Methods for the study of protein-protein interactions

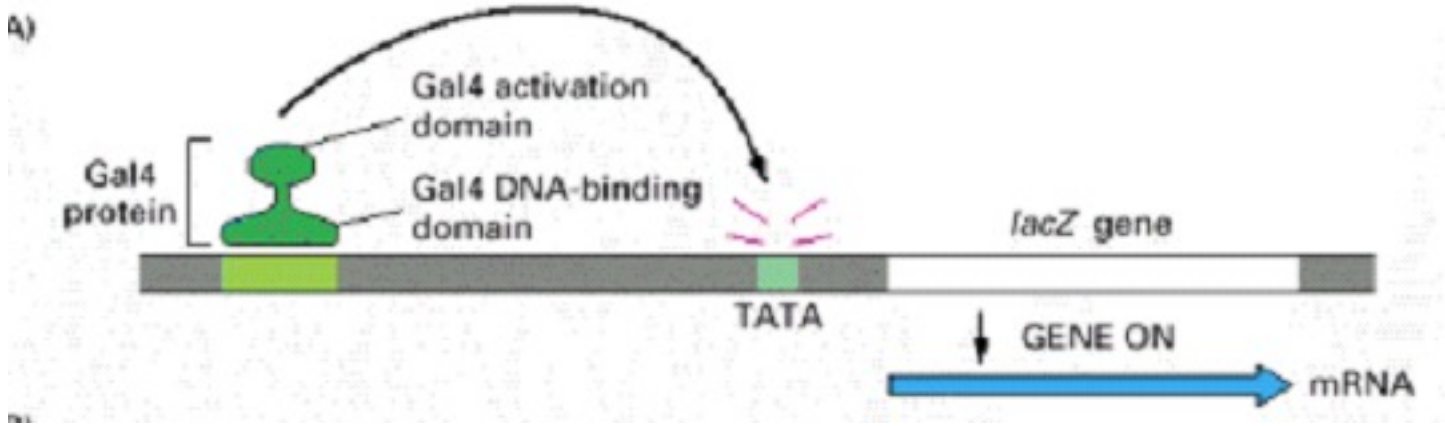
Genetic methods based on split proteins

Y2H: yeast two-hybrid system

- It is a genetic method based on transcriptional activity to measure protein-protein interactions ***in vivo***

- Transcriptional activators are proteins that present a DNA-binding domain (DBD) and a transcription activating domain (AD)

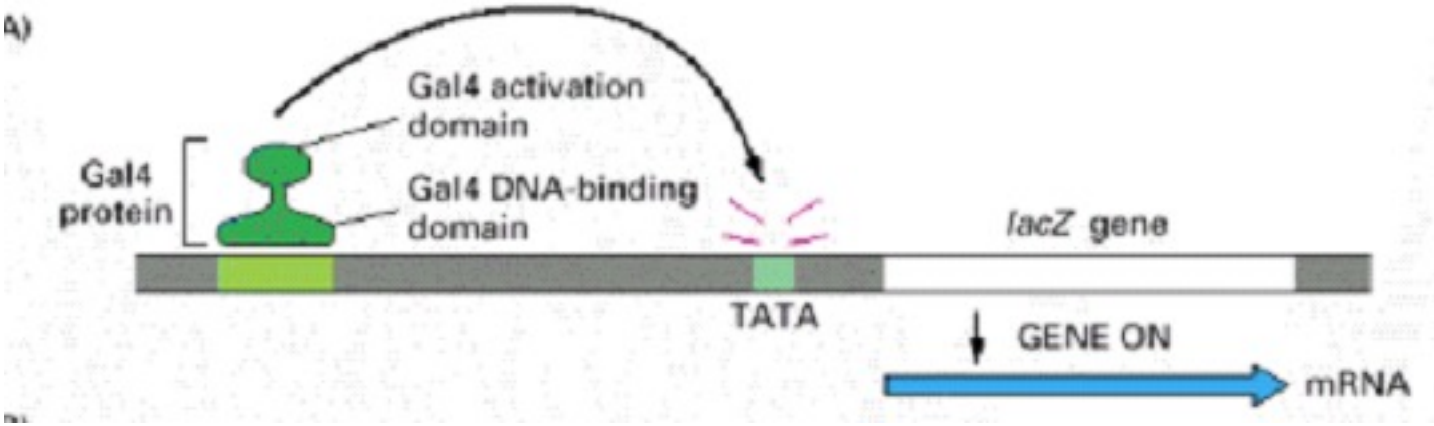
- The DNA-binding domain is required to position the activator domain on the promoter of the gene that must be transcribed. The activator domain recruits the transcriptional machinery.



To activate transcription, the two domains do not necessarily need to be covalently bound, but their association may be mediated by interactions between two proteins X and Y (Field and Song, 1989)

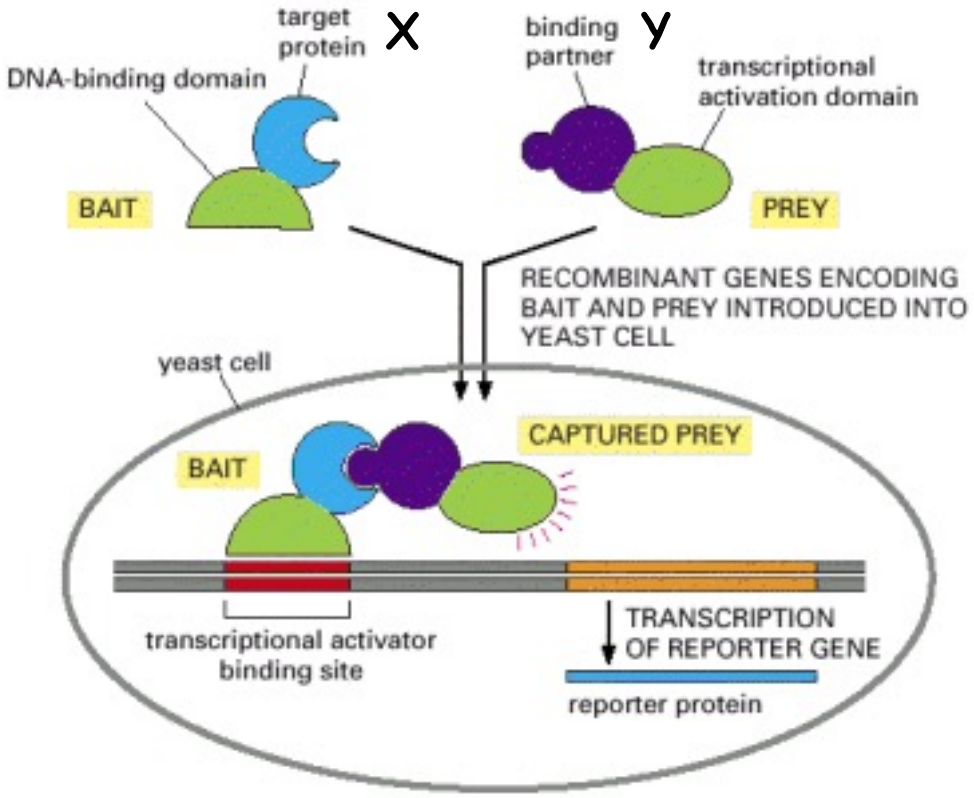
The yeast transcription factor **Gal4** can be split in a DNA-binding domain and transcription activator domain.

- The N-terminal 147 residues encode the DNA-binding domain (DBD)
- The C-terminal 114 residues encode the activator domain (AD)



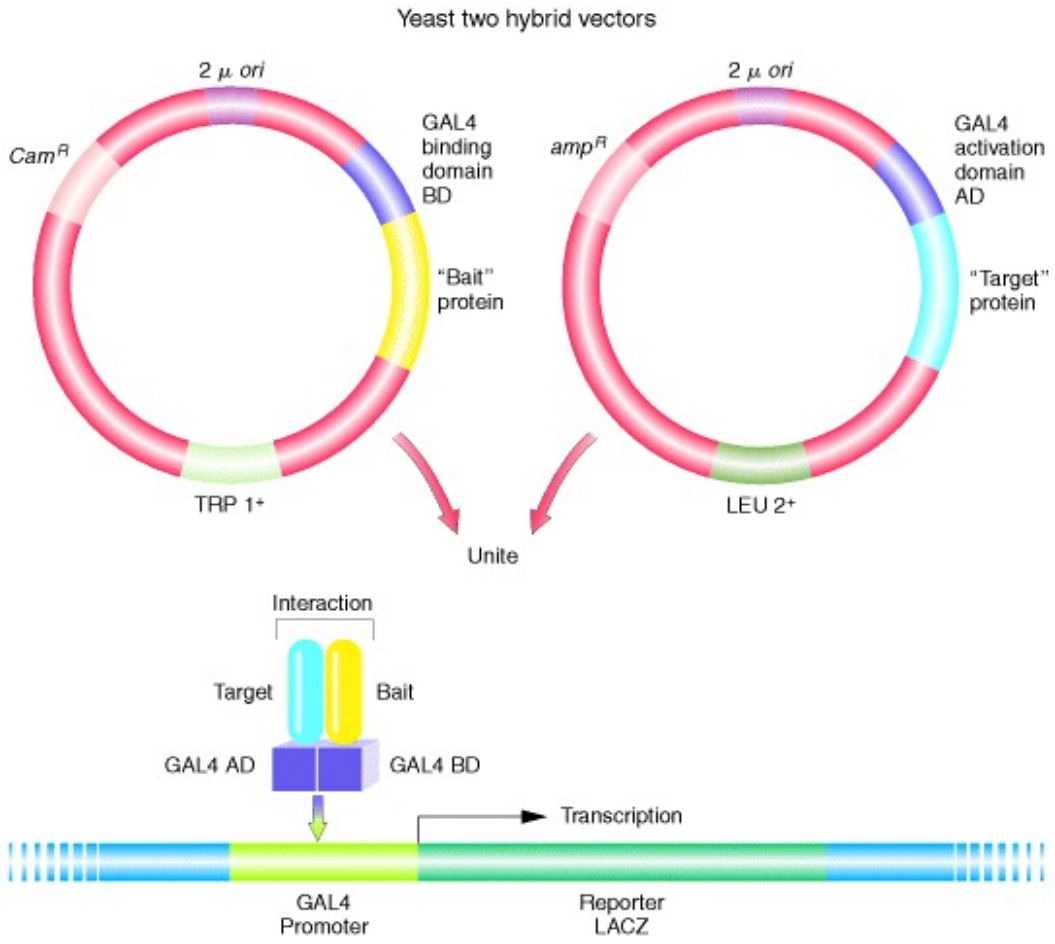
Construction of the two hybrids

- Gal4 DNA binding domain cDNA is fused to cDNA of bait protein
- Gal4 activation domain cDNA is fused to cDNA of prey protein
- The two chimeric proteins are expressed in a cell harbouring one or more reporter genes under control of the Gal promoter



If bait and prey proteins interact, functional reconstitution of Gal4 takes place and reporter genes are expressed

The two vectors that contain Gal4 binding domain fused to 'Bait' protein and Gal4 activation domain fused to 'Target/prey' protein



The reporter gene usually complements an auxotrophy of the host cell or it is an enzyme that can be easily detected (es. HIS3, lacZ)

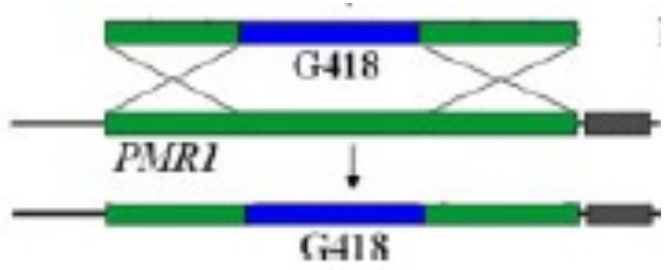
The host yeast strain must

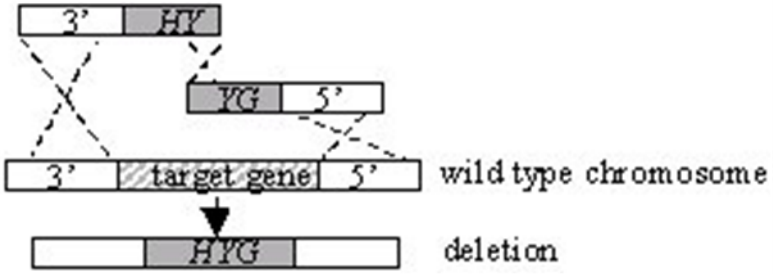
- Have the reporter gene(s) under control of the Gal promoter
- Have an inactive endogenous *GAL4* gene
- Be auxotrophic for the selection markers of the two vectors for expression of the two hybrids, and for the (eventual) reporter gene

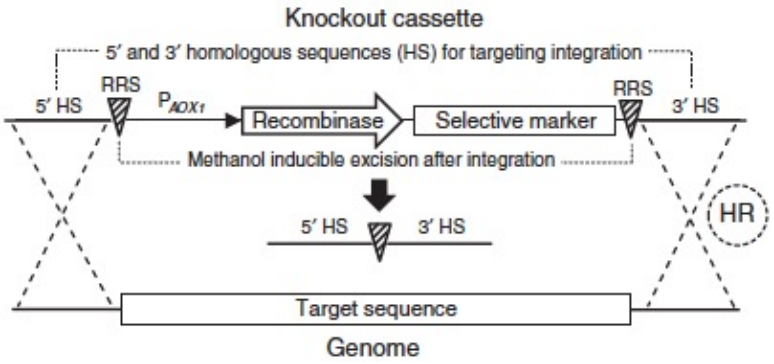
How do you inactivate a gene in yeast?

How do you inactivate a gene in yeast?

- Chemical mutagens (aspecific)
- Cassettes for homologous recombination (gene-specific)







Verification ➔ phenotypic screening and genomic DNA analysis (PCR)

Applications of Y2H

Possibility to identify rapidly proteins that interact with the 'bait' protein.
It is possible to prepare cDNA libraries fused to the activation domain AD that wil be introduced in a strain that expresses the DNA-binding domain DBD fused to the 'bait' protein

Limits

It is impossible to study membrane protein interactions or DNA binding proteins that are able to activate transcription

Protein Complementation Assay

- The reporter protein is split in two inactive fragments
- The fragments are genetically fused to the (potentially) interacting partners
- Recovery of function of the reporter protein is mediated by interaction between the partner bait and prey proteins

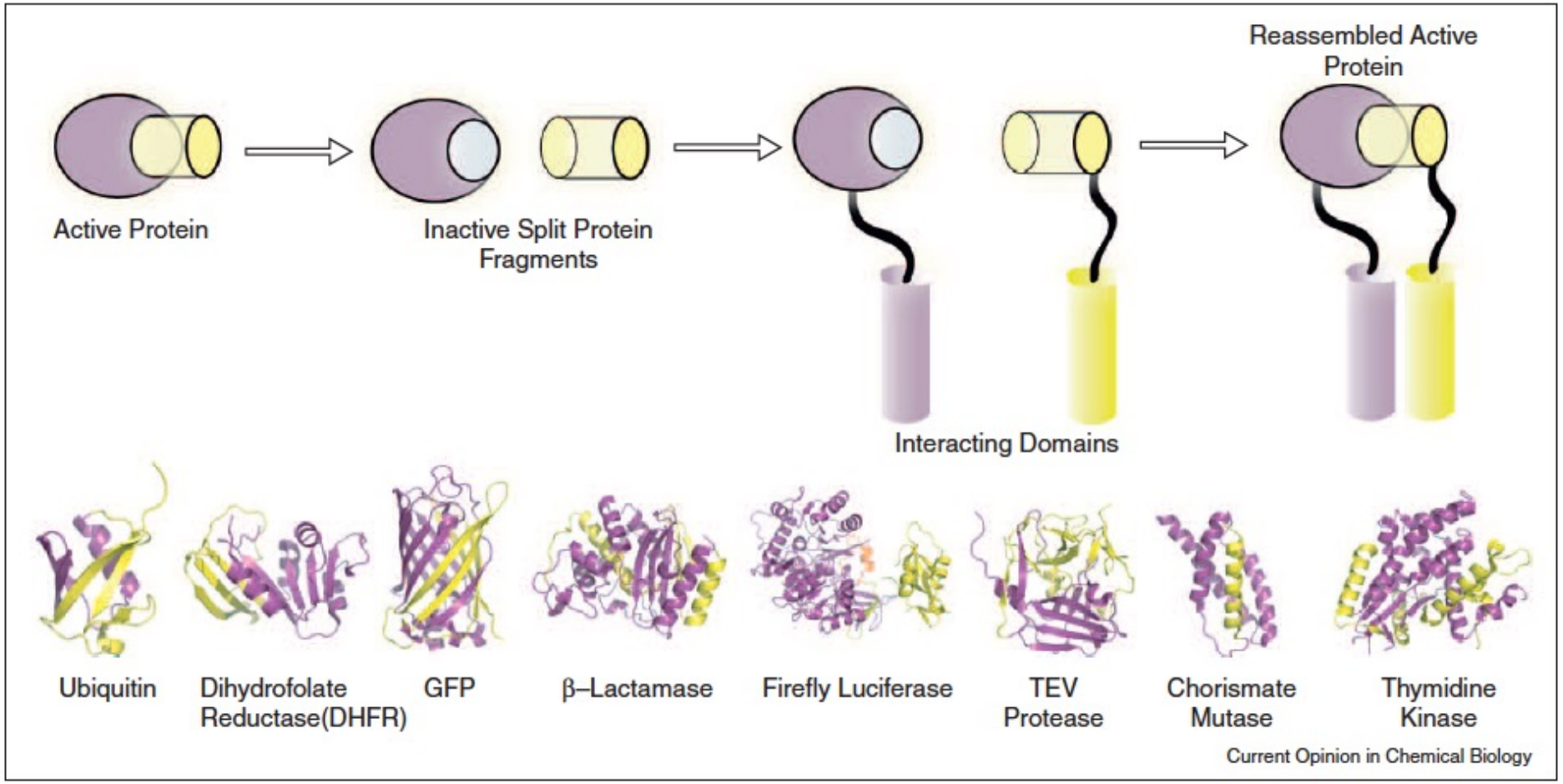


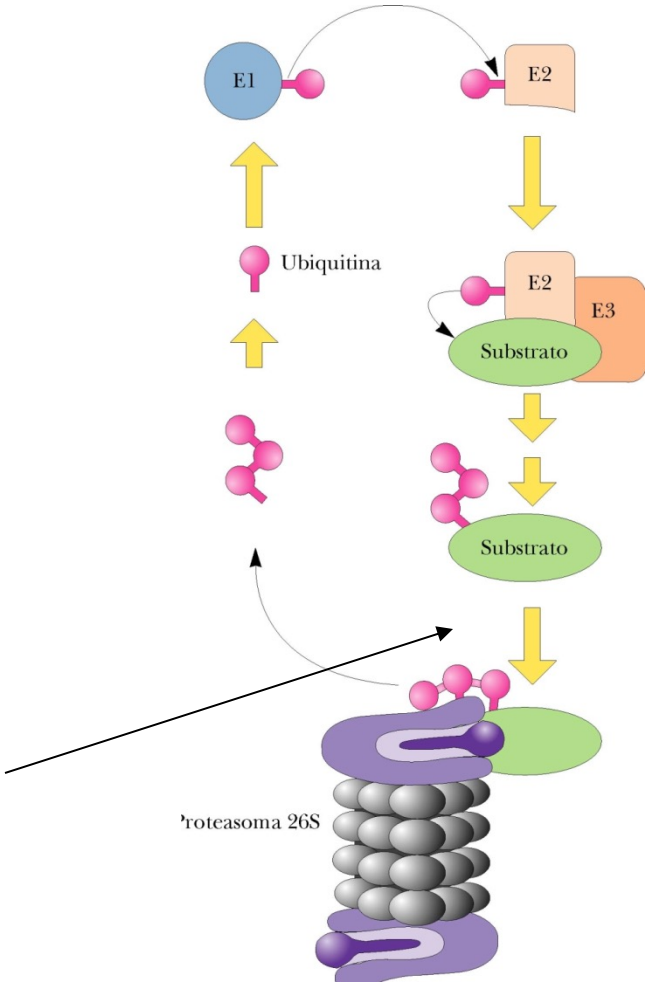
Illustration of conditional split-protein reassembly. A generic split-protein system is shown where a functional protein is dissected into two inactive fragments, purple and yellow. The attachment of two interacting proteins or protein domains brings the inactive fragments into close proximity and overcomes the entropic cost of fragmentation. This leads to the reassembly or complementation of the fragments thus providing a direct readout for the partnership between the interacting domains. Crystal structures of representative proteins which have been shown to be amenable to interaction dependent reassembly, where the N-terminal and C-terminal fragments are shown as purple and yellow respectively.

Split ubiquitin membrane-based yeast two hybrid system

Ubiquitin is 76 amino acid protein involved in protein turnover

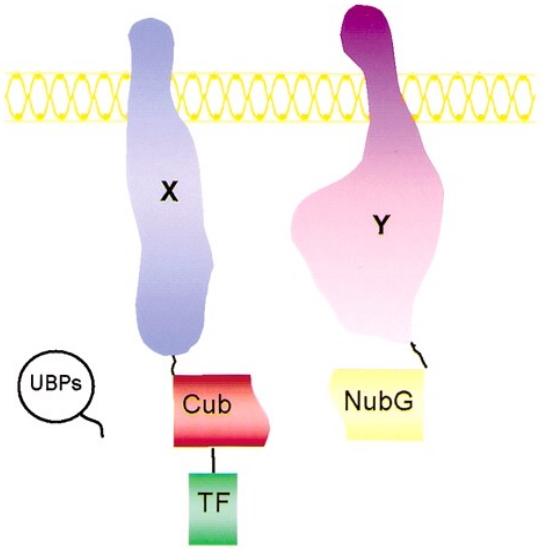
The C-terminal end of ubiquitin forms an amide bond with the protein that must be degraded

UBP ubiquitin-specific proteases are cytosolic proteases that hydrolyze the peptide bond between ubiquitin and the target protein



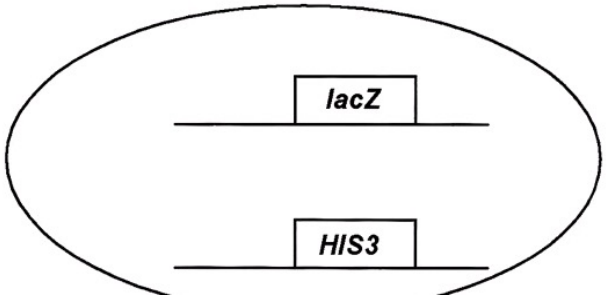
Ubiquitin can be split in two fragments:
Nub N-terminal fragment (1-34)
Cub C-terminal fragment (35-76)

NubG has a point mutation (Ile13Gly) that abrogates spontaneous reassociation with Cub. **Association takes place only when the the fused proteins (X e Y) interact**

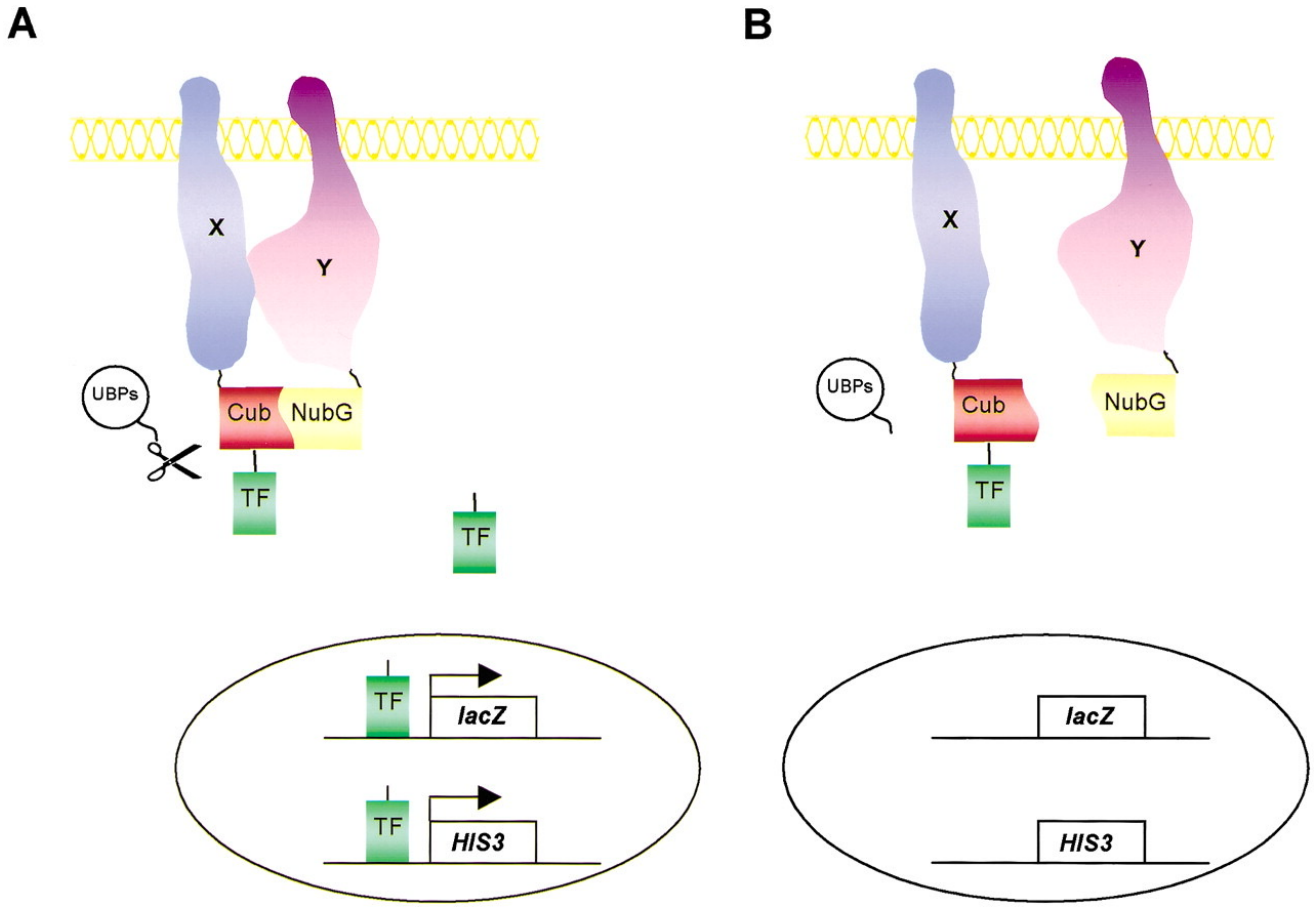


Construction of the two hybrids

At the 5′ of Cub cDNA the cDNA of bait protein (X) is fused
At the 3′ of Cub cDNA the cDNA of a transcription factor TF is fused
At the 5′ of NubG cDNA the cDNA of prey protein (Y) is fused



Association between X and Y reconstitutes ubiquitin and allows cleavage by the UBP protease to release TF LexA. LexA regulates expression of easily detectable reporter genes (HIS3 and lacZ).



Requirements and limits of the split-ubiquitin system

NubG and Cub must be fused to the region of the membrane protein that localizes in the cytosol because the UBP protease is found in the cytosol

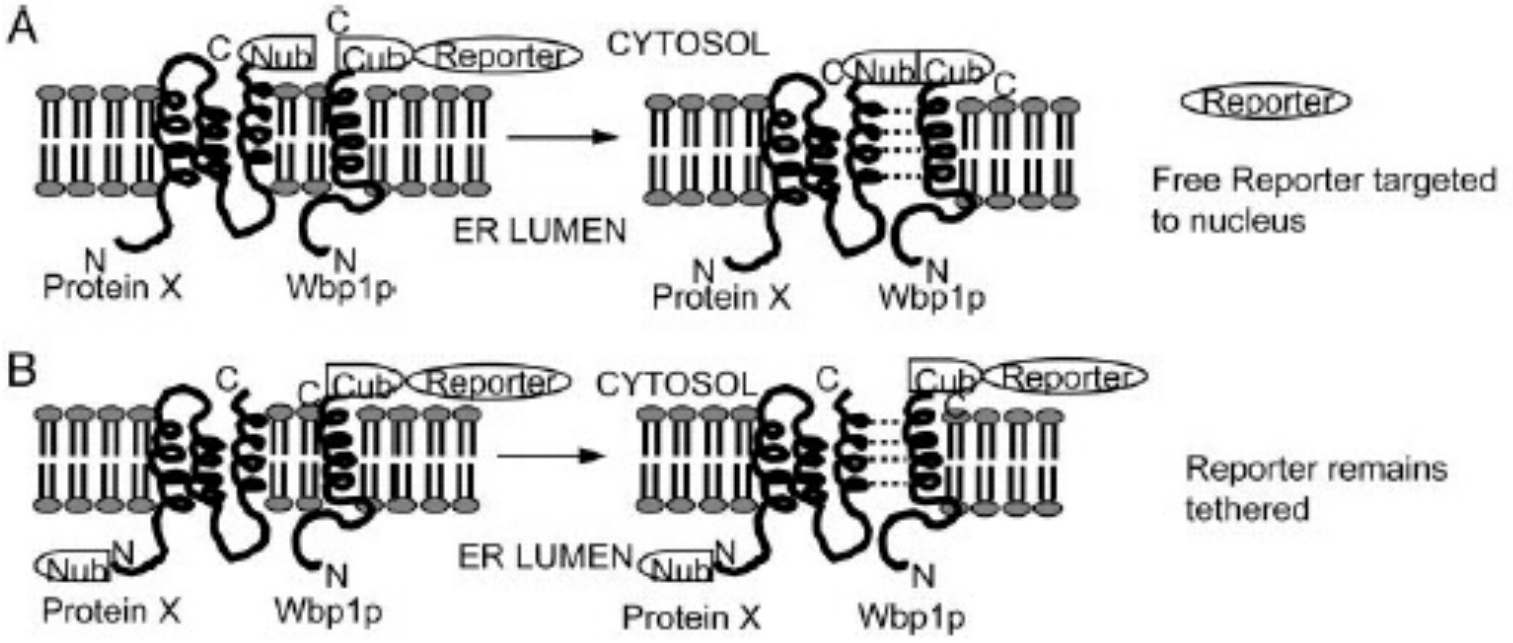
Disadvantage: in a screening to identify new interaction partners of the bait protein some interactors may be missed (false negatives)

Advantage: it is possible to obtain information on the orientation of the N and C terminal end of membrane proteins

➔ NubG-X and X-NubG fusions

An example…
Interaction Wbp1p-Cub + fusion protein X-Nub
➔ A signal is produced
➔ The C-terminus of protein X is in the citosol

Interaction Wbp1p-Cub + fusion Nub-protein X
➔ A signal is NOT produced even if the two proteins interact
➔ The N-terminus of protein X is not in the citosol



Protein Complementation Assay
Enzymes as reporters

Folding induced by interaction of the two partner proteins leads to recovery of enzymatic activity

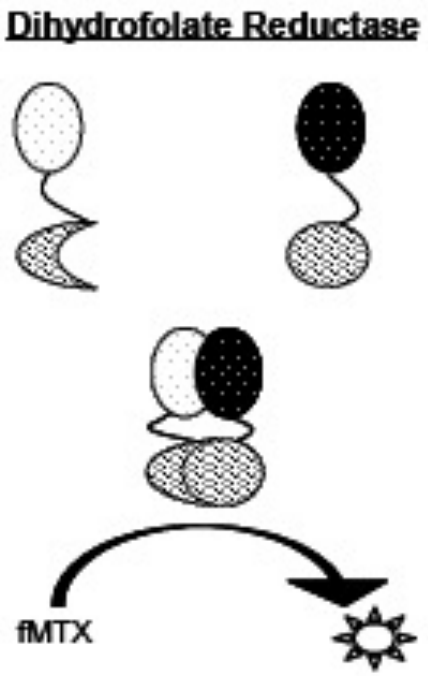
- beta-galactosidase (lacZ)
- dihydrofolate reductase (DHFR)
- beta-lactamase (bla)

- Availability of membrane-permeable chromogenic substrates that allow easy detection of enzymatic activity
- Survival assays

Dihydrofolate reductase catalyzes reduction of dihydrofolate to tetrahydrofolate, required for nucleotide biosynthesis

Murine DHFR is split in two fragments: 1-105 and 106-186
Interaction between partner proteins is evidenced:

1. Survival of $DHFR^-$ cells seeded on media lacking nucleotides
2. Measure of fMTX (Methotrexate labelled with a fluorescent probe) fluorescence



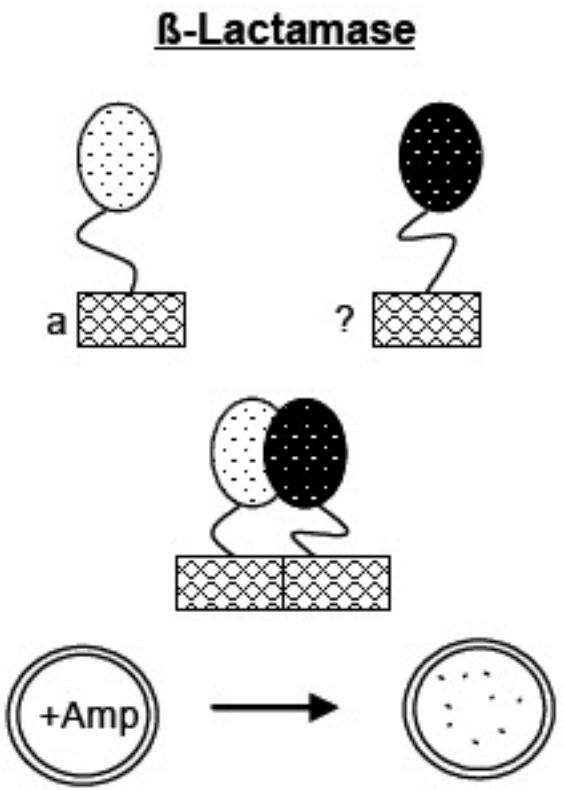
fMTX binds DHFR with high affinity

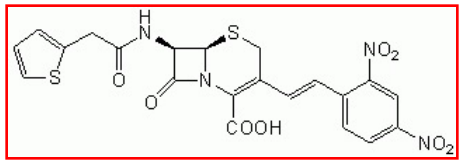
K_d = 540 pM

Limits of the system: response time is long (hours to days)

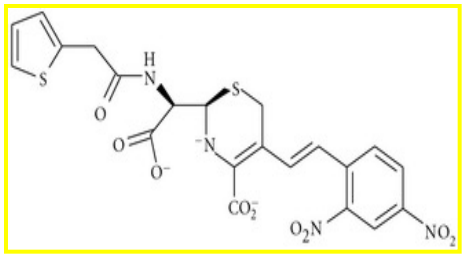
Beta-lactamase monomeric enzyme 29 kDa
Can be split in two fragments 24-215 and 216-286
Interaction between partner proteins can be evidenced:

- resistance to ampicillin
- colorimetric assay with nitrocefin



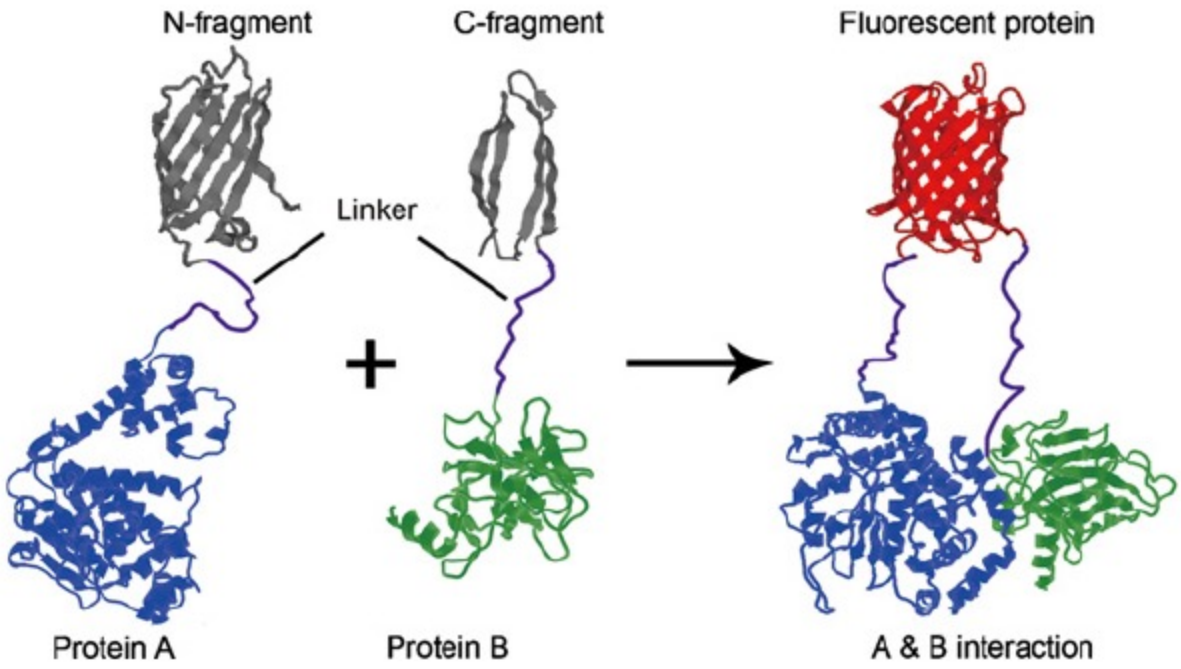


Beta lactamase



Bimolecular Fluorescent Complementation (BiFC)

Folding of the fluorescent protein induced by interaction of the partner proteins leads to formation of the chromophore



Fluorescence

- Fluorescence is the emission of light due to transition of electrons from a higher to a lower energy level
- The *wavelength* of the absorbed radiation is lower than the emitted radiation (Stokes *shift*)
- The *energy difference* between absorbed and emitted radiation is dissipated as heat (molecular collisions)

$E = h\nu = hc / \lambda$

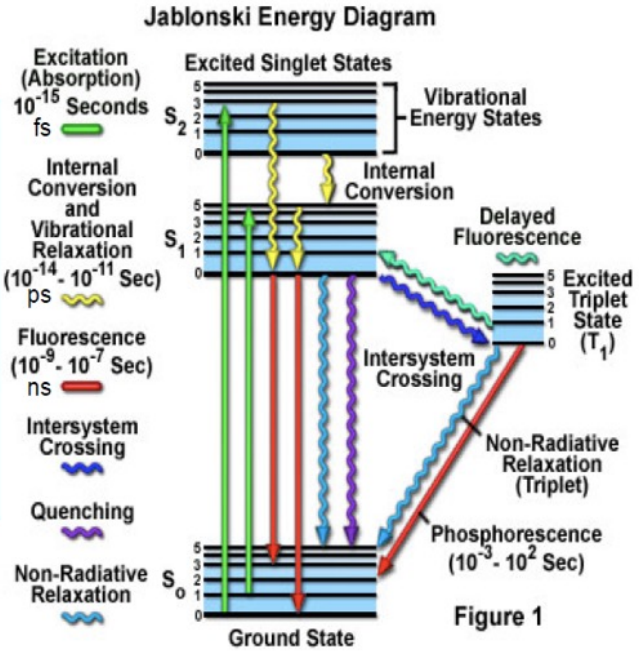
E – *energy*

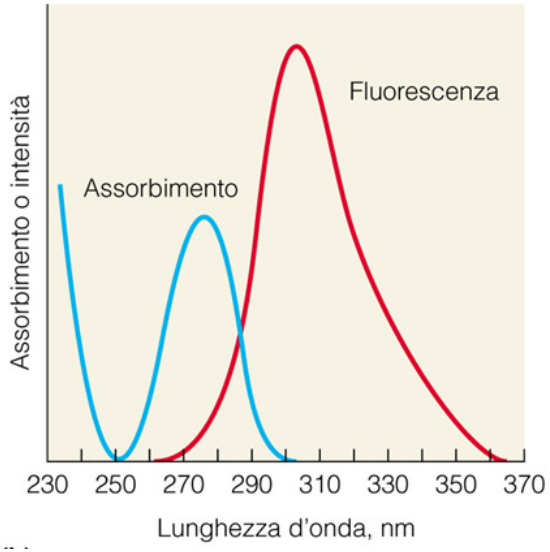
ν – *frequency*

λ – *wavelength*

h – *plank*

c – *speed*





Fluorescence

- *QUANTUM YIELD* is defined as

$$Q = \frac{quanti \cdot di \cdot fluorescenza \cdot emessi}{quanti \cdot assorbiti}$$

- Quantum yield is *independent* of excitation wavelength
- At low concentration of the *fluorophore* fluorescence intensity (I_f) is proportional to concentration:

$$I_f = 2.3 I_0 \varepsilon_\lambda c d Q$$

c is the fluorophore concentration
d is the optical pathlength
ε_λ is the molar extinction coefficient of the fluorophore
I_0 is the intensity of incident radiation

Structure of *Aequorea victoria* GFP (Green Fluorescent Protein)

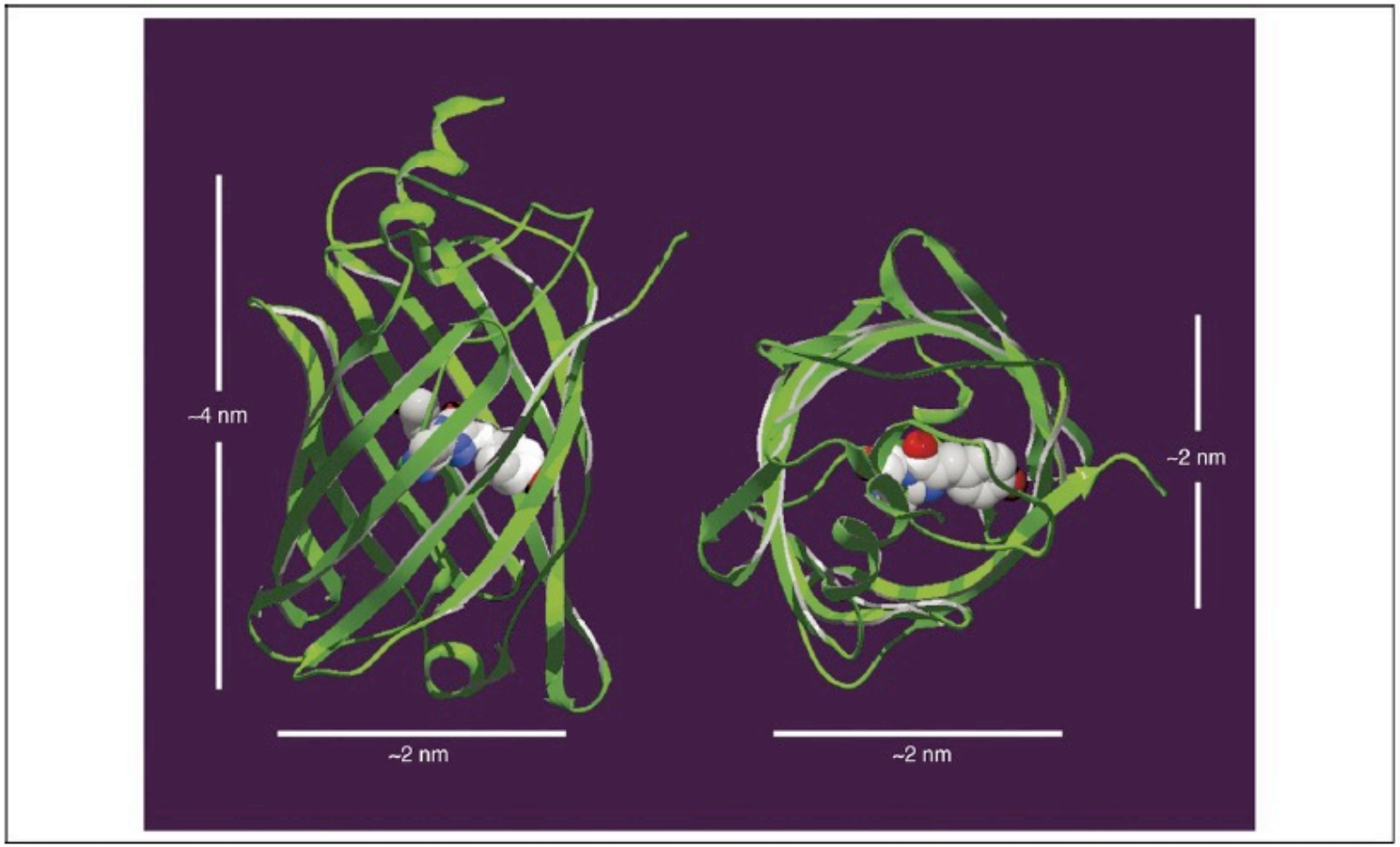
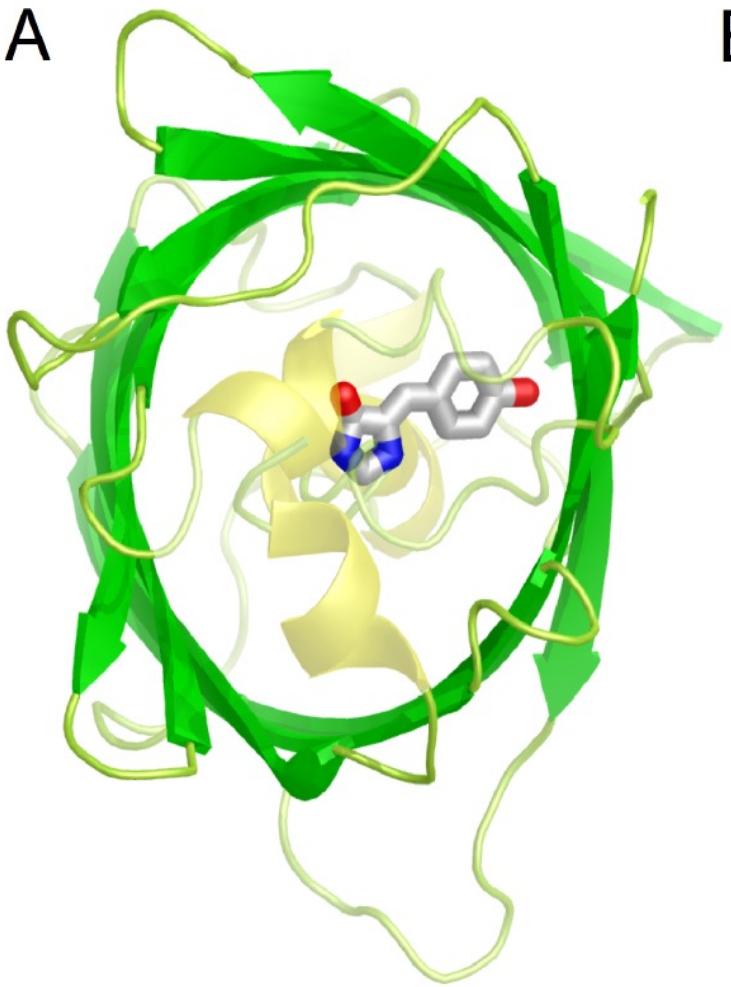
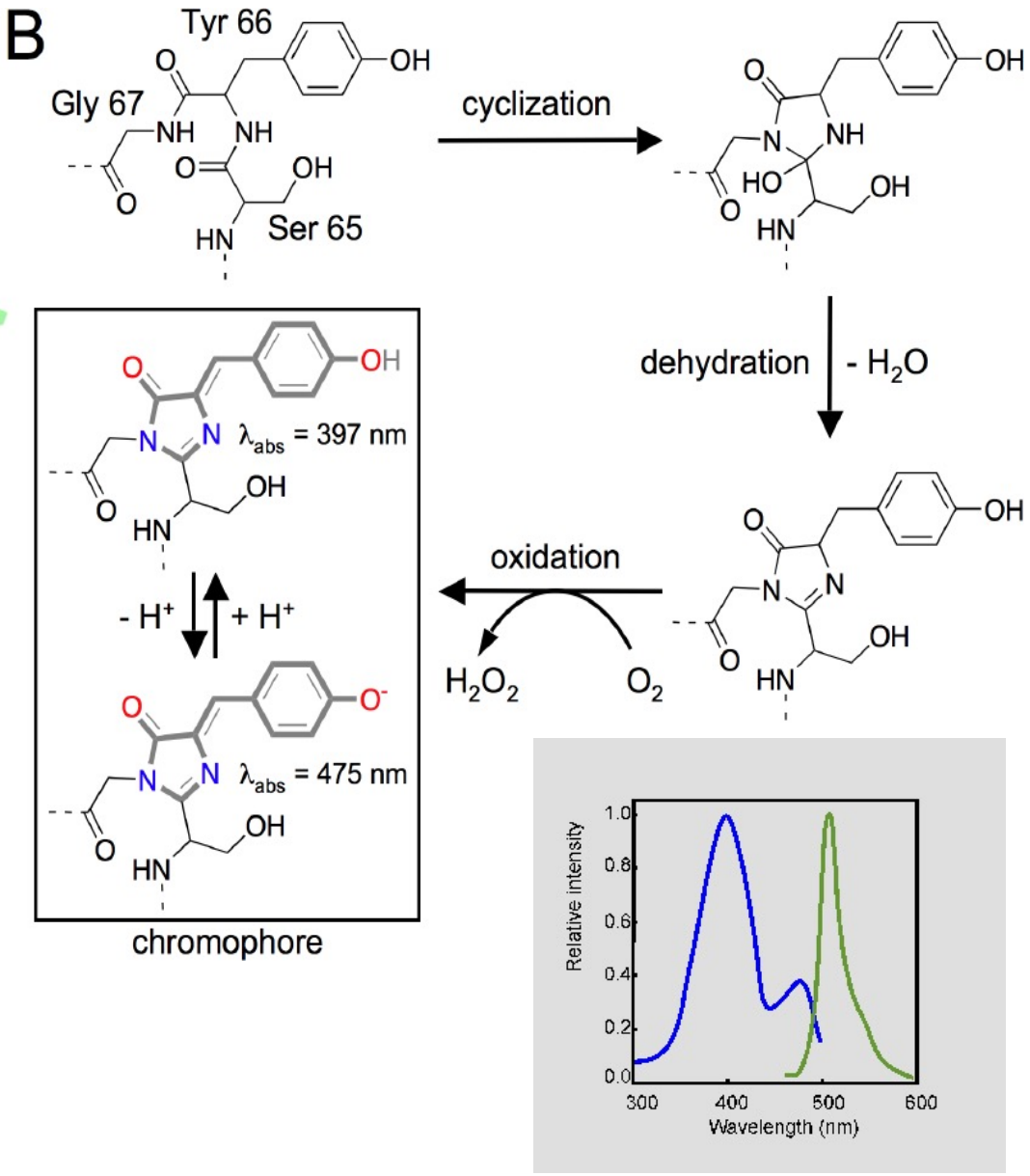


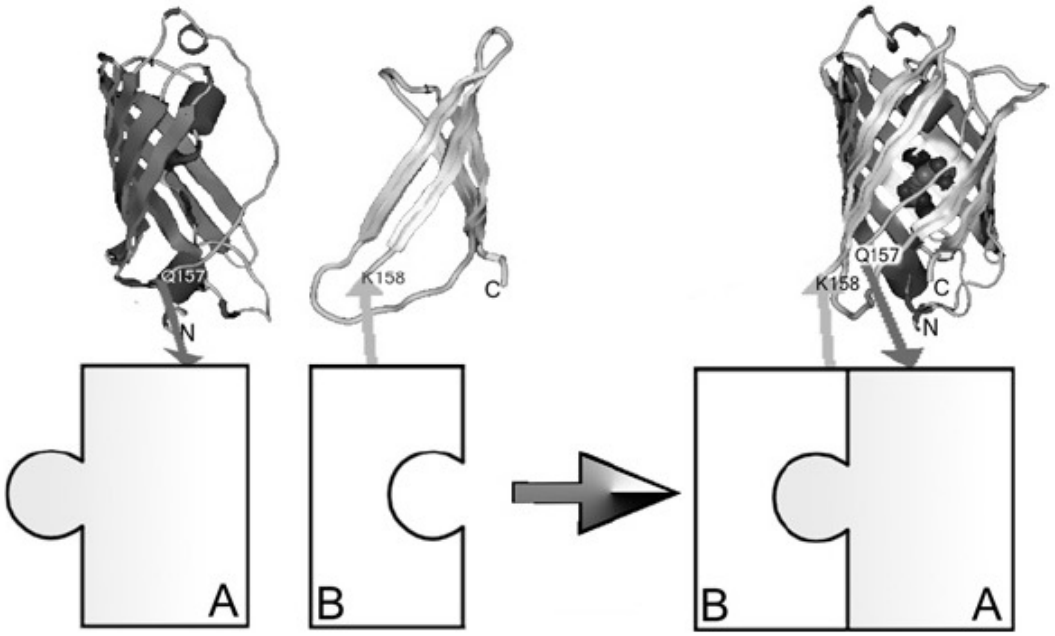
Figure 3. Crystal structure of GFP. The protein is cylindrical, with a diameter of ~2.4 nm and a length of 4.2 nm. The cylinder consists of 11 β strands with a single α helix running along its axis. The chromophore is located in the α helix at the center of the protein. All fluorescent proteins have a similar structure.

Formation of the chromophore



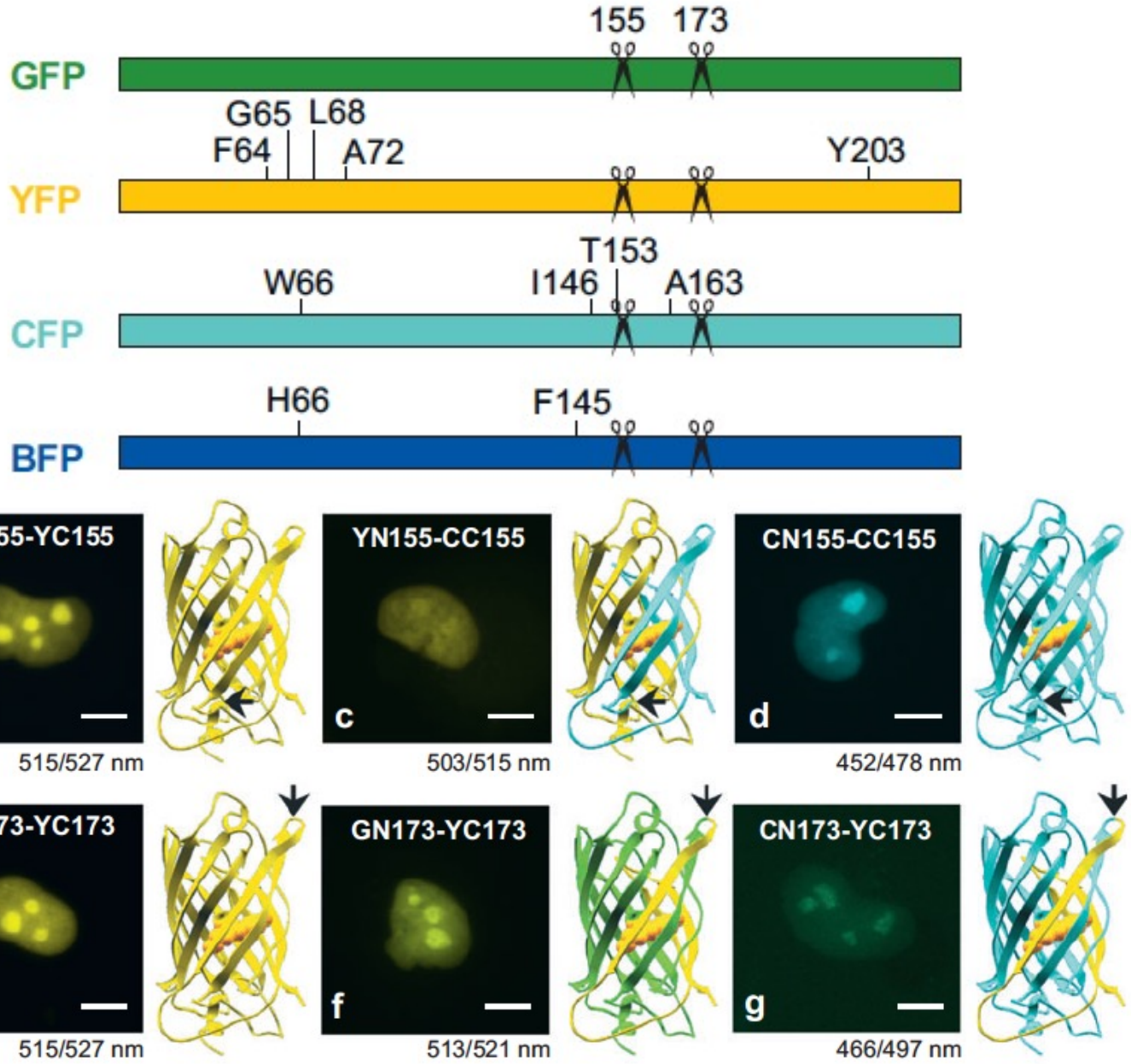


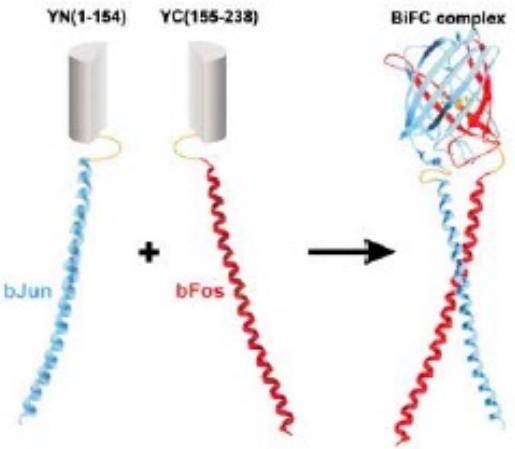
GFP can be split between residues 157 and 158 to produce two non-fluorescent fragments N-GFP and C-GFP.



Multicolor bimolecular complementation assay allows simultaneous visualization of protein complexes in the same cell

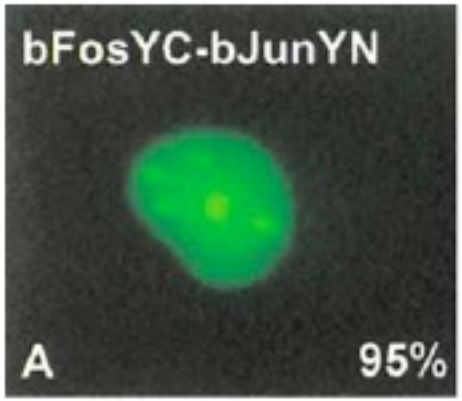
Other split-GFP variants

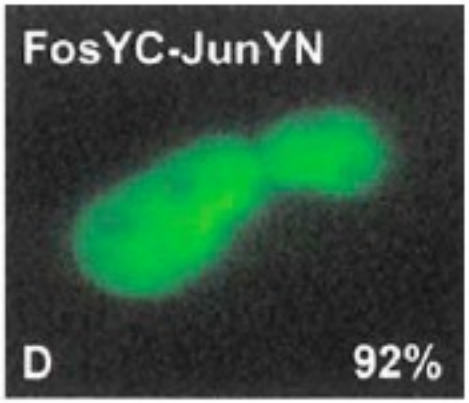




Molecular Cell, Vol. 9, 789–798, April, 2002, Copyright ©2002 by Cell Press

Visualization of Interactions among bZIP and Rel Family Proteins in Living Cells Using Bimolecular Fluorescence Complementation





FRET

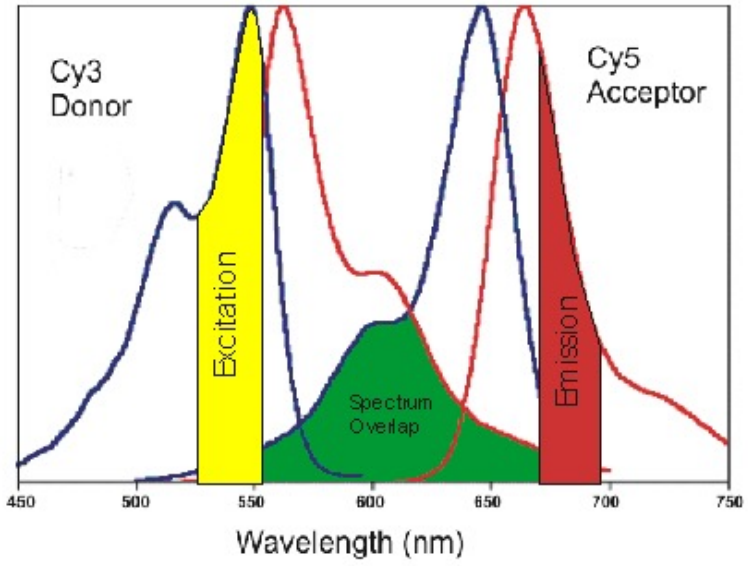
Fluorescence Resonance Energy Transfer

Non-radiative energy transfer (dipole-dipole) from a fluorescent donor to a fluorescent acceptor.

Energy transfer depends on:

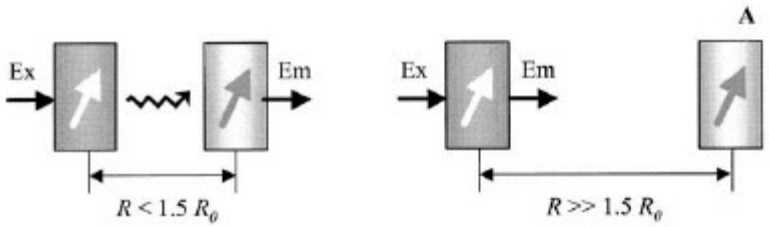
- spectral properties
- orientation
- distance between donor and acceptor (1-10 nm)

Donor emission spectrum and acceptor excitation spectrum must partially overlap



$$E_{FRET} = \frac{F_A}{F_A + \gamma F_D} = \frac{R_0^6}{R_0^6 + r_{DA}^6}, \quad \gamma = \frac{\eta_A \phi_A}{\eta_D \phi_D}$$

E = transfer efficiency
r = distance between chromophores
Φ = quantum yield
η = counting efficiency
R_0 = distance of 50% energy transfer (it is specific for each donor-acceptor pair)



FRET is observed if excitation at donor λ produces emission by the acceptor (at acceptor λ)

Applications of FRET

- Protein-protein interactions *in vitro* and conformational changes in proteins
- Protein-protein interactions *in vivo*

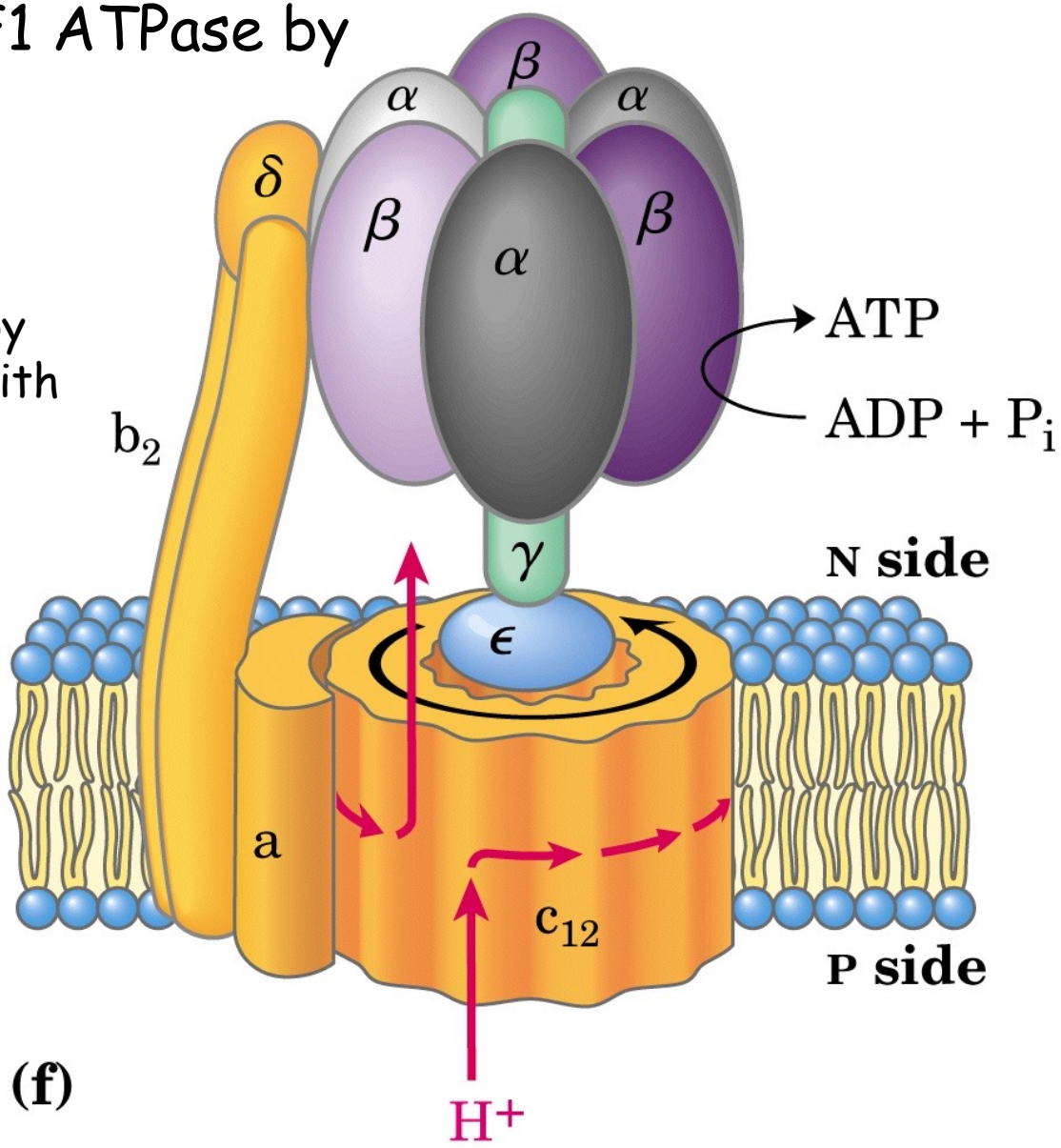
- Donor and acceptor fluorophores can be small molecules or proteins

Analysis of conformational modifications of the ε subunit of F0 F1 ATPase by FRET

F0 is a proton channel formed by three types of subunits **a**, **b**, **c** with stoichiometry $ab_2c_{10\text{-}12}$

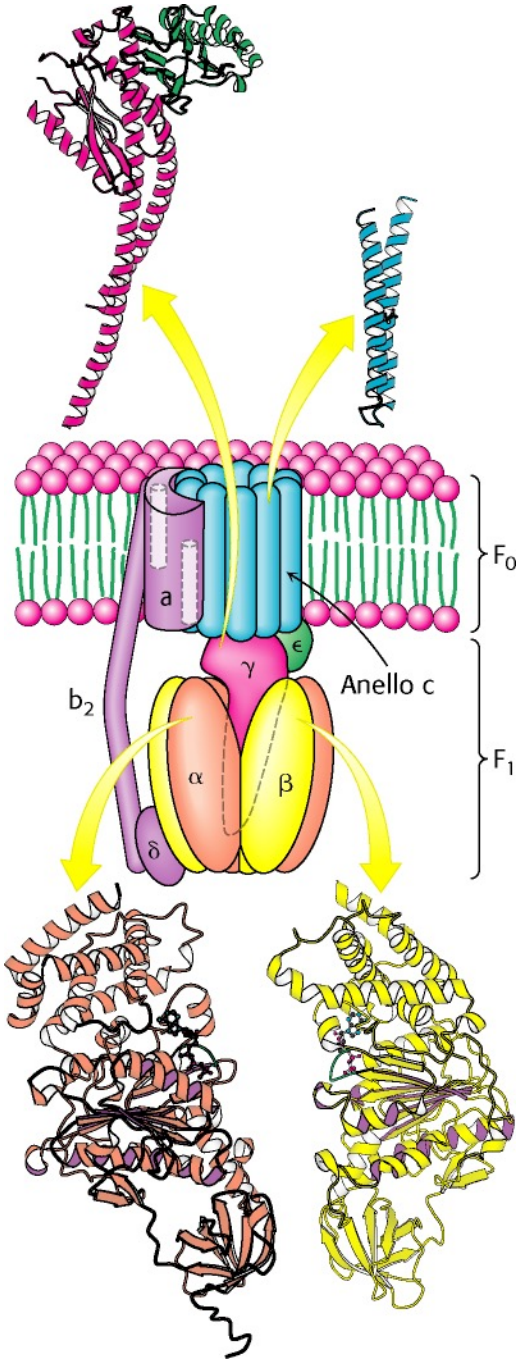
F1 catalyzes synthesis of ATP and it is formed by nine subunits $\alpha_3\beta_3\gamma\delta\varepsilon$

Each β subunit has a catalytic site for synthesis of ATP



The enzyme is constituted by two parts: a mobile unit formed by c and γε that rotates and a static unit or stator constituted by the rest of the molecule.

The γ subunit is the central axis of the molecular motor and it interacts with c and ε of the F0 portion and the αβ dimers of the F1 portion.



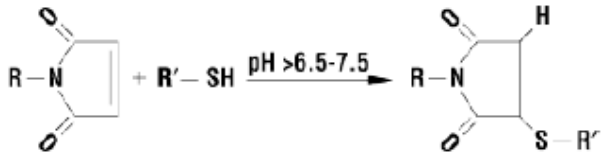
Recombinant expression of $\alpha_3\beta_3\varepsilon\gamma$

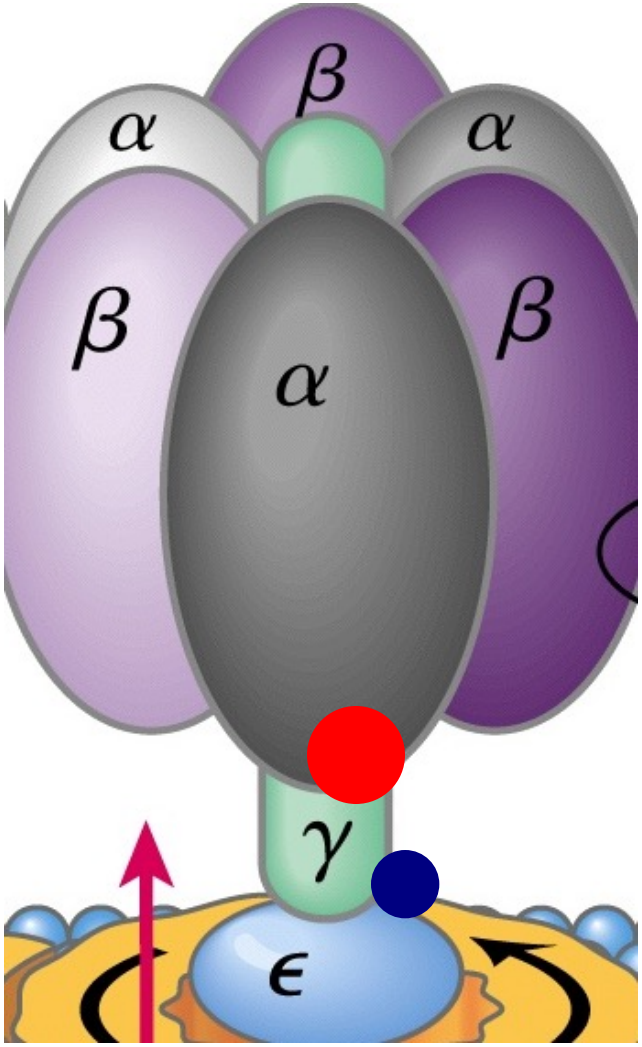
Site-specific labelling of the γ and ε subunits with fluorescent probes Cy3 and Cy5

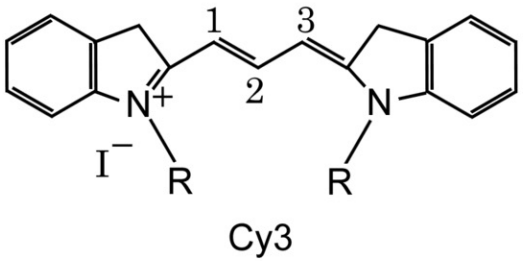
$\alpha_3\beta_3\gamma$(S3C-Cy3)

ϵ(134C-Cy5)

The two fluorophores were conjugated to cysteines of the protein via maleimide coupling

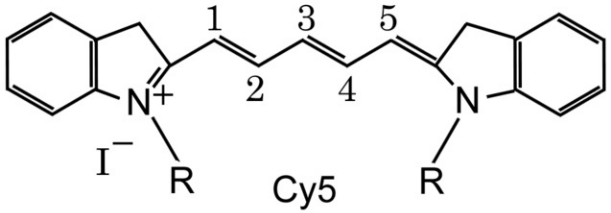






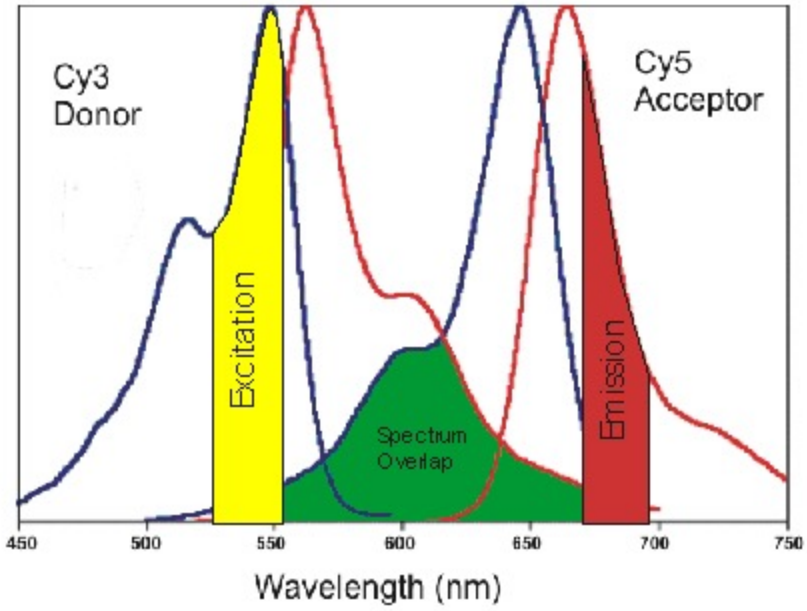
Cy3

Absorbs at 530 nm and emits at 570 (615) nm



Cy5

Absorbs at 625 nm and emits at 670 nm

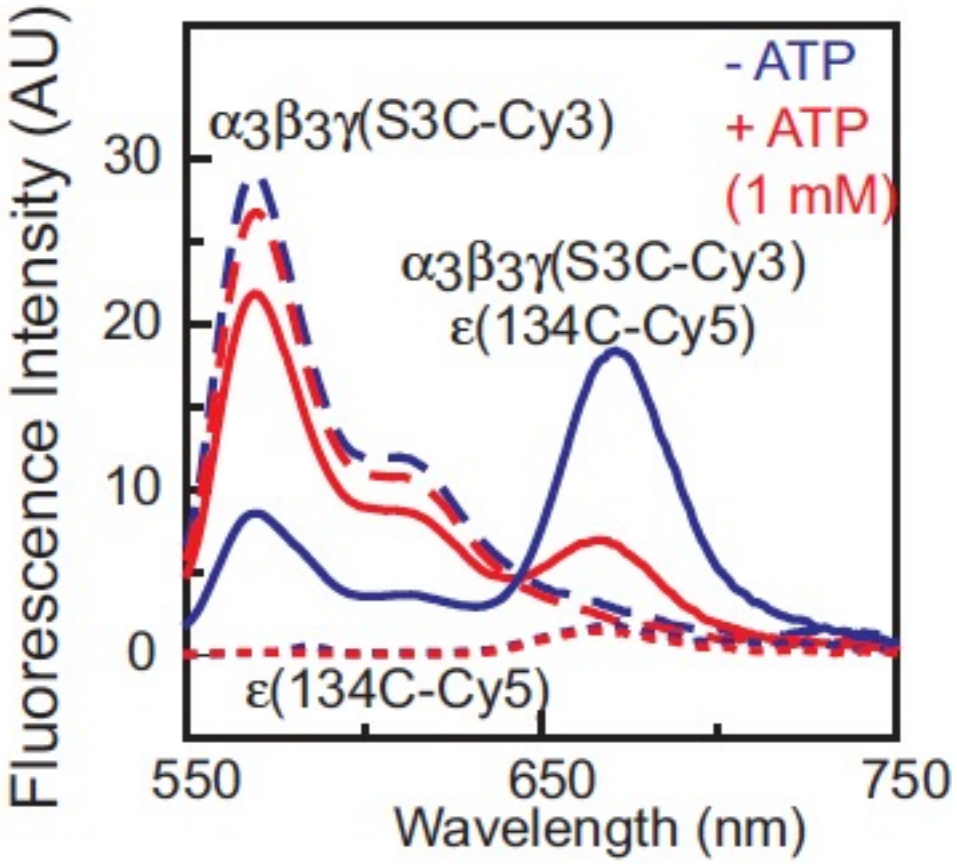


Fluorescence spectra (λ_{exc} 532 nm) emission 550-750 nm

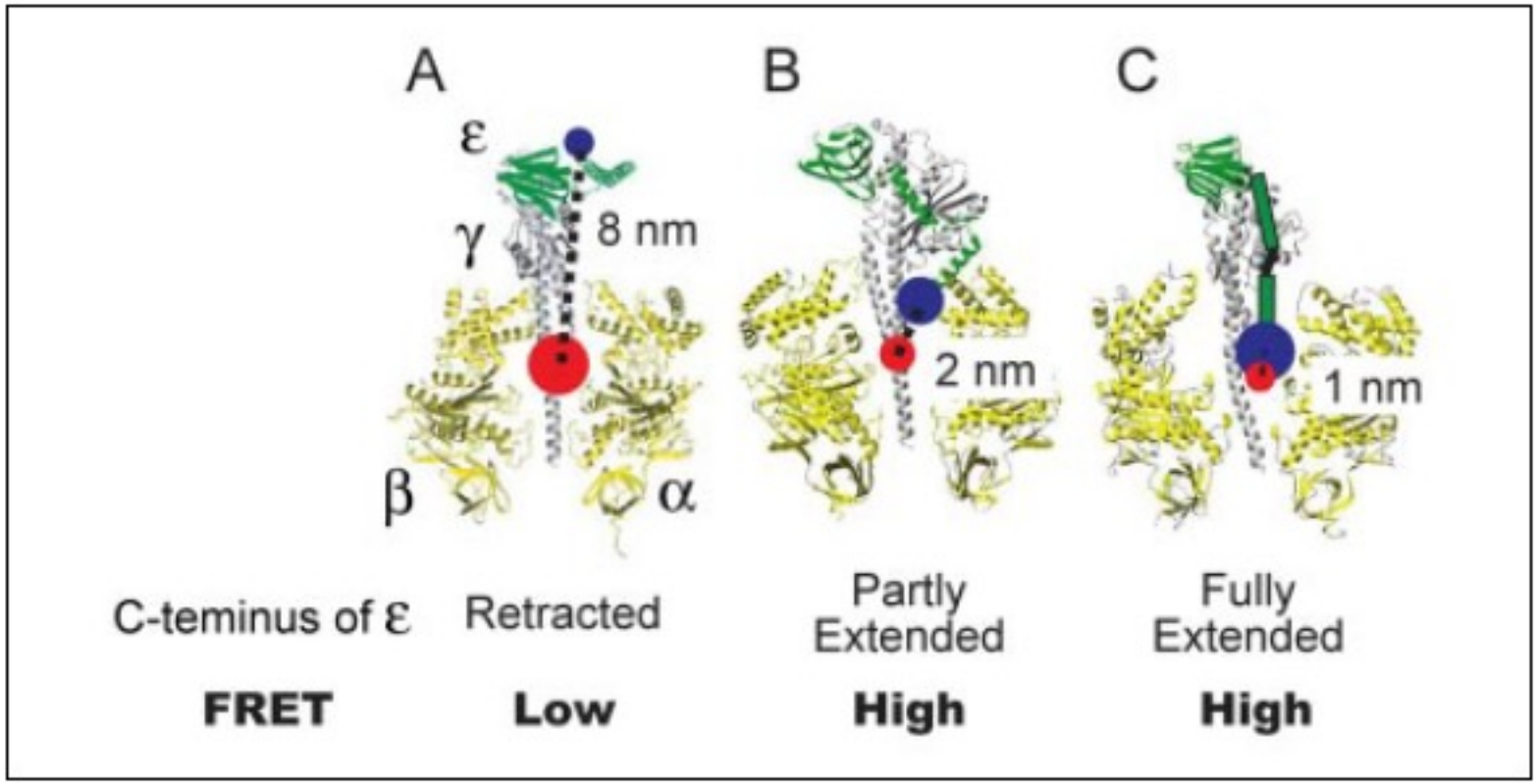
$\alpha_3\beta_3\gamma$(S3C-Cy3) broken line — Absorbs at 530 nm emits at 570 nm

ϵ(134C-Cy5), dotted line — Absorbs at 649 nm emits at 670 nm

$\alpha_3\beta_3\gamma$(S3C-Cy3) ϵ(134C-Cy5), **continuous line**



Demonstration by FRET that the two alpha helices of ε change conformation from a HIGH FRET extended form to a LOW FRET retracted form in the presence of ATP



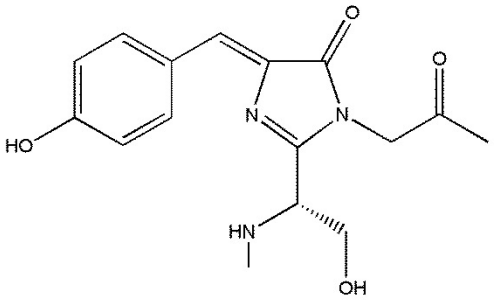
GFP variants with different spectroscopic properties

Table 1. Adapted from Tsien (1998).

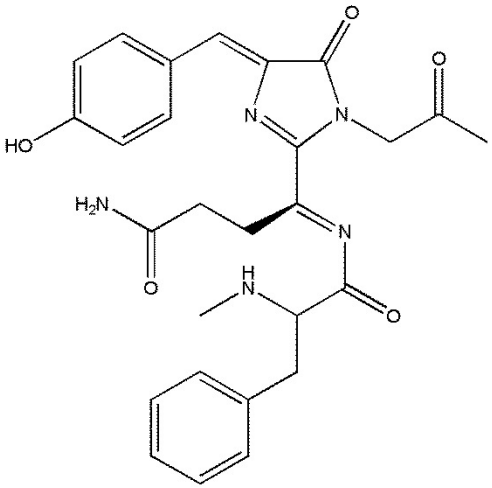
Fluorescent protein	Amino acid substitutions	Absorbance/emission
GFP (green fluorescent protein)	–	396, 488/505 nm
EGFP (enhanced GFP)	Ser65Thr, Phe64Leu	488/509 nm
CFP (cyan fluorescent protein)	Phe64Leu, Ser65Thr Tyr66Trp, Asn146Ile Met153Thr, Val163Ala Asn212Lys	434, 452/476, 505 nm
YFP (yellow fluorescent protein)	Ser65Gly, Ser72Ala Thr203Tyr	514/527 nm
(less pH-sensitive version of YFP) (Miyawaki)	Ser65Gly, Val68Leu, Gln69Lys, Ser72Ala Thr203Tyr	516/529 nm
BFP (blue fluorescent protein)	Tyr66His, Tyr145 Ser	434, 452/476, 505 nm
DsRed	–	558/583 nm

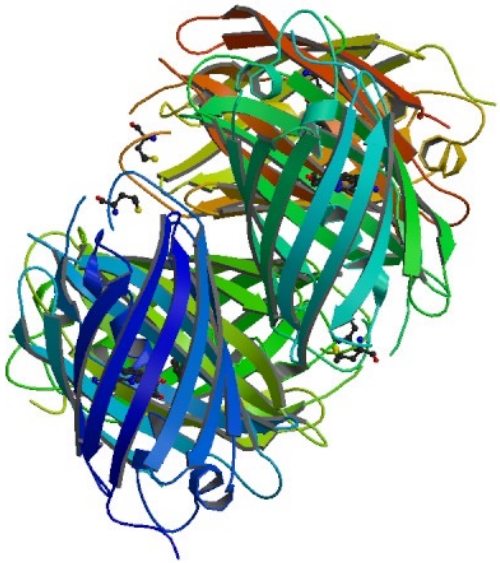
DsRed, red fluorescent protein from Discosoma sp.

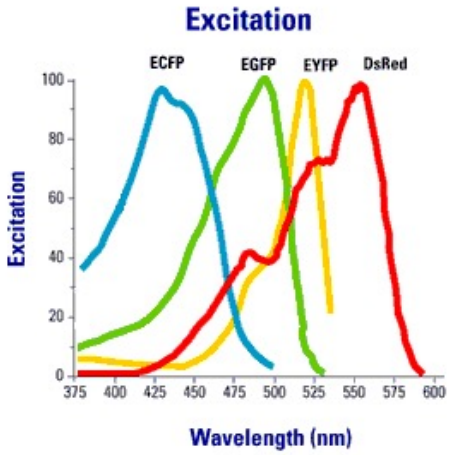
GFP

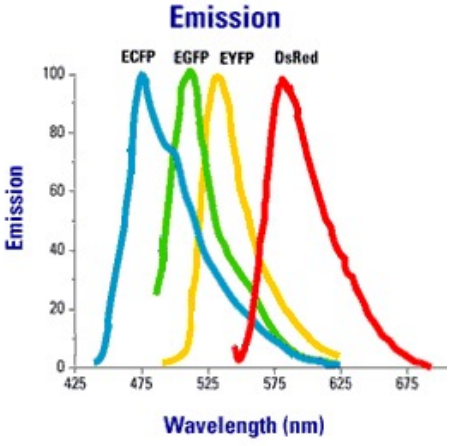


DsRED

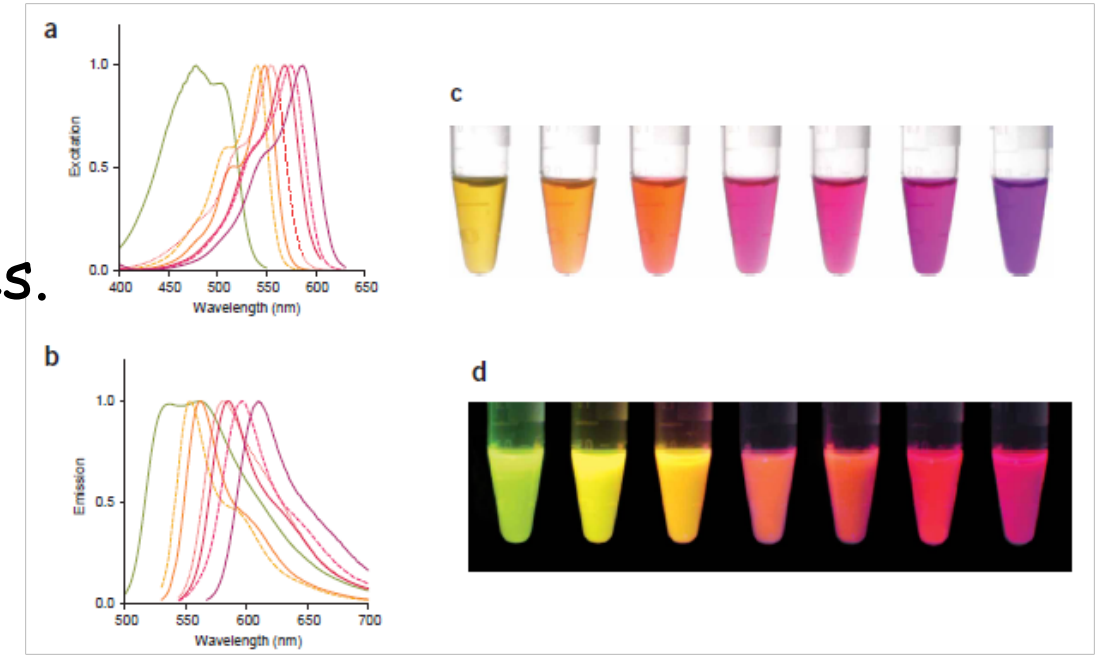


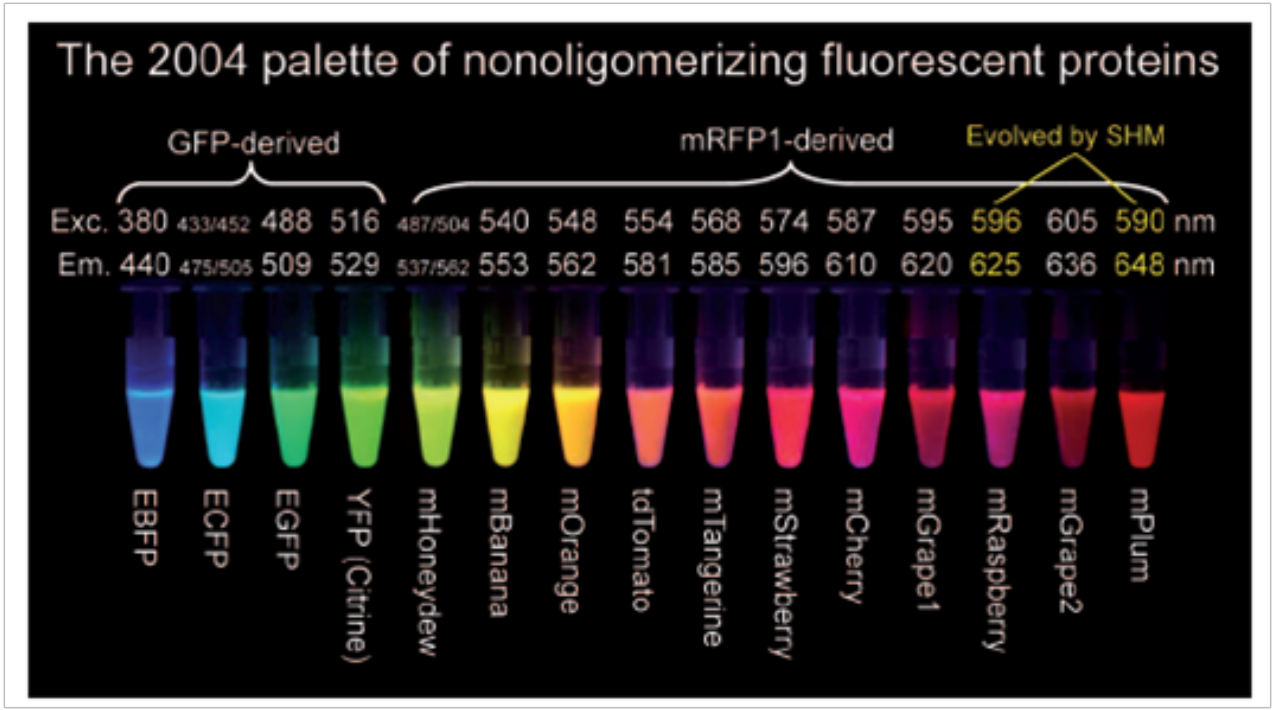






GFP and DsRed variants with different spectroscopic properties. These proteins were obtained by random and site-specific mutagenesis.

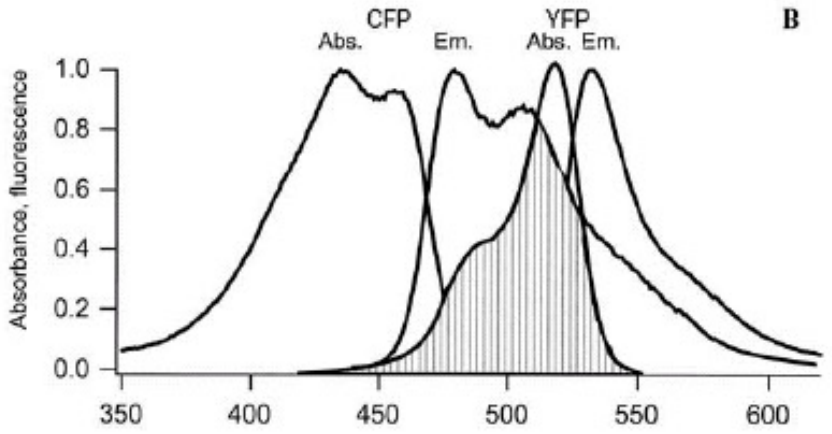




FRET CFP-YFP

CFP and YFP are employed as donor and acceptor in FRET experiments even if emission spectra overlap is high. The fluorescence ratio CFP/YFP is analyzed.

It is possible to perform FRET *in vivo* on fusion proteins.



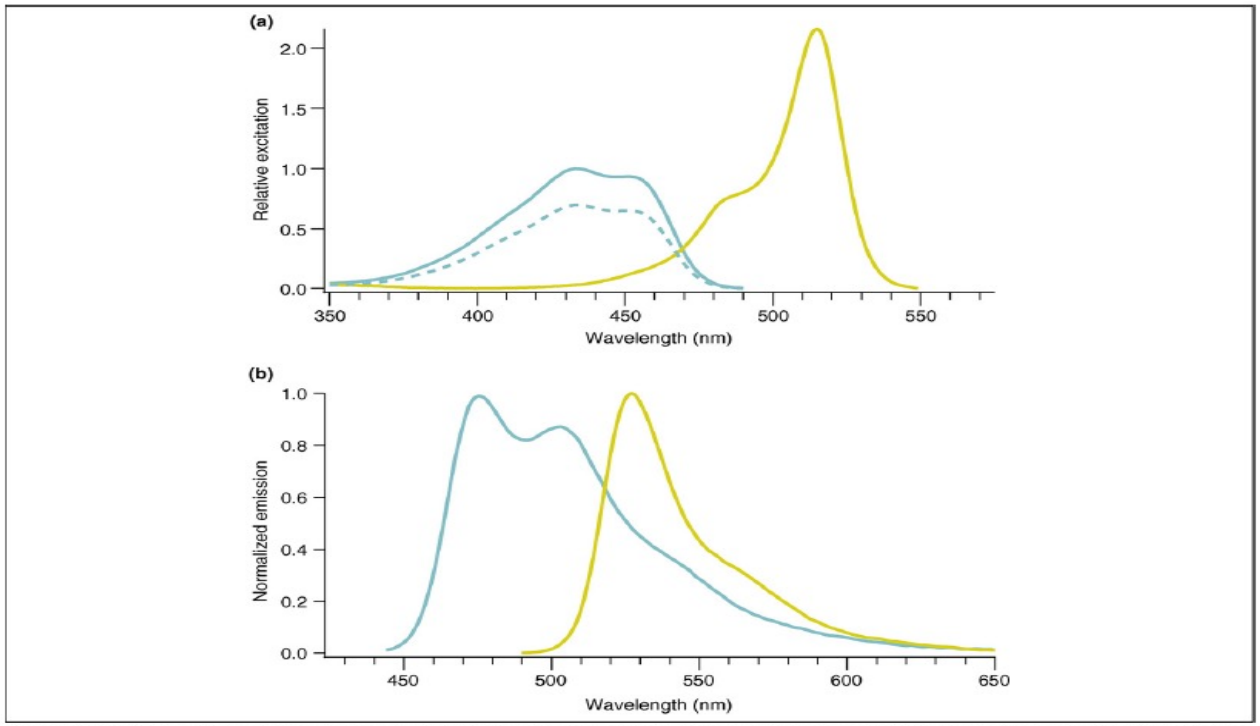
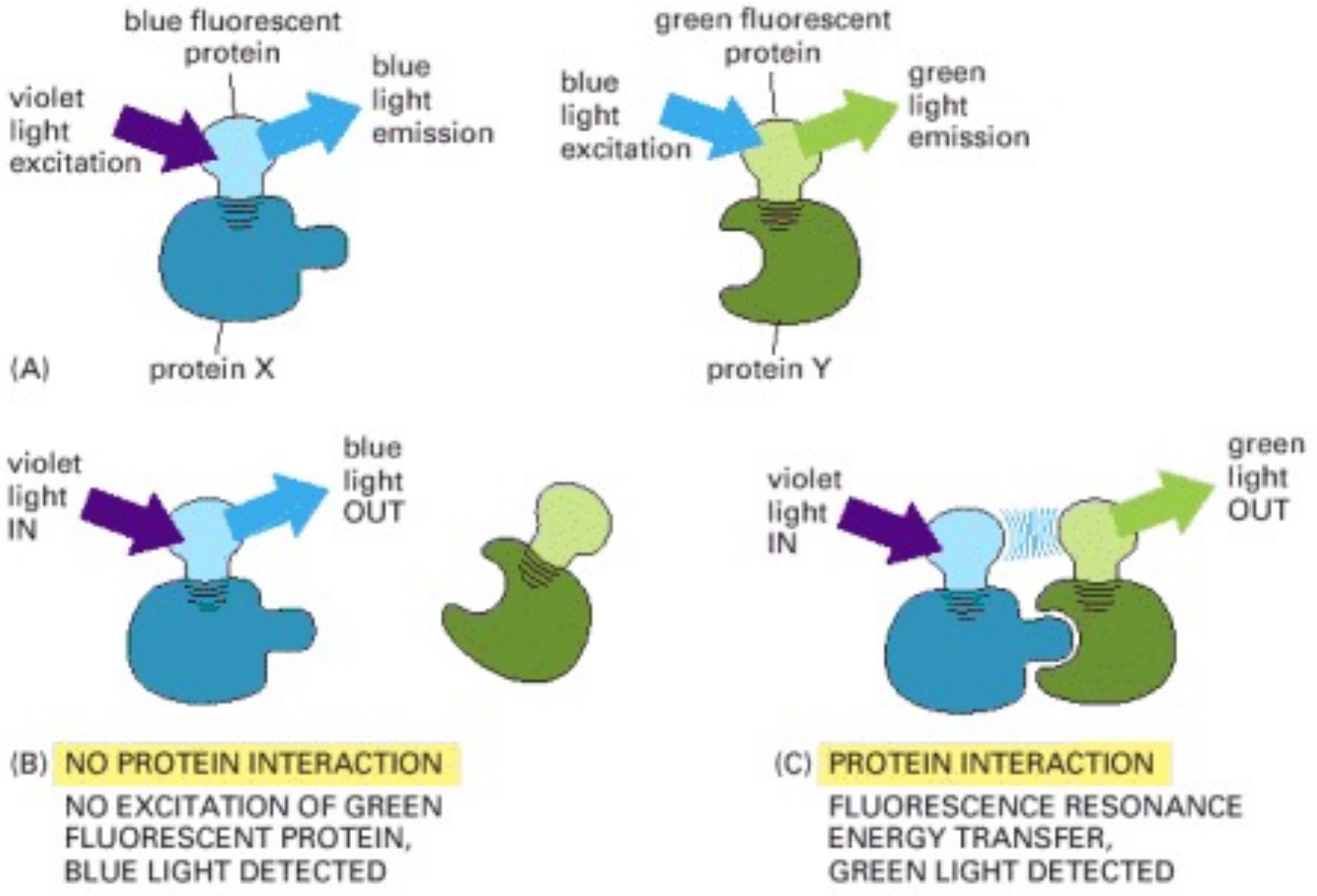


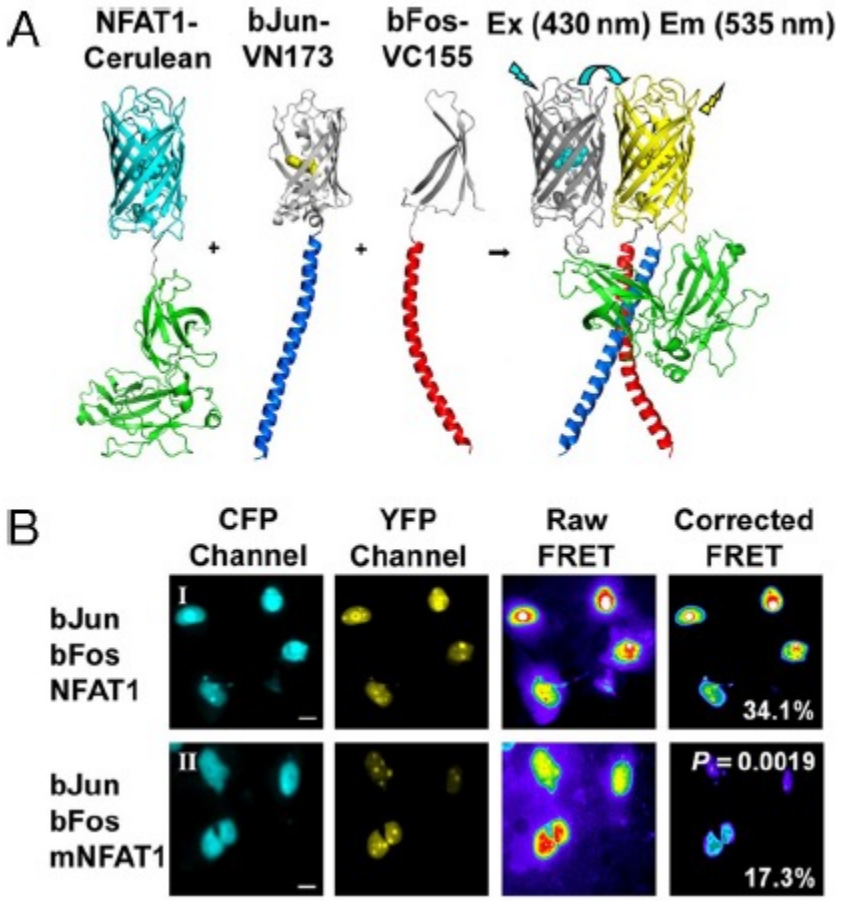
Figure 4. Overlap in the excitation and emission spectra of CFP and YFP. These two proteins have considerable overlap in both excitation **(a)** and emission **(b)** spectra. Direct excitation of the acceptor (YFP, yellow line) can be significant depending on the wavelength used for excitation of the donor (CFP, blue lines) owing to the higher extinction coefficient of YFP as compared with CFP. This overlap is especially problematic when enhanced CFP (dotted blue line) is used as the donor and can be partially overcome by using CFP variants with higher extinction coefficients such as mCerulean (unbroken blue line) or SCFP3A. The broad fluorescence emission spectrum of CFP shows considerable intensity in the region of YFP emission.

FRET BFP-GFP mediated by interaction between proteins X and Y



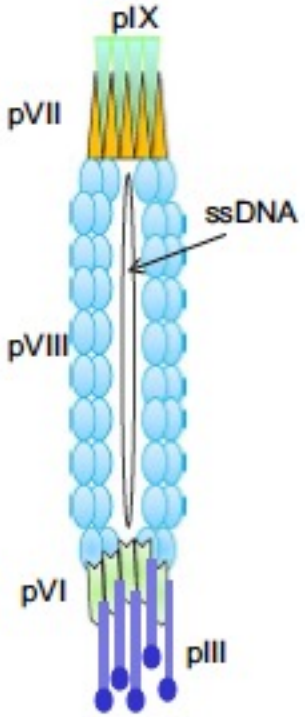
Analysis of the ternary complex of transcription factors AP-1 and NFAT

BiFC-FRET between Cerulean and Venus, two CFP and YFP variants.
NFAT1 is a transcription factor required for T cell development, it forms a complex with AP-1 (Jun-Fos heterodimer).
The 'leucine zipper' regions of Jun and Fos are fused to N- and C-terminal fragments of Venus (BiFC).
Interaction with NFAT1 produces FRET between Cerulean and reconstituted Venus.

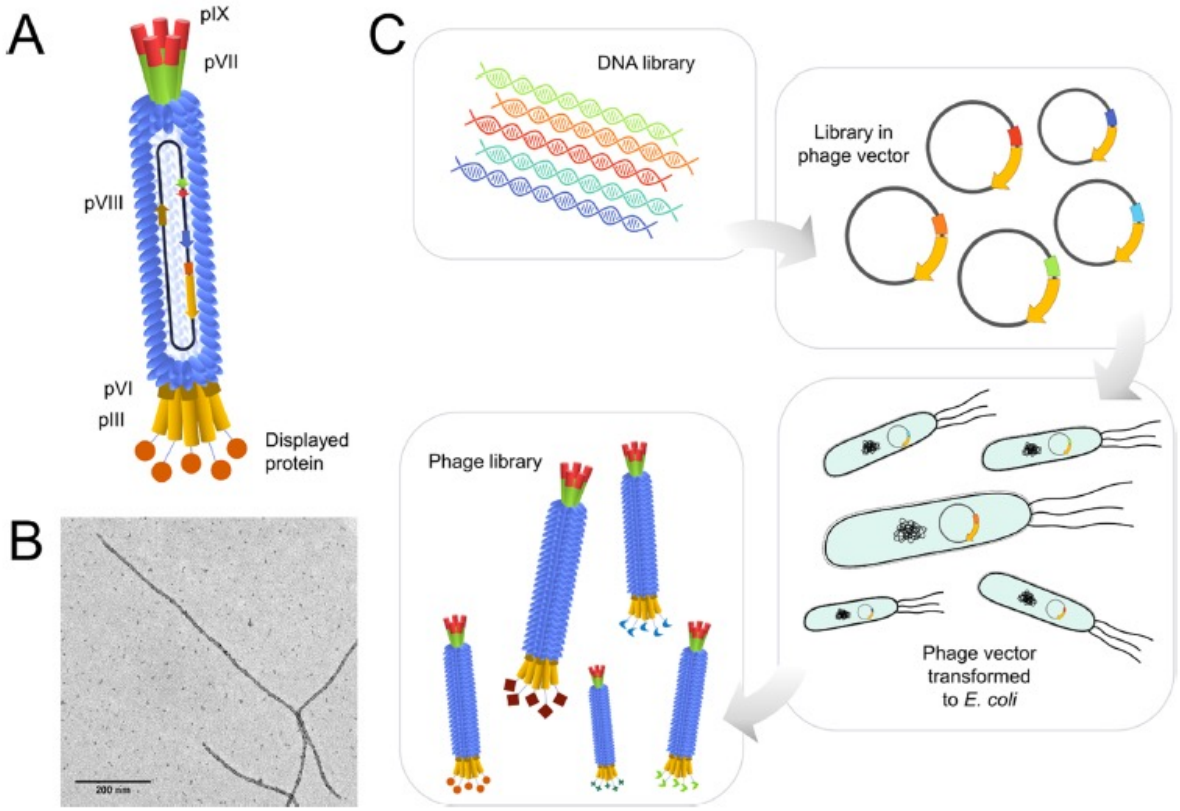


Phage display

- Technology described in 1985 by George Smith (Nobel for Chemistry in 2018) to select peptides/proteins that bind a target molecule
- Requires exposure of peptides/proteins on the surface of a phage
- Peptide or protein coding sequences are fused to the sequence a protein of the phage coat
- The lysogenic filamentous phage M13 and its pIII protein are the most utilized



Phage display



Fusion at 5' end of pIII protein

Construction of a phagemide and cotransformation in *E. coli* with a helper phage

Figure 1. (A) Structure of filamentous phage M13 which consists of a protein coat made of the major coat protein (pVIII) and the minor coat proteins (pIII and pVI on one end and pVII and pIX on the other). The genomic DNA encoding for the coat proteins is enclosed within the protein coat. By introducing modified segments into the genomic DNA, the phage can be engineered to display foreign peptides or proteins as a fusion with one of the coat proteins, most commonly pIII. (B) Transmission electron micrograph of the filamentous phage M13 acquired with negative staining using a JEOL JEM-1400PLUS instrument operating at 120 kV, with a LaB6 electron source and a GATAN US1000 CCD camera (2k × 2k). (C) Construction of phage-displayed libraries includes generation of the DNA library encoding for the different variants and introduction of the variable sequences in the phage DNA (typically a phage vector or phagemid system). After transforming the DNA to bacteria, phages are amplified and will display an individual protein or peptide variants outside the virion.

Phage display: selection strategy of phages that bind the immobilized target molecule

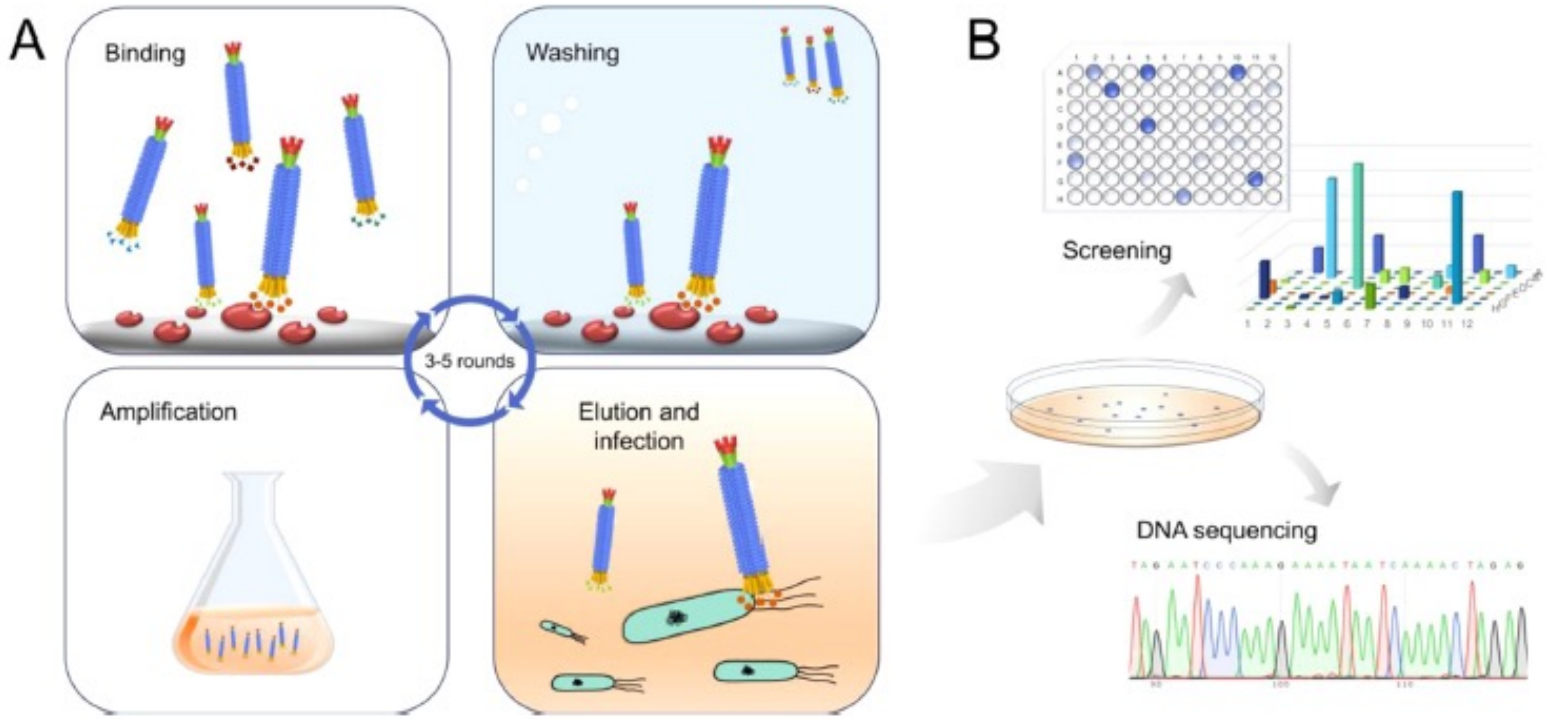


Figure 2. Selection of affinity binders from a phage-displayed library. (A) Selection process consists of binding, washing, elution, and amplification steps, which are usually repeated to three to five times to enrich target-specific binders. (B) Target specificity of individual clones can then be determined by screening the monoclonal clones from single colonies in enzyme-linked immunosorbent assay (ELISA), and the positive clones can be identified by DNA sequencing.

Phage display: limits in protein-protein interaction analyses

- Length of inserts < 1500 bp
- Export in periplasm for assembly of phage particles
- Construction of cDNA/ORF libraries is complex: it is necessary that inserts are *in frame* with pIII with no stop codons
- Direct or indirect fusions to pIII
- jun-fos system: pIII-jun fusion and fos-cDNA library. The jun-fos interaction is stabilized by disulphide bridges introduced ad hoc