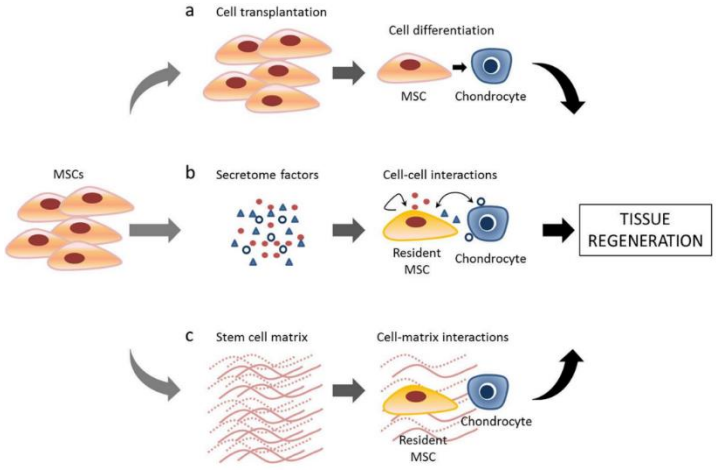
PROTEIN-PROTEIN INTERACTIONS: the "Interactome"

Secretomics:

Responsible to cell-to-cell interaction, used for basic science and diagnostics, also being a challenge to proteomics. As it has its application in both science & therapeutics, specifically in regenerative medicine.

Mesenchymal stem cells (MSCs) demonstrate great promise and clinical efficacy in cartilage regeneration. With a deeper understanding of stem cell biology, new therapeutics and new bioengineering approaches have emerged and showed potential for further developments. There has been a paradigm shift in applying MSCs for tissue regeneration from the use of stem cells for transplantation to the use of stem cell derived matrix and secretome components as therapeutic tools and agents for cartilage regeneration.



MSCs can be utilized for cartilage repair in various fashions. **(a)**MSCs can be used in direct transplantation. Alternatively, the cells may be pre-treated/pre-conditioned and processed to harness specific **(b)** secretome factors and **(c)** matrix molecules as stem cell-based therapeutics for the treatment.

INTERACTOMICS: importance of protein interactions

Protein-protein interaction are intrinsic to virtually every cellular process; cell growth, cell cycle, metabolic pathway, signal transduction

- Understanding of how proteins work within the cell;
- Gene mutation → protein interaction confusion → disease;
- New drug development by protein function analysis;
- Unknown protein may be discovered by known protein in protein signal pathway;

Now that we know how to identify and quantify proteins, we still don't have all the clues for understanding pathological issues.

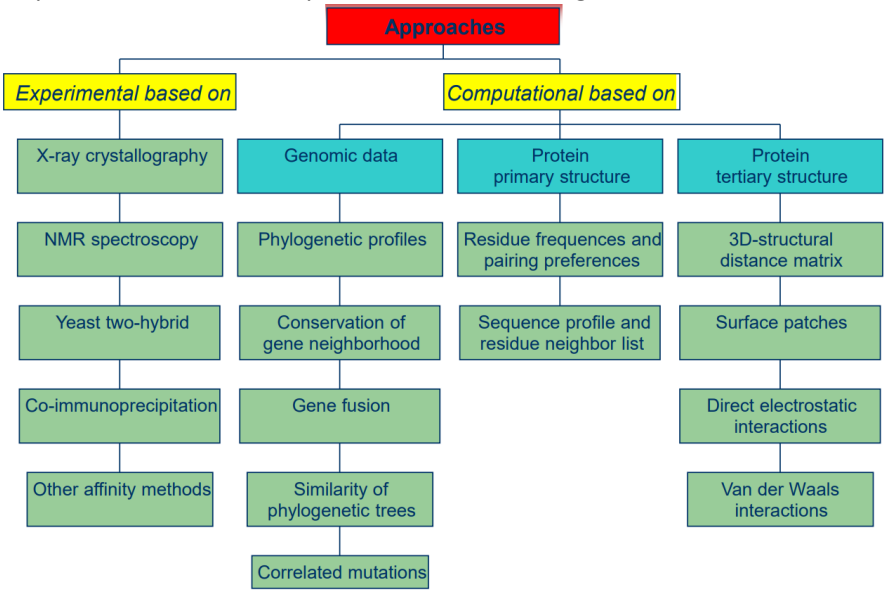
Interactomics has two **main goals:**

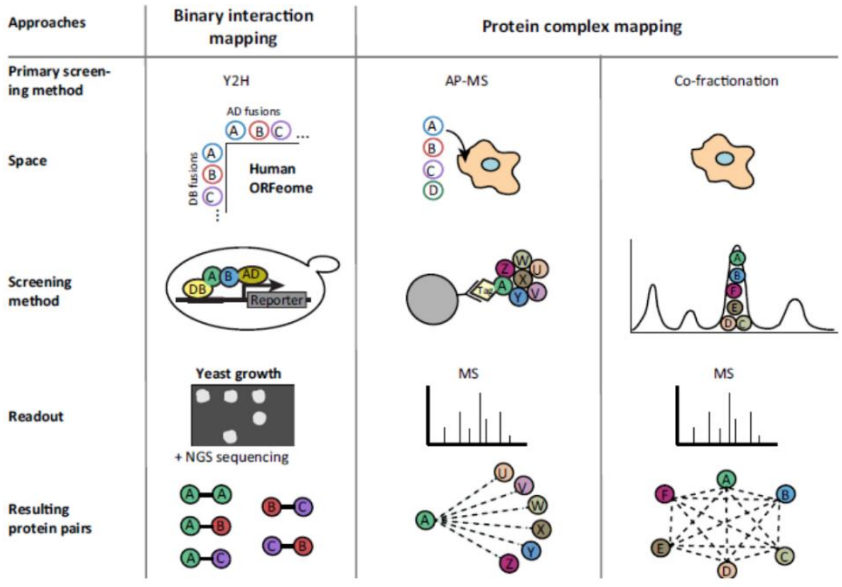
- **identify which are the interactive proteins (the partners)**

- **Identify the interface (interacting residues)**

Also, here we have an increase of complexity (we can't apply a single methodology to cover all the targets)

We can take advantage of several methodologies both **experimental and computational**. Computational ones are useful to predict the result of experimental methodologies, and vice versa.





In this approach we have to consider that sometimes we are dealing with a one-to-one interaction (binary interaction), other times we have multiple interactions.

There are approximately 2-4 million proteins per cubic micron in bacteria, yeast, and mammalian cells (Milo, 2013). The number of interactions between these proteins is hard to imagine yet alone study.

1. Experimental methods: Physical interaction between protein binding domains

In vitro

- **Co-immunoprecipitation**
- GST-pull down assays
- Protein arrays
- FRET, BRET

In vivo

- **Yeast two-hybrid system**
- **Phage display**
- FRET, BRET
- Proximity labelling

All the methods rely on the bait-prey model--> does the bait bind to the prey? All methods have to answer to this question. In all the approaches, you are taking a photograph of the moment, as in proteomics.

In some approaches we test the interaction between a bait and MORE preys deriving by a library.

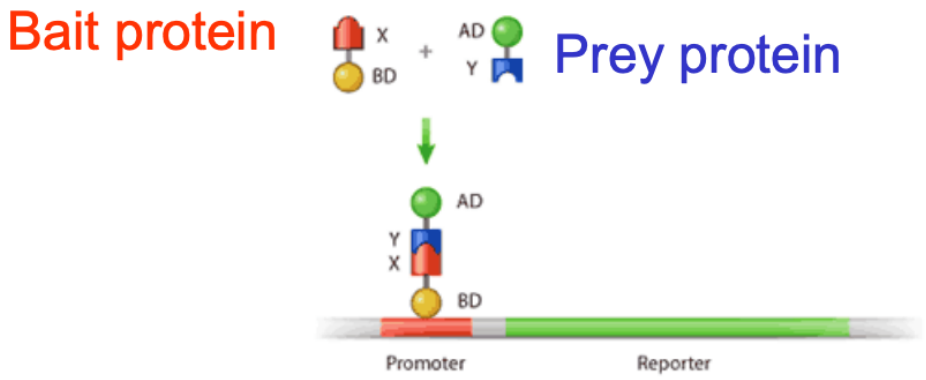
There are methods used in vitro (monitor with the purified proteins), in vivo and in living laboratory animals. Some of them can be used monitoring the different interaction at different levels.

Yeast two-hybrid system

Everything started with the yeast two-hybrid system--> based on the fact that eukaryotic transcriptional activators (naturally present in yeast) consist of two individual domains. In these cells, for the transcription of this gene, we have two domains:

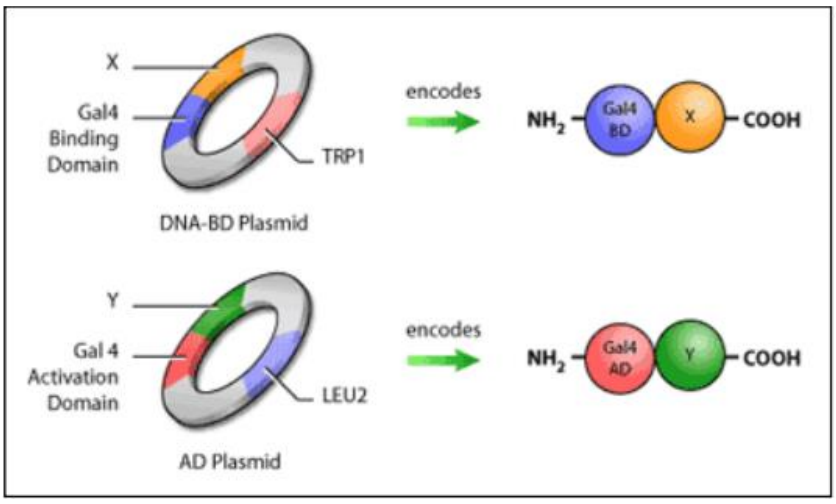
1. **The DNA binding domain** (DBD), recognizes a specific DNA sequence
2. **The Activation domain** (AD). The DBD coordinates the assembly of the elements required for transcription and enables RNA polymerase II to transcribe a specific reporter gene downstream of the DBD domain.

The protein of interest (X) is expressed as a fusion protein to the DBD (DBD-X; also known as the “bait” protein) and the activation domain is fused to the second protein of interest (Y), (AD-Y; also known as the “prey” protein).



If protein X and protein Y interact, then their DNA-binding domain and activation domain will combine to form a **functional transcriptional activator (TA)**. The TA will then proceed to transcribe the reporter gene that is paired with its promoter.

In this way I can monitor the interaction, at the end I measure the expression of the reporter.



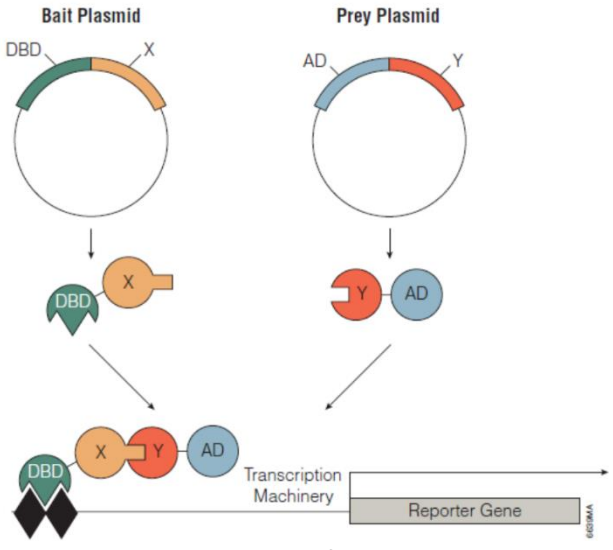
This is a primary and easy way (most of the methods rely on this principle); we can use plasmid--> The 'bait' DNA is isolated and inserted into a plasmid adjacent to the GAL4 BD DNA. When this DNA is transcribed, the 'bait' protein will now contain the GAL4 DNA-binding domain as well. The 'hunter' fusion protein contains the GAL4 AD.

*TRP1 and LEU2 genes are needed to select the cells, empower the cells for growth without essential amino acid (tryptophan and leucine, respectively). We can't use antibiotics as in bacteria because we are using yeast.

We test the interaction of both N or C terminal fusions, it's a trial-and-error method, we have to test both possibilities.

Yeast two-hybrid system: formats

There are several varieties of the yeast two hybrid system. The two most commonly used systems **differ in** the nature of the **DBD used to express the bait** fusion protein (**GAL4 or LexA**), and **the AD used to generate the prey** fusion protein (**GAL4, VP16 or B42**).



<u>When to use the yeast two-hybrid system?</u>

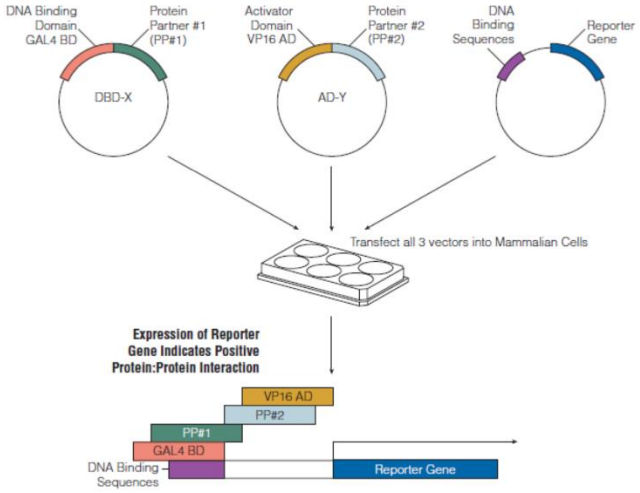
The yeast two-hybrid system is often **the first method used to identify protein interactions**. It provides an ideal format for screening one individual bait protein against large prey cDNA libraries, which can be generated by the user or purchased from commercial sources.
Once putative partners have been identified, other methods are used to confirm and characterize the pair.

Primary reagent requirements for the yeast two-hybrid system:

- Yeast expression vector containing **DNA binding domain** (e.g., GAL4) and sequence coding for the **bait protein**
- Yeast expression vector containing **activator domain** (e.g., VP16) and sequence coding for the **prey protein**
- Appropriate yeast strains
- Appropriate media for yeast growth and detection of interactions

In some cases we rather use mammalian cells, for doing exactly the previous thing.

1. One plasmid contains a transcriptional activation domain upstream of coding sequences for the prey protein (AD-Y).
2. The second vector contains a DBD upstream of coding sequences of the bait protein (DBD-X).
3. The third vector contains five DNA binding sites upstream of a minimal TATA box (promoter), which is upstream of a specific reporter gene.



Why mammalian cell lines?

- In yeast we have PTMs, but the pattern may be different.
- Because not all the interactions occur in the cytoplasm, but also in other organelles, for this reason sometimes it is better to use human cells or mammalian cells.
- In these cells there are all the actors needed, while in yeast cells we have to co-transform to express the different factors.
- Another advantage of the mammalian two-hybrid system is that **the assay is less time-consuming than the yeast two-hybrid system.** Instead of waiting 3–4 days for yeast colonies to grow to a reasonable size for a blue-color assay, typical reporter assays in the mammalian system can be performed within 48 hours of transfection.

For multiple partners interactions, we need to use mammalian cells. Otherwise the interaction will not take place anymore.

The mammalian two-hybrid system allows characterization of mammalian protein-protein interactions within a cellular environment that mimics native conditions. **Yeast and mammalian cells differ in patterns of post-translational modification**, such as glycosylation, phosphorylation and acylation, as well as in the intracellular localization of proteins. These types of protein modifications, as well as other unique factors or modulators present in mammalian cells, **may influence the ability of protein domains to interact**.

Thanks to the reporter gene technology, we get an amplification of the signals, it means that we can detect even weak interactions. Depending on the reporter gene used, the results will be different.

The choice of the reporter gene is not easy, because sometimes you want to have an increased sensitivity, and some others you want to have a very easy signal, such as colorimetric signal, in which you have your agar-plate with single clone and so you want just to visualize it. Therefore you can select different reporters with different detections, including colorimetric, fluorescent and luminescent. As general rule **the colorimetric detection is the less sensitive.** Most of the essays use luciferase because is the most sensitive. The following image shows the reaction catalyzed by the

firefly Luciferase: it requires ATP and oxidates a small molecule that is luciferin, so we have an adenylation followed by an oxidation with the production of photons.

So if you want to get an high sensitivity, you can use this reporter gene. In some cases there are different forms of the same reporter that can be used and in some cases it's an advantage to use the destabilized forms. The last one is for studying the dynamics of different promoters.

$$\mathrm{Luc} + \mathrm{LH_2} + \mathrm{ATP} \overset{\mathrm{Mg^{2+}}}{\leftrightarrows} \mathrm{Luc \cdot LH_2\text{-}AMP} + \mathrm{Ppi},$$

$$\mathrm{Luc \cdot LH_2\text{-}AMP} + \mathrm{O_2} \rightarrow \mathrm{Luc \cdot Oxyluciferin^*} + \mathrm{AMP} + \mathrm{CO_2},$$

$$\mathrm{Luc \cdot Oxyluciferin^*} \rightarrow \mathrm{Luc \cdot Oxyluciferin} + h\upsilon.$$

The three most commonly used reporter genes are:

- **Luciferase**
- **β-galactosidase**: the most used as a first screening. Thanks to the use of different substrates, it is very easy to detect with fluorescent, chemiluminescent and colorimetric detection methods. The fluorescent and chemiluminescent methods are 20- to 1000- fold more sensitive than the standard colorimetric method.
- **Secreted Alkaline Phosphatase** (SEAP): in other cases it's more useful to use reporter genes that are expressed outside the cell, so the ones that are secreted. Sometimes we have the same or a similar protein (that is endogenous) in the sample-->an example: for alkaline phosphatase it is used a thermostable enzyme, in this way it is easier to denature the endogenous enzyme just leaving that into the sample at 65°C for 20 minutes and then measuring the stable one. SEAP activity can be measured using chemiluminescence or fluorescence detection methods.

<u>When to use the mammalian two-hybrid system</u>

In conjunction with other techniques, the mammalian two-hybrid system is used to characterize protein-protein interactions in a true mammalian environment. Due to the format, this technique is not usually used to screen large numbers of prey proteins. Usually the interaction is NOT proportional to the production of recombinant protein: usually acts as a trigger, all-or-nothing.

Oral exam:

Know only the general idea of the different methods

Co-immunoprecipitation:

One of the most common and rigorous demonstrations of protein-protein interaction is the co-immunoprecipitation of suspected complexes from cell extracts.

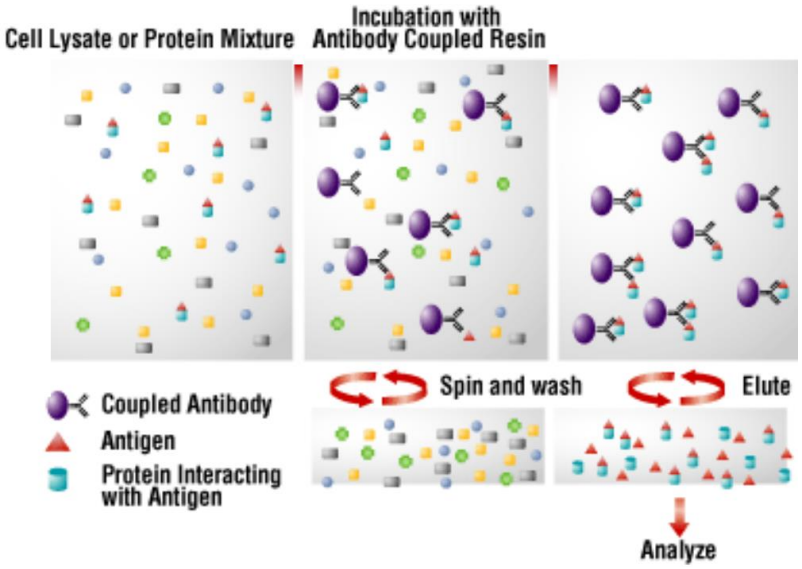
Co-immunoprecipitation (CoIP) confirms interactions utilizing a whole cell extract where proteins are present in their native conformation (they can't be denatured because to target a protein with an antibody, that protein should be in the correct folding, otherwise the antibody will not be able to bind the protein anymore) in a complex mixture of cellular components that may be required for successful interactions. In addition, **use of eukaryotic cells** enables post-

translational modification which may be required for interaction and which would not occur using prokaryotic expression systems.

The basic assumption is that, if I use antibody that is specific to an antigen, all the proteins that will be able to bind that antigen will bind the antigen-antibody complex.

In a typical experiment, **cells are lysed and a whole cell extract is prepared under non-denaturing conditions.** It is critical to use non-denaturing conditions in order to maintain any interaction that occurs. **A bait-specific antibody** is then added to the extract, forming a new complex. **This protein-protein complex is then immobilized on protein A or protein G sepharose beads. Proteins that do not bind are removed by a series of washes.** The protein complex is then eluted from the beads and dissociated by SDS sample buffer. Samples are then evaluated by SDS-PAGE followed by Western blotting with specific antibodies for the bait or prey partners. If the complex is not stable, it could be eluted after the washing.

Co-IP is the easiest method to perform if you have high quality antibody, with high affinity towards your antigen. There are usually some problems of linking from the beads, so the complex is not so stable and during the washing steps, in some cases, we have also the elution of the complex.

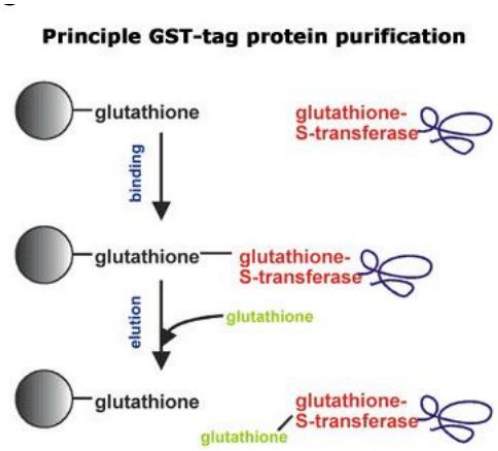


The assumption that is usually made when associated proteins are coprecipitated is that **these proteins are related to the function of the target antigen at the cellular level.** This is only an assumption, however, that is subject to further verification.

Pull-down assays

Pull-down assays are a form of affinity purification and are **similar to co-immunoprecipitation**, except that a **"bait" protein** is used instead of an antibody. The principle is the same, but we use the GST (Glutathione sulfo-transferase) which is bound to the protein and beads with the tripeptide glutathione. It is based on the same principle of purification: we will have the binding between Glutathione and **GST tag, that is genetically linked to the bait protein**. This is followed by immobilization on particles that contain reduced glutathione which binds to the GST tag of the fusion protein.

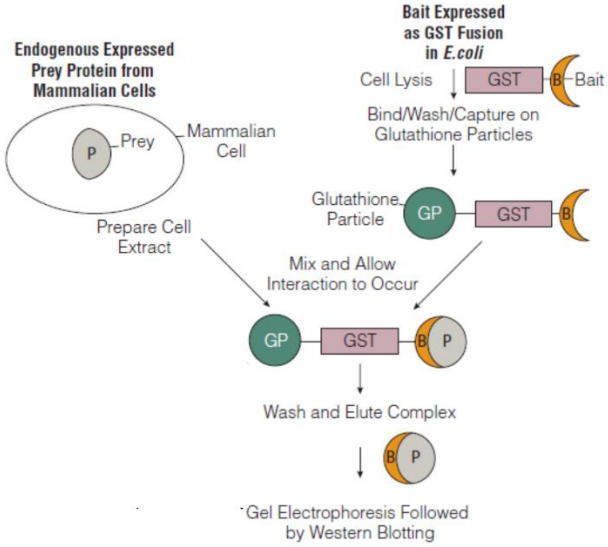
- The primary advantage of a GST tag is that it can increase the solubility of insoluble or semisoluble proteins expressed in E. coli (GST acts as a chaperone to facilitate protein folding).



Pull-down assays probe interactions between a protein of interest that is expressed as a fusion protein (e.g., bait) and the potential interacting partner (prey). In a pull-down assay one protein partner is expressed as a fusion protein (e.g., bait protein) in E. coli and then immobilized using an affinity ligand specific for the fusion tag.

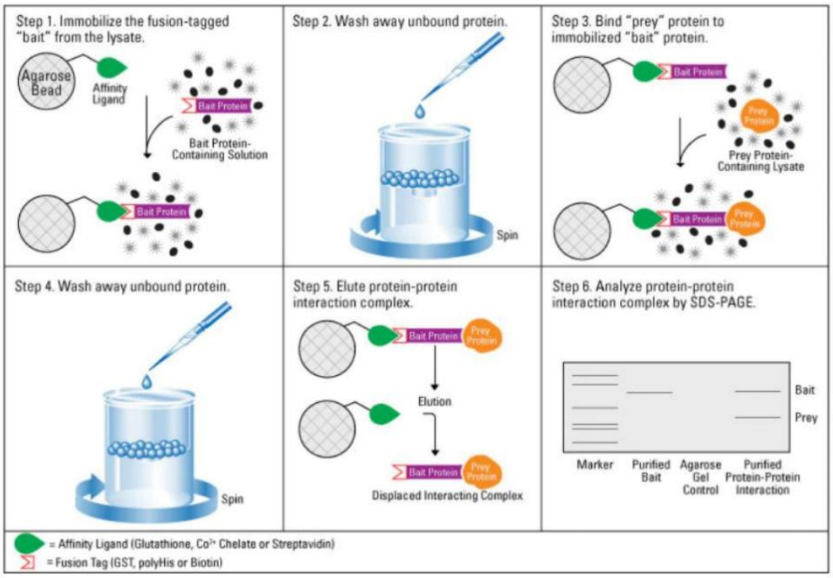
The immobilized bait protein can then be incubated with the prey protein. After a series of wash steps the entire complex can be eluted from the affinity support using competitive analytes, low pH or reducing buffers, then evaluated by SDS-PAGE.

Successful interactions can be detected by Western blotting with specific antibodies to both the prey and bait proteins, or measurement of radioactivity from a [35S] prey protein.



[Description of the last image] The sequence of the bait protein is cloned into a vector containing a GST tag. Following expression, the GST bait fusion protein is purified **using glutathione particles, which bind to the GST tag.** A cellular extract is prepared from mammalian cells containing the prey protein. An aliquot of the cell extract is then allowed to interact with the bound GST bait fusion protein for several hours. After washing away non-specifically bound proteins the complex is eluted by adding reduced glutathione and analyzed by Western blotting (also by mass spectrometry).

This is not for monitoring protein-protein interaction. Thanks to this we can increase the solubility of proteins expressed in E.coli.



PROTEIN ARRAYS

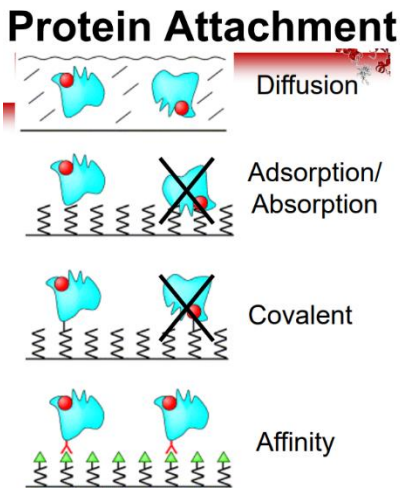
Another possibility is to exploit protein arrays. They are very used for monitoring the interactions between proteins and other proteins or proteins and other ligands, because we can functionalize substrate with a number of different chemical groups or biomolecules.

In the literature you can find basically with the same meaning both arrays and chips. The main difference is that chips rely on microfluid technology.

What can we immobilize into microarrays? We can have the capturing antigen, for example an antibody. This is the easiest way, in which we simply analyze the target proteins that are captured by antibody. Or we can immobilize other proteins to monitor their interactions; we can also immobilize much more complex samples.

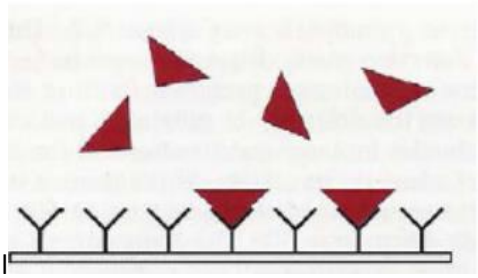
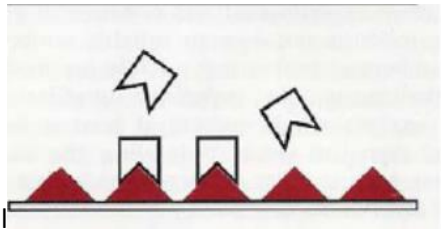
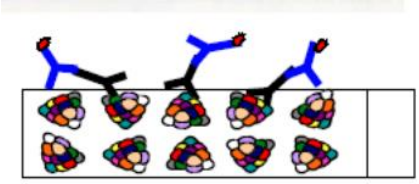
In principle we have several protocols to obtain protein chips in which we can immobilize either the protein or the interacting analytes with a number of different substrates. The problem is that **proteins are delicate, so when you immobilize the protein, most of the time you will lose the folding** (if you work with an enzyme you will lose its catalytical activity) and in this way the protein will be no more functional and probably the binding with other proteins will not occur anymore. For this reason the software approaches are more favorable because you can keep your protein with the right folding. The covalent approaches are less used because they are more invasive while with the physical absorption approaches you don't have any control about the orientation of your protein (it could attach with a part that is required in the interaction).

Most of the cases you will opt for physical absorption method.



In some cases, if we have antibodies and we know where the antibodies bind, the affinity method would be the best option.

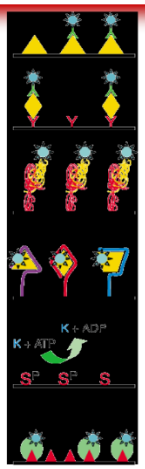
ARRAYS

- Analytical 
- Functional 
- reverse 

PROTEIN INTERACTION--> also with other molecules, different interactions require different capture molecules.

Protein Interactions

- Different capture molecules must be used to study different interactions
- Examples
  - Antibodies (or antigens) for detection
  - Proteins for protein-protein interaction
  - Enzyme-substrate for biochemical function



Antigen–antibody

Protein–protein

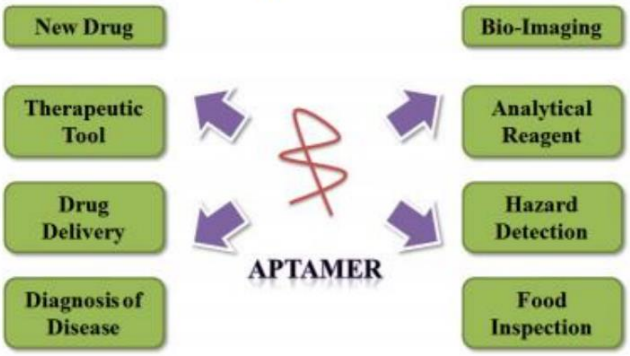
Aptamers*

Enzyme-substrate

Receptor–ligand

APTAMERS: Aptamers are oligonucleotides, such as ribonucleic acid (RNA) and single-strand deoxyribonucleic acid (ssDNA) or peptide molecules that can bind to their targets with high affinity (not as high as antibody but very high) and specificity due to their specific tridimensional structures (not oligonucleotides binding). Especially, RNA and ssDNA aptamers can differ from each other in sequence and folding pattern, although they bind to the same target.

Figure 1. Various application fields of aptamers.

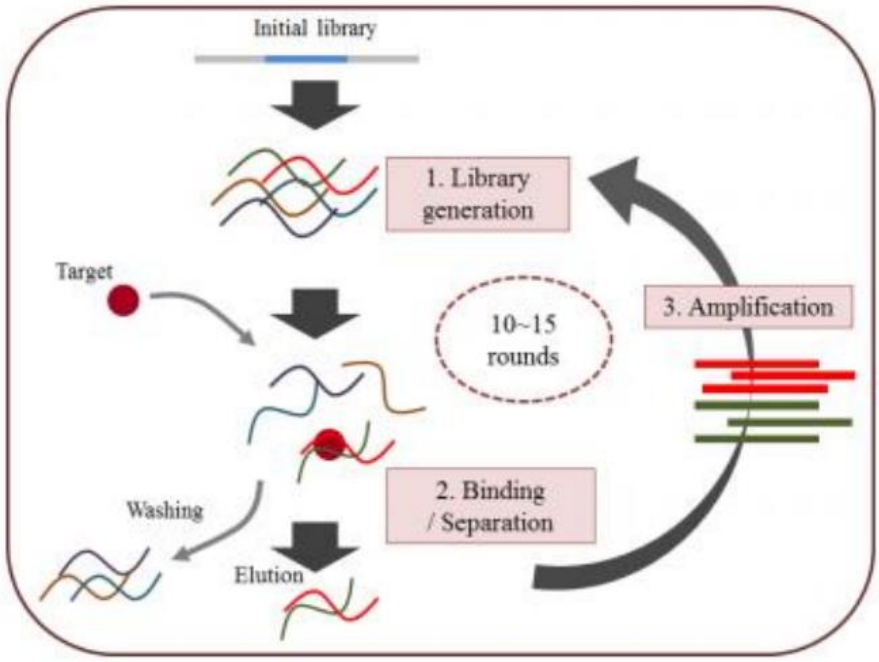


They are used to replace antibodies; for detecting target analytes (for example for food quality control); to capture toxic agents. Why these oligonucleotides interact with a target molecule that has no DNA structure? The reason behind this affinity (an affinity not present in nature, because in nature oligonucleotides are not used to recognize a target, except for transcription factors). This is used by recreating in the lab what happened during millions of years of evolution. How is it possible?

- [image below] Let's start with a library of random mixture of DNA molecules (shown in the picture with different color of the strands)
- We mix this library with a target and incubate them
- Wash everything and collect only oligonucleotides that bind the target
- **Amplify these positive oligonucleotides**--> second library obtained
- Second library is the starting point for a second round of selection and amplification (in every round I will eliminate the oligonucleotides that don't bind the target: this is what mimic in some way the evolution, because in every round of amplification we have a sort of "in vitro evolution", choosing just those that are able to bind the target)
- We perform the same procedure for 10-15 rounds

This is a very powerful approach to obtain non-natural molecules that have very high affinity to certain targets. This technique is very used also in other approaches (such as Phage Display).

With DNA the protocol is faster.



Issues:

- You could eliminate something that is important.
- We are in vitro without a buffer and the condition may vary from the physiological condition (because maybe that high affinity could be present just in that condition, and in the physiological conditions the affinity could change) --> when you go back to a practice and try to use this oligo for analysis, it may not be able to bind the target because the conditions are different.

Peptide nucleic acids (PNAs):

PNAs are synthetic mimics of DNA in which the deoxyribose phosphate backbone is replaced by a pseudo-peptide polymer to which the nucleobases are linked.

PNAs hybridize with complementary DNAs or RNAs with remarkably high affinity and specificity, essentially because of their uncharged and flexible polyamide backbone.

Main advantage of this molecules--> **they are not present in nature so they can't be recognized neither by proteases and nucleases.**

With these molecules we can build expression or interaction arrays in which we can:

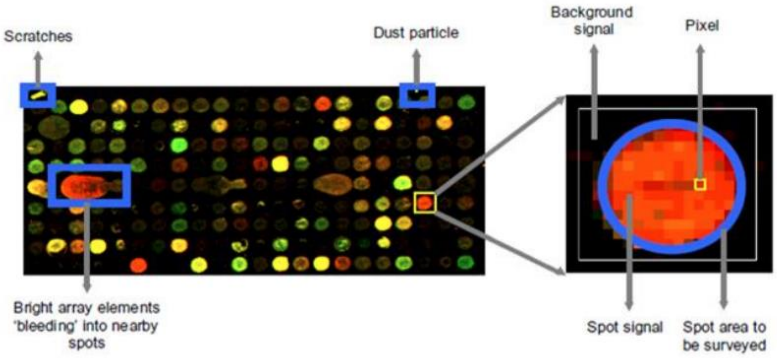
With **Expression Array**

- Probes (antibody) on surface recognize target proteins.
- Identification of expressed proteins from samples (like in SELDI)

- Typical quantification method for large number of expressed proteins.

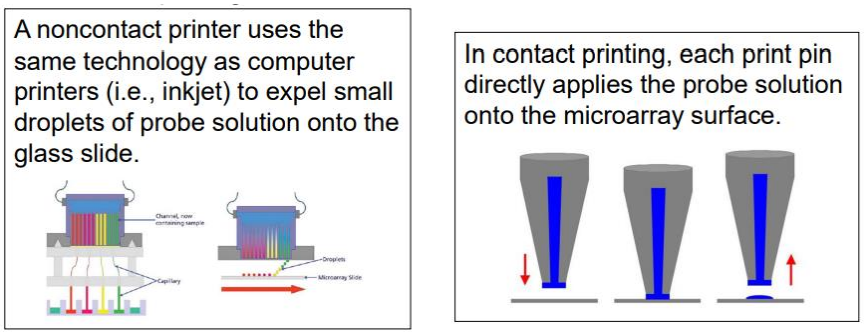
With **Interaction Arrays**

- Probes (proteins, peptides, lipids) on surface interact with target proteins.
- Identification of protein interactions.
- High throughput discovery of interactions.



However the practice is quite different: there are different issues related to the very unscaled control of orientation of proteins. As previously said, when you immobilize a protein you lose folding, 50-60% of its catalytical activity and the proteins are most likely denatured.

For these kind of arrays everything is performed via robotic dispensing. The image is an example of robotic printing, a quick overview in which we can have printers that can be **contact or noncontact.** They create arrays with very high density.



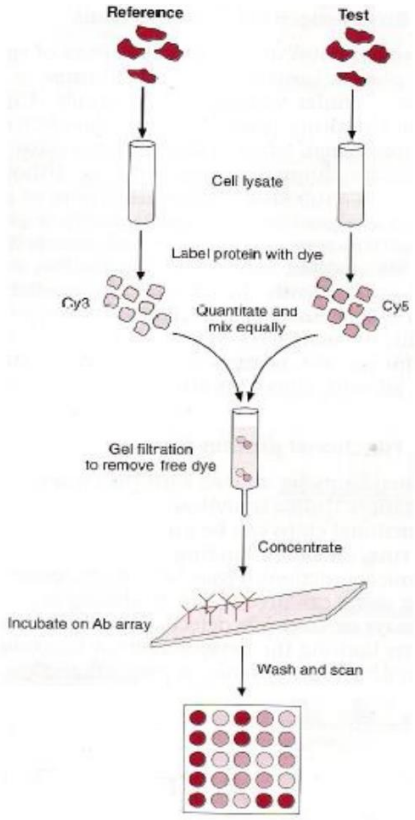
The result in both cases is the application of a few nanoliters of probe solution per spot to create an array of 100- to 150µm features

<u>Noncontact vs contact printing</u>

- Both contact and noncontact ink-jet printing allow the spotting of virtually any biological molecule of interest, including cDNA, genomic DNA, antibodies and small molecules.
- In contrast to contact printing, ink-jet printing avoids direct surface contact, introducing surface-contact feature anomalies, and resulting in consistent spot uniformity and traceability.
- They can't manufacture as dense as those prepared by photolithography.

One of the basic applications of these arrays is the **"Differential profiling",** in which you try to identify a different signature from two samples (e.g. normal vs disease). Proteins are labelled with dye (Cy3, Cy5) so we can incubate them

with antibodies, and then wash and scan. The algorithms for the signal analysis are crucial because depending on the cut-off you choose, you will obtain very different results.



Surface-enhanced laser desorption ionization (SELDI)-TOF MS

The same approach is used in SELDI in which we use this kind of chips to enrich our proteins. SELDI-TOF MS is an innovative microarray approach and offers on-chip purification of unlabeled target proteins (through on-chip chemical and/or biological affinity chromatography) followed by subsequent ionization and MS detection of the retained molecules.

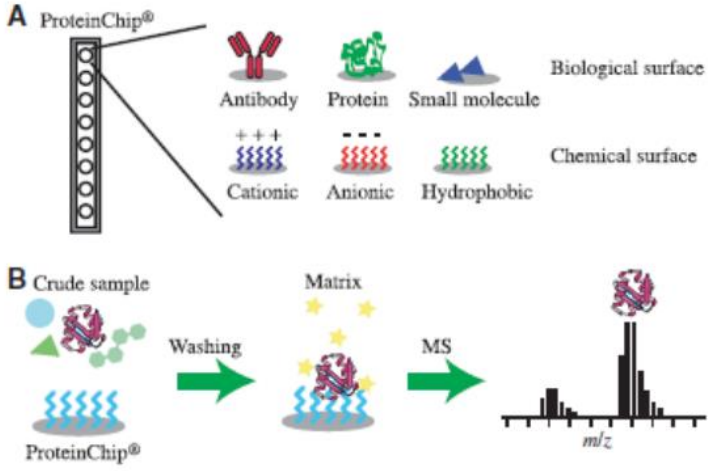


Fig. 5. (A) The types of ProteinChip. (B) The detection process using SELDI-TOF MS technology [55].

PHAGE DISPLAY

How can it be used for protein-protein interaction?

We are using phages (virus that affect bacteria): on the coat surface of a phage we express peptides or small proteins as a fusion with a phage coat protein.

We use the infection of phages as an amplification, producing library of phages each one exposing different peptides on the coat--> and we can infect bacteria cells and in this way we obtain a library of cells that are infected with different phage clones, so each phage will expose different peptides.

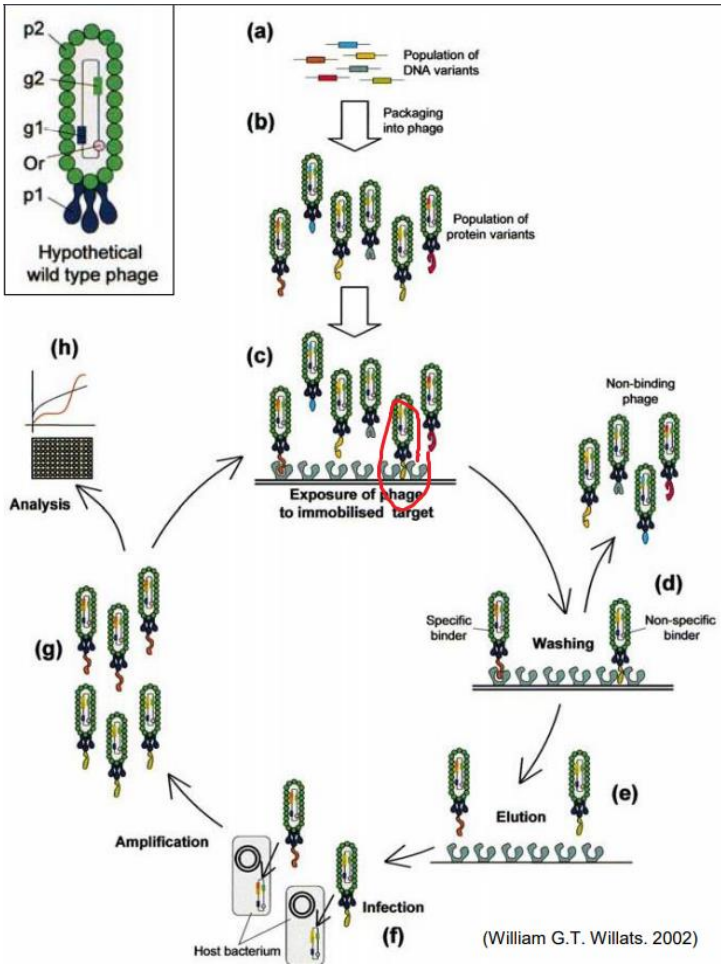
E. coli phages used due to ease of culture and quick regeneration. E. coli can be infected to multiply the number of bacteriophages. It can **quickly create large libraries of phage clones displaying different peptides.**

Steps for developing phage display libraries:

- Creation of vector (plasmid)
- Binding/Selection
- Washing steps
- Elution
- Amplification

We start with a library of DNA variants encoding for different peptides and we put them into phages; we obtain a population of phages, each one expresses a different protein on the external surface; we expose this population of phages to an immobilized target (a protein, a small organic molecule, whatever) and let them incubate; then we perform washing steps to eliminate all the phages that were not able to bind the target; we elute and just recover the “positive phages”, we let the phages infect the bacteria cells and then perform the amplification. In this way we enrich the library with only the phages that are able to bind our target.

At this point we repeat the steps: we incubate the new library phages with the immobilized target and so on. Perform this for 10-20 rounds. The yellow phage is binding non-specifically with the surface and will be discarded with the new cycles. Hopefully there will be just the red phage.



This approach is used in several applications, even for **drug discovery** to identify new targets, new ligands; to create **new antibodies libraries; Epitope discovery** – new vaccines (use antibodies as a receptor to select peptide that is an antigen mimic, can bypass animal immunization by mimicking immune selection).

There are also phage displays with **nanobodies.** found in camelids

Nanobodies (or single-domain antibodies) are antibody fragments consisting of a single monomeric variable antibody domain. Like a whole antibody, a nanobody can bind selectively to a specific antigen. With a molecular weight of only 12–15 kDa, single-domain antibodies are much smaller than common antibodies (150–160 kDa). They **are easily isolated** using the same phage panning procedure used for antibodies, allowing them to be cultured in vitro in large concentrations. The smaller size and single domain make these antibodies easier to transform into bacterial cells for bulk production, making them ideal for research purposes.

Phage display -advantages:

an antibody is too big for a phage, but a nanobody is small enough!

- Easy to screen large library of clones >10^9
- Easy to amplify selected phages in E. coli
- Selection process easy and already in use in various forms.
- It can create phage library variation by inducing mutations, using error prone PCR, etc

Limitations:

- Might not have long enough peptide insert so **critical folding can be disrupted** (we can't express big proteins)
- Could lose phage variations if first bind/wash step too stringent: we need to check that the same conditions apply also to the practical uses
- Affinities or binding that results during selection might not work in vivo

Method	Advantage	Disadvantage
Co-immunoprecipitation	Independent of cloning and ectopic gene expression Rapid procedures	Cross-reactivity of antibody Antibody bleeding from column
GST-pull down assays	Applicable to very weak protein interactions	Complex formation in-vitro Competition with in-vivo pre-assembled complex
Protein arrays	High-throughput assay Disease diagnosis	Difficulty of protein chip production
Yeast two-hybrid system	Highly sensitive detection Applicable to a wide range of protein interactions No biochemical purification	Stability of folding and activity in yeast Not post-transcriptional modification
phage display	Random library screening of many cDNAs through panning cycle	Size of limitation of protein sequence Incorrect folding or modification

Oral exam:

- Main advantages and disadvantages

Pubmed---> covers most of the clinical related stuff

Scopus--> multidisciplinary database, contains ''peer reviewed'' journals, books and conference papers-allows citation analysis

Web of science: multidisciplinary platform

They are not comprehensive--> miss the part related to the market SO LOOK ALSO THE PATENT DATABASE: ESPACENET

Impact factor--> the highest the better (most important parameter for judging the journal)

Impact factor per se is not the most important parameter, there are multidisciplinary journals, so you have to look at the quartile (the best journals are in the first quartile) in which the journal is positioned, it depends on the subject you are interested in

H-index: assigned to an author, it is the number *n* of articles cited at least *n* times. For example, an H-index of 50 means that a certain author has 50 articles cited at least 50 times

Scopus will provide the citation score instead of the impact factor.

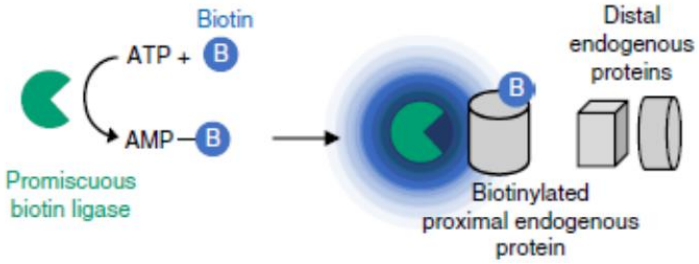
With Scopus you can also sort the publications of the researcher by the most cited, and see which articles contributed the most to the h-index

Look always at patent, relevant for discover what we have already done in the field--> sometimes they provide more info than the publication (if you don't provide sufficient details the patent will be rejected)

With espacenet you can find all the patents, not only the europeans. You can see also the storylines of the patent (if the patent expired ecc.)

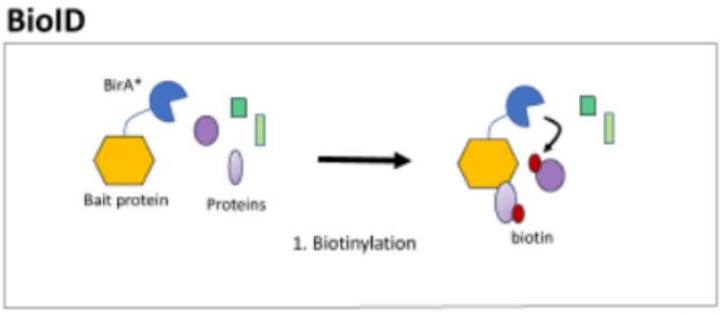
PROXIMITY ASSAYS (interactions usually occur in a 1-10 nanometers between two proteins for instance)

Enzyme-catalyzed proximity labeling (PL) has emerged as a new approach to study the spatial and interaction characteristics of proteins in living cells. In PL, a promiscuous (= it binds any protein that is in few nanometers of distance) labeling enzyme is targeted by genetic fusion to a specific protein (bait) or subcellular compartment. Addition of a small-molecule substrate, such as biotin, initiates covalent tagging of endogenous proteins within a few nanometers (range 1-10 nm) of the promiscuous enzyme. Subsequently, the biotinylated proteins are harvested using streptavidin-coated beads, and identified by mass spectrometry.



Biotin-based proximity labelling

The first biotin-based proximity labelling technique, BioID, was developed in 2012 (Roux et al., 2012). **BioID utilizes the E. coli biotin ligase, BirA*, with a site mutation (R11G)**. This site mutation destabilizes the ligase, allowing for active biotin molecules to dissociate from the ligase and bind to exposed lysines on neighboring proteins. Thus, to capture protein complexes, scientists can simply express a bait protein fused to BirA* in cells and then incubate those cells with exogenous biotin. After a few hours biotinylated proteins can be captured on streptavidin affinity beads or matrices. The bait protein can even target BirA* to specific subcellular locations, such as the nuclear envelope, allowing scientists to probe interactions in locations of interest.

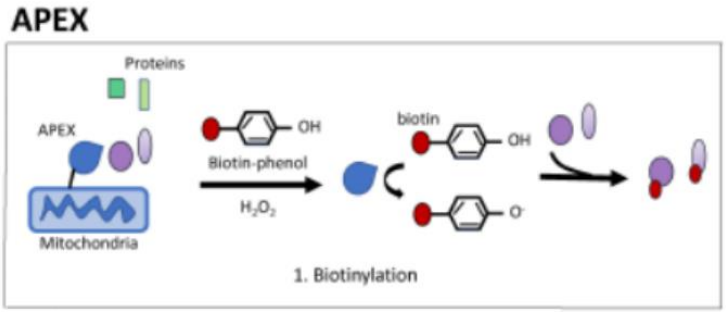


Problem: first configuration of this assay the biotinylating requires hours

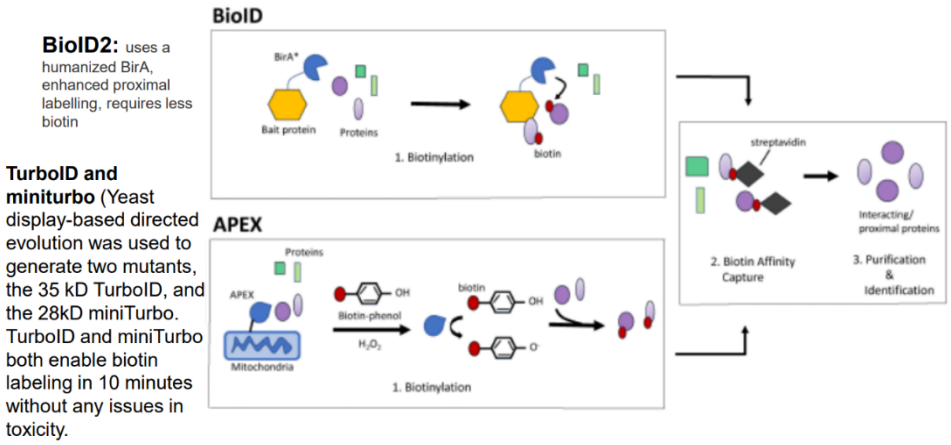
Alternatives: APEX

APEX: engineered ascorbate peroxidase that can rapidly biotin label proteins

Engineered ascorbate peroxidase (APEX), used as a genetic tag for electron microscopy, can be used for efficient proximity labeling (Rhee et al., 2013). APEX catalyzes the oxidation of **biotin-phenol** to the shortlived biotin-phenoxyl radical in the presence of **hydrogen peroxide**. This radical reacts with electron-rich amino acids such as tyrosine on neighboring proteins resulting in their biotinylation. The advantage of APEX is its speed, labeling neighboring proteins in **minutes** rather than hours like BioID. The biotin-phenol reagent can be toxic to living cells thus APEX is not suitable for organoid or in vivo studies.



Biotinylating with a different method:



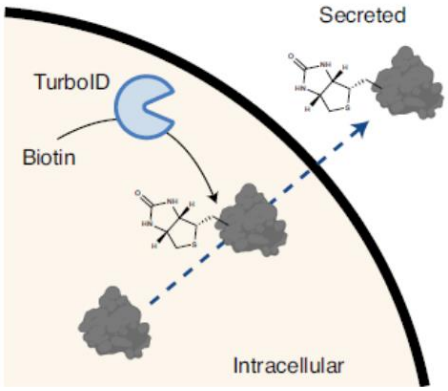
There are plenty of different applications, among them:

- Proximity biotinylation for secretomics
- Peroxidase-catalyzed proximity labeling to track GPCR G-protein-coupled receptors (there are different proteins interacting with GPCR, and we can biotinylate all of them; this is not a binary interaction)

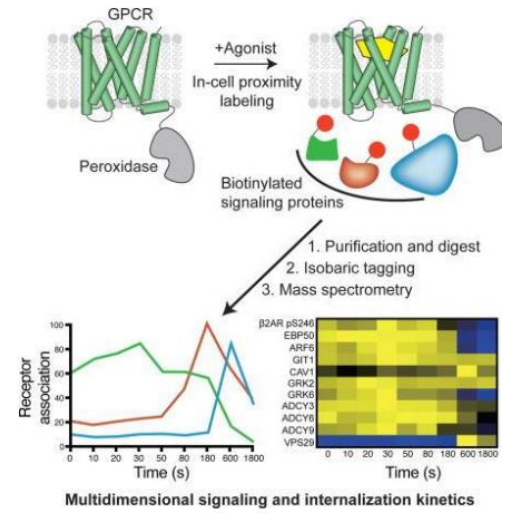
Proximity biotinylation for secretomics

A proximity biotinylation strategy that enables labeling, detection and enrichment of secreted polypeptides in a cell type-selective manner in mice was described and a proteomic atlas of hepatocyte, myocyte, pericyte and myeloid cell secretomes by direct purification of biotinylated secreted proteins from blood plasma was obtained.

This secretome dataset validates known cell type-protein pairs, reveals secreted polypeptides that distinguish between cell types and identifies new cellular sources for classical plasma proteins.



Peroxidase-catalyzed proximity labeling to track GPCR



GPCR is highly hydrophobic and will never go in solution, so this approach can solve this:

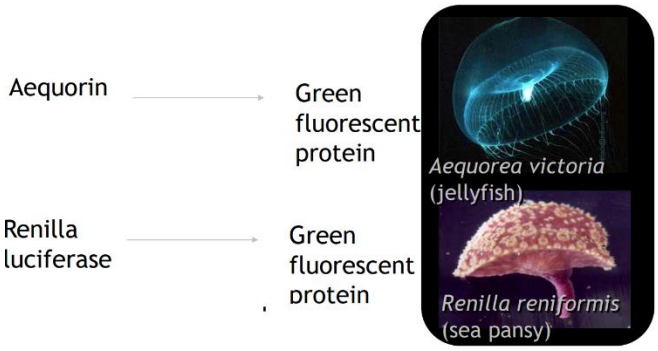
We have different varieties in these approaches, where we look into the ones that can perform biotinylation in a more rapid way.

RET: Resonance Energy Transfer–based techniques to understand protein chatter

A useful approach for studying protein-protein interactions is based on Resonance Energy-Transfer Phenomenon

- **BRET** (Bioluminescence Resonance Energy Transfer) --> Bioluminescence (emission of light that occurs in living species) donor and fluorescent receptor.

BRET is a naturally occurring phenomenon in marine environment whereby a **light-emitting protein** non-radiatively transfers energy to a **fluorescent protein** (protein within 1-10 nanometers of distance).



- **FRET** (Fluorescence Resonance Energy Transfer) --> if the two proteins are close, there is a transfer through the dipoles that results in the light emission of the receptor.

They both work in a range of 100 angstroms.

RET (RESONANCE ENERGY TRANSFER)

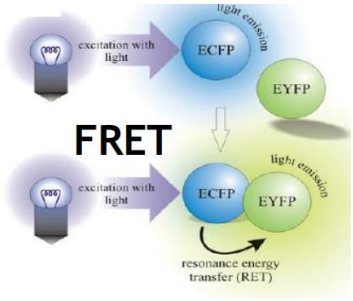
Energy transferred from a donor to an acceptor molecule in a non-radiative manner as a result of dipole–dipole coupling.

→ If the two proteins are far apart, if we excite at the wavelength of absorption of the donor, we'll just see the emission coming from the donor.
→ If the two proteins are close within a hundred Å, by exciting the donor, there will be the emission of the acceptor, since a transfer between the dipoles of the two molecules has occurred = the distance is below a certain limit, therefore there is an overlapping of the emission spectrum of the donor and of adsorption of the acceptor and a correct orientation of the dipoles.

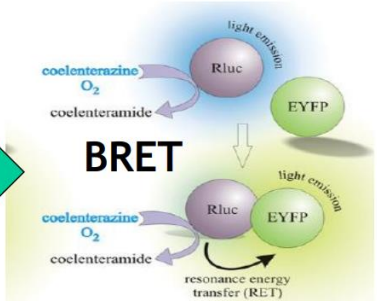
If the donor is a fluorescent molecule (FRET), exposure to light of a characteristic wavelength will result in excitation with a subsequent energy transfer to a fluorescent acceptor molecule.

In BRET, instead of providing a light source, we will add a chemical compound that will be oxidated by the donor. If the donor and the acceptor are close to each other, there will be the emission of the acceptor, which is fluorescent and since we haven't provided any light source, we can be pretty sure that the emission is due to an energy transfer (no fluorescent protein can emit fluorescent without excitation).

If the donor is a fluorescent molecule, exposure to light of a characteristic wavelength will result in excitation with a subsequent energy transfer to a fluorescent acceptor molecule.



Alternatively, the donor molecule can be a luciferase which causes energy to be released as a light upon oxidization of a suitable substrate. Resultant energy transfer to a fluorescent acceptor molecule is then referred to as BRET.



RET efficiency: depends on several factors described by the Förster rate equation (do not have to remember the equation)

$$E = \frac{1}{(1 + (r/R_0)^6)}$$

Ro is the Förster distance of donor and acceptor at which the RET efficiency is 50%

$$R_0^{\,6} = 8.8 \times 10^{23}\ \kappa^2 n^{-4} Q_0 J$$

Where:
κ2 is the dipole orientation factor,
***n* is the refractive index of the medium,**
***Q*o is the fluorescence quantum yield of the donor in the absence of the acceptor,**
***J* is the spectral overlap integral calculated as**

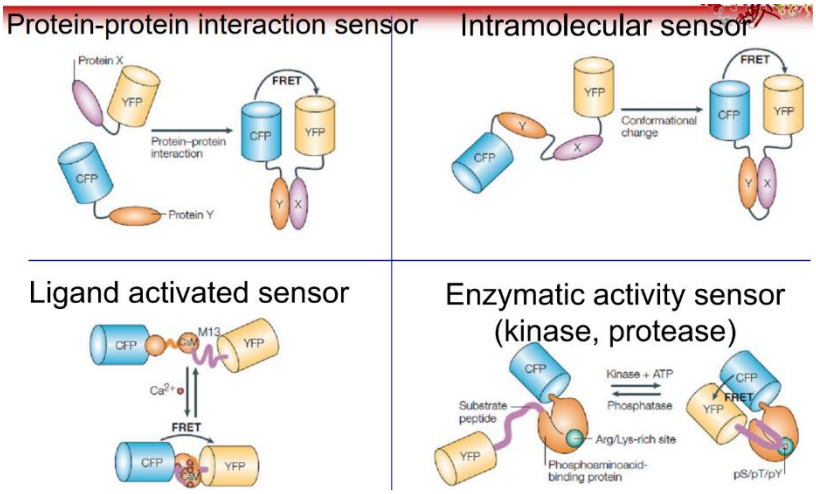
$$J = \int f_D(\lambda)\,\epsilon_A(\lambda)\,\lambda^4\,d\lambda$$

where *f*D is the normalized donor emission spectrum, and εA is the acceptor extinction coefficient.

-

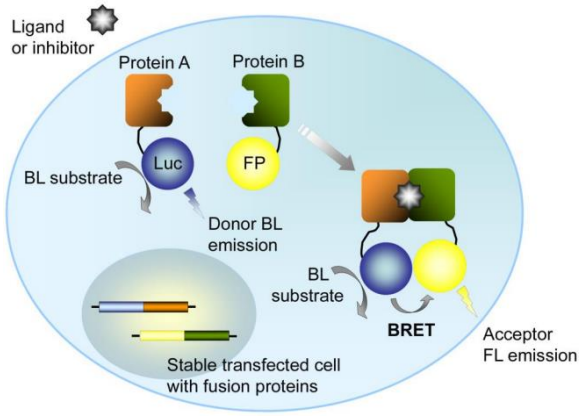
1- Distance: The distance between the donor and the acceptor (typically 1–10 nm); the efficiency is **dependent on the inverse sixth power of the intermolecular separation =>** by increasing the distance of just few nanometers, there is a drop of the efficiency of the energy transfer.
2- The relative orientation of the donor emission dipole moment and the acceptor absorption dipole moment (the orientations must be approximately parallel);
3- The spectral overlap of the donor emission spectrum and the acceptor absorption spectrum.

With energy transfer we can monitor different interactions and molecular events:

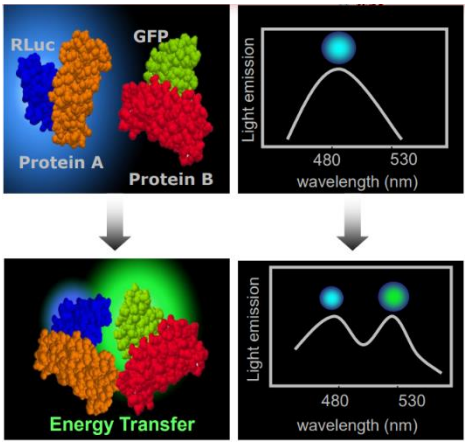


- Intermolecular transfer
- Intramolecular (e.g., different folding of ligand binding domain that get the donor and acceptor to the right distance)
- Ligand activated sensor
- Enzymatic activity sensor (kinase, protease): the donor and the acceptor are in the same fusion protein, in the middle you add the sequence recognized by a protease, so that the protease will recognize this site and cleave it. This is significant since proteases are markers of tumors and of tumor aggressiveness, because they are related to the process of metastasis => you can develop a sensor that can detect the activity of proteases not just their quantity.

Prepare plasmids, protein fused with the donor and another fused with the receptor.



Analytical signal



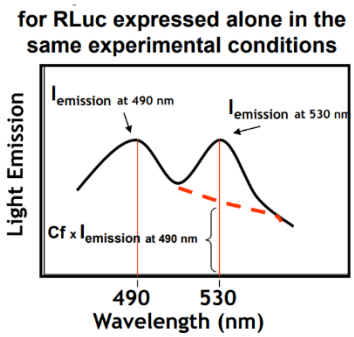
- **No interaction** --> just one peak (of the donor)
- **Energy transfer** --> decreasing of the donor emission and appearing of the receptor emission. The more the energy transfer is, the higher this peak will be.

RET signal is a **ratiometric measurement,** because is the ratio of the intensity of the donor over the intensity of the acceptor, minus a correction factor; every time you have a ratiometric signal, you are normalizing the signal according to experimental conditions (n° cells, temperature, etc, …) => this provides a more robust method, because it eliminates data variability.

BRET signal

$$\text{BRET ratio} = \frac{I_{\text{emission at 530 nm}} - Cf \times I_{\text{emission at 490 nm}}}{I_{\text{emission at 490 nm}}}$$

$$Cf = \frac{I_{\text{emission at 530 nm}}}{I_{\text{emission at 490 nm}}}$$



Emission spectra normalized by defining the donor peak emission intensity as 1.

BRET ratio is calculated as the intensity to which the acceptor emission contributes minus the corresponding area observed when only the donor is present.

FRET vs BRET

- Both allow real-time monitoring of protein-protein interactions in live cells, cell extracts or purified preparations

FRET	BRET
Energy transfer between two fluorophores (need for a light source)	Energy transfer between a luminescent donor and a fluorophore (no need for external excitation)
Photobleaching	No photodamaging to cell
Autofluorescence	No autofluorescence background
Can be monitored using both donor and acceptor fluorophores conjugated to target proteins.	It requires at least the donor t be part of a fusion protein.

The one that doesn't require any light excitation is better simply because there's no background fluorescence and any time you are providing a light source, it could cause damage to the cell (photobleaching).

Both are suitable both for in vitro and in vivo experiments.

BRET applications:

- Protein-protein interactions
- Receptor-ligand binding
- Cell-based assays/Signal transduction
- cAMP activity
- G-Protein coupled receptors
- Enzyme-based assays (kinases, proteases)
- Quantum dots
- BRET imaging from single cells

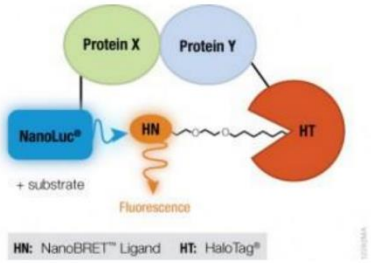
Once you have to design such an experiment and you have a donor and an acceptor that are fluorescent or bioluminescent, you need to perform genetic fusion; although with small organic molecules it could be much easier to conjugate them, you cannot test within living cells, because you can't conjugate protein that are in the cytoplasm, so you can just work with purified proteins.

There is a method that mixes advantages of both sides:

NanoBRET technology (Promega)

The donor is a luciferase (NanoLuc), the acceptor is genetically fused to an HaloTag, a modified haloalkane dehalogenase, which is able to bind to a chloroalkane linker through a nucleophilic attack.

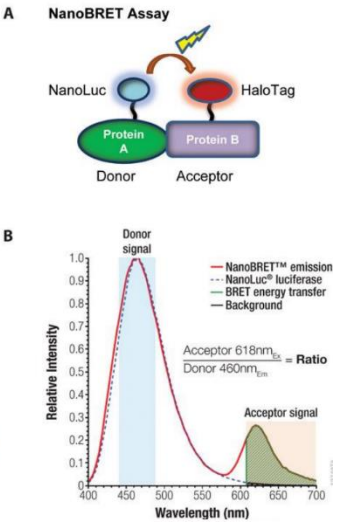
NanoBRET™ assays allow real-time measurement of protein interactions in living cells.



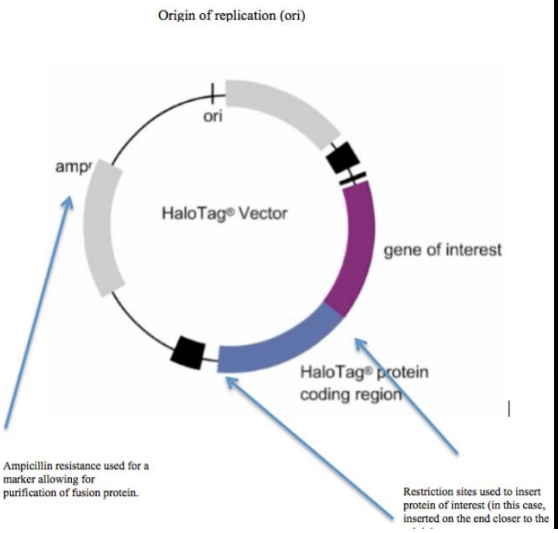
Donor: NanoLuc® luciferase

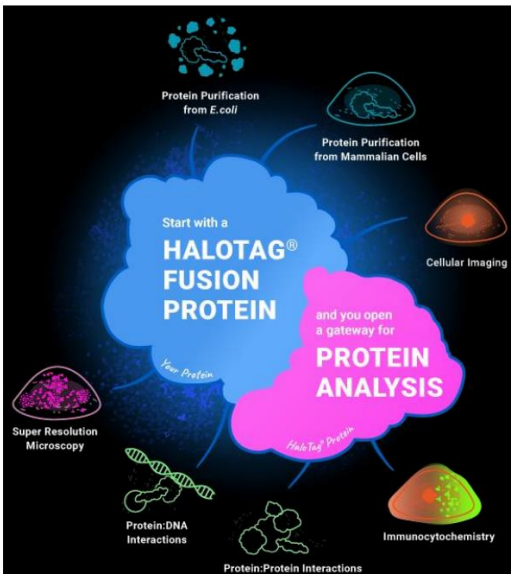
Acceptor: HaloTag® Protein Labeled with the NanoBRET® 618 fluorophore

This assay offers an increased signal and decreased spectral overlap compared to other BRET technologies.



HaloTag: a self-labeling protein tag designed to covalently bind to a synthetic ligand. It is a 297 residue peptide (33 kDa) derived from a bacterial enzyme (haloalkane dehalogenase)

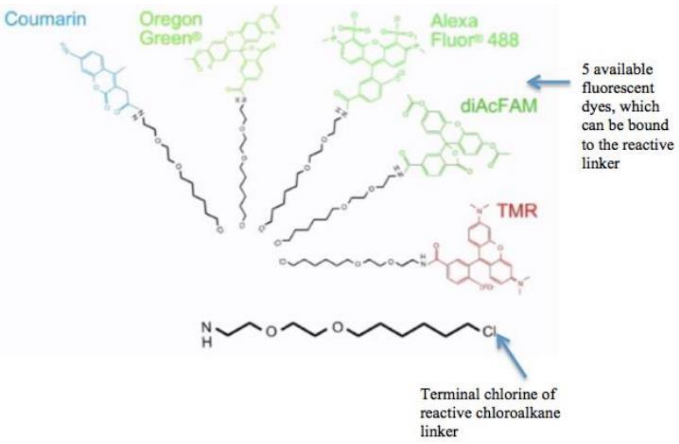




<u>The HaloTag is a genetically modified dehalogenase</u> which binds the reactive chloroalkane linker. You can attach different fluorescent dyes to the reactive linker.

The bond between the protein tag and chloroalkane linker is fast and irreversible under physiological conditions due to the terminal chlorine of the linker portion. A nucleophilic attack of the chloroalkane reactive linker causes displacement of the halogen with an amino acid residue, which results in the formation of a covalent alkyl-enzyme intermediate. This intermediate would then be hydrolyzed by an amino acid residue within the wildtype hydrolase (enzyme regeneration).

But in the HaloTag the reaction intermediate cannot proceed through a subsequent reaction because it cannot be hydrolyzed due to the mutation in the enzyme. This causes the intermediate to persist as a stable covalent adduct with which there is no associated back reaction.

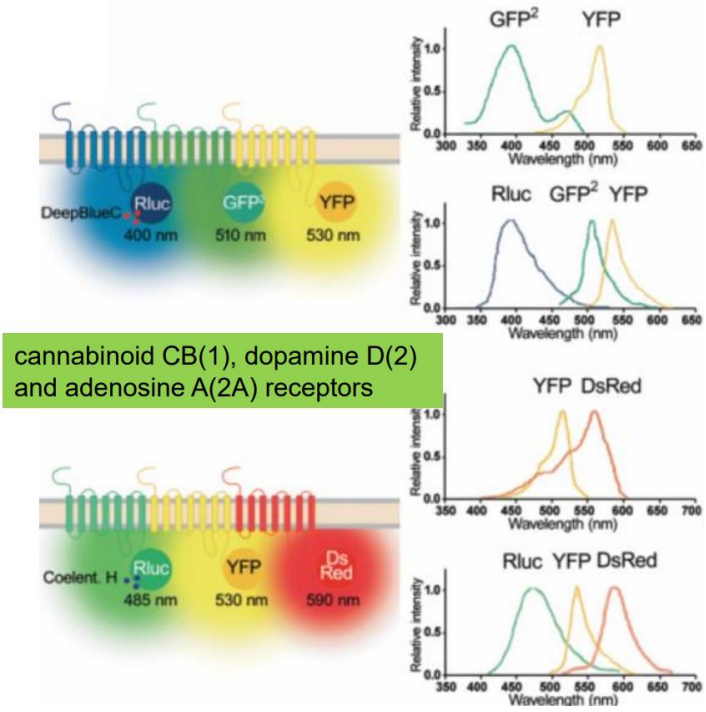


If we have more proteins interacting--> sequential BRET and FRET for monitor multiple interactions

The acceptor will act as a donor for a secondary energy transfer.

- **And for monitoring nonbinary protein interactions?**

Combining BRET and FRET technologies --> **SRET: Sequential BRET - FRET**



The oxidation of a Renilla luciferase (Rluc) substrate by an Rluc fusion protein triggers acceptor excitation of a second fusion protein by BRET and a subsequent FRET to a third fusion protein. SRET provides a valuable technique to identify compound that alter the GPCR–GPCR interactions (homo-oligomerization or hetero-oligomerization).

Limitation of RET technology:

- You can plan your experiment, but you never know if it will work: fusion receptors can act in an artifactual manner and behave in a not reliable way;

- The fused donor and acceptor moieties can cause a steric hindrance thus abolishing the protein-protein interaction.

With living cells, you overexpress fusion constructs to increase the sensitivity--> may lead to artifacts since PPI occurs in non-physiological intracellular environment.

Limitations of FRET

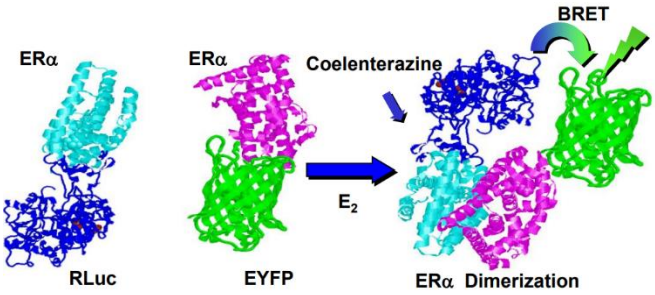
- external light source required
- photobleaching of donor fluorophore
- simultaneous excitation of both the donor and the acceptor
- autofluorescence which occurs within the cells.

Limitations of BRET

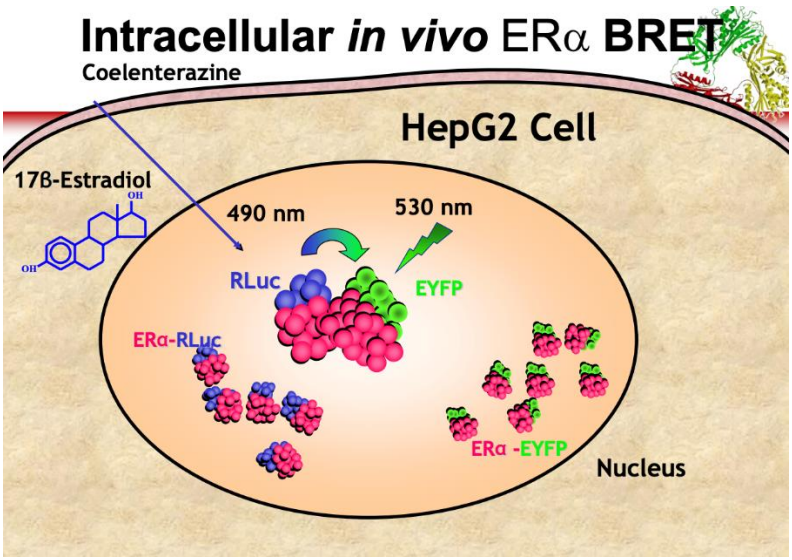
- requires the addition of an exogenous substrate for the BL reaction
- only few donor BL proteins are available (Rluc, Gluc, Nanoluc, …)

BRET ASSAY FOR HUMAN ESTROGEN RECEPTOR α DIMERIZATION

During ER activation, the homodimerization of ERα brings the C-termini of each monomer in the dimer interface with a symmetric 'head-to-head' arrangement--> It is possible to fuse the C-termini of ERα to either a RLuc or a EYFP for a BRET assay.

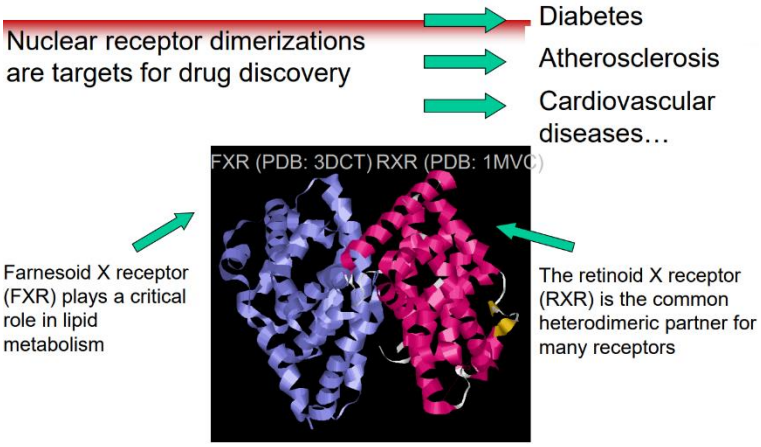


we have expression plasmids

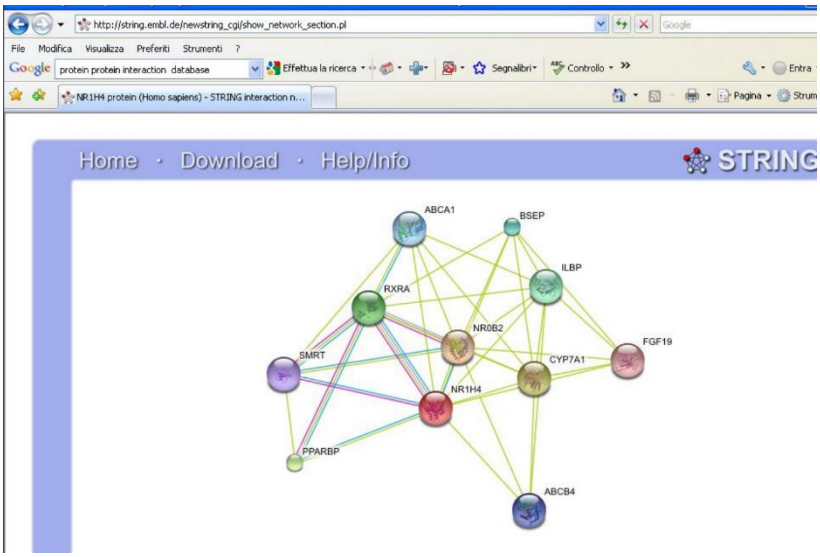


Issues: we could have a homodimerization that won't lead to energy transfer, also you have to select that doesn't express endogenous receptors.

New BRET assays for protein-protein interaction



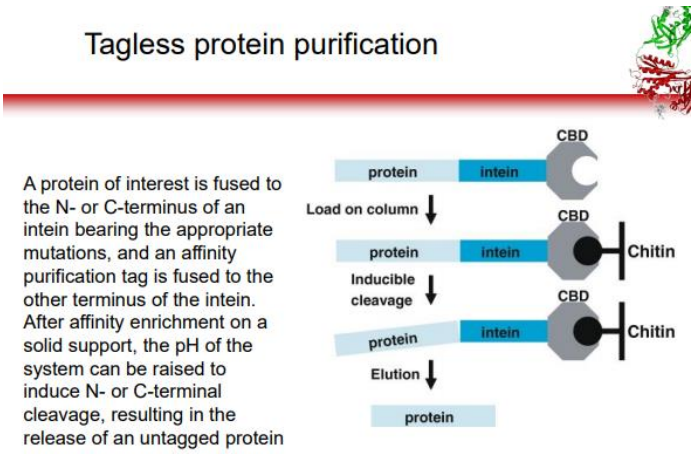
Database **string**: to see what is already available, it contains interactions that have been proved experimentally and those that are predicted (instead of looking in scientific papers)



INTEINS

Another possibility is to exploit **inteins** (proteins that do **self-cleavage**) that can ally to express protein without any tag (tagless protein purification), for example to produce a recombinant hormone.

This rearrangement occurs post-translationally (or possibly co-translationally), as intein genes are found embedded in frame within other protein-coding genes. Furthermore, intein-mediated protein splicing is spontaneous; it requires no external factor or energy source, only the folding of the intein domain.



PCAs: Protein complementation assays

reporter protein= protein that can be quantitatively measured

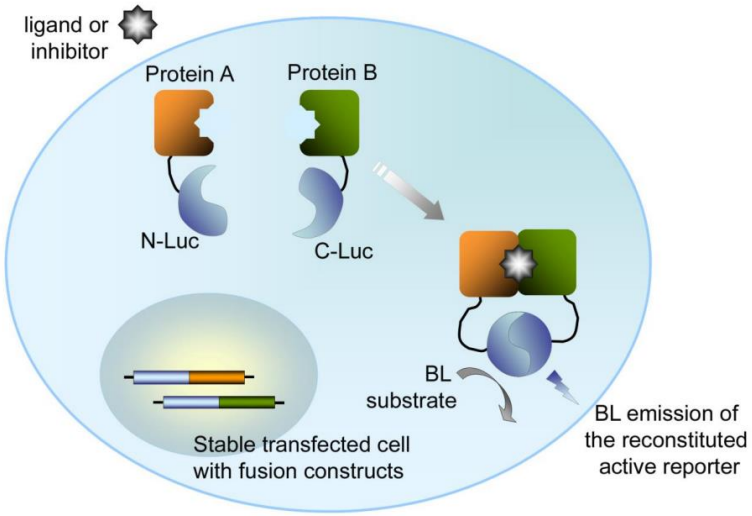
It's a method that relies on the **functional reconstitution of a splitted reporter**.

Protein–fragment complementation assays (PCAs) are based on the fusion of hypothetical binding partners to two rationally designed fragments of a reporter protein. The interaction between bait and prey proteins brings the split reporter fragments close enough to enable their non-covalent and specific reassembly followed by recovery of its native structure and activity.

We can split the reporter into two parts that are not able to reconstitute. The interaction between bait and prey proteins brings the split reporter fragments close enough to enable their non-covalent and specific reassembly followed by recovery of its native structure and activity.

In the first configuration the problem was that, even in absence of interaction, the two parts were likely to come together because of their high affinity, whereas, then adding some mutations, researchers were able to obtain two parts that are not able to come together in absence of the interaction between protein A and protein B.

There are not parameters that you cannot control, so this method is more likely to work in respect to energy transfer.



You can have a picture of what is happening in the cell, monitoring the reconstitution of the reporter that is already inside of the cell--> Differently from gene reporting techniques, it's not necessary to wait for transcription and translation immediate, there is an immediate reconstruction of the reporter with the emission of the signal (= **real time result**)

Next generation split luciferase: NanoBiT

NanoBit luciferase is a binary splitted reporter protein based on a large N-terminal (18 kDa polypeptide) and small C-terminal (1.3 kDa peptide) of NanoLuc luciferase. The main advantage is that they are not able to reconstruct a functional protein unless their labeled proteins interact: the NanoBit subunit are weakly associated with an intrinsic affinity.

BRET VS PCA

BRET	PCA
Real-time measurements	Real-time
Can be adapted for studying protein interactions in living animal models by molecular imaging	Can be adapted for studying protein interactions in living animal models by molecular imaging
No false positives	Possibility of false positives (dissected fragments can spontaneously associate in the absence of the binding proteins)
Low dynamic range	High dynamic range (due to high cooperativity of the process)
Low multiplexing capability	Possibility of multiplexing (by using split reporters with different properties and substrate specificity)
The ratiometric signal minimises interferences from assay conditions	

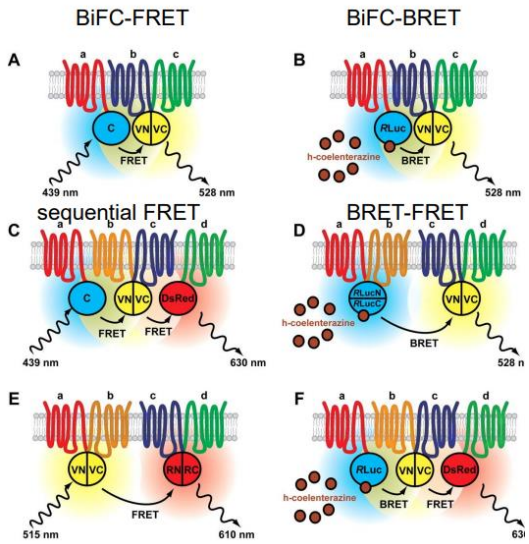
They both rely on fusion protein, therefore in both cases there is the risk of artifacts.

With BRET, the signal is ratiometric (more robust); since you are dealing with two signals there is a low multiplexing capability.

But both BRET and PCA can be applied to living cells

They can also be COMBINED:

Combining protein complementation assays with resonance energy transfer



Trimeric protein complexes (a, b, c) detected by BiFC-FRET (A) or BiFC-BRET (B)
In both approaches, a reconstituted FP serves as donor or acceptor in a RET pair. In this case, complemented Venus serves as acceptor for Cerulean (A) or *R*Luc (B).

Complemented fluorescent (or luminescent) proteins replace a donor or acceptor in sequential FRET (C) or BRET-FRET (F) assays. Bioluminescence (D) or fluorescence (E) resonance energy transfer between two complemented luminescent or fluorescent proteins also allows the detection of tetramers.

NON-INVASIVE IN VIVO IMAGING

Right now, in Europe the idea is to use animal according to **3R principle**:

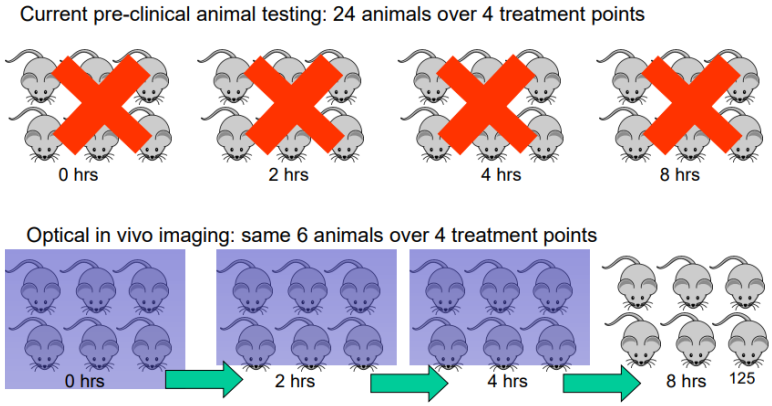
- **Replacement**--> every time is feasible replace animal testing with in vitro/in silico test

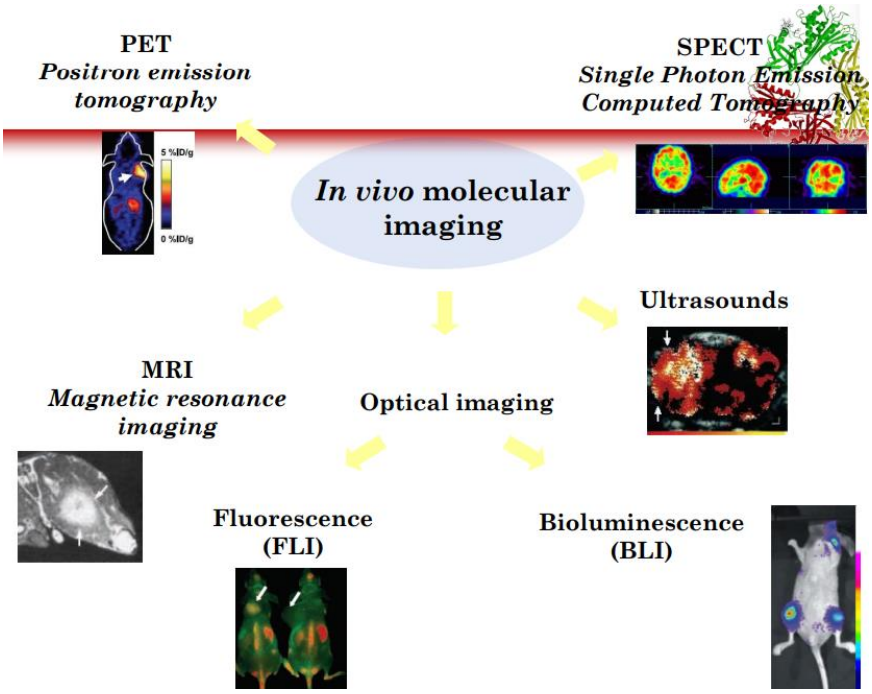
- **Reduction** --> of the number of animals used

- **Refinement** --> achieve the most accurate model which mimics more faithfully the pathological or physiological conditions that you're looking at, choosing reliable ones based on better understanding.

The main goal is to reduce at minimum the use of animals, or at least to keep the same animal over different treatment points and perform *longitudinal monitoring*.

For example, optical imaging is the way to avoid the excessive use of animals.

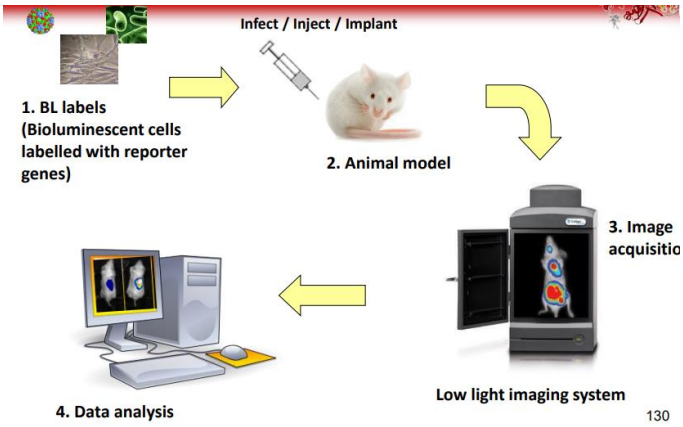




Which is the optimal light wavelength for in vivo imaging?

Infrared (IR) because it goes in deep, is less absorbed and scattered by the tissue and we can improve the resolution.

In this case, as in all methods that rely on luminescent approaches, it's not important the total light output, but rather the signal to noise ratio. The goal is to achieve the highest sensitivity and the highest resolution. With bioluminescent and fluorescent tools, you can monitor different pathological and physiological conditions just by tagging the protein with the specific reporter. For example, you can engineer the cell with a reporter which is under the regulation of an inducible promoter (or constitutive or tissue-specific), you inject the cells in the animal model and perform the imaging anesthetizing the animal.



It's also possible to inject the animals with labeled bacteria or viruses in order to monitor the efficacy of a treatment after the infection.

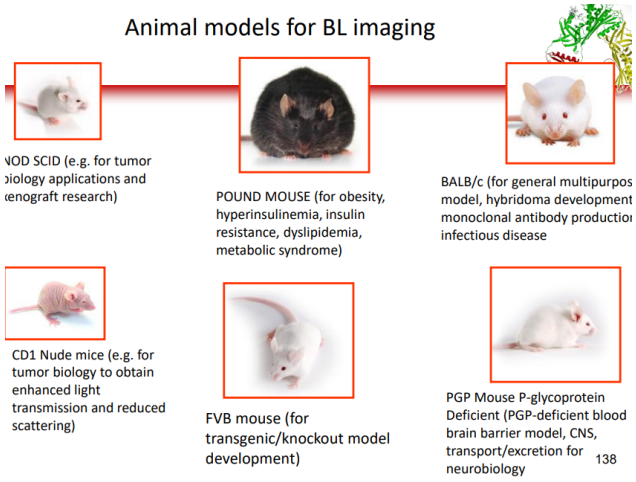
Mice carrying BL reporter genes

i. Transgenic mouse models: animals that contain additional, artificially-introduced genetic material *in every cell*. This often confers a gain of function, for example the mouse may produce a new protein, but a loss of function may occur if the integrated DNA interrupts another gene (e.g., knock-in/knock-out mice)

ii. Xenograft mouse models: animals with transplanted human tumors or other tissues obtained by injection or implantation of cells (e.g., tumor cells, stem cells)

iii. Syngeneic mouse models: animals with transplanted tumors or other tissues obtained by injection or implantation of cells obtained from the same species

It is better to use the most suitable animal model:



The issue is to identify the best probes, that should emit from 600 nanometers and beyond in order to have the right sensitivity. So, the **red emitting probes are the best choice**, just because a lot of biological molecules absorb in the green-yellow part of the visible spectrum.

It's also possible to push further the red shifting by using nanomaterials:

- Single-walled carbon nanotubes (SWNTs)
- Quantum dots (QDs)
- Polymethine dyes

Single-walled carbon nanotubes

They are nanometers diameter cylinders consisting of a single graphene sheet wrapped up to form a tube.

SWNTs as optical sensors are photostable and fluorescence in the near-infrared where blood and tissue absorption and autofluorescence is minimal.

Quantum dots

They are nanocrystals or semiconductors nanocrystals, composed of CdSE, CdS, ZnSe etc.

Depending on their size, they have different emission properties, so it's possible to tune the emission color depending on the fabrication process. Therefore, it's possible to use the same excitation wavelength to obtain different emission from different quantum dots.

There are plenty of applications of quantum dots and in general nanomaterial, the problem is the safety: most of these nanomaterials have been applied in animal models for imaging, drug delivery, etc.., but very few studies have investigated their toxicity and pharmacokinetics.

The same protein-protein interaction methods can be applied with mice or another animal model. By using the same engineered cells.

(FBbase: fluorescent protein database)