

Fundamentals of Protein Chemistry
Prof. Swagata Dasgupta
Department of Chemistry
Indian Institute of Technology, Kharagpur

Module - 11
Protein Macromolecule Interactions II
Lecture - 52
Protein Protein Interactions - II

We continue our discussion on protein-protein interactions, where we have looked at the geometric and the chemical complementarity of these protein-protein interactions in the previous lecture.

(Refer Slide Time: 00:27)

CONCEPTS COVERED

- Protein-protein binding domains
- Regulation of protein-protein interactions
- Methods to detect protein protein interactions
- PPIs and diseases

The slide features a video inset of Prof. Swagata Dasgupta in the bottom right corner. The background is light yellow with a blue and green geometric design on the right side. Logos of IIT Kharagpur and NPTEL are visible at the bottom.

In this lecture we will be looking at protein-protein domains, the regulation of these protein-protein interactions, a specific method to detect these protein interactions and how the PPIs and diseases are related.

In the identification of the domains, in the understanding of how a protein-protein interaction is going to take place, we realize that the recognition is the most important. In the previous lecture we looked at the interfaces, the specific amino acids that may occupy the interface, the loss in the surface area on the interactions and how there may be induced fit.

(Refer Slide Time: 01:12)

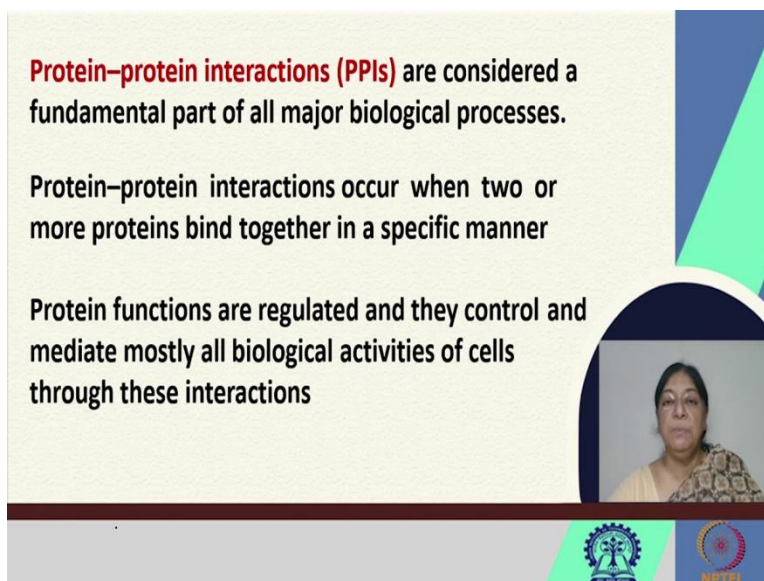


KEYWORDS

- Src-homology
- Phosphotyrosine
- Zinc-finger motif
- Databases

The slide features a light beige background with a dark blue header bar containing the word 'KEYWORDS' in white. Below the header, four bullet points are listed, each preceded by a dark blue right-pointing arrow. The text is in a black sans-serif font. In the bottom right corner, there is a circular inset video feed showing a woman with dark hair wearing a patterned sari. At the very bottom of the slide, there are two logos: the Indian Institute of Technology (IIT) logo on the left and the NPTEL logo on the right.

(Refer Slide Time: 01:14)



Protein-protein interactions (PPIs) are considered a fundamental part of all major biological processes.

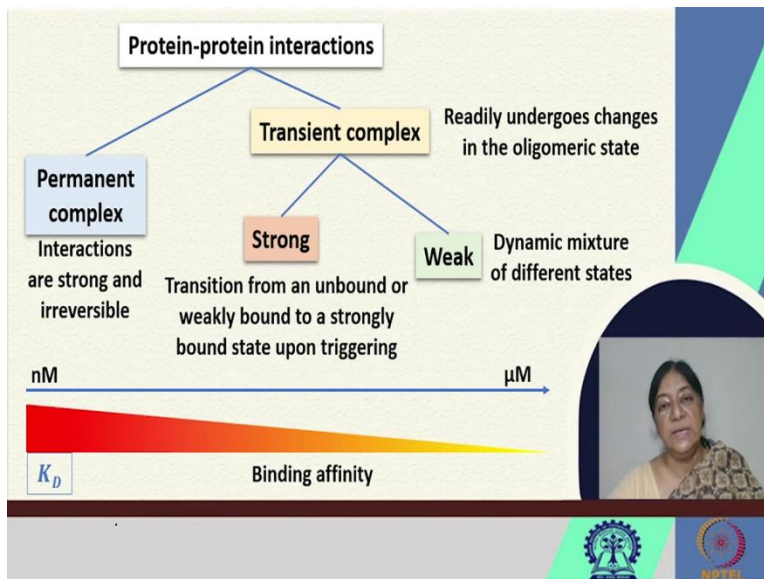
Protein-protein interactions occur when two or more proteins bind together in a specific manner

Protein functions are regulated and they control and mediate mostly all biological activities of cells through these interactions

The slide has a light beige background with a dark blue header bar. The first line of text is in bold black font, with 'Protein-protein interactions (PPIs)' in red. The subsequent two lines are in regular black font. In the bottom right corner, there is a circular inset video feed showing the same woman as in the previous slide. At the very bottom, the IIT and NPTEL logos are displayed.

When we look at the specific protein-protein interactions, we realize their importance in all biological processes and we have these interactions when two or more of them will bind in a specific manner, with high or low affinity; depending on the specific biological activity that they are involved in.

(Refer Slide Time: 01:34)



So for the protein-protein interactions we can have a permanent complex where the interactions are strong and irreversible, like in specific subunit proteins where we have the quaternary structure involved which is a permanent complex of the monomeric units.

We can have transient complex where we have either a weak type, depending upon the specific conditions, the specific reaction that has to occur and also a strong type, where we can have a transition from an unbound weakly bound to a strongly bound state that is triggering on the action of any cofactor or ATP hydrolysis. If we look at the specific dissociation constants associated with this, we see that the binding affinity for the permanent complexes are much smaller, the K_D values are smaller, indicating a tighter complex.

(Refer Slide Time: 02:38)

Protein-protein binding domains

Structural aspects of protein complexes formed reversibly and /or transiently show

diversity of interactions – with respect to interfacial size and nature of interactions

Proteins possess distinct structural domains that allow interactions with and bind to specific sequences on other proteins

A small inset video of a speaker is visible on the right side of the slide.

When we consider the protein-protein binding domains, we try to understand what kind of residues are involved in these interactions and whether that are common motifs that are involved in the recognition where we have protein-protein interactions.

The structural aspects of the protein complexes formed are reversibly or even transiently show a diversity of interactions, with respect to the interfacial size, as well as the nature of the interactions. So they possess these distinct structural domains, that allow interactions with specific sequences of other proteins.

(Refer Slide Time: 03:27)

Protein-protein binding domains

- Domains that bind sequences with phosphotyrosine
 - Src-homology 2 (SH2)
 - PTB (Phosphotyrosine binding domains)

The slide features a video inset of a woman in the bottom right corner. At the bottom, there are logos for IIT Bombay and NPTEL.

If we look at the protein-protein binding domains, there are domains that bind sequences with phosphotyrosine. So, this one example is the Src homology 2 the SH2 type or a phosphotyrosine binding domains that may be present on specific proteins.




(Refer Slide Time: 03:45)

Src homology domain is one of the two small protein binding domains found in the Src oncoprotein.

Src is a non-receptor tyrosine kinase protein which is encoded in humans by the SRC gene.

It includes an SH2 domain, an SH3 domain and a tyrosine kinase domain.

The proto-oncogene has a role in the regulation of embryonic development and cell growth.






The Src homology domain is one of the two small protein binding domains, that are found in the Src oncoprotein. Now Src is a non-receptor tyrosine kinase protein, which is encoded in humans by the Src gene. It includes an SH2 domain, an SH3 domain and a tyrosine kinase domain. The proto-oncogene has an important role in the regulation of embryonic development as well as cell growth.

(Refer Slide Time: 04:19)

Function of Src

Activated Src

- promotes survival, angiogenesis, proliferation and invasion pathways
- participates in the regulation of angiogenic factors and vascular permeability
- regulates matrix metalloproteinase-9 activity




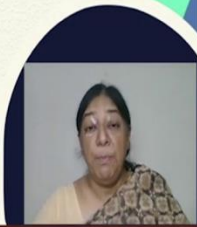
So the function of the Src, is when it is activated it promotes survival, angiogenesis, proliferation and invasion pathways. It participates in the regulation of angiogenic factors and vascular permeability and regulates the activities of many proteins including the mmp-9 activity.

(Refer Slide Time: 04:50)

Activation of the cellular Src pathway is observed in about 50% of tumors from colon, liver, lung, breast and the pancreas.

Use as drug target:

Many tyrosine kinase inhibitors that target c-Src (cellular Src) tyrosine kinase or any related tyrosine kinases have been developed for therapeutic use.

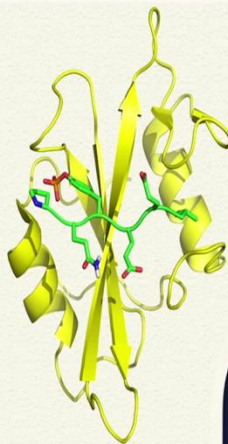


So the activation of the cellular Src pathway is observed in many tumors, as such it is used as a drug target. Many tyrosine kinase inhibitors that target the cellular Src tyrosine kinase or any related tyrosine kinases have been developed for therapeutic use.



(Refer Slide Time: 05:11)

Src-homology 2 (SH2)

- Common in signaling proteins
- Binds **phosphotyrosine** residues (pY) in the target protein
- The peptide usually has a **β conformation**

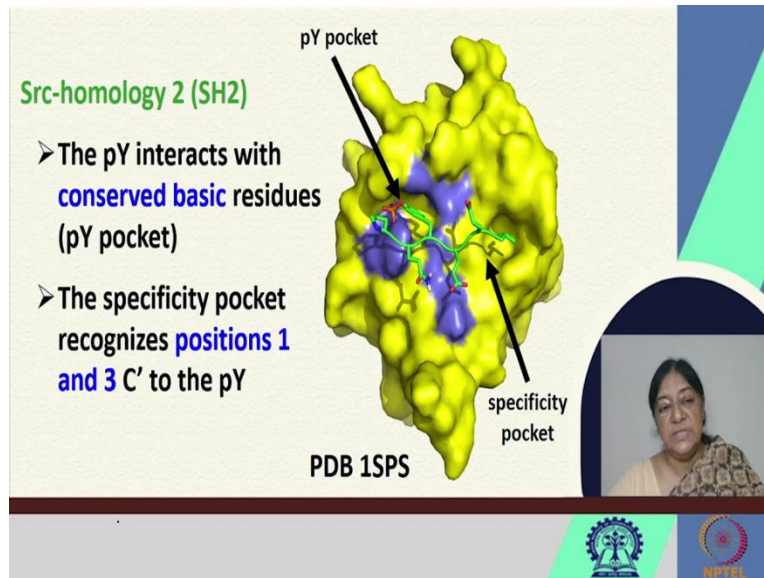


PDB 1SPS



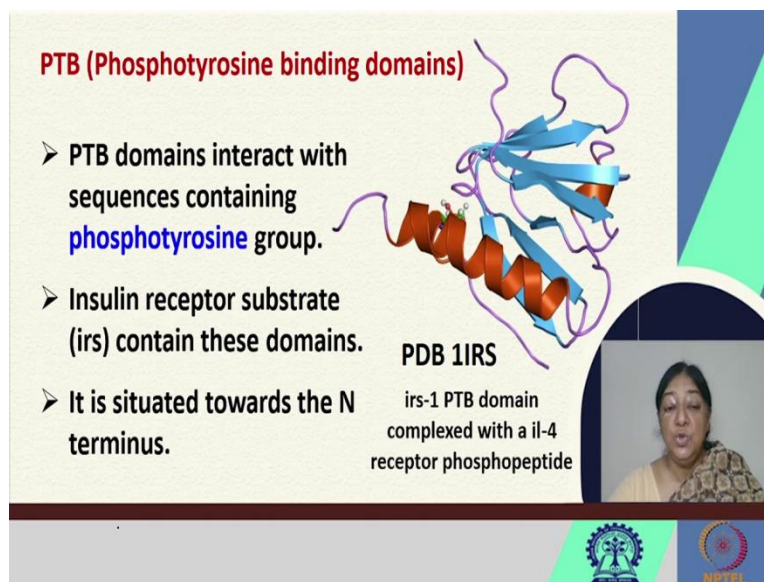
If we look at the SH2 protein and the recognition it has, it is very common in signaling proteins and it binds the phosphotyrosine residues in the target protein. The peptide usually has a β conformation.

(Refer Slide Time: 05:30)



Now, the importance of this [refer to slide] protein is in the pocket or the cleft that is seen in the specific receptor protein. There is a pY pocket, that is the phosphotyrosine pocket and we have a specificity pocket, related to the geometric and the chemical complementarity that is required for the strong affinity. So the pY interacts with conserved basic residues and the specificity pocket recognizes specific positions of the receptor protein.

(Refer Slide Time: 06:04)




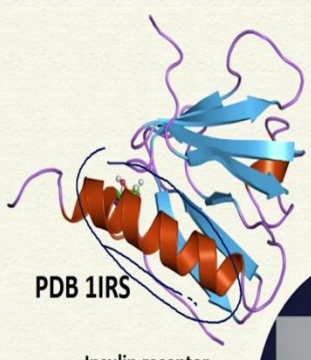
If we look at other phosphotyrosine binding domains, the importance of these domains lies in their geometry, in their overall structure and in the residues that are required for the specific recognition. So in the PTB domains they interact with sequences again that contain the phosphotyrosine group and one such is the insulin receptor substrate, the IRS that contains these domains. This is situated towards the N terminus.

(Refer Slide Time: 06:35)

- **PTB**
 - It has a compact, 7-stranded β -sandwich structure, capped by a C-terminal helix.
 - The substrate peptide fits into an L-shaped surface cleft formed from the C-terminal helix and strands 5 and 6.

PDB 1IRS


Insulin receptor substrate-1 also contains a pleckstrin homology domain



The PTB, as we can see [refer to slide] it has a compact 7 stranded β -sandwich structure that is capped by a C terminal helix and the substrate peptides fit into the L shaped surface cleft, that is formed by this C terminal helix and the strands 5 and 6.

(Refer Slide Time: 07:04)

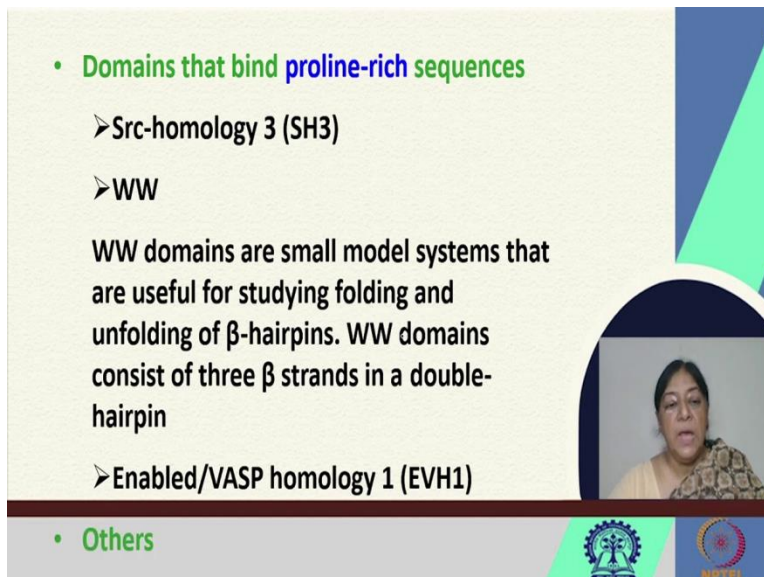
- **Domains that bind proline-rich sequences**
 - Src-homology 3 (SH3)
 - WW
 - Enabled/VASP homology 1 (EVH1)
- **Others**
 - PDZ: recognizes the [S/T]XV motif
 - LIM: binds e.g. to helix-loop-helix motifs



If we look at the other types of domains, there are domains that can bind proline rich sequences. This is the Src-homology 3, the SH3 type domain, a WW domain and the enabled VASP for a EVH1 domain and there are others such as the PDZ domain, that recognizes as we can see serine, threonine, any amino acid and a valine motif.

We also have the LIM domain that binds to helix-loop-helix motifs. So the specific recognition sites have been defined by these domains.

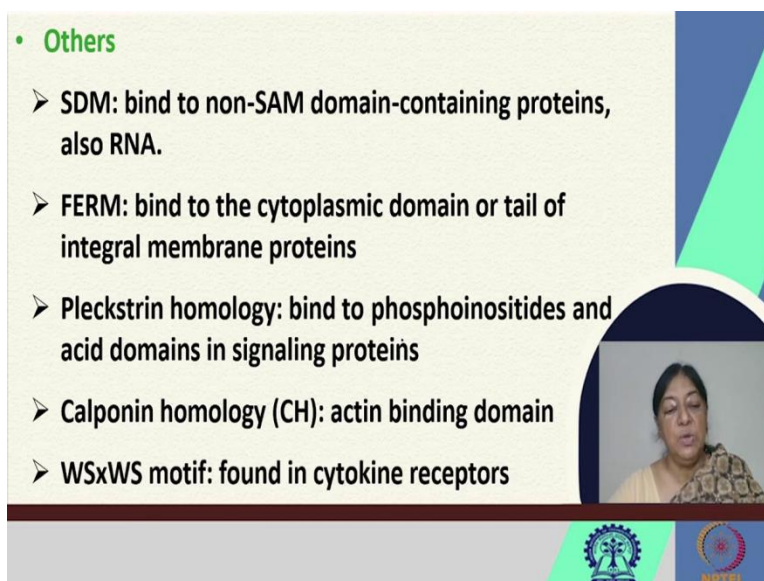
(Refer Slide Time: 07:49)



- **Domains that bind proline-rich sequences**
 - Src-homology 3 (SH3)
 - WW
 - WW domains are small model systems that are useful for studying folding and unfolding of β -hairpins. WW domains consist of three β strands in a double-hairpin
 - Enabled/VASP homology 1 (EVH1)
- **Others**

The domains that bind proline rich sequences, an example is these WW domains. The WW domains are small model systems that are useful for studying folding and unfolding of β -hairpins and they consist of 3 β strands in a double hairpin and we also have the enabled VASP homology 1 (EVH1).

(Refer Slide Time: 08:14)



- **Others**
 - **SDM**: bind to non-SAM domain-containing proteins, also RNA.
 - **FERM**: bind to the cytoplasmic domain or tail of integral membrane proteins
 - **Pleckstrin homology**: bind to phosphoinositides and acid domains in signaling proteins
 - **Calponin homology (CH)**: actin binding domain
 - **WSxWS motif**: found in cytokine receptors

Now when we look at the other types that we have. We have the SDM, this binds to a non-SAM domain containing protein also RNA, the FERM type that binds to the cytoplasmic domain or the

pleckstrin homology type that binds to phosphoinositides and acid domains in signaling proteins and a CH type, a calponin homology actin binding type.

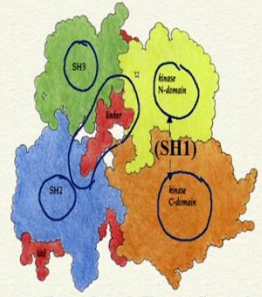
But what we need to recognize or understand that there is specific recognition involved in the protein-protein interactions for the specificity and the high affinity that is associated with the specific processes, specific biological action of each of these domains. Then we have another the WS and WS motif found in cytokine receptors.

(Refer Slide Time: 09:07)

SH2 and SH3 domains

➤ Originally identified as regions lying outside the kinase (SH1) domain of cytoplasmic tyrosine kinases (PTK)

Typically found in proteins that are involved in **growth factor signaling** (PTKs, PLC, adaptor proteins)



(Branden & Tooze, Introduction to protein structure)

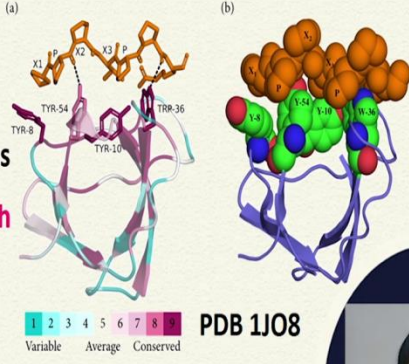
The diagram illustrates the structure of a cytoplasmic tyrosine kinase (PTK) with four domains: SH3 (green), SH2 (blue), SH1 (orange), and Kinase N-domain (yellow). The SH3 domain is located at the top left, SH2 at the bottom left, SH1 at the bottom right, and the Kinase N-domain at the top right. A linker connects the SH2 and SH1 domains. The SH3 domain is shown interacting with the SH2 domain. The Kinase N-domain is shown interacting with the SH1 domain. The SH1 domain is labeled as the kinase domain. The SH2 domain is labeled as the phosphotyrosine binding domain. The SH3 domain is labeled as the proline-rich domain. The Kinase N-domain is labeled as the catalytic domain. The SH1 domain is labeled as the kinase domain. The SH2 domain is labeled as the phosphotyrosine binding domain. The SH3 domain is labeled as the proline-rich domain. The Kinase N-domain is labeled as the catalytic domain.

So, when we look [refer to slide] at the SH2 and SH3 domains, they are originally identified as regions lying outside the kinase domain of the cytoplasmic tyrosine kinase. So if you look at the overall structure that are typically found in proteins that are involved in growth factor signaling, we have the SH2 and SH3 domains here and here is the kinase N domain and the kinase C domain. These are connected through a linker. This structure is important as we can see, in the formation of this specific complex for the necessary action to occur.

(Refer Slide Time: 09:48)

- **Src-homology 3 (SH3)**

- Common in signaling proteins
- Binds **proline-rich** sequences in target proteins



The interaction of an SH3 domain with the polypeptide that has a PPII conformation

PDB 1JO8

Variable Average Conserved

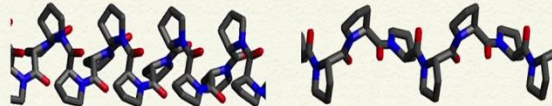
NPTEL

If we look at the SH3 domain, that is a domain that recognizes the polypeptide. This is common again in signaling proteins and it binds proline rich sequences in target proteins and we see that it binds polypeptides in a PPII conformation.

(Refer Slide Time: 10:13)

Polyproline helix - protein secondary structure

Occurs in proteins comprising repeating proline residues

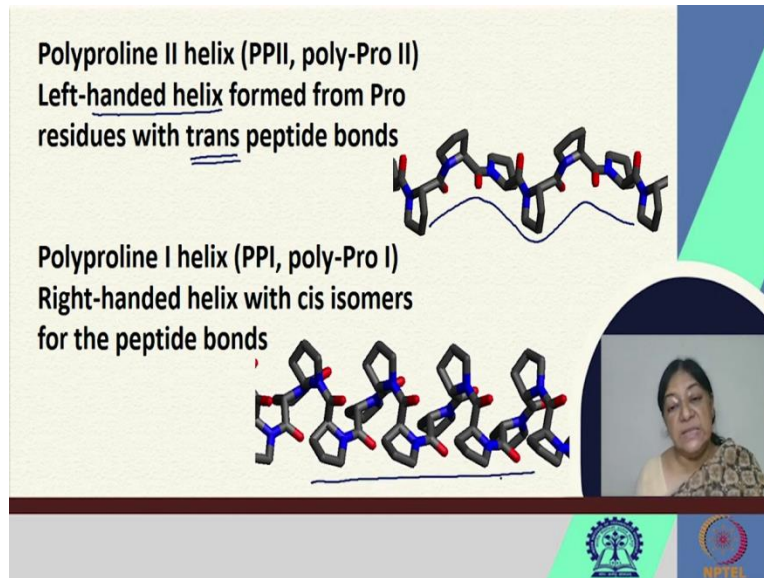


Polyproline I helix (PPI, poly-Pro I)
Polyproline II helix (PPII, poly-Pro II)

NPTEL

If we try and see what the polyproline helix means, this is an adaptation of a specific type of protein secondary structure involved with a polyproline type of residue involvement, where we have proteins that are comprised of repeating proline residues. So, we have the polyproline I type and the polyproline II type.

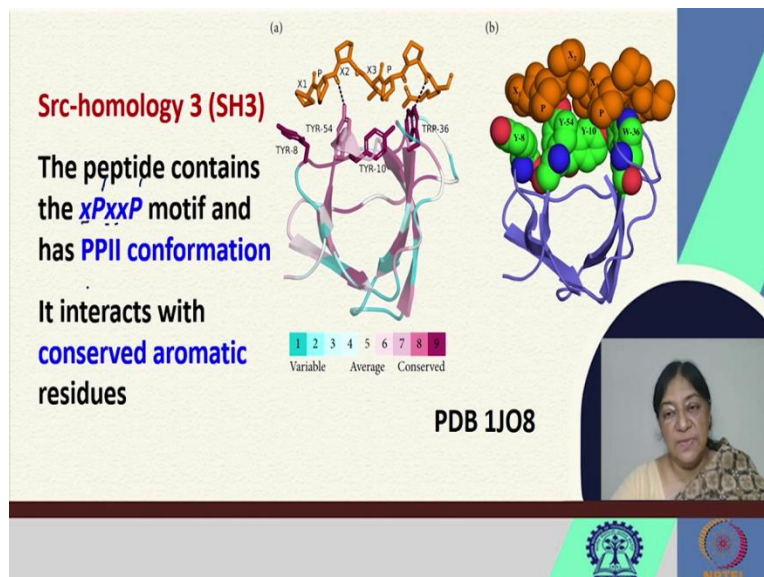
(Refer Slide Time: 10:41)



When we look at the polyproline II, the PPII types of helix we will see that they have a specific characteristic. These are left handed helices that are formed mainly from trans peptide bonds and they appear in this [refer to slide] fashion.

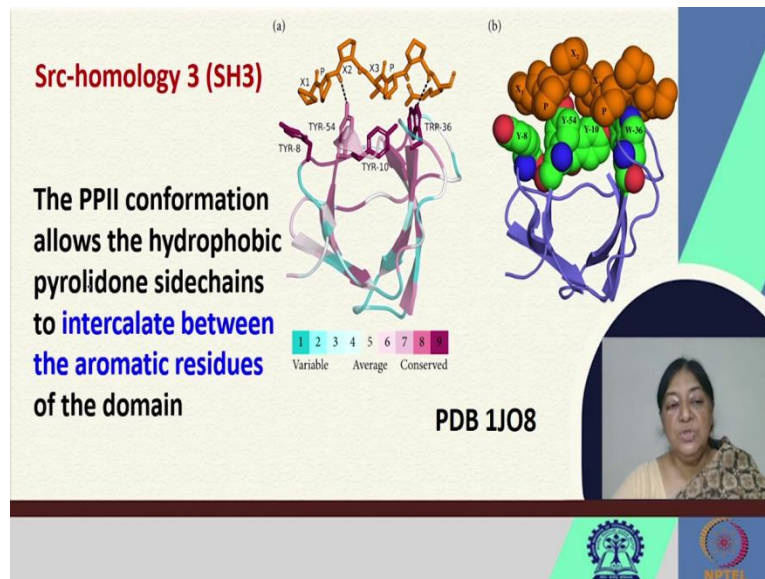
This is what is recognized by these specific types of domains that recognize a polyproline PPII type helix. For the polyproline I type helix, the PPI, we have right handed helices with cis isomers for the peptide bonds. The polyproline helix in this type appears in this [refer to slide] fashion.

(Refer Slide Time: 11:24)



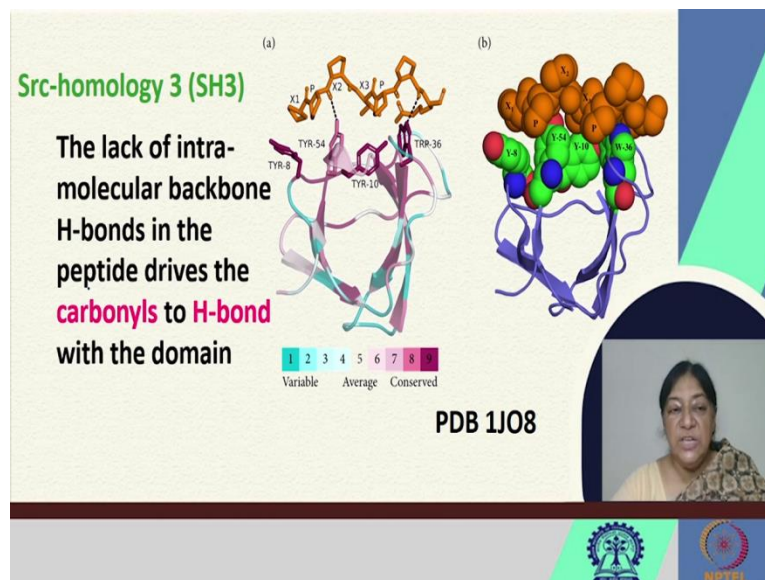
If we look [refer to slide] at the SH3 type of domain, this peptide contains the *xPxxP* motif and has the PPII conformation. It interacts with conserved aromatic residues that are there in the receptor.

(Refer Slide Time: 11:44)



The PPII conformation allows the hydrophobic pyrrolidone sidechains to interconnect between the aromatic residues of the domain.

(Refer Slide Time: 11:54)



As such, there is a lack of intramolecular backbone hydrogen bonds in the peptide and this drives the carbonyls to hydrogen bond with the domain, creating an affinity site for the specific recognition required.

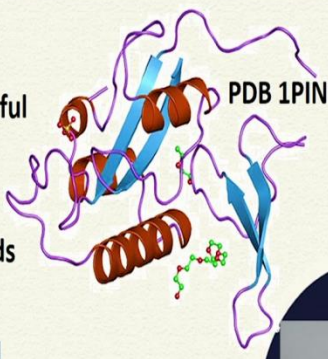

(Refer Slide Time: 12:09)

WW domains

- Small model systems useful for studying folding and unfolding of β -hairpins
- Consists of three β strands in a double-hairpin
- Binds to proline enriched sequences and to pS – pT containing motifs

PDB 1PIN

Structure of the human mitotic rotamase

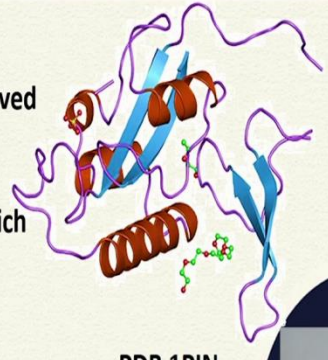

The WW domains are small model systems, that are useful for studying folding and unfolding of β -hairpins. These consist of three β strands in a double hairpin mode and again bind to proline enriched sequence and also to phosphoserine and phosphothreonine containing motifs. So, we have a specific structure associated with this for the human mitotic rotamase, which has the WW domain for the recognition in a protein-protein interaction specifics.

(Refer Slide Time: 12:42)

WW domains

- Structural module involved in protein-protein interactions through recognition of proline-rich peptide motifs
- Mediates regulatory protein complexes in various signaling networks

PDB 1PIN

The structural module involved in the protein-protein interactions, is again through the recognition of proline rich peptide motifs and this mediates regulatory protein complexes in various signaling networks. So it is important in the signaling.

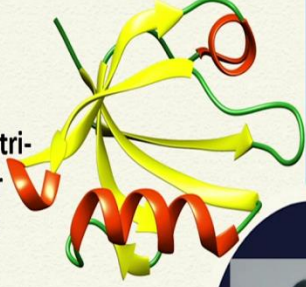


(Refer Slide Time: 13:05)

PDZ domain

- Structural domain of 80-90 amino-acids
- Recognizes carboxy-terminal tripeptide motifs (S/TXV), other PDZ domains or LIM domains
- Binds via a short peptide sequence which has a C-terminal hydrophobic residue

PDB 2DC2

Solution Structure of PDZ Domain

The PDZ domain is another domain that is involved in recognizing the carboxy terminal tripeptide motifs and also other PDZ domains or LIM domains. Thus, this structural domain consists of 80 to 90 amino acids in the recognition site. This binds via short peptide sequence, which has a C terminal hydrophobic residue, that is required for the recognition.

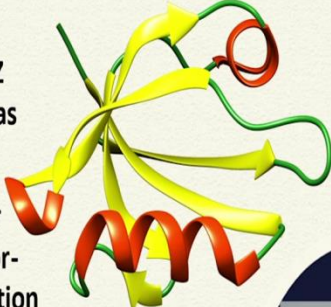


(Refer Slide Time: 13:43)

PDZ domain

- Some proteins with PDZ domains are identified as scaffolding proteins
- Involved in ion receptor assembling and receptor-enzyme complex formation
- PDZ domains are directly involved in the regulation of different cellular pathways

Solution Structure of PDZ Domain

PDB 2DC2

Some proteins with this PDZ domains are identified as scaffolding proteins and they are involved in ion receptor assembling and receptor enzyme complex formation and these PDZ domains are also directly involved in the regulation of different cellular pathways. So, we realize the importance of the types of residues involved in this recognition.

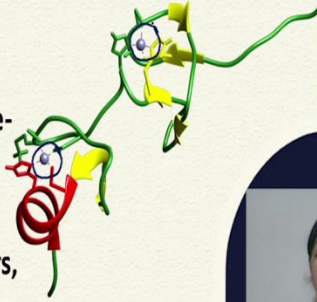



(Refer Slide Time: 14:10)

LIM domain

- Conserved cysteine-rich sequence found in eukaryotic proteins
- Contains a tandem cysteine-rich Zn^{2+} finger motif
- Binds to PDZ domains, specific transcription factors, and other LIM domains.

Structure of the 4th LIM domain of Pinch protein. Zinc atoms are shown in purple

PDB 1NYP

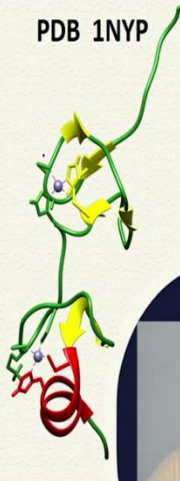



In the last domain that we are going to look at is the LIM domain. There are several other domains that are also known, but we will look at these specific ones. This has a conserved cysteine rich sequence, that is found in eukaryotic proteins. It contains a tandem cysteine rich zinc finger motif, where we see the zinc attached to this specific LIM site. This binds to PDZ domains, specific transcription factors and also other LIM domains.

(Refer Slide Time: 14:45)

LIM domain

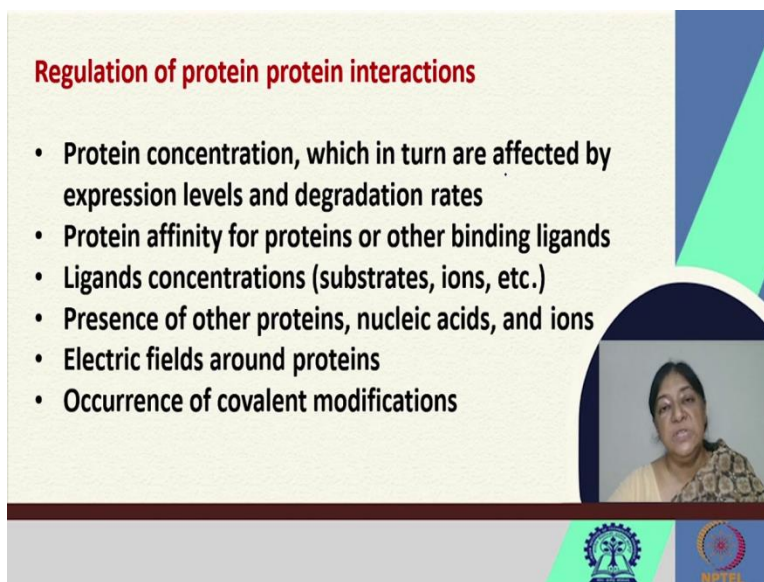
- Divergent amino acid sequences, that facilitate the formation of a stable protein core and tertiary fold.
- Possible high-affinity binding to a wide variety of protein targets due to the double zinc-finger structure
- Plays roles in cytoskeletal organization, organ development, regulation of gene transcription

PDB 1NYP

In this specific helix-loop-helix recognition, there are divergent amino acid sequences that facilitate the confirmation of a stable protein core and a tertiary fold that is required for the recognition. Then there is possible high affinity binding to the wide variety of protein targets, due to the double zinc finger structure observed. It plays a very important role in cytoskeletal organization, organ development and the regulation of gene transcription.

(Refer Slide Time: 15:19)



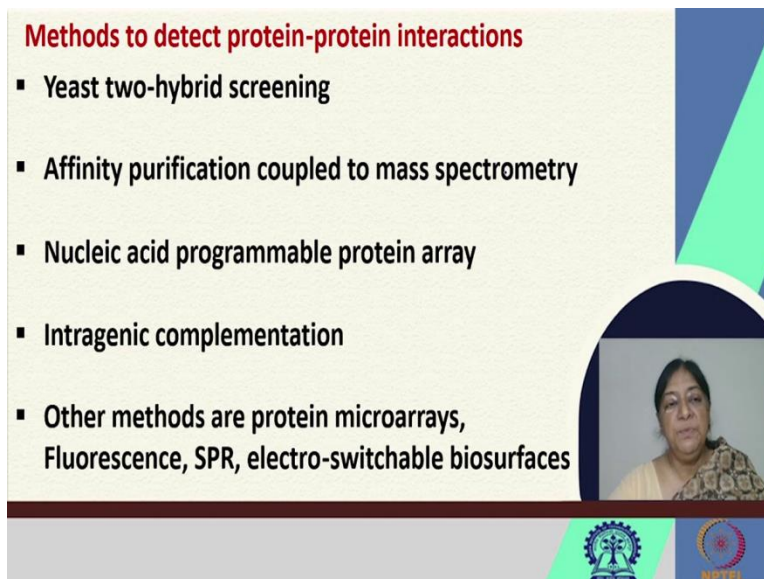
Regulation of protein protein interactions

- Protein concentration, which in turn are affected by expression levels and degradation rates
- Protein affinity for proteins or other binding ligands
- Ligands concentrations (substrates, ions, etc.)
- Presence of other proteins, nucleic acids, and ions
- Electric fields around proteins
- Occurrence of covalent modifications

The slide features a video inset of a woman in the bottom right corner and logos for IIT Bombay and NPTEL at the bottom.

A look at these different types of protein-protein interactions means, there has to be a regulation of the protein-protein interactions. This can be regulated by the protein concentration, which in turn is affected by the expression levels and also the degradation rates. The specific protein affinity for other proteins or other binding ligands. The ligand concentrations, the substrates or the ions. The presence of other proteins, nucleic acids and ions and also the electric fields around the proteins.

(Refer Slide Time: 16:28)



Methods to detect protein-protein interactions

- Yeast two-hybrid screening
- Affinity purification coupled to mass spectrometry
- Nucleic acid programmable protein array
- Intragenic complementation
- Other methods are protein microarrays, Fluorescence, SPR, electro-switchable biosurfaces

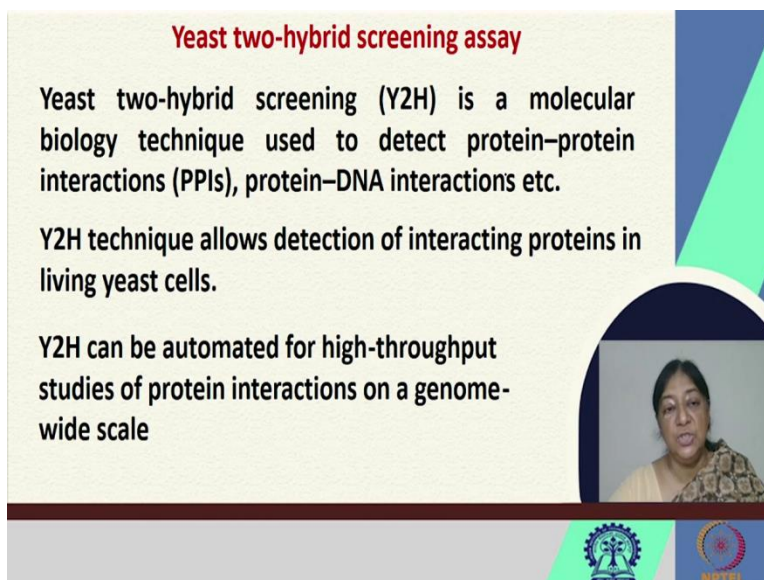
The slide features a video inset of a woman in the bottom right corner and logos for IIT Bombay and NPTEL at the bottom.

So, there are several factors that are going to regulate or affect the protein-protein interactions that are necessary to be understood and there is the occurrence of covalent modifications, that

may break or possibly interrupt the protein-protein interaction site, affected in a manner that it would not be able to perform its specific biological function.

There are several methods that are required to detect these protein-protein interactions. If we know that a specific protein-protein is interacting, there are several tests that can tell us whether there is a specific protein-protein interaction. There is the yeast two-hybrid screening, known as the Y2H test; the affinity purification that can be coupled with mass spectrometry to check for specific affinity of two proteins and we have different protein array methods and other methods including fluorescence SPR and electro-switchable biosurfaces.

(Refer Slide Time: 17:11)



Yeast two-hybrid screening assay

Yeast two-hybrid screening (Y2H) is a molecular biology technique used to detect protein-protein interactions (PPIs), protein-DNA interactions etc.

Y2H technique allows detection of interacting proteins in living yeast cells.

Y2H can be automated for high-throughput studies of protein interactions on a genome-wide scale

The slide features a video inset of a woman speaking in the bottom right corner. At the bottom of the slide, there are two logos: the Indian Institute of Technology (IIT) Bombay logo on the left and the NPTEL logo on the right.

We will briefly be looking at the Y2H screening. In this screening assay, the yeast two-hybrid screening assay, this is a molecular biology technique that can be used to detect protein-protein interactions and protein DNA interactions. It allows the detection of interacting proteins in living yeast cells and it can also be automated for high throughput studies to understand protein interactions on a genome wide scale.

(Refer Slide Time: 17:41)

Y2H method :

Transcription can be activated in eukaryotic organisms
Interaction between two proteins, called bait and prey,
activates reporter genes that enable growth on specific
media or a color reaction

DNA-Binding Domain (BD), the bait is physically
associated with an Activating Domain (AD), the prey

A protein of interest is fused to BD, while another
protein is fused to AD.

Outcome: Transcription of reporter gene activated

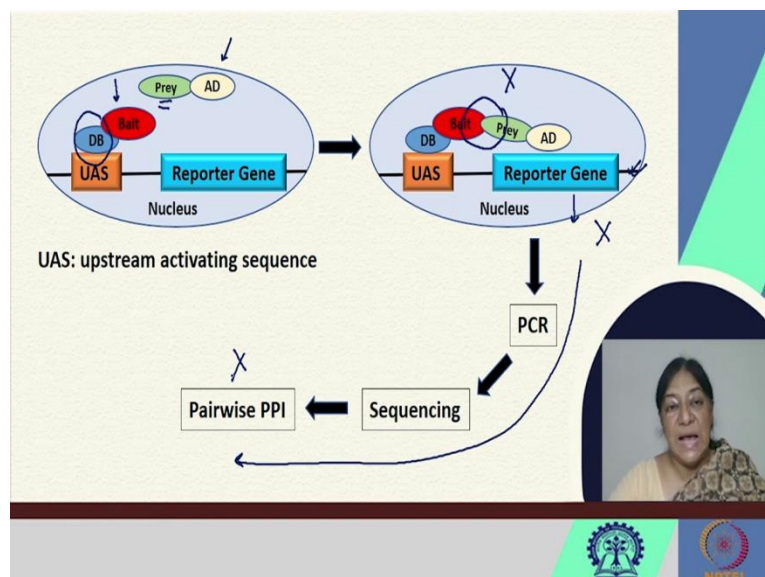
Result: The proteins physically interact



This method is used because there is a transcription that can be activated in eukaryotic organisms. This is taken into consideration and there are the interactions between two proteins, the bait and the prey, that activates the reported genes that enable the growth that could be on a specific media or a specific colored reaction, to indicate that there has been a protein-protein interaction involved.

There is a DNA binding domain, the BD which is the bait. If it is physically associated with an activative domain the prey, then a protein of interest is fused to the BD, while another is fused to the AD. If these interact with each other, there is the transcription of the reporter gene that is activated. So the result indicates that these proteins, physically interact with each other.

(Refer Slide Time: 18:39)



So just for a cartoon representation [refer to slide] a schematic, we have the DNA binding domain where we have the bait attached, we have the activating domain where we have the prey attached. If there is an interaction between these two proteins, then the reporter gene is activated and as such we will see a pair wise protein-protein interaction. However, if there is no interaction between the bait and the prey, then this reported gene will not be activated and there will be no pair wise protein-protein interaction involved.

(Refer Slide Time: 19:20)

Computational methods

- Genomic Context Methods
- Text mining methods
- Machine learning methods

The slide features a video feed of a woman in the bottom right corner. At the bottom, there are logos for IIT Bombay and NPTEL.




This means with this assay possible, there are the possibilities to get a large number of protein-protein interactions. So specific computational methods have been developed to look at these protein-protein interactions from different points of view, where we can look at the genomic context methods, the text mining methods or machine learning methods.

(Refer Slide Time: 19:44)

PPIs

Studied from different perspectives and with many methods
Biochemical point of view, quantum chemistry,
molecular dynamics, signal transduction, etc.

Information enables the creation of large protein
interaction networks that can help discover putative
protein targets of therapeutic interest.






In an understanding of the PPIs that can be looked at from different perspectives, from a biochemical point of view, looking at a molecular dynamics or signal transduction, there is a lot of information available that of the creation of large protein interaction networks that can actually help discover other putative protein targets that may be of therapeutic interest.

(Refer Slide Time: 20:10)

❑ Databases:
Identification of a large number of PPIs generates
many interactions - collected together in specialized
biological databases that are continuously updated
to provide complete information

❑ Interaction networks:
Information found in PPIs databases supports the
construction of interaction networks






This is done by looking at specific databases. These databases are generated from the information of specific protein-protein interactions that are collated and collected from biological databases, that are continuously updated to provide complete information. There are specific interaction networks that are possible, where construction of these interaction networks gives us an idea about how the interactions occur and which proteins can interact with which other proteins.

(Refer Slide Time: 20:47)

Protein-protein interaction databases

Molecular interaction databases aim to fulfill the need of producing PPI data by extracting information from scientific publications or, in some cases, by including protein-protein interaction predictions found using computational methods.

The storage of interactions in publicly available databases allows access to a large volume of interaction data and subsequent analysis of the interactome.






Based on that, there are protein-protein interaction databases that are obtained from extracting information from scientific publications. Other protein-protein interaction predictions can be found using these computational methods. So, we have the storage of these interactions in publicly available databases, that allow access to a large volume of interaction data that can give us insights into protein-protein interactions.

(Refer Slide Time: 21:21)

The first of these databases was the Database of Interacting Proteins (DIP).

Primary databases are those that collect experimental molecular interaction data exclusively from peer-reviewed scientific publications.
Ex: IntAct, MINT and MatrixDB

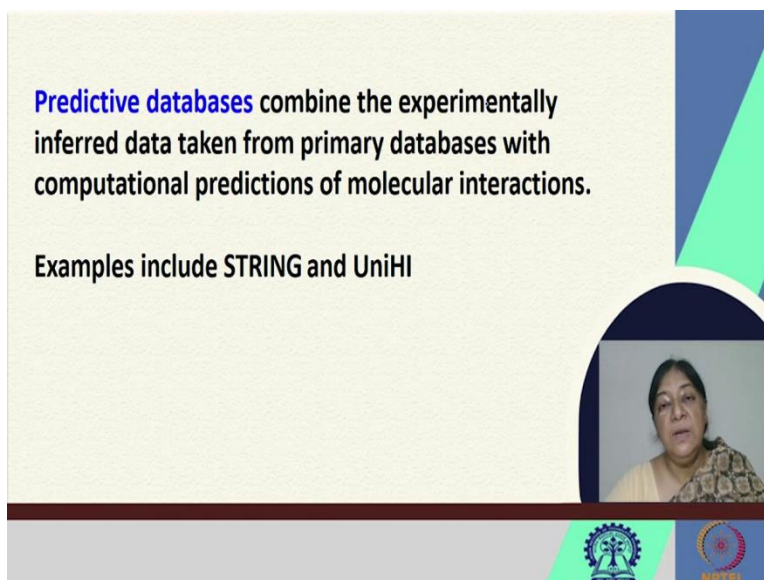
Secondary databases or meta-databases, seek to integrate the data curated by several primary databases in one, integrated repository.
Ex: APID and PINA



To look at this, there is the database of interacting proteins. There is a primary database which is thus the collection of experimental molecular interaction data, exclusively from peer reviewed scientific publications. So protein-protein interactions reported will get a place in this matrix DB

or the primary database. The secondary database is a meta-database that integrates the data that has been curated from the primary database into an integrated repository.

(Refer Slide Time: 21:57)



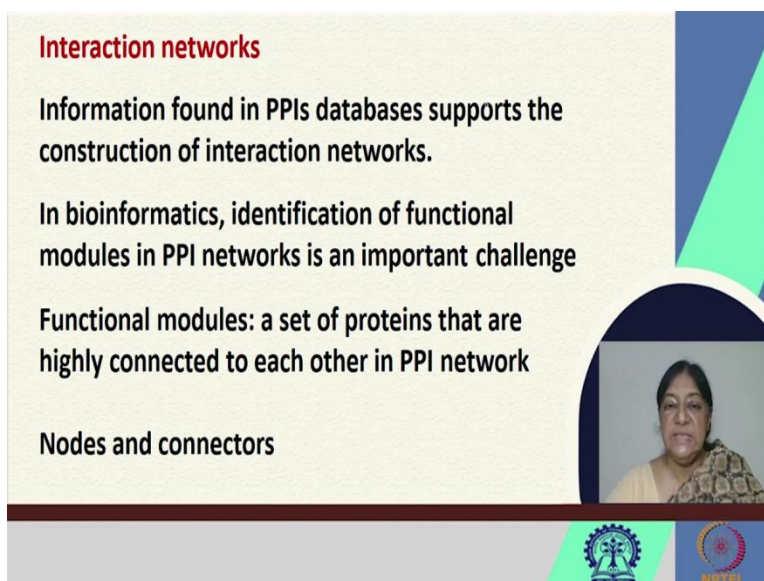
Predictive databases combine the experimentally inferred data taken from primary databases with computational predictions of molecular interactions.

Examples include STRING and UniHI

The slide features a light beige background with a blue and green geometric design on the right. A circular inset in the bottom right corner shows a woman speaking. At the bottom, there are logos for a university and NPTEL.

Then we have the predictive database, that then combines the experimentally inferred data taken from the primary database with computational predictions of possible molecular interactions. Now, this is important in identification of specific networks and see whether we have interactions between proteins, between networks that were not known or cannot be determined experimentally.

(Refer Slide Time: 22:31)



Interaction networks

Information found in PPIs databases supports the construction of interaction networks.

In bioinformatics, identification of functional modules in PPI networks is an important challenge

Functional modules: a set of proteins that are highly connected to each other in PPI network

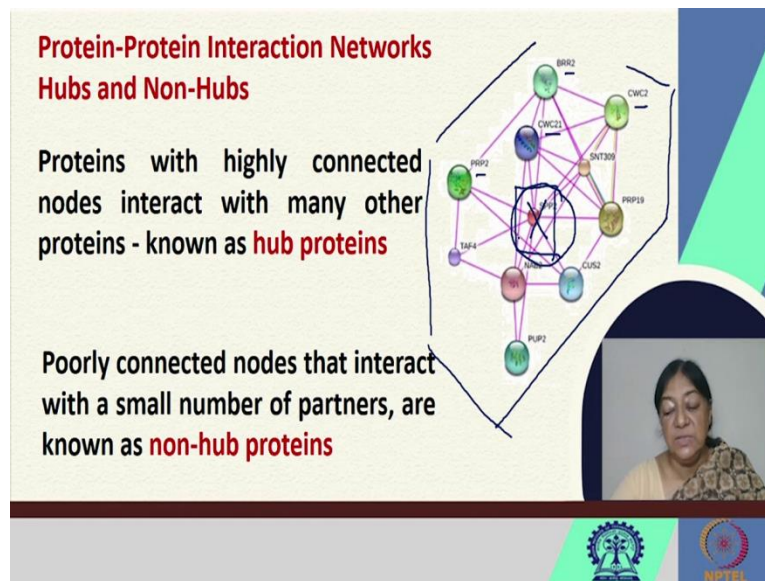
Nodes and connectors

The slide features a light beige background with a blue and green geometric design on the right. A circular inset in the bottom right corner shows a woman speaking. At the bottom, there are logos for a university and NPTEL.

The information found in the PPI databases, supports the construction of these interaction networks and we have functional modules of the PPI networks that help us try to connect different proteins together, that are involved say in a specific process or involved in a specific disease because we realize that having the specific biological process or any biochemical reaction, is a cascade of effects that require a large number of proteins in the cascade reactions that take place, to interact with each other for the specific process to occur.

In this case we have what are called nodes and connectors, that connect these different proteins.

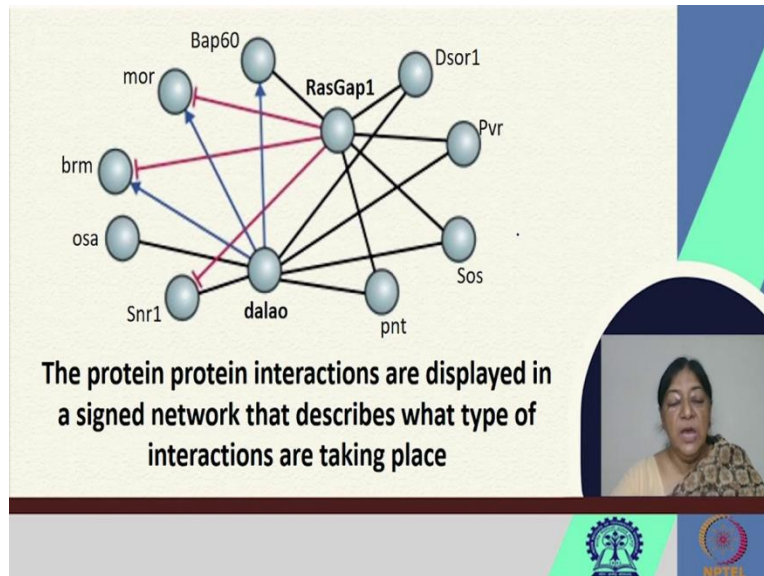
(Refer Slide Time: 23:17)



In the protein-protein interaction networks, we have what are called hub proteins and non-hub proteins. Proteins with highly connected nodes that interact with other proteins are known as the hub proteins and the non-hub proteins are those that are poorly connected, that interact with only a small number of partners.

For example this [refer to slide] interaction network shows us that this particular protein has a larger network. So this would be a hub protein that has highly connected nodes, that are going to tell us that if we are going to target a network that is involved in a manner where this set of proteins are involved in a specific reaction, then it would be good to target this specific protein to see if we can inhibit its action, which would mean that we would in turn inhibit its interaction with the other set of proteins, that it is associated with.

(Refer Slide Time: 24:23)



So similarly the protein-protein interactions can be displayed [refer to slide] in this network, that describe what types of interactions are actually taking place.

(Refer Slide Time: 24:36)

Hub proteins

- special biological properties
- could serve as interesting drug targets
- more essential than non-hub proteins
- role in modular organization of protein-protein interaction network

Hubs are larger than non-hubs
 Substitution of aliphatic, hydrophobic and aromatic amino acids with acidic, charged and hydrophilic amino acids from non-hub to hub proteins can lead to increased connectivity

So, we can look at these hub proteins that have special biological properties and they could serve as interesting drug targets because they seem to be more essential than the non-hub proteins and they have a large role in modular organization of the protein-protein interaction network. The hubs are larger than non-hubs and the substitution can change the connectivity.

(Refer Slide Time: 25:03)

PPIs and neurodegenerative diseases

Interaction of tau protein with amyloid- β and their role in neurodegenerative diseases such as Alzheimer's or Parkinson's disease is a subject of ongoing research. These are associated with the prion-like propagation and aggregation of toxic proteins.

As therapeutic targets

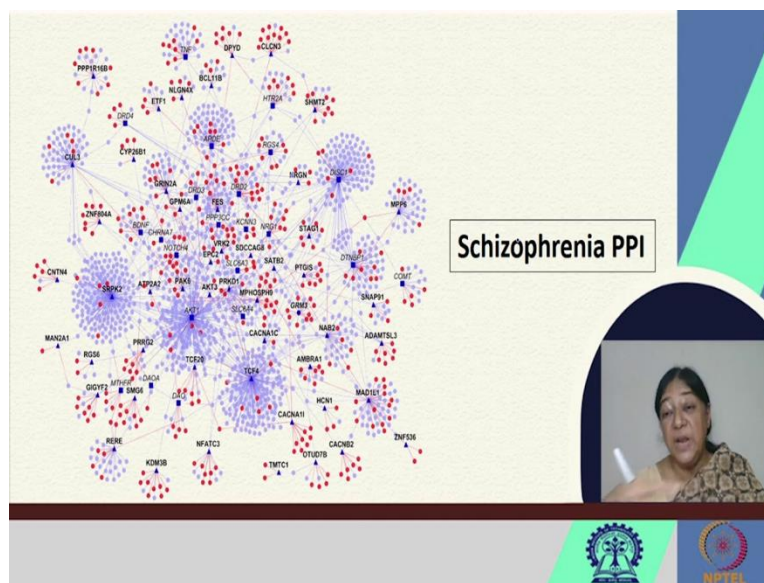
Several properties of PPI have been incorporated into drug-design strategies, e.g. allosteric sites and hotspots.



When we look at PPIs and neurodegenerative diseases for example, we can look at the interaction of the tau protein with the amyloid- β and their specific role in neurodegenerative diseases, such as the Alzheimer's or the Parkinson's disease. This is a subject of ongoing research and these are associated with prion like propagation and aggregation of toxic proteins.

So they can be used as therapeutic targets. Several properties of the PPI have been incorporated into drug design strategies, where we look at the specific hotspots, the conserved regions that are required for the interactions and allosteric sites that might affect the protein-protein interaction per se.

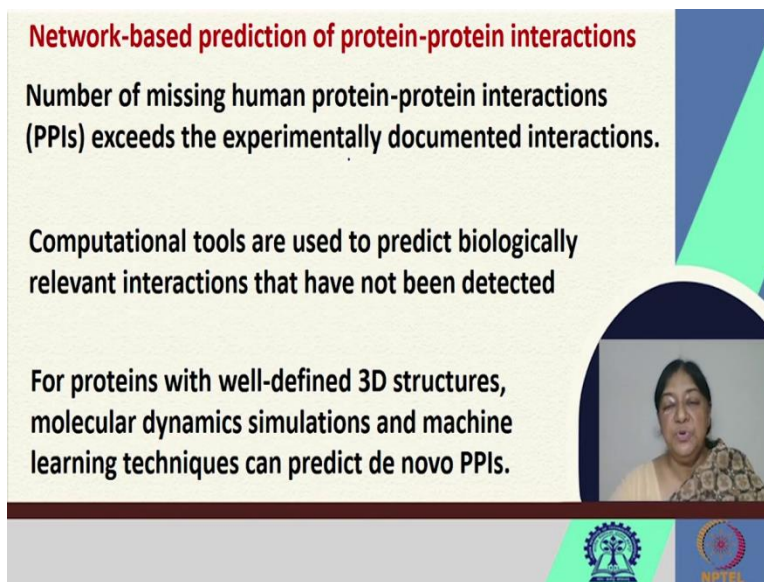
(Refer Slide Time: 25:51)



So if we look [refer to slide] at an example of a schizophrenia PPI, this is what the PPI network looks like. So we can see the location of the specific hub proteins, but we realize the whole

cascade of interactions and how large this interaction network is, that tells us about the protein-protein interactions involved in this specific disease. These protein-protein interactions can give us an idea of which specific proteins to target in an attempt to combat disease.

(Refer Slide Time: 26:26)



Network-based prediction of protein-protein interactions

Number of missing human protein-protein interactions (PPIs) exceeds the experimentally documented interactions.

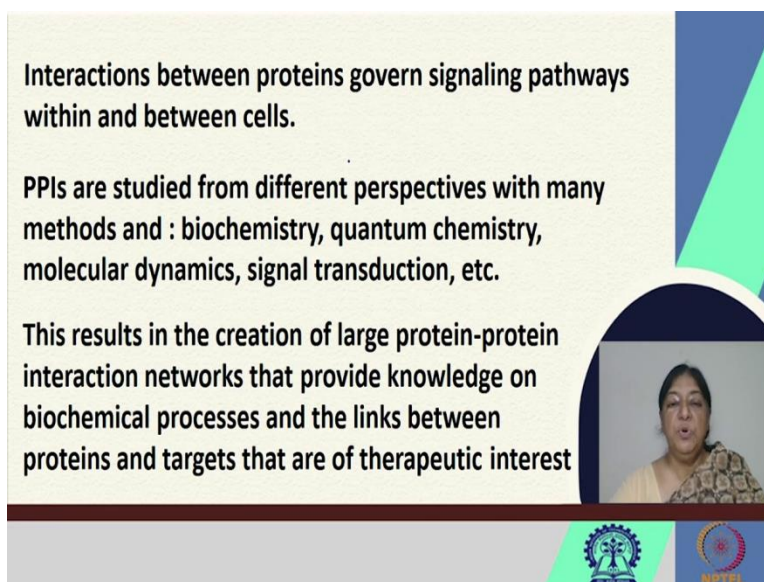
Computational tools are used to predict biologically relevant interactions that have not been detected

For proteins with well-defined 3D structures, molecular dynamics simulations and machine learning techniques can predict de novo PPIs.

The slide features a yellow background with a blue and green geometric design on the right. A circular inset on the right shows a woman speaking. At the bottom, there are logos for IIT Bombay and NPTEL.

So, the network based prediction of protein-protein interactions thus is extremely important because the number of missing human protein-protein interactions, far exceeds the experimentally documented interactions. These computation tools are used to predict biological relevant interactions that have yet to be detected. For proteins with well defined 3D structures, there are MD simulations that may be possible and machine learning techniques that can predict de novo protein-protein interactions.

(Refer Slide Time: 27:04)



Interactions between proteins govern signaling pathways within and between cells.

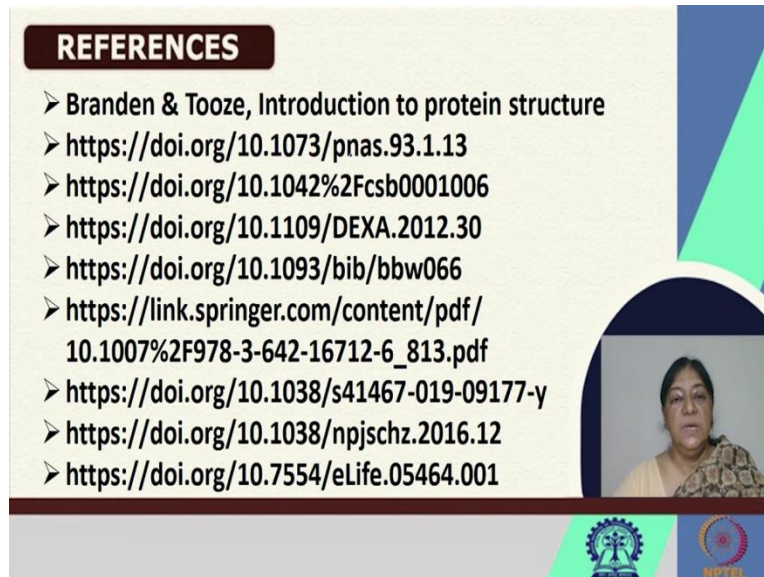
PPIs are studied from different perspectives with many methods and : biochemistry, quantum chemistry, molecular dynamics, signal transduction, etc.

This results in the creation of large protein-protein interaction networks that provide knowledge on biochemical processes and the links between proteins and targets that are of therapeutic interest

The slide features a yellow background with a blue and green geometric design on the right. A circular inset on the right shows a woman speaking. At the bottom, there are logos for IIT Bombay and NPTEL.

The interactions between proteins that govern signaling pathways between and within cells, can be studied from different perspectives and we see the results. This results in a creation of large protein-protein interaction networks, that provide us with knowledge on the biochemical process and the links between proteins and targets that are of therapeutic interest.

(Refer Slide Time: 27:29)



REFERENCES

- Branden & Tooze, Introduction to protein structure
- <https://doi.org/10.1073/pnas.93.1.13>
- <https://doi.org/10.1042%2Fcsb0001006>
- <https://doi.org/10.1109/DEXA.2012.30>
- <https://doi.org/10.1093/bib/bbw066>
- https://link.springer.com/content/pdf/10.1007%2F978-3-642-16712-6_813.pdf
- <https://doi.org/10.1038/s41467-019-09177-y>
- <https://doi.org/10.1038/npj-schz.2016.12>
- <https://doi.org/10.7554/eLife.05464.001>

These [refer to slide] are the references.

Thank you.