IJPSR (2019), Volume 10, Issue 9 (Review Article)

INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES AND RESEARCH

Received on 21 January 2019; received in revised form, 06 May 2019; accepted, 13 July 2019; published 01 September 2019

DIFFERENT APPROACHES FOR STUDYING PROTEIN-PROTEIN INTERACTION AND ANALYSIS

Shivani Joshi [1], Preeti Gupta [2] and Gopaljee Gopal [*1]

C. G. Bhakta Institute of Biotechnology [1], Uka Tarsadia University, Maliba Campus, Bardoli, Surat - 394350, Gujarat, India.
Independent Researcher in Biological Sciences [2], New Delhi - 110070, Delhi, India.

Keywords:

Protein-protein interaction (PPI's), *In-vitro*, *In-vivo*, and Interactome

Correspondence to Author:
Dr. Gopaljee Gopal

Assistant Professor,
C. G. Bhakta Institute of Biotechnology, Uka Tarsadia University, Maliba Campus, Bardoli, Surat - 394350, Gujarat, India.

E-mail: gopaljigopal@rediffmail.com

ABSTRACT: Proteins interact in the cellular arena to carry out a majority of the vital functions of the cells. After exploring the transcriptome, the researchers are now targeting to investigate the interactome of a cell. This can help to design more efficient drug and vaccine formulations and therapeutic targets benefiting humankind in every possible way. Advancement in protein and protein networking databases gives the first insight to study such macromolecular interactions paving the way for their *in-vitro* and *in-vivo* analysis. However, none of the technique can promise 100% efficiency alone due to some limitations like time, cost, and false positive results. Combination of one or more techniques can enhance data quality and reliability. Recent advancements in the studies of the protein-protein interaction using computational methods have led the identification of different signaling pathways and protein complex involved in various specific diseases. However, the major challenge faced is to verify the results obtained from the virtual prediction software's. In this review, we have focused on how different techniques have been accustomed so far to determine the interaction between and among proteins of the target complex or same metabolic/signaling pathway.

INTRODUCTION: Proteins are essential pillars of primary structural organization of an organism and perform a wide variety of functions like- DNA replication, cellular networking, metabolic attributes, growth, transportation and signaling, apoptosis [1, 2]. Most of the cellular functions are accomplished by protein-protein interactions (PPI) with other proteinaceous and non-proteinaceous moieties [3]. The function of a protein is represented by the functional group present like- amide, thiol, carboxyl-amine, hydroxyl, *etc*., which also modulates the activity of a protein [4].

QUICK RESPONSE CODE

DOI: 10.13040/IJPSR.0975-8232.10(9).4006-25

The article can be accessed online on www.ijpsr.com

DOI link: http://dx.doi.org/10.13040/IJPSR.0975-8232.10(9).4006-25

Human genome acts as the blueprint for the information to be translated into proteins. Vidal *et al*., (2011), reported three types of studies in which cellular dynamics can be scored up: gene-based, metabolic pathway based, and protein-protein interaction (PPI) based [5]. Physical interaction between biomolecules forms the basis for biological processes.

Among all biomolecules, proteins form the core moiety around which the cellular dynamics revolve. The study of protein-protein interactions (PPI's) is summarized under the title of 'Interactome Analysis' [6]. The interactome studies have drawn the attention of the majority of system biologists due to its crucial role in cellular networking and signal transduction. It is reported that more than 80% of proteins do not function alone, and they are inevitably found in close associations with other proteins or non-protein

molecules *in-vivo* conditions [7, 8]. Based on their surface properties, the interactions can be transient (occurring for short duration) or permanent (stable for a longer time) having either homo or hetero-oligomeric residues. Transient PPI's will serve the purpose in signaling pathways, for instance, where protein-receptor complex assembles, they pass the signal/information and then dissociate back to individual components once the signal is being relayed. Permanent PPI's are either fixed or at least stand for a considerable time duration, contributing to their higher significance and non-substitutable nature. Based on stability factors, the size of PPI's complex may be larger or smaller to facilitate important cellular functions. Those proteins that serve as the nucleus for multiple associations have been reported to function as either enzyme complexes, central molecules in the signaling cascade or as transcriptional mediators. PPI's have been known to have the following attributes:

1. It can muffle the active site on proteins.
2. Alters the protein-substrate binding site.
3. Performs the regulatory functions in the cellular environment.
4. Diversify the enzymatic reactions due to changes in the enzyme kinetics requirements.
5. Forms the local ground for cellular communication.
6. Helps to construct a detailed map of pathways involved in the variety of biological processes.
7. To characterize an unknown protein found in association with an already known protein or that forms a part of the cellular walkway.

PPI's studies assist in identify the leading drug targets and to characterize the unknown protein moieties based on their networking with other cellular proteins [9, 10]. Based on *in-vitro, in-vivo,* and *in-silico* techniques, multiple experimental approaches have been devised. In this review, we have focused on the *in-vitro* and *in-vivo* based techniques employed to identify and analyze PPI's.

In in-vitro based approach, the experimental procedures are carried outside the cell under a controlled environment, and it includes the techniques like Co-immunoprecipitation, phage display, circular dichroism (CD-spectra), cross-linking, density gradient centrifugation, gel filtration chromatography, native page, and pull-down assays (see **Fig. 1**). Each technique has been discussed in the review in detail along with their variants and modifications developed under the heading of *in-vitro* analysis of PPI's. These approaches aid in detecting a PPI's in the cellular extract and further isolation of proteins by combining different techniques as one technique alone cannot promise 100 % resolution with high throughput.

The advantage of the *in-vitro* based methodology is that the working state can be controlled and modulated as per the necessity of the experiment. Many multi-subunit enzyme complexes and cellular macromolecular protein assemblies can be studied using pull-down assay, co-immunoprecipitation and cross-linking like methodologies. Gel filtration chromatography, density gradient centrifugation, and Native page like techniques can be used in affiliation to the techniques described to study the interaction between multi-unit macromolecular complexes. When individual proteins undergo some association, their native conformation alters, making their structure different from their single monomeric and non-associated state.

Using spectrometry, finest change in the 3-dimensional conformation of a protein which can be attributed due to PPI can also be picked up using circular dichroism, also referred to as CD-spectrometry. Phage display technique, on the other hand, uses the phage's genome to display the target protein on the surface. This target protein also called as the fusion or hybrid protein can then be screened out using specific antibodies. Once the fusion protein-antibody pair is established, an entire library can be produced and then can be channelized for large-scale production of monoclonal antibodies or can be used to serve the purpose in vaccine development.

In *in-vivo* based approach, the whole organism is used to carry out the experimental procedure to study PPI's in the cellular premise in their native interacting state. It includes techniques like Yeast Two-Hybrid systems (Y2H and Y3H) and Fluorescence Resonance Energy Transfer (FRET). Yeast Two-Hybrid system lies on the principle of

binding between the interacting proteins, which in turn can activate the transcription of the reporter gene. The reporter gene product can be a fluorescent protein or chemiluminescent moiety. Different variants of the Yeast Two-Hybrid system have been worked out, giving a better answer to the complex puzzle of the functioning of the cellular arena. Fluorescence Resonance Energy Transfer (FRET) is another variant of Y2H system which uses Fluorescent surface epitopes that absorb light of one wavelength and emits it at another wavelength [11]. The different wavelength emitted falls in the distinct spectra of light (mostly in the visible range) which can be observed either directly or using computer-based software.

Certain *in-silico* based techniques like sequence-based approach, gene expression-based approach, and phylogenetic tree analysis, and few more have been designed using the concepts of bioinformatics. All three modes, *i.e.*, *in-vivo*, *in-vitro*, and *in-silico*, work hand in hand to form an efficient system to study Interactome and its analysis [12]. Proteins and other biomolecules come together to give birth to some novel molecules having potentials far beyond the scope of individual monomer. Understanding the concepts of PPI's will help to resolve the challenges faced in cellular biology and biochemistry and structural biology [13].

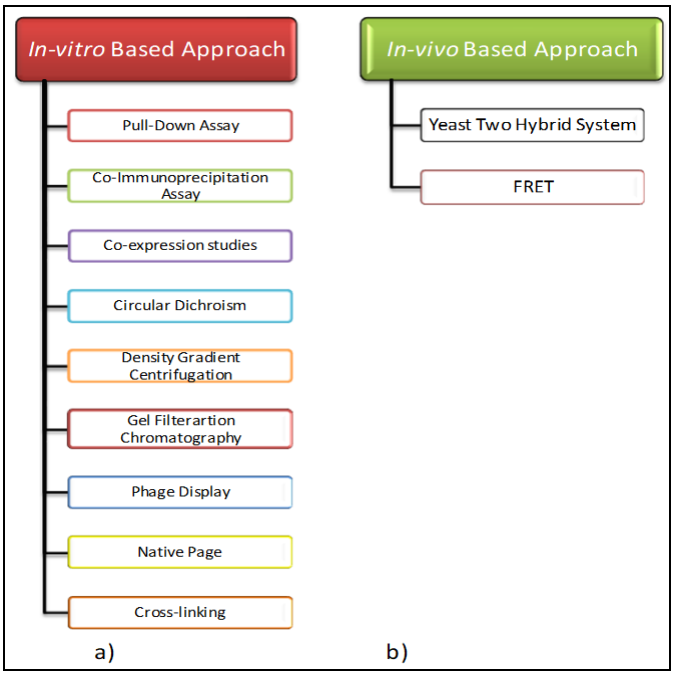

FIG. 1: AN OVERVIEW OF DIFFERENT TECHNIQUES FOR STUDYPPI'S (a) TECHNIQUES INVOLVED IN *IN-VITRO* APPROACH. (b) TECHNIQUES INVOLVED IN *IN-VIVO* APPROACH

***In-vivo* Approach:**

A) Pull-Down Assay: In a regular functional cell, different cellular processes are carried out by interacting with multi-subunit complexes of proteins. During signal transduction pathways, different chemical and the physical stimuli's interact with the effector molecules to govern the proper cellular functions [14].

Several efficient protein interaction studies have been carried out using the pull-down assay [15]. They identified strategies to characterize individual macromolecular complexes by Affinity purification - Mass Spectrometry (AP-MS) by either high-density data acquisition (indirect approach) or by biochemical coupling of AP-MS (direct approach). Moreover, studies in *S. cerevisiae* showed that AP-MS could be used efficiently for high throughput interactome studies [15]. The cellular fraction contains all the protein complexes, which can be purified using the concept of affinity purification [16]. Currently, the major challenge faced is not of protein characterization but protein purification. Bioinformatics, along with proteomics, can help to identify the molecular dimensions of the proteins with high accuracy. Proteins vary based on criteria's like; organization of the domain and motifs, active site present, polarity associated, and many more. These parameters can aid in purifying a protein of interest from the total cellular fraction if used appropriately [17].

The protein interactome studies deal with the identification of the target protein from the total cell fraction and its purification by a variety of approaches, one being pull-down assay (see **Fig. 2**). The proteins interacting with the prey protein can also be co-precipitated along with it, thus giving an insight into the structural organization of the prey protein complex additional to the prey protein interactome assembly. The bait protein can be attached to the column in two different ways; - (i) using a ligand or (ii) using an Antibody specific for the bait protein.

Ligand must be firmly bound covalently to the bait protein so that it does not get dissociated with the bait protein at the time of elution. Also, it should not alter the specificity of the active site in the bait protein. Various polymers can effectively serve the purpose of a linker. Antibodies can be used to

immobilize the bait protein in the column. The specificity provided by antibodies is more to ligands used. A group of researchers experimentally expressed MBP-tagged CagX protein from *H. pylori* in *E. coil* strain DH10β. Specific antibodies namely, Anti-His and Anti-MBP were used to find out the His-CagX (prey protein) and another MBP/MBP-tag (used as a bait protein) [18]. Points to keep in consideration while dealing with antibodies is that chances of cross-reactivity are high.

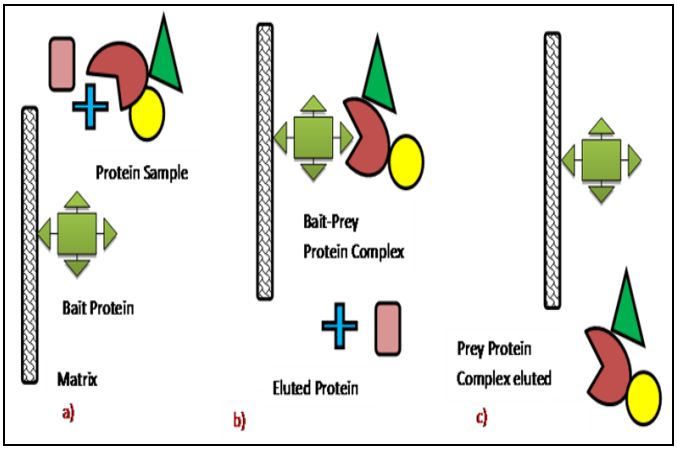

FIG. 2: GENERAL STEPS INVOLVED IN PULLDOWN ASSAY. a) Binding of bait protein to the matrix. This binding is the generally strong covalent type where the Bait protein is attached to the matrix before the experiment. b) Formation of Bait-Prey Protein complex. The bait protein has a selective affinity towards the target protein, which causes it to stop and stand by the matrix. c) Elution of the Prey Protein Complex. Once the monomeric protein gets eluted out, the targeted complex can be obtained and then purified. The pattern of elution varies when the proteins exist in the native state from when it lies in conjugated form, and this variation caused can be detected using Pull down Assay.

Point to keep in consideration while dealing with antibodies is that the chance of cross-reactivity is high. Moreover, the type of antibody used will also influence the purification process. Once the bait protein is appropriately immobilized, the cellular fraction is allowed to pass through the column. The prey protein has a specific binding site in this domain, and as a result, it can bind covalently to the bait protein. The target protein may or may not interact with some other proteins. In a pull-down assay, if the prey protein interacts with other protein moieties, these interacting proteins can also be purified together in a single run. The eluted fractionate can be then analyzed using high throughput techniques like Mass spectrometry and Column chromatography. Low throughput techniques like SDS-PAGE analysis and Western Blotting can be conjugated to study the expression levels of the isolated target protein.

Recently, a group of researchers has developed a technique called single-molecule pull-down or SiMPull assay in which the traditional pull-down assay has been combined with single-molecule fluorescence imaging (TIRF microscopy), enabling direct molecular visualization of the single cellular protein complex directly from cellular or tissue extract [19]. The cellular extract, when applied to the specific antibody, passivated microscopic slides, the antibody holds the protein along with its interacting partners. Washing of unbound components followed by its probing can then be visualized either by antibody labeling or by fluorescence protein tags genetically encoded. SiMPull utilizes less time and reagents, giving more efficient quantitative data. SiMPull can also distinguish between proteins having multiple associated states and applies to the range of endogenous proteins [20]. Limitation of this method is that the weak interactions with dissociation rate constants > 0.01 s^{-1}s, which may not be accessible for the detection [19].

Thus the Pull -down assay can help to resolve many challenges faced till date which includes;

- Functional activity of the protein can be studied
- Used for those protein complexes that cannot be studied using recombinant methods
- Promises to improve sensitivity, time consumption and accuracy in term of western blot analysis
- Cost-effective.

Despite having such a high efficiency of purification, this technique is not entirely free from flaws. Since it relies on fluorescent signals emitted, any contamination or similar protein can give false results. Moreover, the separation results are based on affinity interaction and not size and density. Also, the data obtained from software analysis requires correct interpretations. Overcoming these hurdles can make pull-down assay as an ideal technique for the study and analysis of PPI's.

B) Co-Immunoprecipitation: Cellular dynamics are chiefly governed by the macro biomolecules

like proteins that function within to carry out a specific function. Some protein functions independently while other works in close proximities to support a biological activity [21]. Expression of the protein depends on the external stimuli received, and consequently, the protein expression pattern becomes dynamic. Also, not all the proteins function as a housekeeping protein; some are expressed in a cell-dependent manner. Such challenges make the task difficult while studying the proteins in their native or biologically active state [22].

With the discovery of the human genome project and advances in the field of genomics and proteomics, the study of proteins became a vital aspect. Traditional techniques focused on a single protein, but with advancements in protein studies, the emphasis has been done on analyses of interactions between the proteins, which is the crucial point for PPI's.

Co-Immunoprecipitation (Co-P) is an analogs technique to immunoprecipitation. A specific antibody is used to bind appropriately to the domains of the target protein. Upon binding of prey protein to the antibody, other interacting proteins also elute out together. Washing steps result in the isolation of the desired binding complex (see **Fig. 3**). The binding complex can then be resolved using SDS-PAGE and Western Blotting technique [6].

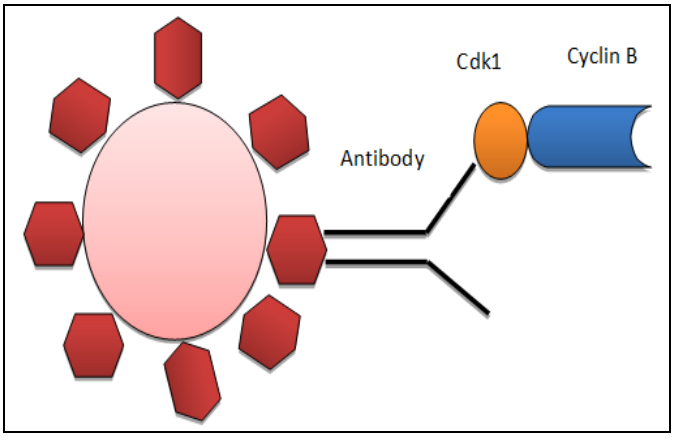

FIG. 3: CO-IMMUNOPRECIPITATION OF Cyclin B USING AN ANTIBODY SPECIFIC FOR Cdk1. Cyclin B INTERACTS WITH Cdk1 IN THE CELL, AND THIS PROPERTY CAN BE USED TO ISOLATE SELECTIVELY Cyclin B MOIETY FROM THE CELLULAR POOL USING Cdk1-SPECIFIC ANTIBODY ATTACHED TO A SOLID SUBSTRATUM (HERE MAGNETIC BEAD). NOT ONLY THIS, OTHER PROTEINS WHICH INTERACT WITH Cyclin B CAN ALSO BE PRECIPITATED OUT USING THIS TECHNIQUE

Co-P technique is subject to the availability of specific antibodies against the targeted proteins. Because of this, the technique cannot be replicated for large-scale analysis, and researchers are trying hard to overcome this limitation. One such strategy proposed is to use short sequences of proteins called "tags," which can conjugate with the target protein. Tag specific antibodies are used instead of protein-specific antibody to carry out Co-P with a higher degree of affinity. One such experiment using the concept was carried out for isolation of CagX and CagT proteins from type IV secretion system (TFSS) in *Helicobacter pylori* [23]. For example, binding of CyclinB and Cdk1 in the cell, where Cdk1 antibody (as bait) is used to bind cyclinB (as prey) and its associated binding partners which can be co-immunoprecipitated using a sepharose or magnetic bead as a stable base for binding. Geer *et al.* also reported a well-studied PPI in a signal transduction pathway mediated by protein-tyrosine kinases receptors [24].

Advantages:

- Used to confirm novel protein-protein interactions.
- Analyze the effect of mutant protein on a normal binding.
- Being cost-effective stands as a gold standard to study protein interactions.

Major limitation faced include; to confront the transient type of protein interactions, sometimes protein interacting site may be overlapping with the antibody binding site, which gives a false negative result for no binding. Variations in antibody binding fashion to the same protein due to different affinities and domains present. Epitope tags can hinder in complex formation or the appropriate antibody binding. Additionally, the technique cannot function in the presence of strong ionic detergents like SDS. When interacting proteins Expression levels are low, a false signal can be obtained due to low binding or weak binding [24]. Efforts are desired to enhance the Co-P technique so that rapid and number of protein or peptide bindings can be studied.

C) Co-expression Studies: Cellular dynamics have active networks where different biomolecules interact to carry out a specific function. Proteins

have different motifs and domains contributing to its solubility in the aqueous medium. When a protein has hydrophilic residues in the motif, they become readily soluble in the aqueous medium, but it is not the same case with hydrophobic proteins. Hydrophobic residues with long polypeptide chains in the domain structure make protein insoluble, which then forms unique structures called inclusion bodies in the cell [25]. The inclusion body protects the hydrophobic residues of the protein. Factors like a reduced cytoplasmic environment in the host cell, lack of molecular chaperones in eukaryotes, improper post-translational machinery and presence of partially or misfolded proteins all contribute towards the formation of more inclusion bodies [25, 26]. The expression of such proteins may last for only a short duration in the cell premise. Analysis of interacting partners and their further purification from within these inclusion bodies poses a challenging task [27]. To showcase some light on this problem, an innovative technique has been proposed, known as Co-Expression studies (see **Fig. 4**).

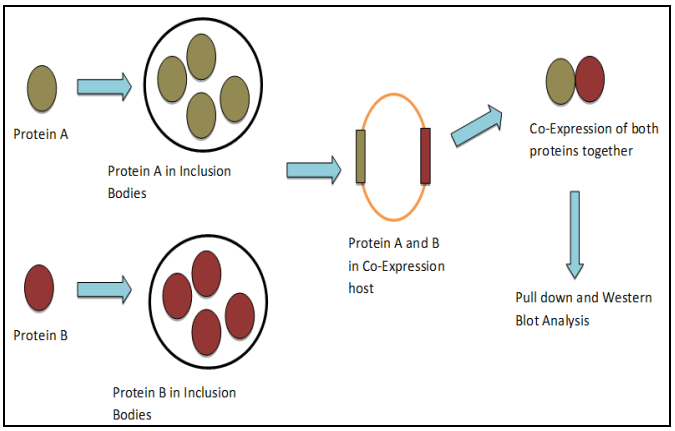

FIG. 4: CO-EXPRESSION STUDIES TO RULE OUT THE PROBLEM OF INCLUSION BODY FORMATION. Protein A and B initially forms inclusion bodies, thus hampering in their studies. When the genes for particular proteins are co-expressed in a common vector, the product formed is in a conjugated form masking the hydrophobic residues or making it soluble in the solution which can then be isolated using Pull down or Western Blot analysis.

A detailed protocol has also been published a decade back in Nat. Protocols for the Co-expression of heterologous proteins in *E. coil* host, which till date faced a lot of hurdles. Many systems were used including Mammalian cell culture, yeast, and viral-based strategies but had sore reputation due to complicated and expensive nature. With the development in dual vector system along with commercial plasmids more efficient affinity tags like maltose binding protein (MBP) and glutathione S-transferase (GST), aided the researchers for better recovery and solubility of desired proteins [28]. In different methods, interactions between the protein molecules may be stable, strong, or transient, as shown in **Table 1**.

TABLE 1: COMMON METHODS TO ANALYZE THE VARIOUS TYPES OF PROTEIN INTERACTIONS

Method	Protein-protein interactions
Co-immunoprecipitation (Co-P)	Stable or strong
Pull-down assay	Stable or strong
Crosslinking protein interaction analysis	Transient or weak

D) Circular Dichroism: In the cell, proteins not only interact with other proteins but also to some non-proteinaceous molecules making their analysis and isolation even more difficult. Rapid detection and characterization of novel proteins are of great importance in the field of Genomics and Proteomics. High throughput techniques like X-ray crystallography and NMR sometimes becomes impractical or difficult in application. In such a situation, Circular Dichroism (CD) can demonstrate its potential to resolve the complex 3-D and 2-D structure of a protein [29].

To understand the concept of CD prior idea about polarization is essential. Transverse wave possesses electric (E) and magnetic (B) fields that lie perpendicular to one another and also to the direction of propagation and oscillate [30]. Those light waves whose oscillation is confined to a particular plane, are called linearly polarized light. When two different lights interact, each at a right angle to one another (one being horizontal and other being vertical) and are in-phase, such light rays behave as linearly polarized light.

In the case when they go out-of-phase, by a quarter wave, they form a helix, and the resultant wave is called circularly polarized light (CPL). The helix can turn either in right-hand fashion (R-CPL) or in left-hand fashion (L-CPL) and interestingly do not superimpose on each other. The difference in the absorbance pattern of L-CPL and R-CPL forms the premise for Circular Dichroism. Distinct from the optical rotation, Circular Dichroism (CD) occurs at a specific wavelength at which the chiral molecule can absorb the light. CD is measured as a function of wavelength and forms the basis for Circular

Dichroism (CD) spectroscopy [31, 32]. The prominent use of CD lies not only in proteomics and genomics but also in biochemistry and structural biology. The technique is used to resolve the complex high order structures of Proteins and DNA molecules and also to study interaction patterns. Each macromolecular structure has a unique CS-signature, and this can be accustomed to identifying essential structural elements and also to note the changes in the structure of those essential elements if any. Secondary structures of proteins like α-helix and the β sheet are widely used as CD- signatures [33].

The gist of CD-spectrum obtained depends on the confirmation of the proteins and hence can be used to identify the conformation of unknown molecules. Not only this, conformational variations caused in a known protein due to reaction conditions like temperature, pH, mutation, heat or interactions are leading to an alteration in overall morphology can also be analyzed. Also, when chromophores domains of the protein interact with the respective ligands, they develop strong extrinsic CD bands as a signal which can assist in tracking the binding zones. For example, aromatic chromophores of a target protein will have bands in near UV-range, but due to mutations, the band signal will be altered, giving an idea that to what extent a variation or alteration can distort the native protein conformation.

In CD-spectroscopy, the sample to be analyzed should be at least 95% pure based on the standards of HPLC, MS, or Gel electrophoresis. For determination of the protein structure, sample concentration (conc.) should be between 0.005-5 mg/mL based on the path length to be used [34]. The hurdle faced in CD is of appropriate protein concentration estimation. The commonly used methods like Folin- Lowery and Bradford Assay are not suitable for analysis of protein concentration for samples to be analyzed using CD spectrometry.

Of all the known techniques, most accurate protein concentration can be determined using quantitative amino acid assay. In this technique, the concentration of stable amino acids (mostly lysine or alanine) is used as a measure to detect the concentration of intact protein present in the sample.

Challenges faced in CD-spectrum are that the 2-D structure of proteins cannot be analyzed. Moreover, more sophisticated techniques like Fourier Transform Infrared Spectroscopy (FTIR) and Raman Spectroscopy can be used over CD-spectrometry [35, 36]. Also, analysis of membrane proteins is possible using CD, but due to high distortion caused by differential scattering of light, the use is limited [37]. Still, the CD can be used efficiently as the collected data can be analyzed in a few hours itself. CD-spectrum can be analyzed using programs [38]: CDSSTR, CONTIN, and SELCON3 from Dichroweb [39]. CDSSTR program gives better results over other programs used [29].

Advantages:

- To compare macromolecules similar in structure.
- To check the correct folding of a new or purified protein.
- Folding pattern in a mutant protein concerning the wild-type.
- To check the activity of a biopharma product.
- Can resolve the conformation of soluble proteins [40].

E) Density Gradient Centrifugation: In the world of genomics and proteomics, large macromolecular complexes can be sorted into their monomeric units based on their molecular masses and the order of varying densities [41]. Density gradient centrifugation (DGC) involves centrifugation of biomolecules in a medium having a spectrum of densities, ranging from lighter to denser. When the interacting proteins are subjected to graded densities, under the influence of the centrifugal field, they migrate through the medium and localize themselves according to their densities [42] (see **Fig. 5**).

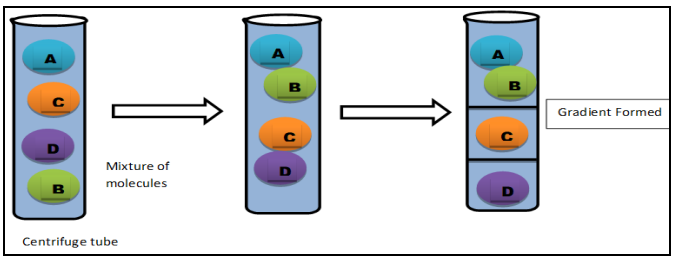

FIG. 5: DENSITY GRADIENT CENTRIFUGATION. From the sample of molecules present in a mixture with graded centrifugation, the interacting and non-interacting molecules get parked in different zones of solution, thereby making their separation possible.

Centrifugation technique did not find its relevance outside the biochemistry lab till the 1930s. This powerful separating approach came into the spotlight when the researcher's attention shifted from the structural basis of biology to separation of cellular fractions and enzymology. As a result, analytical ultracentrifuges were designed to carry out Qualitative analysis while preparative centrifuges served to solve the problems of Quantitative analysis. For the very first time "mitochondrial fraction" was resolved into a 'heavy fraction' possessing mitochondrial pellet and the 'light fraction' enriched with lysosomal content [43]. Based on the differences in the principle of separation, DGC can be classified into:

Rate Zonal Centrifugation: The era of the 1950's witnessed the emergence of a new discontinuous gradient system which was designed for separating non-proteinaceous moieties from the protein complexes. The gradient medium used is made up of different densities of the same solution and hence is referred to as Rate zonal. This method isolates protein complexes based on size and shape. The most commonly used gradients were of sucrose and glycerol (see **Table 2**). The technique was employed to understand the interactive nature between receptor-associated proteins (RAP) and its binding partner Lipoprotein Lipase (LPL) which is thought to have importance in LPL maturation [44].

Based on the application, gradients makers are presently used for large-scale sorting and manual gradients for small-scale separations [45]. Also, the uniformity and linearity of the gradient prepared are checked using a refractometer. Up to > 5mg of total protein on a 40 ml gradient has been successfully resolved [46]. Purification and detailed analysis of fractionates obtained can be done further using even more sophisticated techniques like Mass Spectrometry. Binding of proteins causes a change in the active site of the molecule resulting in an alteration in its physical and chemical parameters, which can be identified using DGC. The zone thus differs where the singular and the associated state can be obtained from, making it an excellent strategy to separate interacting protein complexes from the solution.

Isopycnic Separation: Isopycnic separation is also referred to as equilibrium or buoyancy separation as the separating particles get distributed solely based on their densities [47]. In here, the size of the molecule influences until the time it reaches the zone of equilibrium density. Also, this technique is very valuable for separating isoforms of proteins, having similar shape and structure [48]. The interacting partners become heavy and thus do not penetrate deep inside the gradient, while the free forms migrate according to their densities and line up in the medium, creating a good window for studying proteins and their binding partners.

The primary function of any density medium is to assist in sorting of macromolecular complexes either based on buoyancy phenomena or due to the rate of sedimentation. An ideal density medium should have sufficient solubility, not very viscous, does not affect the biology of the sample and moreover, is non-toxic.

TABLE 2: TYPES OF GRADIENT MEDIUM USED IN RATE ZONAL CENTRIFUGATION

Type of Gradient Medium	Principle Use
Polyhydric Alcohols	
Sucrose	Organelles, Membrane Vesicles, Viruses, Protein and Ribosomes
Glycerol	Mammalian Cells (not widely used), protein separation
Inorganic Salts	
CsCl	DNA, Proteins, Viruses
KBr	Plasma lipoproteins
Iodinated Gradient Media	
Iodixanol	Mammalian cells, Protein complex in cell, DNA
Polysaccharides	
Ficoll	Mammalian subcellular particles (not widely used)

Major limitations in DGC are of sample dilution following lower resolution. Also, extended durations of spins pose time challenge [49]. If accompanied by other separation techniques like chromatography and Mass Spectrometry, better and efficient results can be obtained.

Advantages:

- Used to separating Nucleic acids and its potential binding partners
- Assist in characterizing proteins with similar molecular weights but different densities
- An important technique to study PPI's at the organelle level as well as the cellular level.
- Proteins conjugated with lipids can be efficiently characterized [50].

F) Gel Filtration Chromatography: One of the ancient techniques used among all is of Gel Filtration Chromatography [51] (see **Fig. 6**). It lies on the principle of the degree of covalent interaction present between the interacting moieties. A similar concept was used by Pfleiderer (1964) to study the interaction between lactic dehydrogenase and reduced diphosphopyridine [52, 53]. Hummel and Dreyer in 1962 came up with an alternate technique using gel filtration to evaluate the binding between ribonuclease and cytidine 2'-phosphate. The technique was known as Hummel–Dreyer Method of Equilibrium Gel Filtration [54]. The activity of many multi-subunit proteins involved in metabolic regulation was initially exploited by gel filtration technique.

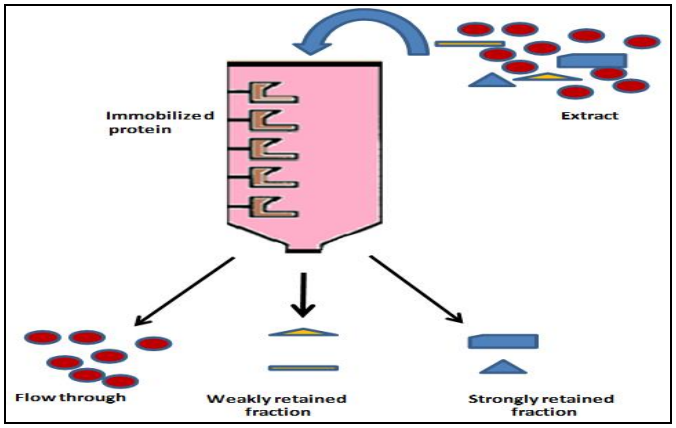

FIG. 6: DIAGRAMMATIC REPRESENTATION OF GEL FILTRATION CHROMATOGRAPHY. The mixture, when loaded into the column, the affinity towards the matrix allows a better screen for the proteins that exist in a conjugated state from those lying in a monomeric state. Stringency conditions maintained during the flow-through and the conformational attributes makes the monomeric proteins to elute out first followed by the elution of the conjugated proteins.

Not only this, using gel filtration chromatography, protein-peptide interactions were also validated. A well worked tryptophan-serum albumin system was developed to appraise the interactome [55, 56, 57]. Using Dialysis-equilibrium method, BSA and human serum albumin were found to strongly bind to 1 mole of L-tryptophan/mole of protein. Also, the study was carried out using different tryptophan derivatives (L-tryptophan methyl ester, tryptamine, acetyl-L-tryptophan, D-tryptophan methyl, and few more) [58].

Active Enzyme Gel Chromatography: Principle lies in running the column using all the substrates of one particular enzyme, one at a time and then monitoring the wavelength of the effluent obtained. When two or more substrates interact, the pattern of wavelength obtained will vary. Different substrates will interact with the same enzyme differently giving altered peaks at different wavelengths, which can be gauged accordingly.

Example: Two mitochondrial enzymes, citrate synthase and pyruvate dehydrogenase (PDH) were visualized using this technique [57].

Hummel–Dreyer Method of Equilibrium Gel Filtration: In this variant, those protein interactions are studied that are very dynamic and last for a short period. Due to which the interaction cannot be obtained in the flow-through, giving a false negative result. The problem can be overcome using Hummel–Dreyer Method targeting such reversible type of ligand-protein binding [54].

Example: Interaction between citrate synthase (100 kDa) and pyruvate dehydrogenase complex (7000 kDa) from pig heart mitochondria was studied using Hummel–Dryer technique [57].

Gel Filtration Equilibrium Analysis: Being a variant of classic gel filtration equilibrium assay its principle is analogous to that of equilibrium dialysis.

Example: Study of the interaction between glutamate dehydrogenase and aspartate aminotransferase [59].

Centrifugal Gel Filtration: In this variant, the fractionation range of gel is chosen such that target protein does not enter the beads while the other one does. By using this technique, large proteins interacting with the small ones can be analyzed,

and if they elute out together, the presence of interaction can be confirmed.

Example: Interaction between α-ketoglutarate dehydrogenase complex (higher molecular weight complex) and NAD-isocitrate dehydrogenase [60].

Being a classical technique, more variant can be developed combining high throughput technology, to analyze PPI and to study the interactome with a higher degree of precision. It can lay light on various untold aspects of proteomics' and protein molecular biology.

G) Phage Display Technology: Proteins form the base of the structural pillars on which the architecture of an organism stands. Thousands of cellular and extracellular proteins work in association to carry out a mutual function. By using traditional methods, it is difficult to screen such a large number of interacting pairs. Stands as the most versatile technique developed, Phage display is used to resolve this problem. The technique is used to study PPI, Protein-DNA interaction, and protein-peptide interaction [61]. The foreign insert upon translation produces the protein of interest, which is showcased on the phage surface. This lays an apparent physical link between the genotype and the phenotype of the phage (see **Fig. 7**).

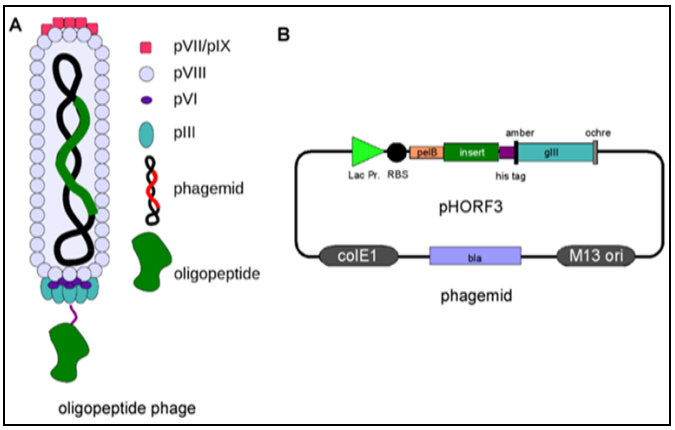

FIG. 7: A) MOLECULAR STRUCTURE OF PHAGE GENOME SHOWING THE PROTEINS IT CAN ENCODE FOR B) PLASMID SHOWING THE ORIENTATION OF PHAGEMID GENES IN A VECTOR [4, 62]

Phage is well adapted parasitic machinery, which has DNA or RNA as its genetic blueprint encapsulated in a proteinaceous coat. The significant attribute of a phage is, it can accommodate a foreign stretch of nucleic acid also called as an 'insert' within its genome. The targeted gene is inserted near the genes coding for coat protein where it gets translated, giving rise to a hybrid proteinaceous coat. The foreign peptide thus gets expressed on the coated surface of the phage aiding in its efficient selection. Because the foreign peptide or protein gets displayed on the surface, this technique is called "Phage display." Phage Genome carrying additional information makes it an excellent vehicle for the intentional delivery of the targeted protein. Phagemids are termed so become of their phage genomic origin. Expression plasmids generally have MCS region, one or more antibiotic resistance gene, epitope tags like His or c-myc and lacZ promoter marker for direct visualization of the r-phagemid [63].

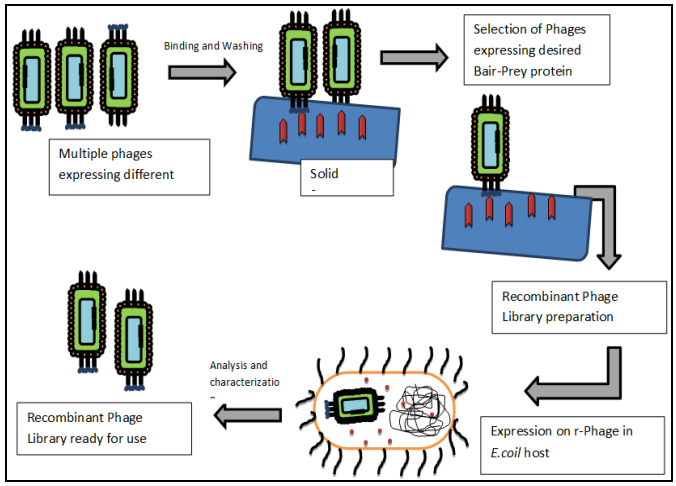

FIG. 8: PHAGE LIBRARY PREPARATION. Multi-step process in which the group of phages is designed expressing different proteins on their surface. Solid support (also called cassata) is immobilized with the bait protein and used to selectively capture the prey protein displayed on the phage surface. Once the bait-prey protein complex is obtained, it is isolated and made to express in a bacterial system like *E. coil*. Amplification and characterization are done to obtain large scale production of prey protein [64].

Once the phage library is prepared, solid support or cassette immobilized with the target prey protein is used to provide the ground for bait-prey binding. Elution step is carried out to remove the non-specific binding, and the desired bait-prey complex is then eluted out (see **Fig. 8**). The phage displaying the correct fusion peptide or protein is then selected and made to infect the host cell; this step is called the amplification. The steps can be repeated multiple times, the process of biopanning. Then after the amplified phage genome is sequenced to detect the gene responsible for the production of the correct fusion protein.

In this way, thousands of bait-prey complexes can be analyzed, and studies using limited resources with a higher degree of accuracy. Phagemids are used to express the target protein but cannot mediate the assembly of the virion on its own and hence requires helper phage to provide the genes essential for the assembly viable phage. However, with the advancement in technology, the need for helper phage can be eliminated by using bacterial packaging cell line technology [65].

E. coil is the standard host for a majority of r-DNA technological protocols. Different bacteriophages used are filamentous phages including strains f1, Fd, M13 and ft, Lambda phage (λ), T4 and T7 phage [66, 67]. These ss-DNA bearing bacteriophages infect a wide range of gram-negative bacteria mediated by their pill [68]. Due to their enhanced potential as a cloning vehicle and for its simplicity of accurate assembly of longer phage particles, they are preferred choice. Most of the peptides and the recombinant antibody fragments are showcased at protein pIII and pVIII. pVIII is a product of 8 gene expression giving rise to nearly 3000 protein copies and coding for major coat protein (see **Fig. 9**). The major coat protein of FF phage binds the peptide using the N-terminal domain [69]. Despite producing a huge number of protein copies to be displayed on the coated surface, it can be fused for only 6-8 amino-acid long residues [70]. Size constraint can be resolved using artificial coat protein (APC) [71].

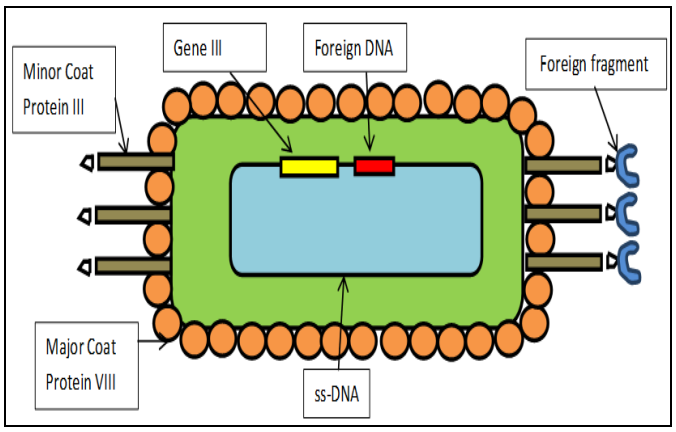

FIG. 9: SCHEMATIC DIAGRAM FOR REPRESENTATION OF COMPONENTS OF A PHAGEMID [72]

The infectivity of the virion is determined by pIII protein, which is composed of three domains, namely; N1, N2, and CT, each of which is linked by linkers rich in glycine amino-acid [69]. The F pill binds to the N2 domain during infection, and the N1 is accountable for binding to the TolA protein present on the bacterial surface. One edge provided by pIII over pVIII is that it provides space for monovalent display of larger protein sequences (> 100 amino acid) and is more receptive for carrying such large inserts [71]. The drawback associated with pIII-fusion protein is its reduced infectivity, which results in ineptitude to infect the host. The phage fusion protein assembly can be displayed as C-terminal fusion with pIII, pVIorpVIII proteins or N-terminal fusion with pIII, pVII, pVIII or pIX for large proteins [73, 74]. The very 1st phage display was carried out in filamentous phage where pIII was conjugated with the targeted peptide [61].

The disadvantage of using FF phage is that it requires fusion protein to be translocated to the inner membrane of the bacteria before their assembly into a virulent phage. Not all fusion proteins can undergo this event and thus cannot be displayed on the surface. T7 phage stands as an alternative for M13 phage display [67]. T7 phage can be used where no such translocation is required, but they come along with their limitations of not displaying larger peptides or protein fragments [75].

Lambda (λ) phage is competent to display larger molecular weight protein sequences on both N and C-terminal domains without undergoing membrane translocation making it an ideal choice [74, 76]. Regardless of many limitations, M13 and T7 are among the widely used phage display system [77].

Amount of r-protein that will be displayed on the phage depends on factors like;

- Display system used for expression (phagemid or phage).
- Choice of the helper phage.
- Type of coat protein chosen for binding (pIII or pVIII)

Apart from the natural libraries prepared, some commercially available systems of phage display enlist; pSEX8, pCOMB3 or its derivatives and pCANTAB 5E [78, 79].

Advantages:

- Easy to screen a huge population of phages (~109 clones).

- iAmplification can be carried out using a PCR based method.
- Can potentially serve as a variant for Yeast Two-Hybrid system.
- A large amount of M-Ab can be produced in an affordable period
- Ag-Ab specificity can be studied with higher precision [80].
- A Direct link between genome and proteome can be established.

H) Native Page: Cellular microenvironment is richly comprised of proteins, enzymes, and other macromolecules that conjugate to form a multi-subunit complex. Multiple such large macro-molecular complexes exist in a cell to carry out the vital process of an organism like growth, reproduction, metabolism, and many more. Traditional techniques were oriented to obtain more of a quantitative data rather than a qualitative one. Moreover, the majority of assay conditions rendered the proteins in the sample inactive and useless for studying further. To add, the availability of pre-existing data was also crucial for carrying out the assay. The concept to the native page was proposed to overcome the challenges faced along with the issue of loss of protein activity after the assay.

Generally, the gel electrophoresis is used to provide data covering the polarity, molecular weights, structural details, and most importantly, the purity of the protein sample. The techniques proposed till date can be categorized into two main sub-types: (i) use of a detergent to induce charge shift so that the treated sample proteins migrate in a uniform fashion and (ii) migration under the effect of protein's charge and the pH of the gel [81].

Being a variant type of technique, it is important in the analysis of those PPI's where confirmation and the activity of the proteins are vital [82]. Since SDS or other detergent treatment is not given, the 3-D or 2-D conformation of the interacting proteins in the sample is not distorted, giving scope for studying protein activity even after the electrophoresis. Not only this, but the native gel also called Clean Gel Electrophoresis (CN-PAGE) can be used to determine the protein masses, and the oligomeric states associated [83]. The Native PPI's can be retrieved back from the gel by native electroblotting, electroelution, and immunodetection. Native gel electrophoresis can be conjugated with high-end electron microscopy to decipher the complex 2D crystal lattice structure of a protein or interacting peptides [83].

Though, it seems advantageous variant, native gel electrophoresis is full of drawbacks, among them major ones include;

- Protein isoforms with similar isoelectric values cannot be separated efficiently.
- Presence of post-translational modifications like acetylation, phosphorylation, methylation, etc. may hamper in the appropriate PPI studies.
- Uniquely optimized protocols need to be established for each protein sample to be studied.
- Amount of starting sample also pose a hurdle, low volume with high protein concentration is desired for the analysis.
- Difficult to work with membrane protein complexes due to the formation of aggregates [81].
- Low resolution on the gel in the presence of multisubunit complexes.

To fill the gaps of CN-PAGE, Blue Native gel Electrophoresis, also called BN-PAGE technique, was developed. This method uses Coomassie Blue-G250 dye (Coomassie/protein ratios (1:1)) to impart an overall negative charge to the protein [84, 85]. CBB being less ionic to SDS, can be used to separate protein complexes based on charge/mass ratio86. Inhere the protein sample migrate through the gel based on its pore size and is anionic. Moreover, water solubility and ability to bind to the membrane proteins form the essential backbone for CBB to be a good candidatewho can substitute SDS from the assay. The separation range lies between 100kDa -10MDa. Other mild detergents like dodecylmaltoside, Triton X-100, and digitonin can be used in BN-PAGE. Limitations associated with BN-PAGE include; CCB being anionic can distort the protein interactions. Lack of resolution and requirement of robust and clean antibodies to detect the protein in its native form pose another serious challenge.

In order to combine the merits of both CN-Page and BN-PAGE, proteins with similar masses can be isolated using CN-PAGE in the 1st dimension followed by BN-PAGE in 2nd dimension [87].

I) Cross-linking: Various protein interactions have been reported to occur in the cellular arena, some of which are stable and can sustain for a few hours while others are of a transient type. These transient interactions are known to stay only for a short duration, after which the interacting members dissociate. The presence of such a short window period makes it challenging to study and further to isolate such interacting partners appropriately. Most of the efforts are diverted to overcome such a challenge. One such strategy proposed is of inserting some chemical linkage so that when the proteins interact, they hold on to each other stably for a longer duration without undergoing dissociation [88]. The technique of making an artificial linkage between the interacting protein-peptides is called cross-linking. Proteins interact when they lie close to each other, within few angstroms and with the help of crosslinkers, they can sustain in-position for even longer [4].

The binding occurs between the functional group of the protein and the chemically reactive moieties in the crosslinker. Most preferred reactive moieties in a crosslinker include specific groups having a Sulfhydryl-reactive group (Maleimides), Amine-reactive group (N-Hydroxysuccinimide{NHS}) and non-specific photoreactive groups having Diazirines, Benzo-phenones, and aryl azides. Among them, NHS cross-linkers are widely used because of the chemical attributes it has. NHS cross-linkers are preferred since they have high reaction efficiency, can form stable amide bonds with amines in protein under reaction conditions, can work efficiently with Lysine residues in protein structure and crosslinking occurs at physiological pH (7-8) allowing the study of proteins structures in their native state [89].

Once the crosslinking is done, the proteins undergo digestion step yielding peptide fragments. These obtained peptide fragments can be screened for the presence of crosslinkers, and then the final result can be analyzed using SDS-PAGE, immuno-precipitation, or Mass spectrometry (MS) [90] (see **Fig. 10**). An alternative to this can be, activating the crosslinker when required by using the concept of photoactivation [91]. The pattern of binding gives an insight into the 3-D conformation existing within the protein and cross-linking same protein complex in different ways can help to study the functional and structural aspects even deeper. Some naturally occurring crosslinking found in protein is of Di-sulphide linkage allowing the 2-D and 3-D stable conformation.

Crosslinking technique can be used in association with MS to allow mapping and analysis of proteins having a low-resolution topology and also to study interacting proteinaceous partners [92]. Also, the cross-linking method can be conjugated with immunoaffinity purification techniques to study binding partners in living cells and to resolve conformational relationships in a protein complex's [93, 94]. However, to obtain good qualitative data, identification of the specific binding site is crucial. Apart from this, specific binding sites in only well-characterized proteins have been identified yet, making it difficult to study interactions in novel proteins or proteins of the unknown character. The significant conclusion is that the cross-linking technique requires a few more advancements, which can pave its way to becoming an ideal technique for the study of PPI's.

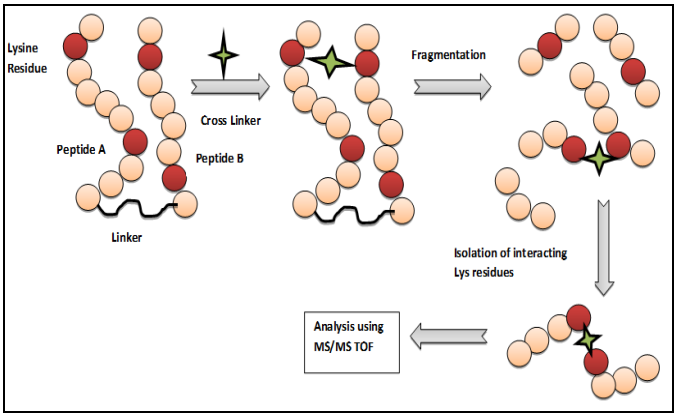

FIG. 10: CROSS-LINKING TECHNIQUE USED TO STUDY PROTEIN-PROTEIN INTERACTIONS. Specific sites can be crosslinked using chemical or physical agents, allowing the study of those particular regions better even after fragmentation of the protein. Crosslinkers used are designed to bind to a particular region aiding in an efficient study of interacting proteins.

***In-vivo* Approaches:**

A) Fluorescence Resonance Energy Transfer: Fluorescent proteins are used in the field of system biology to study the localization of a protein in a

cell and the dynamics associated. Tagged chromophores are used to study how proteinaceous moiety interacts with similar and dissimilar macromolecules [90]. One such technique developed by Theodor Förster was the Förster resonance energy transfers (FRET) or as fluorescence resonance energy transfers or electronic energy transfer (EET) [95]. In FRET there is no actual transfer of fluorescence, and the technique is also designated as the radiationless transfer. Since no photon emission occurs, the acceptor molecule need not be fluorescent. Transfer of energy from donor to acceptor leads to a decrease or quenching of donor's fluorescence and a decline in the lifetime of excited state accompanied by an increase in the fluorescence intensity of the acceptor [96] (see **Fig. 11**).

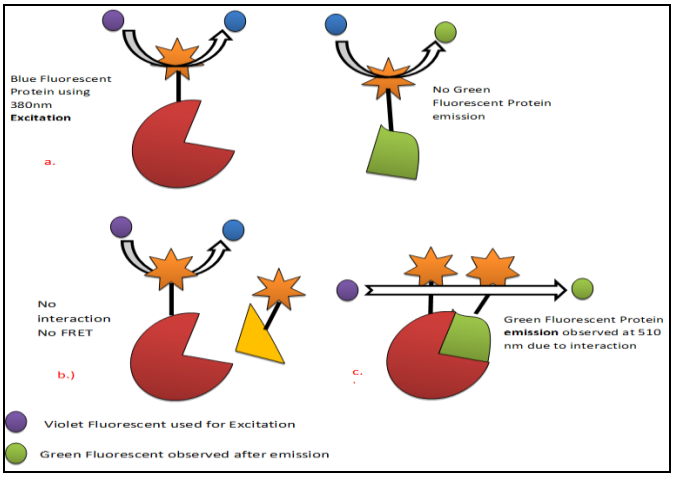

FIG. 11: FRET TO STUDY PROTEIN-PROTEIN INTERACTION. a) No emission of green fluorescence is obtained as there is no protein interaction occurring. b) YFP does not bind with BFP giving any fluorescence. c) BFP binds with GFP to give out green fluorescence at 510 nm)

FRET relies on the distance-dependent mode of transfer [97]. It is comparable to short-distance communication in which the range of contact ought to be smaller than the excitation wavelength used. According to the theory, the excited fluorophore acts as an oscillating dipole transferring its energy to acceptor dipole having the same resonance [96]. Förster, through his experiments, established that the sixth- inverse distance between donor and acceptor gives the efficiency of the process (E) given by the following equation [95]:

$$E = Ro^6 / (Ro^6 + r^6)$$

Ro is the Förster radius, the distance at which energy transfer is 50% efficient, and r represents the actual distance between the donor and the acceptor. The value of Ro is influenced by the spectral features of both the donor and the acceptor.

Criteria important for FRET to function include: [98]

- The interacting fluorophore should be nearby of each other (1-10 nm distance)
- The fluorescence emission spectrum of the donor should significantly overlap with the excitation or emission spectrum of the acceptor.
- The transition orientation of donor and acceptor dipole should be parallel.
- The fluorescence lifetime of the donor should be sufficient enough to allow the FRET to occur [11].

The degree of spectral overlap is referred to as spectral overlap integral (J). Any change in the distance between the interacting fluorophores will directly influence the rate of FRET. Also, this technique is not restricted to just fluorophores-chomes used. FRET working with phosphorescence has also been reported [99].

FRET is used along with fluorescence microscopy and confocal laser scanning fluorescence microscopy to solve the puzzles of molecular biology, biochemistry, and biophysics [99]. It has dramatically revolutionized the area of PPI's. FRET's efficiencies can be measured by calculating the emissions of both the donor and the acceptor fluorophores. To evaluate the complex formation between the donor and the acceptor, both the proteins are fused with fluorophores emitting different colors upon excitation. Since, FRET results in both increases in acceptor's fluorescence and a decrease in donor's fluorescence, a ratiometric determination of the signals obtained can be used to quantify the efficiency of FRET. In contrast to the single-single sensor, ratiometric sensors possess two or more varied types of fluorophores, and the detection is based on the ratio of intensities emitted by the two fluorophores upon excitation [96].

Different fluorophores used in FRET are; CFP-YFP pairs, BRET or bioluminescence resonance energy transfer and Homo-FRET.

Applications:

- ✓ Structure and conformation of proteins and nucleic acids [100].
- ✓ Spatial distribution and assembly of proteins [5].
- ✓ Receptor/ligand interactions [101] Immunoassays [102] Real-time PCR assays and SNP detection [103], Nucleic acid hybridization [104]. Distribution and transport of lipids [105].
- ✓ Membrane fusion assays [116]. Membrane potential sensing [106].
- ✓ Fluorogenic protease assays [107].
- ✓ Cyclic AMP indicators [108].

B) Yeast Two-Hybrid System: The cellular machinery works on the interaction between the biomolecules, especially macromolecules like proteins involved in the basic cellular process and thus very imperative to study the interaction patterns observed [109]. Traditional methods like co-purification, gradient centrifugation, and chromatographic techniques did not prove to be very useful for analyzing such interactions. The problem was solved when a revolutionary concept was proposed in the 1980s by Stanley Fields and Ok-Kyu Song [110]. They developed a strategy to study PPI's and named it as "Yeast Two-Hybrid System" or Y2H Assay. Y2H has contributed to the field of molecular biology to validate protein interactome assembly through activation of the downstream reporter gene (see **Fig. 12**).

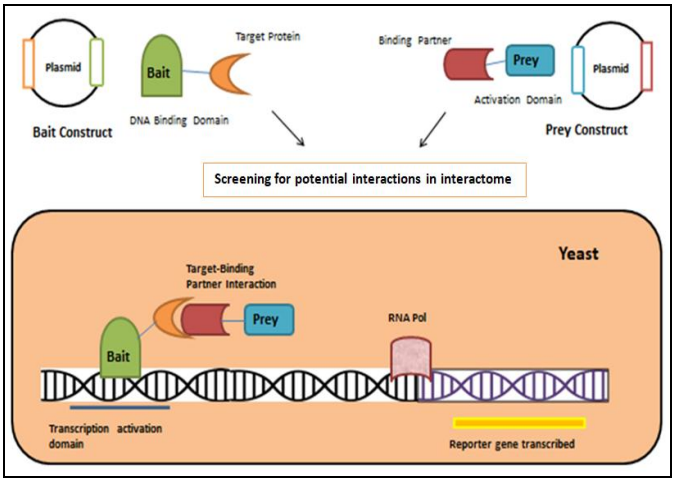

FIG. 12: YEAST TWO-HYBRID SYSTEM. a) Bait plasmid construct expressing bait-target protein. b) Prey plasmid constructs expressing prey-binding partner protein complex. c) Interaction of Bait-target protein-complex on DNA Binding Domain with the prey-binding partner protein on Activation Domain results in active transcription.

One important aspect of the success of this strategy was due to its timeliness [110]. To carry out Y2H assay, an auxotrophic strain of yeast is required, deficient for Tryptophan (TRP) and Leucine (LEU) gene for the biosynthesis pathway. A plasmid construct is developed encoding for fusion proteins with both binding domain and the activation domain, respectively. Such plasmid constructs can be developed using cDNA library preparation. When they are cloned into a host cell having one antibiotic selection marker each, and then made to transform on a trp- and leu- deficient medium, only those cells survive that have both the recombinant plasmid constructs an expression, *i.e.*, trp+ and leu+. The most common used Bait and Prey proteins are residues 263-352 of yeast Gal11P with an N342V mutation and residues 58-97 of yeast Gal4 respectively [111]. The advantage of using these domains is that they can adapt well to both yeasts and bacterial system with a strong interaction [109, 111]. Gal4 Protein is responsible for activation of the transcription of the gene for galactose utilization and as a result, forms the basis for selection. BD-Bait protein complex binds the UAS region on the DNA but cannot initiate transcription.

Similarly, the AD-Prey protein complex has the potential to recruit RNA Pol to carry out transcription but cannot initiate the process alone. When the binding between the bait-prey protein complex results, the BD-AD domains come in close proximity, which causes the transcription of the reporter gene. Variety of reporter genes have been developed suiting the aim of the experiment. Most commonly used reporters are; Green Fluorescent Protein (GFP), GUS and LacZ gene (encoding β-galactosidase). Some AD has been developed for Y2H system which includes, herpes simplex virus-derived AD, VP16 and yeast Gal4 AD giving successful results [109]. Using *E. coil* as the host, IPTG- induced lac promoter is used to carry out the selection [112]. Variants developed for yeast two-hybrid system having such potential to identify protein-protein interactions; some variants have been developed to enhance the sensitivity and the range of study which includes:

Yeast One-Hybrid: The technique is used to study Protein-DNA interaction. Inhere a single fusion protein is used, and the AD directly binds to the BD.

Yeast Three-Hybrid: The technique is used to study Protein-RNA interaction. It involves the use of non-fusion proteins. These proteins do not normally interact, but binding results due to RNA intermediate present on these non-interacting proteins [109].

Yeast One-Two-Hybrid: The technique is to study both Protein-Protein and Protein-DNA interaction. Because both protein and DNA interaction can be analyzed simultaneously, it is called the one-two-hybrid approach. This method enhances the stringency of the screening [109].

Split-Ubiquitin Yeast Two-Hybrid: The technique is used to analyze the interaction between non-soluble proteins present in the membrane. The assay was specifically designed to overcome the challenge faced in simple Y2H [113]. In here, the two different ubiquitin moieties; C-terminal moiety ('Cub', residues from 35-76) and N-terminal moiety ('Nub,' residues from 1-34) are fused with the integral membrane proteins and are known as "bait" and "prey" respectively. Binding of bait and prey protein results in the activation of reporter transcription [114].

Fluorescent Two-Hybrid Assay: Here, the fusion proteins are conjugated to a fluorescent tagged, and the interactions result in the emission of fluorescence, which can be detected either visually or using fluorescent microscopy. The fusion proteins are tagged to the lacI gene, encoding for lac repressor protein [115].

Enzymatic Two-Hybrid Systems: Like Y2H system which uses Transcription factor binding in the nucleus, this technique utilizes the concept to enzymatic activation to study protein-protein interaction. This is also called a Kinase Substrate Sensor ("KISS"), is designed to locate intracellular protein-protein interactions. In this mammalian two-hybrid approach, kinase containing the portion of TYK2 is fused to the bait protein, and the gp130 cytokine receptor fragment is conjugated to prey protein. The interaction of bait-prey protein complex results in phosphorylation of STAT3 docking site on prey chimera by TYK2, leading to reporter gene activation [116].

Dual Bait Yeast Hybrid System: Unlike simple Y2H system, dual bait hybrid system uses two different baits to interact with one putative interacting partner simultaneously. The utility of this method was proved using Ras-superfamily of proteins. Selective interaction between Raf, Krit1 (AD-fused moiety) and Ral-GDS with related Ras family GTPases, Ras (Bait 2) and Krev-1 (Bait 1) proteins were analyzed. The study gave experimental evidence for the interaction between the Raf-Ras, Krit1-Krev-1, and RalGDS-Ras and Krev-1 both [117].

Advantages:

- An easy and relatively speedy technique to detect protein-protein interaction.
- Presence of Technical skills is not a prerequisite criterion.
- Can be completed in an affordable period.
- Futuristic scope for scaling up the available.
- Large quantitative data can be presented within a short time zone.

Being one of the most promising *in-vivo* technique, it has a darker side as well. In here, chances of getting false positive results are relatively high. Post-translational modifications can alter the binding sites in the domain, and the difference in the pattern of modification in yeast system concerning the mammalian system also poses a hurdle. To note, Y2H occurs in the nucleus, and it can be challenging to analyze cytoplasmic proteins as they may act as non-interacting ones. Lastly, the fusion of tag may alter the active site architecture.

Application:

- ✓ Drug discovery [118].
- ✓ Sequence determination of the interacting proteins can be done [109].
- ✓ Study of protein structure can be elucidated [109].
- ✓ Structure of interactome of protein can be analyzed.
- ✓ Zinc finger protein selection [119].

CONCLUSION: The proteins are mischievous moieties that switch their functions almost immediately even with minor changes in the surroundings, and so their native state isolation pose a big issue, and that is what needs to be

addressed. Techniques to isolate these interacting complexes serve as the tool which can be used to dig out the answers to various unsolved puzzles of cell biology. Some techniques may give excellent yield for one set and adverse outcomes for another type of proteins. It does not mean that one technique is better than the other. The choice must depend on the physio-chemical and topological properties of the protein of interest. Based on the knowledge gained, different groups can extend their findings and can help to place a piece of the puzzle, eventually completing the picture of cellular, molecular dynamics which is governed predominantly by the proteins. Also, a better understanding of the principles can help to design new outlets which can promise even better results in the future.

ACKNOWLEDGEMENT: Financial support provided by the Science and Engineering Research Board (SERB), New Delhi under fast track scheme of young scientists gratefully acknowledged.

CONFLICT OF INTEREST: The authors confirm that this article content has no conflicts of interest.

REFERENCES:

1. Aze A and Maiorano D: Recent advances in understanding DNA replication: cell type-specific adaptation of the DNA replication program. F1000 Research 2018; 7.
2. Phizicky EM and Fields S: Protein-protein interactions: methods for detection and analysis. Microbiol Mol Biol Rev 1995; 59(1): 94-23.
3. Ardejani MS, Powers ET and Kelly JW: Using cooperatively folded peptides to measure interaction energies and conformational propensities. Accounts of Chemical Research 2017; 50(8): 1875-82.
4. Kluger R and Alagic A: Chemical cross-linking and protein-protein interactions-a review with illustrative protocols. Bioorganic Chemistry 2004; 32(6): 451-72.
5. Watson BS, Hazlett TL, Eccleston JF, Davis C, Jameson DM and Johnson AE: Macromolecular arrangement in the aminoacyl-tRNA. cntdot. elongation factor Tu. cntdot. GTP ternary complex. A fluorescence energy transfer study. Biochemistry 1995; 34(24): 7904-12.
6. Ng SK and Tan SH: Discovering protein-protein interactions. Journal of Bioinformatics and Computational Biology 2004; 1(04): 711-41.
7. Berggård T, Linse S and James P: Methods for the detection and analysis of protein-protein interactions. Proteomics 2007; 7(16): 2833-42.
8. Yanagida M: Functional proteomics; current achievement. Journal of Chromatography B 2002; 771(1-2): 89-06.
9. Pedamallu CS and Posfai J: Open source tool for prediction of genome-wide protein-protein interaction network based on ortholog information. Source Code for Biology and Medicine 2010; 5(1): 8.
10. Sarmady M, Dampier W and Tozeren A: HIV protein sequence hotspots for crosstalk with host hub proteins. PLoS One 2011; 6(8): e23293.
11. Clegg RM: Förster resonance energy transfer-FRET what is it, why to do it, and how it's done. Laboratory Techniques in Biochemistry and Molecular Biology 2009; 33: 1-57.
12. Rao VS, Srinivas K, Sujini G and Kumar G: Protein-protein interaction detection: methods and analysis. International Journal of Proteomics 2014.
13. Zhang A: Protein interaction networks: a computational analysis. 2009.
14. Yamada T and Bork P: Evolution of biomolecular networks-lessons from metabolic and protein interactions. Nature Reviews, Molecular Cell Biology 2009; 10(11): 791.
15. Gingras AC, Gstaiger M, Raught B and Aebersold R: Analysis of protein complexes using mass spectrometry. Nature Reviews, Molecular Cell Biology 2007; 8(8): 645.
16. Puig O, Caspary F, Rigaut G, Rutz B, Bouveret E and Bragado-Nilsson E: The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods 2001; 24(3): 218-29.
17. Deutscher MP: Rethinking your purification procedure. Methods in Enzymology 1990; 779-80.
18. Gopal GJ, Kumar A, Pal J and Mukhopadhyay G: Molecular characterization and polyclonal antibody generation against core component CagX protein of *Helicobacter pylori* type IV secretion system. Bioengineered 2014; 5(2): 107-13.
19. Jain A, Liu R, Ramani B, Arauz E and Ishitsuka Y and Ragunathan K: Probing cellular protein complexes using single-molecule pull-down. Nature 2011; 473(7348): 484.
20. Jain A, Liu R, Xiang YK and Ha T: Single-molecule pull-down for studying protein interactions. Nature Protocols 2012; 7(3): 445.
21. Pawson T and Nash P: Assembly of cell regulatory systems through protein interaction domains. Science 2003; 300(5618): 445-52.
22. Van der Geer P: Analysis of protein-protein interactions by coimmunoprecipitation. Methods in Enzymology 2014; 35-47.
23. Gopal GJ, Pal J, Kumar A and Mukhopadhyay G: C-terminal domain of CagX is responsible for its interaction with CagT protein of Helicobacter pylori type IV secretion system. Biochemical and Biophysical Research Communications 2015; 456(1): 98-03.
24. Van der Geer P, Wiley S, Gish GD and Pawson T: The Shc adaptor protein is highly phosphorylated at conserved, twin tyrosine residues (Y239/240) that mediate protein-protein interactions. Current Biology1996; 6(11): 1435-44.
25. Carrió MM, Cubarsi R and Villaverde A: Fine architecture of bacterial inclusion bodies. FEBS Letters 2000; 471(1): 7-11.
26. Carrió MM and Villaverde A: Localization of chaperones DnaK and GroEL in bacterial inclusion bodies. Journal of Bacteriology 2005; 187(10): 3599-601.
27. Burgess RR: Refolding solubilized inclusion body proteins. Methods in Enzymology 2009; 259-82.
28. Tolia NH and Joshua-Tor L: Strategies for protein coexpression in *Escherichia coli*. Nature Methods 2006; 3(1): 55.
29. Chakraborty H and Lentz BR: A simple method for correction of circular dichroism spectra obtained from membrane-containing samples. Biochemistry 2012; 51(5): 1005-08.

30. Rodger A and Nordén B: Circular dichroism and linear dichroism. Oxford University Press, USA; 1997.
31. Fasman GD: Circular dichroism and the conformational analysis of biomolecules. Springer Science & Business Media 2013.
32. Berova N, Nakanishi K, Woody RW and Woody R: Circular dichroism: principles and applications. John Wiley & Sons; 2000.
33. Greenfield NJ and Fasman GD: Computed circular dichroism spectra for the evaluation of protein conformation. Biochemistry 1969; 8(10): 4108-16.
34. Greenfield NJ: Using circular dichroism spectra to estimate protein secondary structure. Nature Protocols 2006; 1(6): 2876.
35. Cooper E and Knutson K: Fourier Transform Infrared Spectroscopy investigations of protein structure. Physical methods to characterize pharmaceutical proteins. Springer, 1995; 101-43.
36. Pelton JT and McLean LR: Spectroscopic methods for the analysis of protein secondary structure. Analytical Biochemistry 2000; 277(2): 167-76.
37. Denisov IG and Sligar SG: Nanodiscs in membrane biochemistry and biophysics. Chemical Reviews 2017; 117(6): 4669-713.
38. Whitmore L and Wallace BA: Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers: Original Research on Biomolecules 2008; 89(5): 392-00.
39. Whitmore L and Wallace B: DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Research 2004; 32(S-2): W668–W673.
40. Kelly SM, Jess TJ and Price NC: How to study proteins by circular dichroism. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 2005; 1751(2): 119-39.
41. Taulbee D and Maroto-Valer MM: Centrifugation 2000;
42. Lodish H, Berk A, Zipursky S, Matsudaira P, Baltimore D and Darnell J: Purifying, detecting, and characterizing proteins. Molecular cell biology WH Freeman, New York. 2000;
43. De Duve C, Pressman B, Gianetto R, Wattiaux R and Appelmans F: Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue. Biochemical Journal 1955; 60(4): 604.
44. Page S, Judson A, Melford K and Bensadoun A: Interaction of lipoprotein lipase and receptor-associated protein. Journal of Biological Chemistry 2006; 281(20): 13931-8.
45. Fernandez-Martinez J, LaCava J and Rout MP: Density gradient ultracentrifugation to isolate endogenous protein complexes after affinity capture. Cold Spring Harbor Protocols 2016; 7: pdb–prot087957.
46. Miernyk JA and Thelen JJ: Biochemical approaches for discovering protein-protein interactions. The Plant Journal 2008; 53(4): 597-09.
47. Price CA: Centrifugation in density gradients. Academic Press; 2014.
48. Wilson K and Walker J: Principles and techniques of biochemistry and molecular biology. Cambridge university press; 2010.
49. Castle JD: Purification of organelles from mammalian cells. Current Protocols in Immunology 2003; 56(1): 8-11.
50. Edelstein C, Pfaffinger D and Scanu AM: Advantages and limitations of density gradient ultracentrifugation in the fractionation of human serum lipoproteins: role of salts and sucrose. Journal of Lipid Research 1984; 25(6): 630-37.
51. Beeckmans S: Chromatographic methods to study protein-protein interactions. Methods 1999; 19(2): 278-05.
52. Beaven G and Holiday E: Ultraviolet absorption spectra of proteins and amino acids. Advances in Protein Chemistry 1952; 319-86.
53. Pfleiderer G: For binding and activation of nicotinamide adenine dinucleotide by dehydrogenases. Mechanisms of enzymatic reactions. 1964; 300-17.
54. Hummel JP and Dreyer W: Measurement of protein-binding phenomena by gel filtration. Biochimica et Biophysica Acta 1962; 63(3): 530-2.
55. McMenamy RH and Oncley J: The specific binding of L-tryptophan to serum albumin. Journal of Biological Chemistry 1958; 233(6): 1436-47.
56. McMenamy RH and Seder RH: Thermodynamic values related to the association of L-tryptophan analogs to human serum albumin. J Biol Chem 1963; 238: 3241-8.
57. Sümegi B and Alkonyi I: A study on the physical interaction between the pyruvate dehydrogenase complex and citrate synthase. Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology 1983; 749(2): 163-71.
58. Hammerschmidt S, Rohde M and Preissner K: Extracellular matrix interactions with gram-positive pathogens. Microbiology Spectrum 2019; 7(2).
59. Fahien LA and Smith SE: The enzyme-enzyme complex of transaminase and glutamate dehydrogenase. Journal of Biological Chemistry 1974; 249(9): 2696-03.
60. Porpaczy Z, Sumegi B and Alkonyi I: Interaction between NAD-dependent isocitrate dehydrogenase, alpha-ketoglutarate dehydrogenase complex, and NADH: ubiquinone oxidoreductase. Journal of Biological Chemistry 1987; 262(20): 9509-14.
61. Smith GP: Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 1985; 228(4705): 1315-7.
62. Kügler J, Zantow J, Meyer T and Hust M: Oligopeptide m13 phage display in pathogen research. Viruses 2013; 5(10): 2531-45.
63. Mead DA and Kemper B: Chimeric single-stranded DNA phage-plasmid cloning vectors. Vectors 1988; 85-102.
64. Saeed AF, Wang R, Ling S and Wang S: Antibody engineering for pursuing a healthier future. Frontiers in Microbiology 2017; 8: 495.
65. Chasteen L, Ayriss J, Pavlik P and Bradbury A: Eliminating helper phage from phage display. Nucleic Acids Research 2006; 34(21): e145–e145.
66. Kehoe JW and Kay BK: Filamentous phage display in the new millennium. Chemical Rev 2005; 105(11): 4056-72.
67. Krumpe LR, Atkinson AJ, Smythers GW, Kandel A, Schumacher KM and McMahon JB: T7 lytic phage-displayed peptide libraries exhibit less sequence bias than M13 filamentous phage-displayed peptide libraries. Proteomics 2006; 6(15): 4210-22.
68. Arap MA: Phage display technology: applications and innovations. Genetics and Molecular Biology 2005; 28(1): 1-9.
69. Clackson T and Lowman HB: Phage display: a practical approach. Oxford University Press, USA; 2004.
70. Fagerlund A, Myrset AH and Kulseth MA: Construction and characterization of a 9-mer phage display pVIII-library with regulated peptide density. Applied Microbiology and Biotechnology 2008; 80(5): 925-36.
71. Sidhu SS, Weiss GA and Wells JA: High copy display of large proteins on phage for functional selections. Journal of Molecular Biology 2000; 296(2): 487-95.

72. Báb'\i\vcková J, Tóthová L, Boor P and Celec P: *In-vivo* phage display-A discovery tool in molecular biomedicine. Biotechnology Advances 2013; 31(8): 1247-59.
73. Fuh G and Sidhu SS: Efficient phage display of polypeptides fused to the carboxy-terminus of the M13 gene-3 minor coat protein. FEBS Letters 2000; 480(2-3): 231-4.
74. Gao J, Wang Y, Liu Z and Wang Z: Phage display and its application in vaccine design. Annals of Microbiology 2010; 60(1): 13-9.
75. Danner S and Belasco JG: T7 phage display: a novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. Proceedings of the National Academy of Sciences 2001; 98(23): 12954-9.
76. Bazan J, Ca\lkosi'nski I and Gamian A: Phage display-A powerful technique for immunotherapy: 1. Introduction and potential of therapeutic applications. Human Vaccines \& Immunotherapeutics 2012; 8(12): 1817-28.
77. Weiss GA and Sidhu SS: Design and evolution of artificial M13 coat proteins. J of Mol Biol 2000; 300(1): 213-9.
78. Day JW, Kim CH, Smider VV and Schultz PG: Identification of metal ion binding peptides containing unnatural amino acids by phage display. Bioorganic & Medicinal Chemistry Letters 2013; 23(9): 2598-00.
79. LaFleur D, Abramyan D, Kanakaraj P, Smith R, Shah R and Wang G: Monoclonal antibody therapeutics with up to five specificities: functional enhancement through the fusion of target-specific peptides. MAbs 2013; 208-18.
80. Hammers CM and Stanley JR: Antibody phage display: technique and applications. The Journal of Investigative Dermatology 2014; 134(2): e17.
81. Schägger H and von Jagow G: Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. Analytical Biochemistry 1991; 199(2): 223-31.
82. Davis BJ: Disc electrophoresis-II method and application to human serum proteins. Annals of the New York Academy of Sciences 1964; 121(2): 404–27.
83. Wittig I, Braun H-P and Schägger H: Blue native PAGE. Nature Protocols 2006; 1(1): 418.
84. Heuberger EH, Veenhoff LM, Duurkens RH, Friesen RH and Poolman B: Oligomeric state of membrane transport proteins analyzed with blue native electrophoresis and analytical ultracentrifugation. Journal of Molecular Biology 2002; 317(4): 591-00.
85. Krause F: Detection and analysis of protein-protein interactions in organellar and prokaryotic proteomes by native gel electrophoresis: (membrane) protein complexes and supercomplexes. Electrophoresis 2006; 27(13): 2759-81.
86. Arndt C, Koristka S, Bartsch H and Bachmann M: Native polyacrylamide gels. Protein Electrophoresis 2012; 49-53.
87. Darie CC: Investigation of protein-protein interactions by blue native-PAGE \ & Mass Spectrometry. Mod Chem Appl 2013; 1(3): 111.
88. Tang X and Bruce JE: A new cross-linking strategy: protein interaction reporter (PIR) technology for protein-protein interaction studies. Molecular Biosystems 2010; 6(6): 939-47.
89. Tang X and Bruce JE: Chemical cross-linking for protein-protein interaction studies. Mass Spectrometry of Proteins and Peptides 2009; 283-93.
90. Cheng PC: The contrast formation in optical microscopy. Handbook of Biological Confocal Microscopy 2006: 162-206.
91. Fancy DA and Kodadek T: Chemistry for the analysis of protein-protein interactions: rapid and efficient cross-linking triggered by long wavelength light. Proceedings of the National Academy of Sciences 1999; 96(11): 6020-4.
92. Sinz A: Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. Mass Spectrometry Reviews 2006; 25(4): 663-82.
93. Petrotchenko EV, Olkhovik VK and Borchers CH: Isotopically coded cleavable cross-linker for studying protein-protein interaction and protein complexes. Molecular \& Cellular Proteomics 2005; 4(8): 1167-79.
94. Vasilescu J, Guo X and Kast J: Identification of protein-protein interactions using *in-vivo* cross-linking and mass spectrometry. Proteomics 2004; 4(12): 3845-54.
95. Forster T: Intermolecular energy migration & fluorescence (Ger.). Annalen der Physik 1948; 2: 55-75.
96. Hussain SA, Dey D, Chakraborty S, Saha J, Roy AD and Chakraborty S: Fluorescence resonance energy transfer (FRET) sensor. arXiv preprint arXiv:14086559. 2014;
97. Zheng J: Spectroscopy-based quantitative fluorescence resonance energy transfer analysis. Ion Chan 2006; 65-77.
98. Hussain SA: An introduction to fluorescence resonance energy transfer (FRET). arXiv preprint arXiv:09081815. 2009;
99. Valeur B: Molecular fluorescence. Digital Encyclopedia of Applied Physics 2003; 477-31.
100. Jonsson T, Waldburger CD and Sauer RT: Nonlinear free energy relationships in Arc repressor unfolding imply the existence of unstable, native-like folding intermediates. Biochemistry 1996; 35(15): 4795-02.
101. Berger W, Prinz H, Striessnig J, Kang HC, Haugland R and Glossmann H: Complex molecular mechanism for dihydropyridine binding to L-type Ca^{2+}- channels as revealed by fluorescence resonance energy transfer. Biochemistry 1994; 33(39): 11875-83.
102. Khanna PL and Ullman EF: 4′, 5′-Dimethoxy-6-carboxy-fluorescein: a novel dipole-dipole coupled fluorescence energy transfer acceptor useful for fluorescence immunoassays. Analytical Biochemistry 1980; 108(1): 156-61.
103. Myakishev MV, Khripin Y, Hu S and Hamer DH: High-throughput SNP genotyping by allele-specific PCR with universal energy - transfer - labeled primers. Genome Research 2001; 11(1): 163-9.
104. Parkhurst KM and Parkhurst LJ: Kinetic studies by fluorescence resonance energy transfer employing a double-labeled oligonucleotide: hybridization to the oligonucleotide complement and to single-stranded DNA. Biochemistry 1995; 34(1): 285-92.
105. Nichols JW and Pagano RE: Resonance energy transfer assay of protein-mediated lipid transfer between vesicles. Journal of Biological Chemistry 1983; 258(9): 5368-71.
106. Gonzalez JE and Tsien RY: Voltage sensing by fluorescence resonance energy transfer in single cells. Biophysical Journal 1995; 69(4): 1272-80.
107. Matayoshi ED, Wang GT, Krafft GA and Erickson J: Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer. Science 1990; 247(4945): 954-8.
108. Mason WT: Fluorescent and luminescent probes for biological activity: a practical guide to technology for quantitative real-time analysis. Elsevier 1999.
109. Young K: Yeast two-hybrid: so many interactions, (in) so little time. Biology of Reproduction 1998; 58(2): 302-11.
110. Fields S: Interactive learning: lessons from two hybrids over two decades. Proteomics 2009; 9(23): 5209-13.
111. Joung JK, Ramm EI and Pabo CO: A bacterial two-hybrid selection system for studying protein-DNA and protein-

protein interactions. Proceedings of the National Academy of Sciences 2000; 97(13): 7382-7.
112. Whipple FW: Genetic analysis of prokaryotic and eukaryotic DNA-binding proteins in *Escherichia coli*. Nucleic Acids Research 1998; 26(16): 3700-6.
113. Stagljar I, Korostensky C, Johnsson N and te Heesen S: A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. Proceedings of the National Academy of Sciences 1998; 95(9): 5187-92.
114. Snider J, Kittanakom S, Curak J and Stagljar I: Split-ubiquitin based membrane yeast two-hybrid (MYTH) system: a powerful tool for identifying protein-protein interactions. J of Visualized Experiments 2010; 36: e1698.
115. Zolghadr K, Mortusewicz O, Rothbauer U, Kleinhans R, Goehler H and Wanker EE: A fluorescent two-hybrid assay for direct visualization of protein interactions in living cells. Molecular \ & Cellular Proteomics 2008; 7(11): 2279-87.
116. Lievens S, Gerlo S, Lemmens I, De Clercq DJ, Risseeuw MD and Vanderroost N: Kinase Substrate Sensor (KISS), a mammalian in situ protein interaction sensor. Molecular \ & Cellular Proteomics 2014; 13(12): 3332–42.
117. Serebriiskii IG and Kotova E: Analysis of protein-protein interactions utilizing dual bait yeast two-hybrid system. Protein-Protein Interactions 2004; 263-96.
118. Hamdi A and Colas P: Yeast two-hybrid methods and their applications in drug discovery. Trends in Pharmacological Sciences 2012; 33(2): 109-18.
119. Hurt JA, Thibodeau SA, Hirsh AS, Pabo CO and Joung JK: Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection. Proceedings of the National Academy of Sciences 2003; 100(21): 12271-6.

How to cite this article:
Joshi S, Gupta P and Gopal G: Different approaches for studying protein-protein interaction and analysis. Int J Pharm Sci & Res 2019; 10(9): 4006-25. doi: 10.13040/IJPSR.0975-8232.10(9).4006-25.