Protein-protein interaction

By: Dr. Ashish C Patel
Assistant Professor
Vet College, AAU, Anand

- **Protein–protein interactions** occur when two or more proteins bind together.
- Proteins control and mediate many of the biological activities of cells by these interactions.
- Information about PPIs → improves our understanding of diseases and can provide the basis for new therapeutic approaches.
- The structure of a protein influences its function by determining the other molecules with which it can interact and the consequences of those interactions.
- Multi subunit protein
  - hemoglobin, core RNA polymerase, small nuclear ribonucleoproteins and the ribosome etc..

Transient protein interaction: strong and irreversible, it readily undergoes changes in the oligomeric state.

- Interactions of protein kinases, protein phosphatases, glycosyl transferases, acyl transferases, proteases, etc., with their substrate proteins.
  - Proteins for Cell growth, cell cycle, metabolic pathways, and signal transduction
- These interactions are very important in our lives as any disorder in them can lead to fatal diseases such as Alzheimer's and Creutzfeld-Jacob Disease.
- Perhaps the most well known example of Protein-Protein Interaction is between Actin and Myosin while regulating Muscular contraction in our body.

Types of Protein-Protein Interactions

On the basis of their Composition

Homo-Oligomers: These are macromolecule complexes having one type of protein subunits.

e.g. : PPIs in Muscle Contraction

- Hetero-Oligomers: These are macromolecule complexes having multiple types protein subunits.

e.g. : PPI between Cytochrome Oxidase and TRPC3 (Transient receptor potential cation channels

On the basis of duration

- Stable Interactions: These comprise of interactions that last for a long duration. These Interactions carry out Functional or Structural roles.

e.g.: Haemoglobin structure

- Transient Interactions : Interactions that last a short period of time.

  e.g.: Muscle Contraction

Effects of protein interaction

- They can alter the kinetic properties of proteins
- Protein interactions are one common mechanism to allow for substrate channeling.
- Can result in the formation of a new binding site
- Can inactivate a protein
- Can change the specificity of a protein for its substrate

- Identify the different interactions, understand the extent to which they take place in the cell, and determine the consequences of the interaction

PPIs Identification Methods

Experimental ***(In vivo)***	• **Yeast two-hybrid system**
Experimental ***(In vitro)***	• **Co-immunoprecipitation** • Tagged Fusion Proteins • X-ray Diffraction • Phage display
Computational ***(In silico)***	• BIND: Biomolecular Interaction Network Database • DIP: Database of Interacting Proteins • MINT • IntAct

Another way of classification for methods for identification of PPIs

- The first is ‘atomic observation’ in which the protein interaction detected using, for example, X-ray crystallography. These experiments can yield specific information on the atoms or residues involved in the interaction.

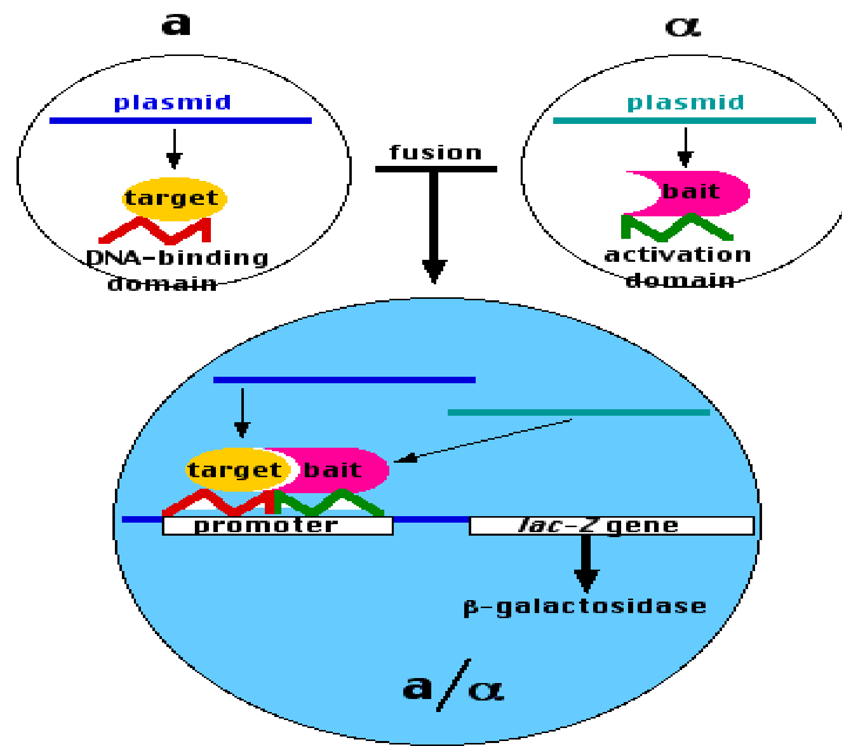
- The second is a ‘direct interaction observation’ where protein interaction between two partners can be detected as in a two-hybrid experiment.

- At a third level of observation, multi-protein complexes can be detected using methods such as immuno-precipitation or mass-specific analysis. This type of experiment does not reveal the chemical detail of the interactions or even reveal which proteins are in direct contact but gives information as to which proteins are found in a complex at a given time.

YEAST TWO HYBRID SYSTEM

- This is method that uses transcriptional activity as a measure of protein-protein interaction
- It relies on the modular nature of many site-specific transcriptional activators, which consist of a DNA-binding domain and a transcriptional activation domain
- DNA-binding domain serves to target the activator to the specific genes that will be expressed
- Activation domain contacts other proteins of the transcriptional machinery to enable transcription to occur

- yeast two-hybrid (Y2H) system has variations involving different reagents and has been adapted to high-throughput screening.
- The strategy interrogates two proteins, called bait and prey, coupled to two halves of a transcription factor and expressed in yeast.
- If the proteins make contact, they reconstitute a transcription factor that activates a reporter gene.



Animation

- Used to detect interactions between candidate proteins whose genes are available by constructing the appropriate hybrids and testing for reporter gene activity
- Point mutations can be assayed to identify specific amino acid residues critical for the interaction
- Can be used to screen libraries of activation domain hybrids to identify proteins that bind to a protein of interest
- Transcriptional activation domains are commonly derived from the Gal4 protein or the herpes simplex virus VP16 protein
- Reporter genes include the *E. coli lacZ* gene and selectable yeast genes such as *HIS3* and *LEU2*

Uses/ Advantages of Yeast Two Hybrid Method

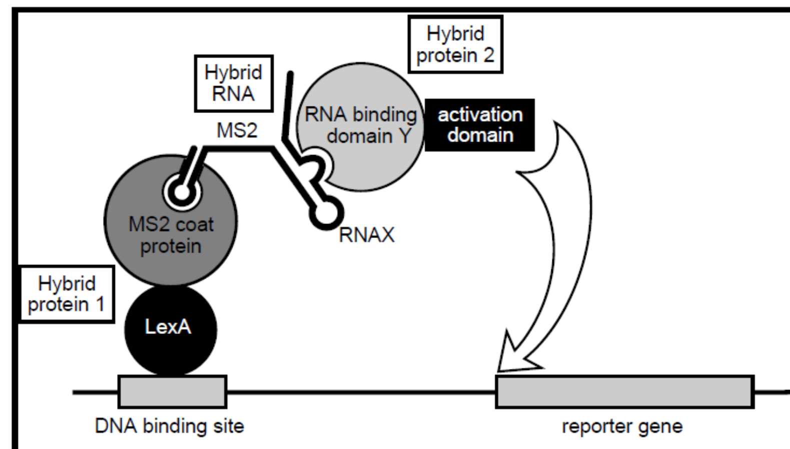
- It is highly sensitive, detecting interactions that are not detected by other methods
- Minimal binding constant required to detect an interaction in two hybrid system is on the order of 1 µM
- Interactions are detected within the native environment of the cell and hence that no biochemical purification is required
- Study of oncogenes and tumor suppressors and the related area of cell cycle control

Limitations

- limited to proteins that can be localized to the nucleus, which may prevent its use with certain extracellular proteins
- Proteins must be able to fold and exist stably in yeast cells and to retain activity as fusion proteins
- Interactions dependent on a posttranslational modification that does not occur in yeast cells will not be detected
- Many proteins, including those not normally involved in transcription, will activate transcription when fused to a DNA-binding domain

Yeast three hybrid

- The three-hybrid system enables the detection of RNA-protein interactions in yeast using simple phenotypic assays.
- It was developed in collaboration with Stan Fields laboratory (University of Washington).
- The LexA DNA binding domain is fused to the MS2 coat protein to form Hybrid Protein 1.
- Hybrid Protein 2 consists of the Gal4 activation domain linked to the RNA binding domain, Y, you wish to test.
- The Hybrid RNA consists of two MS2 RNA binding sites and the RNA sequence you wish to test, RNAX.
- Hybrid Protein 1 and the presence of MS2 sites in the Hybrid RNA are fixed, as is the Gal4 Activation Domain. RNAX and RNA Binding Domain Y vary.



- This method consists in the expression in yeast cells of three chimerical molecules, which assemble in order to activate two reporter genes.
- Thus, using the yeast three-hybrid system, in contrast to other methods, RNA–protein interactions are detected in vivo.
- This system uses a transactivator protein in yeast, such as Gal4p, that is able to recruit the transcriptional machinery and trigger transcription of a gene.
- It consists of a DNA binding domain (DB) and an activation domain (AD) and, importantly, these two domains are functionally independent, meaning that they can be inserted into other molecules.
- In this method, the three components of the system are expressed from two plasmids allowing the use of any previously described yeast strains for two-hybrid system that provide the two reporter genes *HIS3* and *lacZ* under the control of a Gal4 operator.

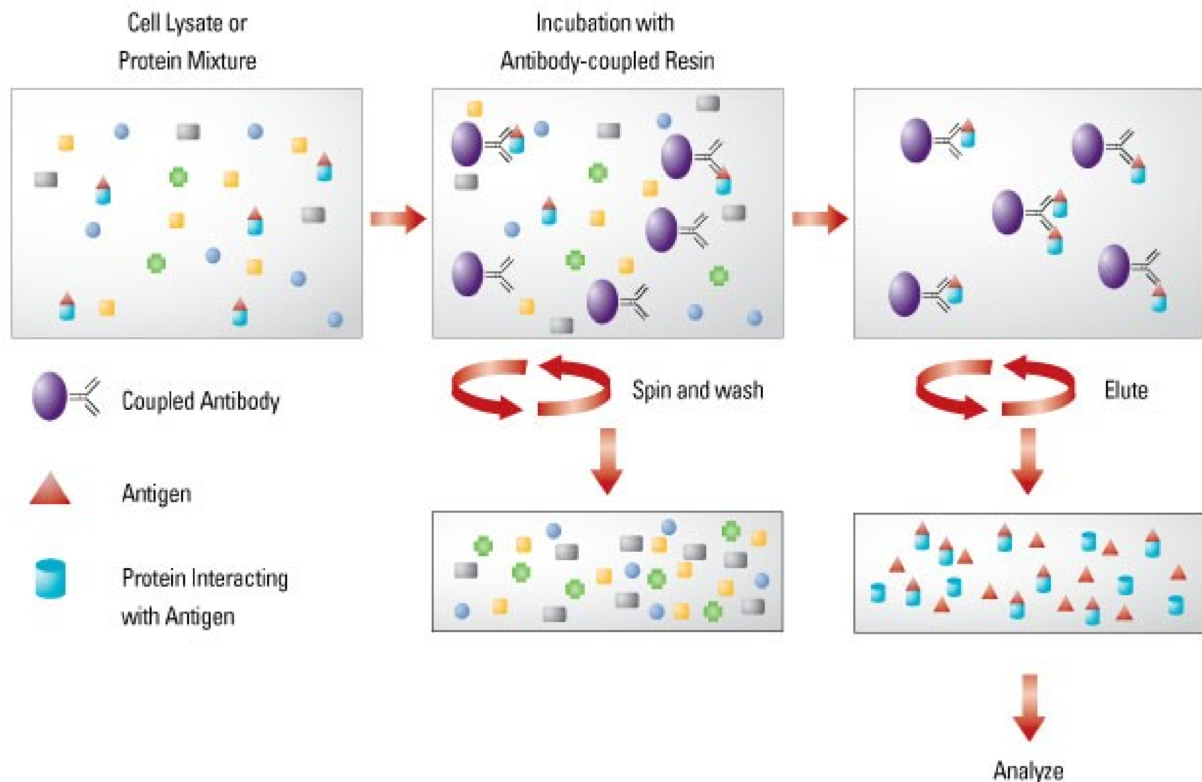
Immunoprecipitation

- This is classical method of detecting protein-protein interactions
- Cell lysates are generated, antibody is added, the antigen is precipitated and washed, and bound proteins are eluted and analyzed
- Coprecipitated protein is precipitated by the antibody itself and not by a contaminating antibody in the preparation
- If the interaction is direct through another protein that contacts both the antigen and the coprecipitated protein determining that the interaction takes place in the cell.
- Adenovirus E1A protein interacts with Rb protein detects the interactions present in a crude lysate
- Both the antigen and the interacting proteins are present in the same relative concentrations as found in the cell.
- Co-immunoprecipitating proteins do not necessarily interact directly, since they can be part of larger complexes
- Co-precipitation is not as sensitive as other methods, such as protein affinity chromatography, because the concentration of the antigen is lower than it is in protein affinity chromatography.

CO-IMMUNOPRECIPITATION (coIP)

- Co-immunoprecipitation (coIP) is the most complex method.
- Co-immunoprecipitation (co-IP) is a popular technique for protein interaction discovery. Co-IP is conducted in essentially the same manner as an immunoprecipitation (IP) of a single protein, except that the target protein precipitated by the antibody, also called the "bait", is used to co-precipitate a binding partner/protein complex, or "prey", from a lysate.

Animation

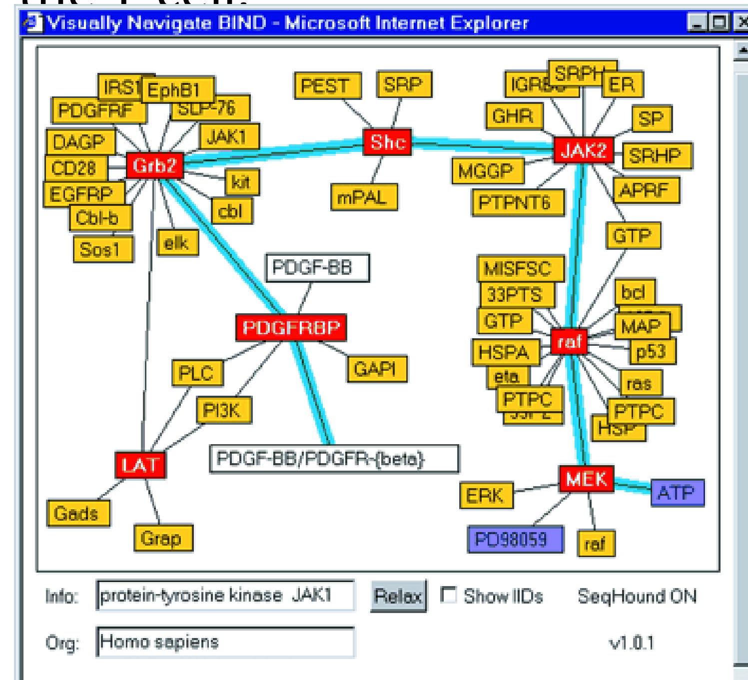


DATABASES

- Protein–protein interactions are only the raw material for networks. To build a network, researchers typically combine interaction data sets with other sources of data. Primary databases that contain protein–protein interactions include DIP (http://dip.doe-mbi.ucla.edu), BioGRID, IntAct (http://www.ebi.ac.uk/intact) and MINT (http://mint.bio.uniroma2.it).
- These databases have committed to making records available through a common language called PSICQUIC, to maximize access.

BIND: (Biomolecular Interaction Network Database)

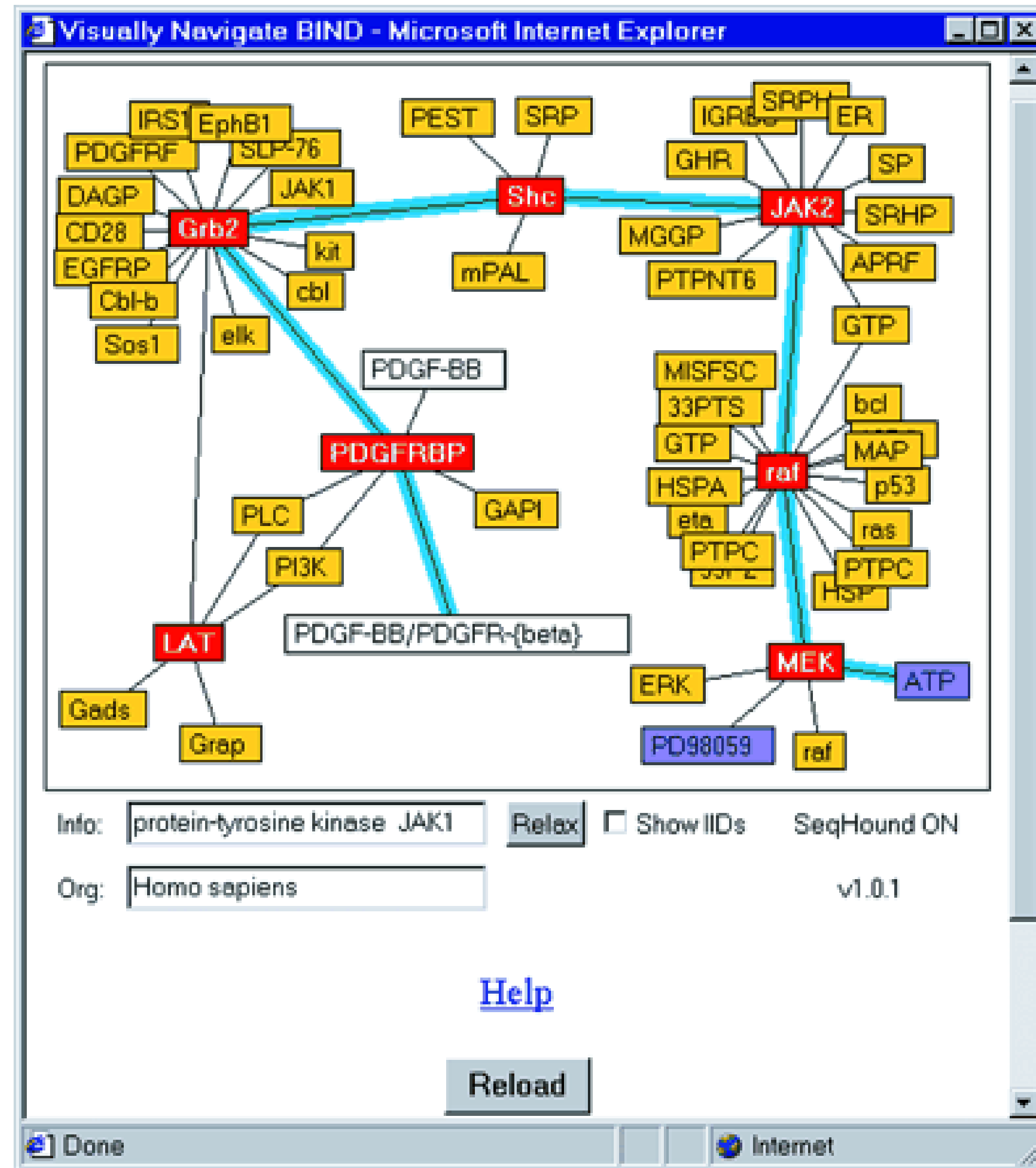
- A free, open-source database for archiving and exchanging molecular assembly information. The database contains Interactions, Molecular complexes and Pathways
- BIND Interaction Viewer Java showing how molecules can be connected in the database from molecular complex to small molecule. Yellow: protein, Purple: small molecule; white: molecular complex;
- This session was seeded by the interaction between human LAT and Grb2 proteins involved in cell signaling in the T-cell.



- BIND Interaction Viewer Java showing how molecules can be connected in the database from molecular complex to small molecule.

- Yellow: protein
- Purple: small molecule;
- white: molecular complex;

- This session was seeded by the interaction between human LAT and Grb2 proteins involved in cell signaling in the T-cell.



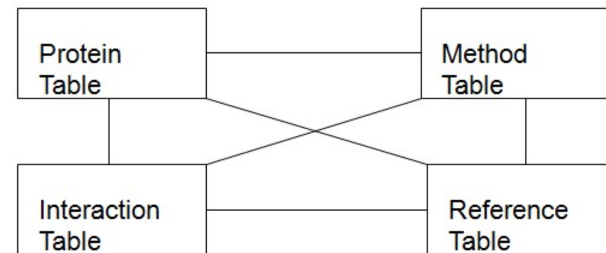
DIP:(Database of Interacting Proteins)

Study: Protein function, Protein-protein relationship, Evolution of protein-protein interaction, The network of interacting proteins, The environments of protein-protein interactions,

Predict: Unknown protein-protein interaction, The best interaction conditions

- Identification numbers from :

  SWISS-Prot, GenBank, PIR



DIP NODE

DIP 1143N		PIR DEBYA	SwissProt ADH1_YEAST	GenBank gi:2144315
Adh1p	Name/Description	alcohol dehydrogenase 1		
CrossRef	G	SGD: ADH1	MIPS: YOL086C	COMPLUYEAST-2DPAGE: P00330
	P	SWISS-2DPAGE: P00330; NCBI: V01291, Z74828, X83121, M38456, V01292	SGD: S0005446	PIR: A92365
	D	Pfam: PF00107		InterPro: IPR002085, IPR002328
	F	PROSITE: PS00059		
Organism		Saccharomyces cerevisiae	Localization	
EC			Function	
Keywords		zinc;NAD;oxidoreductase;acetylated amino end;metalloprotein;alcohol metabolism		

DIP LINK

CORE DIP 15354E				
DIP 1143N	PIR DEBYA	SwissProt ADH1_YEAST		GenBank gi:214431
	Name/Description	alcohol dehydrogenase 1		
DIP 6603N	PIR S67208	SwissProt Q08773		GenBank gi:142067
	Name/Description	hypothetical protein YOR304w		
Evidence	Type	Method	Details	Source
	E	Tandem Affinity Purification (TAP) [GS]	---	PMID:11805826
	V	PVMS(2)	---	---

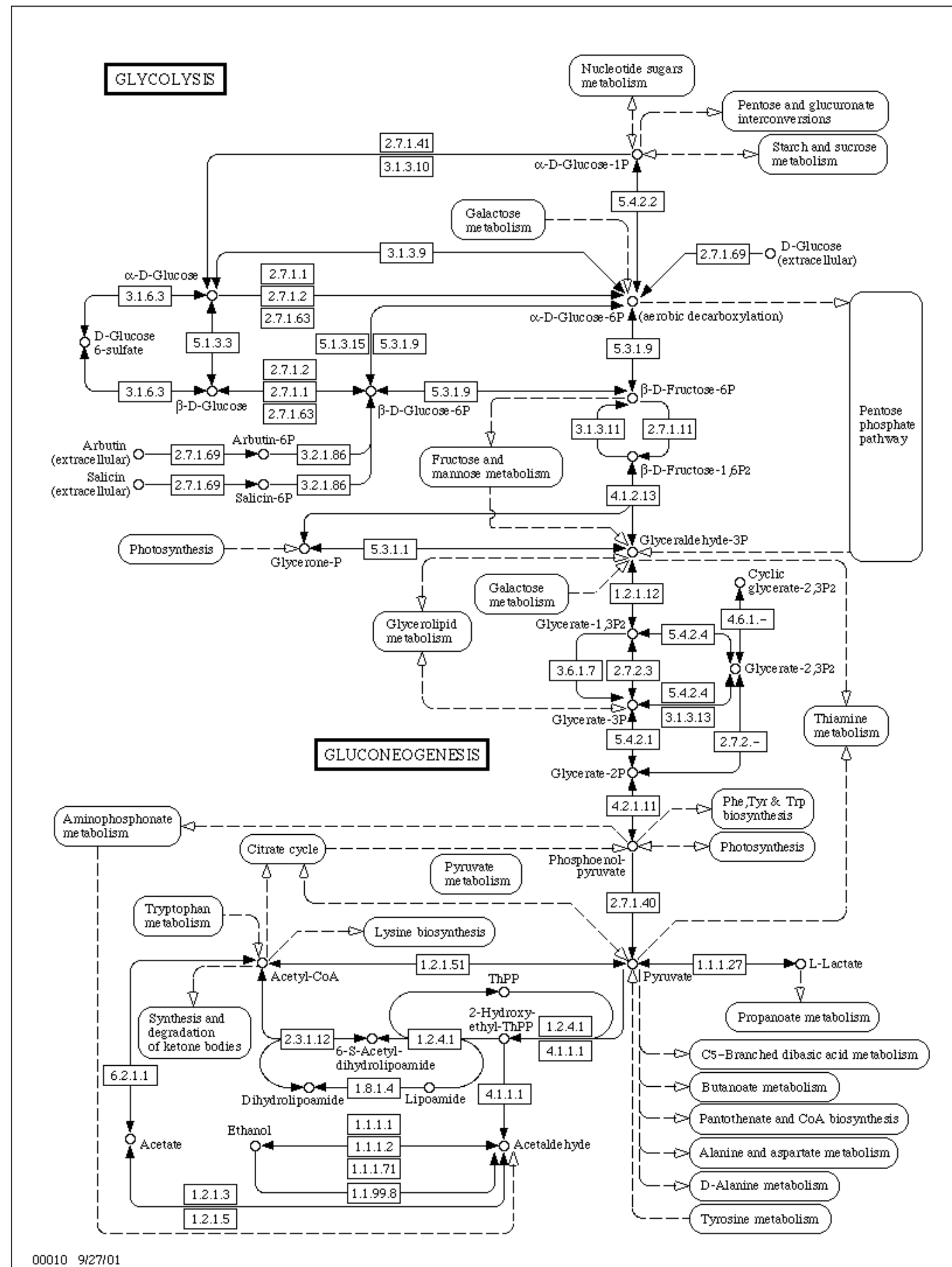
The current status of DIP

- Number of proteins: 6978
- Number of organisms: 101
- Number of interactions:18260
- Number of distinct experiments describing an interaction: 22229
- Number of articles: 2203

	Data Stored	Data Format
BIND	▪ interactions ▪ Molecular Complex ▪ Pathways	▪ ASN.1 ▪ XML
DIP	▪ interactions ▪ Protein information	▪ XIN ▪ tab-delimited

Pathway Databases and Algorithms

- 1) KEGG (Kyoto Encyclopedia of Genes and Genomes): Representation of higher order functions in terms of the network of interaction molecules
- GENES database contains 240 943 entries from the published genomes, including the bacteria, mouse and human.
- KEGG has 3 databases, GENES, PATHWAY and LIGAND databases.
- By matching genes in the genome and gene products in the pathway, KEGG can be used to predict protein interaction networks and associated cellular function.
- KEGG is a network of gene products with three types of interactions or relations: enzyme-enzyme relations which catalyzes the successive reaction steps in the metabolic pathway, direct protein-protein interactions and gene expression relations.
- PATHWAY database contains 5761 entries including 201 pathway diagrams with 14,960 enzyme-enzyme relations.



An example of a pathway entry in KEGG- Glycolysis

Other databases..

- MINT

  http://mint.bio.uniroma2.it/mint
- DLRP (Database of Ligand-Receptor Partners)

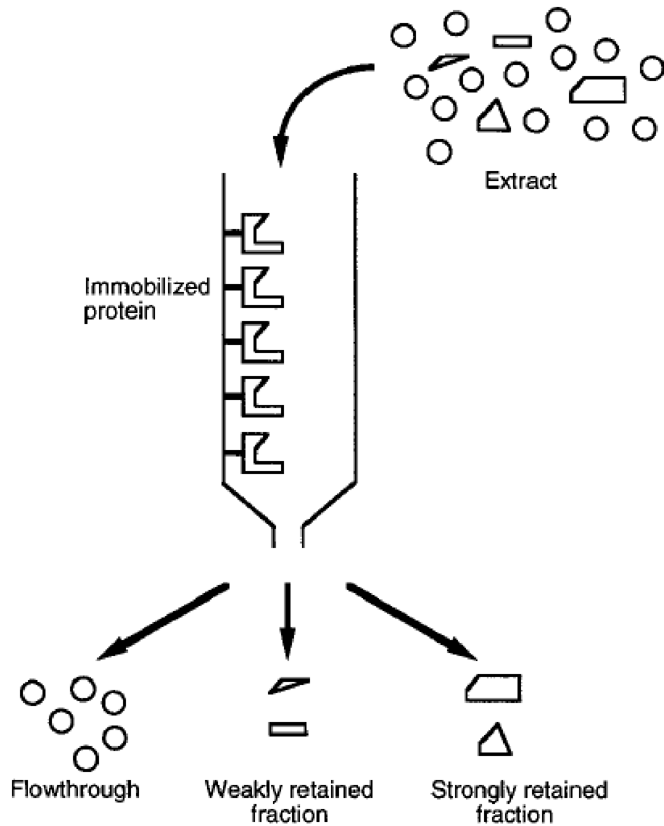
  http://dip.doe-mbi.ucla.edu/dip/DLRP.cgi
- STRING

  http://string-db.org
- Human Protein Interaction Database

  http://www.hpid.org

Protein Affinity Chromatography

- **Affinity Chromatography** is essentially a sample purification technique, used primarily for biological molecules such as proteins.
- It is a method of separating a mixture of proteins or nucleic acids (molecules) by specific interactions of those molecules with a component known as a ligand, which is immobilized on a support.
- If a mixture of proteins is passed over (through) the column, one of the proteins binds to the ligand on the basis of specificity and high affinity (they fit together like a lock and key).
- The other proteins in the solution wash through the column because they were not able to bind to the ligand.
- Protein fusions
  - Glutathione *S*-transferase
  - Staphylococcus protein A
  - Maltose-binding protein



- The interactions of many proteins with their target proteins often depends on the modification state of one or both of the proteins

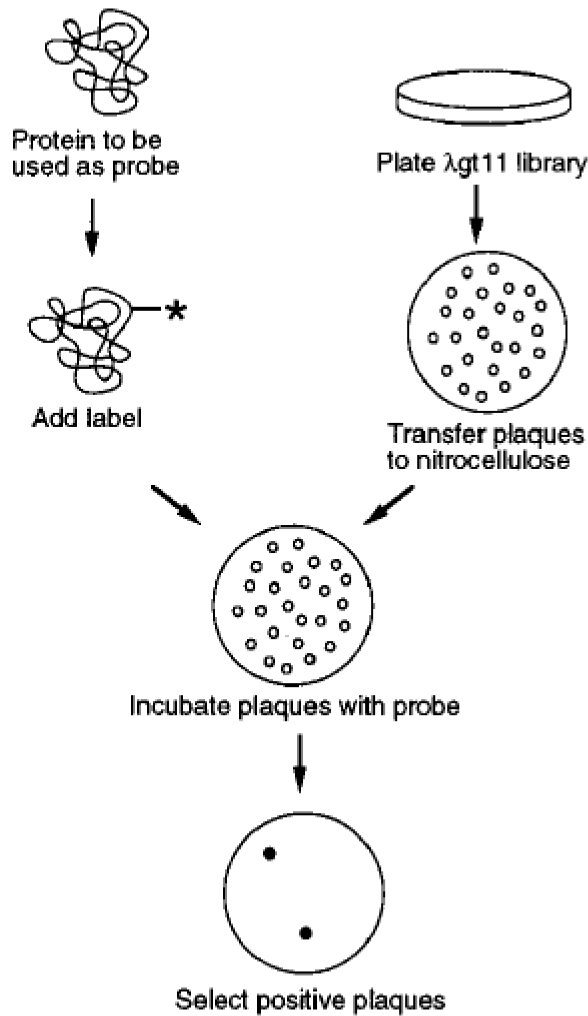
- Incredibly sensitive
  - Can detect interactions with a binding constant as weak as 10^{-5} M
- Tests all proteins in an extract equally
- Easy to examine both the domains of a protein and the critical residues within it that are responsible for a specific interaction, by preparing mutant derivatives
- Interactions that depend on a multi subunit tethered protein can be detected
- Independent criteria must be used to establish that the interaction is physiologically relevant

Affinity Blotting

- Analogous to the use of affinity columns, proteins can be fractionated by PAGE transferred to a nitrocellulose membrane, and identified by their ability to bind a protein, peptide, or other ligand
- Complex mixtures of proteins, such as total-cell lysates, can be analyzed without any purification
- Biological activity of the proteins on the membrane, the preparation of the protein probe, and the method of detection widely used in studies of the regulatory subunit of the type II cAMP-dependent protein kinase with numerous specific anchoring proteins
- Two-dimensional procedures of isoelectric focusing followed by SDS-PAGE have been used to increase the separation of proteins

Protein Probing

- A labeled protein can be used as a probe to screen an expression library in order to identify genes encoding proteins that interact with this probe.
- Interactions occur on nitrocellulose filters between an immobilized protein and the labeled probe protein.
- The method is highly general and therefore widely applicable, in that proteins as diverse as transcription factors and growth factor receptors have been used as probe.
- The method is based on the approach of Young and Davis, who showed that an antibody can be used to screen expression libraries to identify a gene encoding a protein antigen



- The λgt11 libraries typically use an isopropyl-b-D-thiogalactopyranoside (IPTG)-inducible promoter to express proteins fused to b-galactosidase.

- Proteins from the bacteriophage plaques are transferred to nitrocellulose filters, incubated with antibody, and washed to remove non specifically bound antibody. Protein probe can be manipulated in vitro to provide, a specific posttranslational modification or a metal cofactor

- Any protein or protein domain can be specifically labeled for use as a probe

Chromatin immunoprecipitation (ChIP)

- Chromatin immunoprecipitation is a common technique for studying epigenetics, as it allows the researcher to capture a snapshot of specific protein–DNA interactions.
- ChIP include crosslinking the DNA and the protein in live cells, then extracting and shearing the chromatin.
- Finally, the samples are immunoprecipitated with an antibody targeting the protein of interest. The DNA is extracted from the protein and can be evaluated either at specific regions of the genome by quantitative PCR (qPCR)

Choosing an antibody for ChIP

- Monoclonal, oligoclonal (pools of monoclonals) and polyclonal antibodies all can work in ChIP.
- monoclonal antibody is that generally it is more specific but oligoclonal and polyclonal antibodies are better candidates for recognizing target proteins, as they recognize multiple epitopes of the targets.

- There are mainly two types of ChIP, primarily differing in the starting chromatin preparation. The first uses reversibly cross linked chromatin sheared by sonication called cross-linked ChIP (XChIP). Native ChIP (NChIP) uses native chromatin sheared by micrococcal nuclease digestion.

Cross-linked ChIP (XChIP)

- Cross-linked ChIP is mainly suited for mapping the DNA target of transcription factors or other chromatin-associated proteins, and uses reversibly cross linked chromatin as starting material.
- The agent for reversible cross-linking could be formaldehyde or UV light.
- The DNA associated with the complex is then purified and identified by PCR, microarray, molecular cloning and sequencing, or direct high-throughput sequencing (ChIP-seq).

Native ChIP (NChIP)

- Native ChIP is mainly suited for mapping the DNA target of Histone modifiers. Generally, native chromatin is used as starting chromatin.
- Then the chromatin is sheared by micrococcal nuclease digestion, which cuts DNA at the length of the linker, leaving nucleosomes intact and providing DNA fragments of one nucleosome (200bp) to five nucleosomes (1000bp) in length.

- Controls are essential for ChIP so, we need a “no-antibody control” (mock IP) for each IP we are doing.
- To Know a region of DNA that should be enriched in IP and be amplified by qPCR or Not??

Step 1: Crosslinking

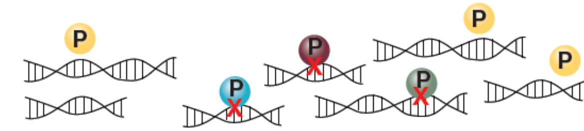
- ChIP assays begin with covalent stabilization of the protein–DNA complexes.
- As there is constant movement of proteins and DNA, ChIP captures a snapshot of the protein–DNA complexes that exist at a specific time.
- In vivo crosslinking covalently stabilizes protein–DNA complexes and performed using a formaldehyde solution.

Step 2: Cell lysis: In this step, cell membranes are dissolved with detergent based lysis solutions to liberate cellular components, and crosslinked protein–DNA complexes are solubilized.

- Because protein–DNA interactions occur primarily in the nuclear compartment, removing cytosolic proteins can help reduce background signal and increase sensitivity. Protease and phosphatase inhibitors are essential at this stage to maintain intact protein–DNA complexes.

Step 3: Chromatin preparation (shearing/digestion): The extraction step yields all nuclear material, which includes unbound nuclear proteins, full-length chromatin, and the crosslinked protein–DNA complexes.

- DNA fragmentation is by sonication or enzymatically by digestion with micrococcal nuclease.
- Ideal chromatin fragment sizes range from 200 to >700 bp; however, DNA shearing is one of the most difficult steps to control. Sonication provides truly randomized fragments, but limitation is Difficulty in maintaining temperature during sonication.



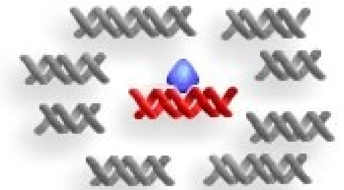
Step 4: Immunoprecipitation

- To isolate a specifically modified histone, transcription factor, or cofactor of interest, ChIP-validated antibodies are used to immunoprecipitate and isolate the target from other nuclear components. This step eliminates all other unrelated cellular material.
- Selection of the appropriate antibody is critical for successful ChIP assays. Numerous ChIP-validated antibodies are available. For target proteins for which qualified antibodies are unavailable, proteins fused to affinity tags such as HA, Myc, His, T7, V5, or GST can be expressed in the biological samples, and then antibodies against the affinity tags can be used to immunoprecipitate the targets.

Step 5: Reversal of crosslinking, and DNA clean-up

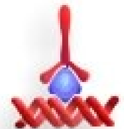
- The Reversal of crosslinking and DNA clean-up done through extensive heat incubations or digestion of the protein component with Proteinase K.

- Treatment with RNase A is recommended as well to obtain a more pure DNA sample.
- A final purification of the DNA from any remaining proteins should be performed using phenolchloroform extraction or spin columns designed for DNA purification.

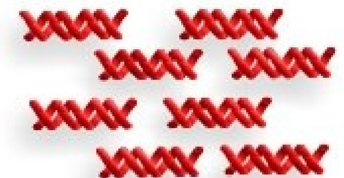


Step 6: DNA quantitation

- We can quantitate the purified DNA products by qPCR.



- qPCR enables analysis of target protein–DNA complex levels in different experimental conditions.
- There is a direct correlation between the amounts of immunoprecipitated complex and bound DNA.



Variation	Advantages	Disadvantages
X-ChIP	Used for studying proteins that are tightly bound to DNA, thus preferred for analyzing histone-chromatin interactions. Can be performed on all cell types, tissues, and organisms. Under optimal conditions, formaldehyde cross-linking can produce relatively small DNA fragments by sonication. Enables DNA-protein, RNA-protein, and protein-protein cross-linking. Reduces the chances of chromatin rearrangement during processing. Requires fewer cells than N-CHIP.	Overfixation can prevent effective sonication. Enzymatic digestion not possible following formaldehyde treatment, probably due to the destruction of the nuclease reactive sites. Formaldehyde can alter the binding properties of the antigens and thus the immunoreactivity. Formaldehyde is a potential carcinogen and extra precautions are needed during handling.
N-ChIP	High antigen specificity for the analysis of histones and their isoforms. PCR amplification of the immunoprecipitated DNA is not needed because DNA recovery is higher. Specificity of the binding is more predictable because the antisera is raised against unfixed proteins.	High concentrations of nuclease may overdigest the chromatin, leading to subnucleosomal proteins, thus hindering the detection of protein-DNA interactions. Not all the nuclease-digested proteins are solubilized. Thus, a fraction of the chromatin is retained with the nuclear pellet and eliminated from the assay. Chances of chromatin rearrangement during processing are likely.

THANK YOU