form a complex with the bait protein through specific binding interactions. The remaining unbound proteins are then washed away. This is followed by cleavage at the TEV site by the TEV protease to release only the protein complex bound to CBP. These reactions constitute the first affinity step.

In the second affinity purification step, the bait-prey complex is bound via the CBP domain to a calmodulin functionalized column in the presence of calcium ions. The column is washed to remove any other unwanted contaminants after which a chelating agent is added, which complexes the calcium ions. Once these are removed, the CBP-calmodulin interaction is weakened and leads to release of the purified protein complex.

b. HIGH-THROUGHPUT TECHNIQUES

Protein Microarrays

Although the traditional methods of studying protein-protein interactions are effective, they are very time consuming and thus, limit the number of proteins that can be studied at a time. For studying large number of proteins simultaneously, high-throughput methods are needed. One of the high-throughput methods developed for studying protein-protein interaction is protein microarrays. It is one of the most convenient methods for studying large number of proteins in a single experiment. Protein microarrays have a solid surface like a chip or a glass slide, on which thousands of proteins are spotted. The array is incubated with specific probe molecules, which are tagged with a reporter molecule such as a fluorescent chromophore. The array is scanned using a laser or any appropriate method of detecting the reporter and the interaction can be identified as the probes bind at the appropriate proteins spots. In this way, thousands of targets can be screened in one go. This high-throughput platform can be used for biomarker discovery, antigen-antibody studies, studying protein-protein interactions, identification of new interactions and for functional characterization of interacting partners.

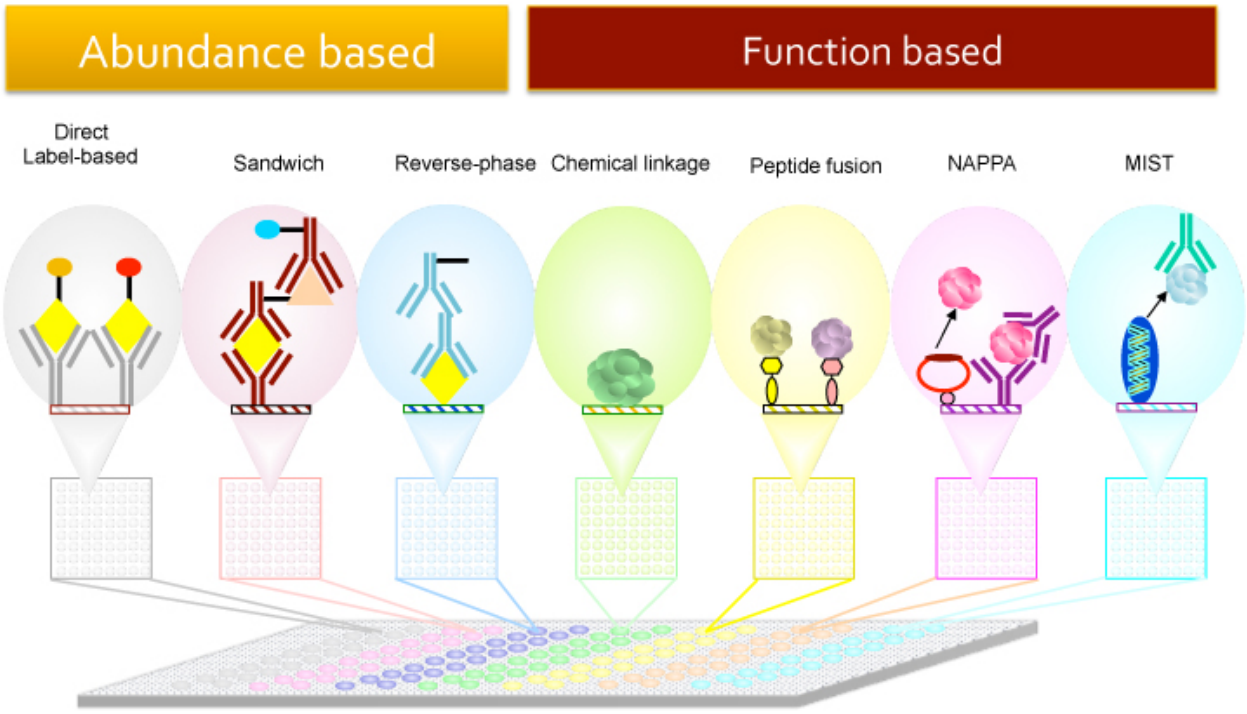

Figure 4. Various types of Protein Microarrays: The protein microarrays can be broadly classified into two kinds, abundance-based and function-based protein microarrays. The abundance-based protein microarrays consist of direct label-based, sandwich and reverse-phase protein microarrays. While the function-based protein microarrays consist of peptide fusion, NAPPA (nucleic acid programmable protein array) and MIST (multiple spotting technique).

Illustration 4: Interactomics and its significance

Once the protein complex has been purified, the components of the complex are separated by electrophoresis under reducing conditions. The SDS gel is then analyzed and the protein components are evaluated thereby providing an understanding about the interactions with the bait protein of interest.

Once the unbound proteins are washed off the array surface, the protein interactions are detected by means of an array scanner. These protein microarrays are extremely useful in studying interactions with other proteins as well as small molecules, DNA or RNA.

III. Applications

The protein-protein interaction studies can be used for various biological applications such as

- Identification of novel proteins
- Identification of function of unknown proteins
- Identification of new binding partners for known proteins
- Functional characterization of protein interactions
- Protein/antibody screening
- Biomarker discovery

III. Challenges

Despite having various distinctive advantages, the techniques involved in interactomics studies have several challenges to overcome and these challenges are summarized in the following table.

Technique	Advantages	Challenges
Yeast Two-Hybrid	• Ability to screen large libraries • Simple protocol • No expensive equipment required • Can be used for DNA-protein, RNA-protein, and protein-protein interactions • Easy to find out gene sequence • Highly sensitive for detection	• High false positive and false negative rates • Proteins must localize and interact in nucleus • Application in a non-yeast environment questionable • Sensitive to toxic gene • Limited to pair-wise interaction • Difficult to detect interactions due to post-translational modifications
Affinity Chromatography	• Proteins in native state • Interactions are natural • Large order complexes can be observed	• Sticky proteins appear regularly • Unclear whether interaction is direct or indirect

		• Expensive • Additional purification may lead to loss of weakly interacting proteins
Protein Microarrays	• Large number of proteins can be assayed in a single experiment • Very small quantities of protein required • Easy screening of the arrays • Protein interaction with nucleic acids, lipids etc can also be studied	• Proteins need to be purified • Protein functionality must be maintained even after binding to the array

CONCLUSIONS

In conclusion, interactomics is a powerful tool to understand the actual interactions networks in the cell – and thus helps in understanding various pathways in the cell. It helps to identify aberrations in protein interactions, which may lead to diseases.

REFERENCES

- Fields, S., Song, O.K. A novel genetic system to detect protein-protein interactions. *Nature* 1989, 340, 245-246.
- Fields, S., Sternglanz, R. The two-hybrid system: an assay for protein-protein interactions. *Trends in genetics* 1994, 10, 286-92.
- Phizicky, E. et al., Protein analysis on a proteomic scale. *Nature* 2003, 422, 208-215.
- Zhu, H. *et al.* Global analysis of protein activities using proteome chips. *Science* 2001, 293:2101-2105.
- Sreekumar A et al: Profiling of cancer cells using protein microarrays: discovery of novel radiation-regulated proteins. Cancer Res 2001, 61:7585-7593.
- Chan, S. M., Ermann, J., Su, L., Fathman, C. G., Utz, P. J. Protein microarrays for multiplex analysis of signal transduction pathways. *Nature medicine* 2004, 10(12), 1390-1396.