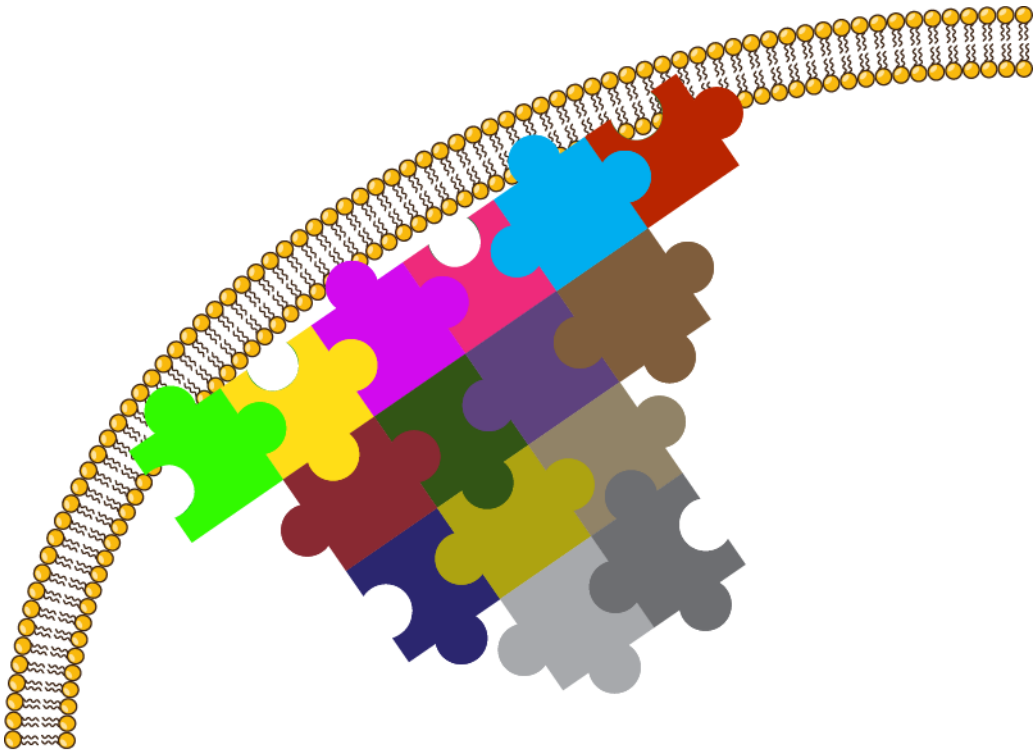

“HOW TO HUNT FOR INTERACTION PARTNERS OF MEMBRANE PROTEINS *IN VIVO*?”

Birol Cabukusta
Student No. 3562425
Master Programme of Molecular and Cellular Life Sciences
Master’s Thesis (7.5 ECTS)
11/06/2012 – 16/07/2012

Examiner: Dr. Toon de Kroon
Utrecht University
Faculty of Science, Dept of Chemistry
Membrane Biochemistry and Biophysics Group
Padualaan 8, Kruytgebouw N602, 3584CH, Utrecht, NL
a.i.p.m.dekroon@uu.nl

Abbreviations:

AbK	3’-azylbutyl-N-carbamoyl-lysine
AD	Activator domain
cAMP	Cyclic adenosyl mono-phosphate
cDNA	Complementary DNA
DBD	DNA-binding domain
EGFR	Epidermal growth factor receptor
ER	Endoplasmic reticulum
GAP	GTPase-activating protein
GEF	Guanine nucleotide exchange factor
GFP	Green fluorescence protein
GPCR	G-protein coupled receptor
MYTH	Membrane yeast two-hybrid
pAzpa	p-azido-L-phenylalanine
pBpa	p-benzoyl-L-phenylalanine
PCA	Protein complementation assay
rRRS	reverse Ras-recruitment system
RRS	Ras-recruitment system
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SRS	Sos-recruitment system
TfmdPhe	4’-[3-(trifluoromethyl)-3H-diazirin-3-yl]-L-phenylalanine
UAS	Upstream activating sequence
UBP	Ubiquitin-specific protease
UPR	Unfolded protein response
YFP	Yellow fluorescence protein

Chapter 1

GENERAL INTRODUCTION

In the early 1970s, the total number of protein-encoding human genes was estimated between 35,000 to over 100,000 based on genetic load arguments, hybridization experiments, the estimated average size of genes and the number of CpG islands [1]. The initial results of the human genome project suggested a total of 30,000 – 40,000 protein-coding genes and today the human genome is known to contain only around 21,000 distinct protein-coding genes [1], [2]. It was thought that about 100,000 protein-coding genes could cope with the complexity of cells, but only one fifth of the initial number (and maybe even less) is obtained today. Thus, the complexity must be mainly at the proteome level, i.e. interaction networks of proteins must be the driving force of complexity. As protein-protein interactions play a major role in cellular processes, they are often directly linked to human diseases. Therefore, knowledge of protein-protein interaction is essential to understand biological processes and molecular mechanisms at the systems biology level.

Lipid membranes serve as barriers between intracellular and extracellular space, and also between intracellular compartments. Proteins on these membranes are associated with many cellular functions. Plasma membrane proteins comprise receptors, pumps, pores, transporters and channels. They are responsible for reacting to changes in the environment; for maintaining homeostasis by exchanging ions and molecules and for transferring action potential. Moreover, cell adhesion proteins on the plasma membrane are required for multicellular formation and involved in metastasis of cancerous tissues. The membrane proteins in the endoplasmic reticulum and Golgi are responsible for folding and modifications of proteins in this tract. Furthermore, the membrane proteins in mitochondria are needed for energy production. Thus, membrane proteins are vital mediators of fundamental cell functions. Recent prediction studies revealed that at least 27% of the human proteome are transmembrane proteins [3],[4] and peripheral membrane proteins may even increase this percentage. Moreover, more than 50% of known drugs are targeted to membrane proteins [5]. Considering their role in cellular processes and homeostasis, the abundance of genes encoding them and their association with diseases, the significance of membrane proteins is not difficult to predict. Therefore, understanding membrane protein interactions is highly desirable.

Biochemical methods revealed many protein-protein interactions over the years, yet they are still useful. These methods include co-immunoprecipitation, tandem affinity purification, protein microarrays, surface Plasmon resonance, non-denaturing gel electrophoresis, blue native poly acrylamide gel electrophoresis and X-ray crystallography. However, these *in vitro* biochemical approaches suffer from several difficulties. First of all, large amounts of starting material are required and detergent-based lysis of the cells and isolation may cause artifacts. Subsequently to cell lysis, the spatial organization of cells might be lost, and two proteins that would not encounter at the cellular level, may seem to be interacting. Secondly, the absence of cellular factors, such as missing of

cofactors, different temperature and having different ionic strength may affect the interaction of interest. Third, biochemical assays have restricted sensitivity and have bias towards high affinity interactions [6]. Moreover, weak and transient interacting proteins can be removed during several washing steps. In addition to all the challenges mentioned, membrane proteins are even more difficult to study *in vitro*. Recombinant expression of membrane proteins is often toxic in *Escherichia coli* [7]. Moreover, subsequent solubilization of these proteins is often problematic, since proper folding of membrane proteins may not be achieved following a solubilization step using detergents. Because of all these difficulties that may be encountered in application of biochemical methods, several optimization steps are required to observe positive readouts. On the other hand, *in vivo* setups to detect protein-protein interactions are valuable, since there is no lysis of cells involved and they require little individual optimization. Furthermore, for *in vivo* methods less amount of starting material is needed, and isolation and solubilization of membrane proteins are not required. Thus, *in vivo* setups are useful to study membrane protein-protein interactions.

Even though *in vitro* approaches helped our knowledge to understand protein interaction networks over the years, they are on the edge of being outdated since they are time consuming and may be ineffective, especially for membrane proteins. On the other hand, *in vivo* techniques became more and more favorable during the last decades, because of several advantages over *in vitro* methods. Here, we present two of the best available techniques, advanced yeast two-hybrid systems and *in vivo* incorporation of photo-crosslinking amino acids into proteins, for studying interactions of membrane proteins with other membrane proteins and/or cytoplasmic proteins. Since both systems are *in vivo*, they do not require considerable optimization for interactions, moreover, weak and transient interactions can be observed in these setups. In addition to their ability to test an interaction between two specific candidate proteins, it is also possible to identify novel interacting partners of membrane proteins by using these methods in combination with cDNA libraries.

Chapter 2

ADVANCED YEAST TWO-HYBRID TECHNOLOGIES

2.1 Introduction

The yeast two-hybrid method is based on the *Saccharomyces cerevisiae* GAL4 transcriptional activator which is responsible for transcription of the genes encoding enzymes required for galactose utilization. The GAL4 transcriptional activator consists of two domains, an N-terminal DNA-binding domain (DBD) which binds to the upstream activating sequence (UAS) and a C-terminal activator domain (AD) recruiting yeast RNA polymerase II and initiating mRNA synthesis. However, either of these domains needs the other to start transcription [9]. Fields and Song constructed plasmids encoding two proteins of interest fused to a GAL4-DBD and a GAL4-AD domain, respectively. Subsequently to transformation of both plasmids into yeast, the interaction between two proteins of interest reconstitutes a functional transcription factor that can initiate transcription (Figure 2.1) [10]. Moreover, the yeast genome can be manipulated to have a reporter gene and GAL4-UAS upstream of that reporter gene. The reporter can be a simple enzyme like *Escherichia coli* β-galactosidase, thus, upon interaction of two hybrid proteins, transcription of β-galactosidase can be initiated and the activity of enzyme can be measured by a simple colorimetric assay. Alternatively, the reporter gene can be an auxotrophic marker whose expression is controlled by the GAL4 promoter activity. In this case, viability of auxotrophic yeast strains on selection medium is only possible when interaction occurs.

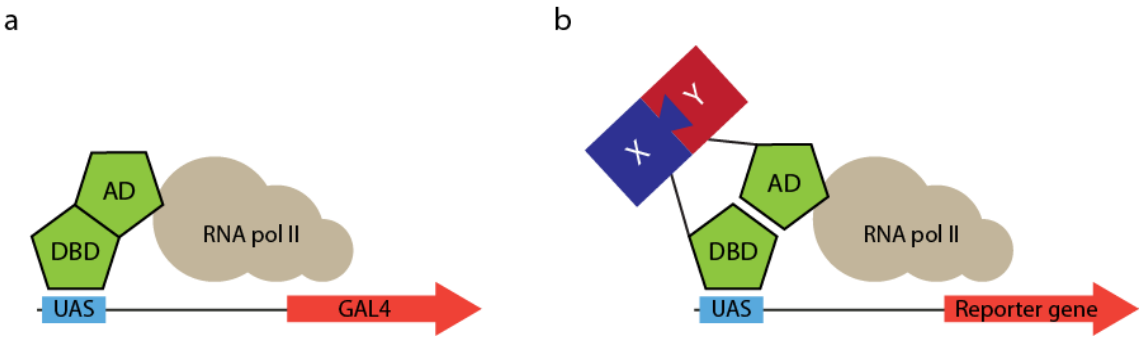

Figure 2.1: The original yeast two-hybrid system. **a.** Findings of Keegan *et al*. GAL4 transcription activator consists of two domains: DBD (DNA-binding domain) binds to the UAS (upstream activating sequence) and AD (activator domain) recruits RNA polymerase II to start GAL4 gene expression. **b.** The original yeast two-hybrid system developed by Fields and Song. Two interacting proteins X and Y (bait and prey, respectively) fused to DBD and AD can recruit transcription machinery. Moreover, a reporter gene can replace GAL4 gene in this system.

Since it is relatively easy to implement and inexpensive, the yeast two-hybrid system rapidly became the system of choice to detect protein-protein interactions. Initially, the yeast two-hybrid system was applied to confirm the interaction between known proteins. As the next step, a library screen was performed, in which plasmids containing yeast genomic fragments were fused to GAL4-

AD (called prey, typically a library of known or unknown proteins or a single known protein) and a protein of interest was fused to GAL4-DBD (called bait, fusion of protein of interest) in another plasmid. Subsequently to the transformation of both plasmids, DNA sequences obtained from reporter-positive yeast are used to identify the interacting partners of the protein of interest[11]. Several variations of the original system have been developed. In the so-called yeast three-hybrid system, the expression of a third protein besides DBD and AD fusion proteins can function as a bridge and the expression of the reporter gene can only occur in the presence of this protein [12]. Alternatively, an RNA strand can function as a bridge between two proteins and RNA-dependent interactions can be observed [13]. Up to date over 98,000 protein-protein interactions were identified in the budding yeast and over 12,000 of these interactions were identified using the yeast two-hybrid systems. Moreover, over 114,000 of 296,000 interactions identified in all organisms were obtained by the original yeast-two hybrid systems and its derivatives (Figure 2.2) [14].

Like any other technology in biology, the yeast two-hybrid system has its limitations. One of the major disadvantages of the system is that two interacting proteins must be targeted to the nucleus. However, integral membrane proteins or membrane-associated proteins cannot be imported into the nucleus. For this reason, this system cannot be applied to membrane proteins. Truncated versions of membrane proteins can still be targeted to the nucleus, however, truncation can lead to misfolding, and the problem remains since the nucleus is not the natural environment of these proteins.

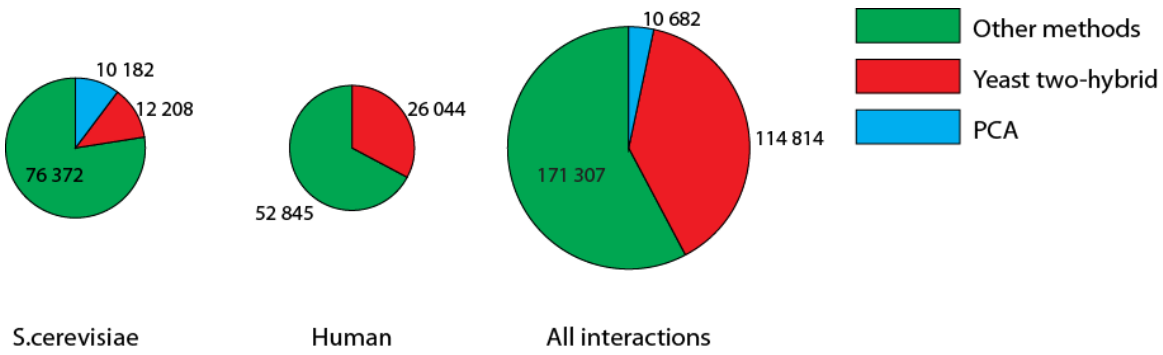

Figure 2.2: Contribution of the yeast two-hybrid methods and protein complementation assay(PCA, See Chapter 2.6) to all known protein interactions, according to IntAct molecular interaction database (http://www.ebi.ac.uk/intact/) [14].

To date, many variants of the yeast two-hybrid system as well as other related methods have been established to be able to determine membrane protein-protein interactions. Here, we will focus on these developed yeast two-hybrid systems and recent molecular biology studies on membrane proteins using these methods.

2.2 Sos-, Ras- and reverse Ras- recruitment systems

Guanine nucleotide-binding proteins (G-proteins) are signaling proteins that are highly conserved among eukaryotes, and they are engaged in several cellular functions such as cell proliferation, differentiation and apoptosis [15], [16]. In their inactive state they are bound to a GDP molecule, and they switch to an active state by exchanging GDP for GTP. They return to their inactive state by hydrolyzing GTP to GDP. G-proteins are divided into two main groups: monomeric G-proteins and heterotrimeric G-proteins [16]. Much knowledge of G-protein signaling in yeast is known since it is less complicated in yeast than in higher eukaryotes.

Ras and Ras-like proteins, abundant small monomeric G-proteins, are found mainly at the inner surface of the plasma membrane. In *S.cerevisiae*, growth and metabolism in response to nutrients is regulated largely by the Ras-cAMP pathway. Activated Ras proteins increase the cAMP levels by activating adenylate cyclase, which synthesizes cAMP. Then, increased cAMP levels cause the activation of cAMP-dependent cytosolic protein kinases, which have an essential role in proliferation [16–18]. Activities of Ras proteins are regulated by two classes of proteins: GTPase-activating proteins (GAPs) and guanine nucleotide exchange factors (GEFs).GAPs are the negative regulators of Ras-cAMP signaling, they accelerate the hydrolysis of GTP to GDP on Ras proteins resulting in a rapid inactivation of GTP-Ras. On the other hand, GEFs are activators of Ras-cAMP signaling pathway. When they are activated, GEFs stimulate the exchange of GDP to GTP in Ras proteins [19]. Cdc25p is a well-known RasGEF and it is involved in cell growth, cell cycling and sporulation. Thus, one of the main positive regulators of yeast Ras proteins is Cdc25p [16–18].

Initially Aronheim *et al.* showed that upon membrane localization, human RasGEF, hSos (human Son of sevenless), can substitute for Cdc25 to allow cell survival and proliferation [20]. In their next study, instead of targeting hSos to the plasma membrane directly, they employed a protein interaction to recruit hSos, whose C-terminal membrane localization signal is deleted, to the plasma membrane [21]. In this approach, a yeast variant strain that has a temperature sensitive cdc25-2 allele is needed. The cdc25-2 strain cannot survive at a certain temperature (36^0C), since the function of Cdc25p to activate Ras is lost, whereas it can grow at lower temperatures, such as 25^0C. In this so-called Sos-recruitment system (SRS), the prey protein must be membrane localized by a lipidation sequence and the bait must be cytosolic and fused to truncated hSos. Upon bait-prey interaction, hSos fusion protein will be recruited to the plasma membrane and yeast Ras signaling will be rescued. Thus, it will result in cell survival and growth of the temperature-sensitive mutant at 36^0C (Figure 2.3a).

Based on a similar principle, the Ras-recruitment system (RRS) is an improved version of the SRS [22]. Like RasGEFs activate Ras, also Ras itself is required to be localized to the plasma membrane for its function [20]. This localization is normally achieved via a lipidation motif at the C-terminus of Ras. In their approach, Broder *et al*. used a mammalian Ras (mRas) lacking its lipidation motif, thus blocking its membrane localization. An interaction between a membrane localized prey protein and a cytoplasmic mRas-fused bait protein must localize mRas to the plasma membrane. Thus, activity of mRas on plasma membrane must result in growth in cdc25-2 yeast strain at 36^0C

(Figure 2.3b). The RRS has the advantages of the SRS and also overcomes some of its limitations: Ras protein is smaller than hSos, thereby overcoming several technical limitations that may be encountered with a hSos fusion. Moreover, the RRS gives less background for many proteins compared to the SRS, resulting a more strict selection [22]. In the initial report, Broder *et al.* also showed that the RRS can be used to screen using a rat pituitary cDNA expression library [22].

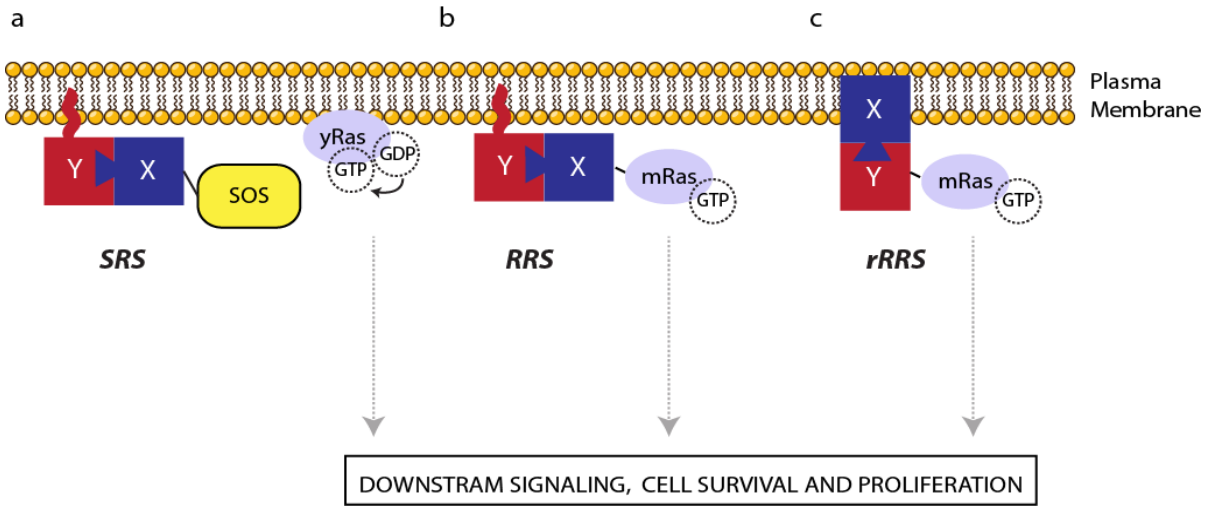

Figure 2.3: Sos-, Ras- and reverse Ras-recruitment systems. **a.** In the Sos-recruitment system (SRS), bait (X) fused to the truncated human Sos protein is recruited to the plasma membrane by the interaction with membrane localized prey (Y). Then, it activates the yeast Ras (yRas) protein. **b.** In the Ras-recruitment system, interaction between membrane localized prey (Y) and mRas-fused bait (X) mediates mRas to the plasma membrane and activates downstream signaling **c.** In the reverse Ras-recruitment system, the bait is a membrane protein, thus prey fused to mRas is mediated to the membrane by the interaction between the bait and prey.

Both Sos- and Ras- recruitment systems are employed when the bait proteins are cytosolic and soluble. Thus, these two methods cannot be applied to identify interacting partners of a membrane localized bait protein. In order to identify soluble interacting partners of membrane proteins, a method called the reverse Ras-recruitment system (rRRS) was developed [23]. Similar to the RRS, expression of mammalian Ras protein in yeast is aimed to complement the mutation in Cdc25p. In the case of the rRRS, however, the bait is a membrane protein and prey proteins are fused to cytoplasmic Ras protein. Upon protein-protein interaction on the plasma membrane, efficient growth occurs in Cdc25-2 yeast cells at higher temperatures (Figure 2.3c) [23].

In a study, Frankel *et al.* identified the interaction between RalA, a member of Ras superfamily of small GTPases, and ZONAB protein, ZO-1-associated nucleic acid-binding protein, using the rRRS [24]. It is known that Ral proteins are required for the activation of phospholipase D1, transcription factor NF-κB and Jun kinase [25–27]. Moreover, it has been reported that RalA protein takes a part in vesicle trafficking through association with exocyst complex via interacting with Sec5 [28]. Since the plasma membrane localization of RalA was already known [29–31], they employed the rRRS to use native RalA as the bait protein to screen a rat pituatory cDNA library attached to the C-

terminus of Ras protein [24]. Subsequently to the extraction of plasmids from Cdc25-2 yeast colonies which can grow at 36^0C, they identified ZONAB, a Y-box transcription factor that binds to the SH3 domain of ZO-1, as an interacting partner of RalA. They also reported that this interaction is GTP-dependent and RalA binding can regulate ZONAB function [24].

In spite of the fact that the rRRS is specifically developed to study interactions of membrane proteins with soluble preys, the RRS was employed more often than the rRRS for membrane proteins. In these studies, soluble cytoplasmic domains or fragments of membrane proteins of interest are fused to Ras protein to be used as bait. This preference is due to an advantage of the RRS over the reverse RRS. When a cDNA library is screened, in the RRS, the library is fused to a small lipidation sequence (less than 20 amino acids), but in the rRRS, it is fused to a Ras protein which is about 200 amino acids. If they are compared, the possibility of steric hindrance in a library screen, which may contain up to more than 100,000 fusion proteins, is less when they are fused to a lipidation peptide. This might be the main reason why the RRS is favored over the rRRS, even for membrane proteins.

In a recent study by Kokkola *et al.*, it has been reported that the somatostatin receptor subtype 5 (SSTR5), a G-protein coupled receptor, interacts with a putative palmitoyl transferase, ZDHHC5 by using the RRS. Even though palmitoylation of many G-protein coupled receptors (GPCRs) has been demonstrated, the enzymes responsible for this modification have been unknown. In this study, soluble C-terminal fragment of rat SSTR5 was fused to Ras protein to be used as bait. After screening the bait with an embryonic mouse brain cDNA library, it is shown that ZDHHC5 interacts with C-terminal end of SSTR5. Moreover, it has been also shown that ZDHHC5 has a positive effect on SSTR5 palmitoylation [32]. Using the same approach, Cao *et al.* used soluble α-kinase domain of mouse transient receptor potential melastatin type 6 (TRPM6), an epithelial Mg+2 channel, as bait for mouse kidney cDNA library and identified its interaction with methionine sulfoxidereductase B1 (MSrB1) [33].

In conclusion, the rRRS, but not the SRS or the RRS, was developed to study interactions of membrane proteins. However, in the rRRS approach, only cytoplasmic proteins can be screened for an interaction with the membrane protein of interest. Moreover, as mentioned above, the RRS is employed more often than the rRRS for membrane proteins. The reason for this preference could be (i) the risk of steric hindrance of proteins in a cDNA library when they are fused to a Ras protein for the rRRS system (ii) the rRRS only allows the interactions between the membrane baits and soluble preys; on the other hand in the RRS, both soluble and membrane preys can be used, if a soluble fragment of a membrane protein can be used as the bait. Another disadvantage of using the rRRS system is that membrane proteins in a cDNA library can localize to the plasma membrane independently during a screen. To avoid that possibility, membrane-associated prey proteins must be excluded from the screen. Moreover, the existence of yeast homologs for mammalian membrane proteins may result in complications with the rRRS due to possible interaction of a prey with the yeast homologs independent of the interaction with the expressed bait. However, in that case, the yeast genome can be manipulated by knockout of genes encoding homologous proteins. A big advantage of choosing the rRRS over the other methods, including the RRS and the SRS, in identifying interactions

is that the bait protein is used in its native form without any requirement for fusion to another protein. Therefore the rRRS is a useful approach for the baits that cause problems due to the steric hindrance when they are fused. However, the disadvantages listed above must be considered before it is employed.

2.3 Split Ubiquitin Yeast two-Hybrid System

The split-ubiquitin system is based on the properties of the ubiquitin protein [34]. Ubiquitin is small, 76 amino acid, highly conserved eukaryotic protein attached to lysine groups of proteins. Attachment of ubiquitin to a protein marks that protein for proteasomal degradation [35]. Proteins that are tagged with ubiquitin moieties become substrates for ubiquitin-specific proteases (UBPs), which cleave ubiquitin from these proteins. Following this cleavage, the ubiquitin is recycled and tagged protein is degraded by the 26S proteasome. Native ubiquitin is composed of two domains that are connected by a linker. Johnsson and Varshavsky showed that inserting artificial linkers of variable length still allowed the folding of the two domains into native ubiquitin [34]. Moreover, when they co-expressed, two domains of ubiquitin, so-called Nub and Cub (N-terminal ubiquitin and C-terminal ubiquitin, respectively) can reassemble spontaneously in the absence of any linker, and the reassembled ubiquitin can be recognized by UBPs to be cleaved. In order to abolish their spontaneous re-association, point mutations were introduced in Nub (I13G or I13A resulting in NubG or NubA, respectively). In these mutants, efficient association is observed only if the two split-ubiquitin moieties are brought into close proximity. Later they employed these findings to identify protein-protein interactions. In the so-called split-ubiquitin system, prey protein is fused to NubG or NubA, while the bait protein and a hybrid transcription factor are fused to Cub. This transcription factor consists of the bacterial LexA (DNA-binding domain of *E.coli* repressor LexA) protein and the *Herpes simplex* VP16 transactivator domain[36]. Cub and NubG reassemble only in the case of bait and prey interaction, and following the formation of native ubiquitin, the transcription factor is released by UBP-mediated cleavage. Therefore, interaction can be detected by the expression of LexA-VP16-regulated reporter genes (Figure 2.4a). The split-ubiquitin system fused to a transcription factor is also called the membrane yeast two-hybrid (MYTH), because localization of the transcription factor-fused bait protein to a membrane is strictly required, since soluble bait fused to a transcription factor can diffuse to the nucleus and initiate transcription without any interaction.

In their original system, Johnsson and Varshavsky used dihydrofolate reductase as reporter protein [34]. Using SDS-PAGE, the release of the reporter protein was detected. However, this readout system was not convenient, since immunoprecipitation and electrophoresis steps were needed. In order to have a better readout system, Ura3p protein, an enzyme involved in the synthesis of pyrimidine ribonucleotides, has been employed [37]. Activity of Ura3p leads to uracil prototrophy and sensitivity to 5-fluoroorotic acid (5-FOA), because 5-FOA is converted into the toxic compound 5-fluorouracil causing cell death. As the yeast two-hybrid reporter, rUra3p, a variant of Ura3p which is N-terminally modified for rapid degradation according to the N-end rule, is used [38]. Thus, if interaction between bait and prey occurs,ubiquitin is cleaved from rUra3p resulting in rapid

degradation of rUra3p and allowing the cells grow on medium containing 5-FOA. Using this method, cells expressing two interacting proteins are identified by their ability to survive on 5-FOA, and interactions between several integral membrane proteins were mapped [37].On the other hand, for membrane proteins, LexA-VP16 reporter protein is the most employed readout technique [39], since transcriptional readout causes amplification, thus offering more sensitivity for transient interactions.

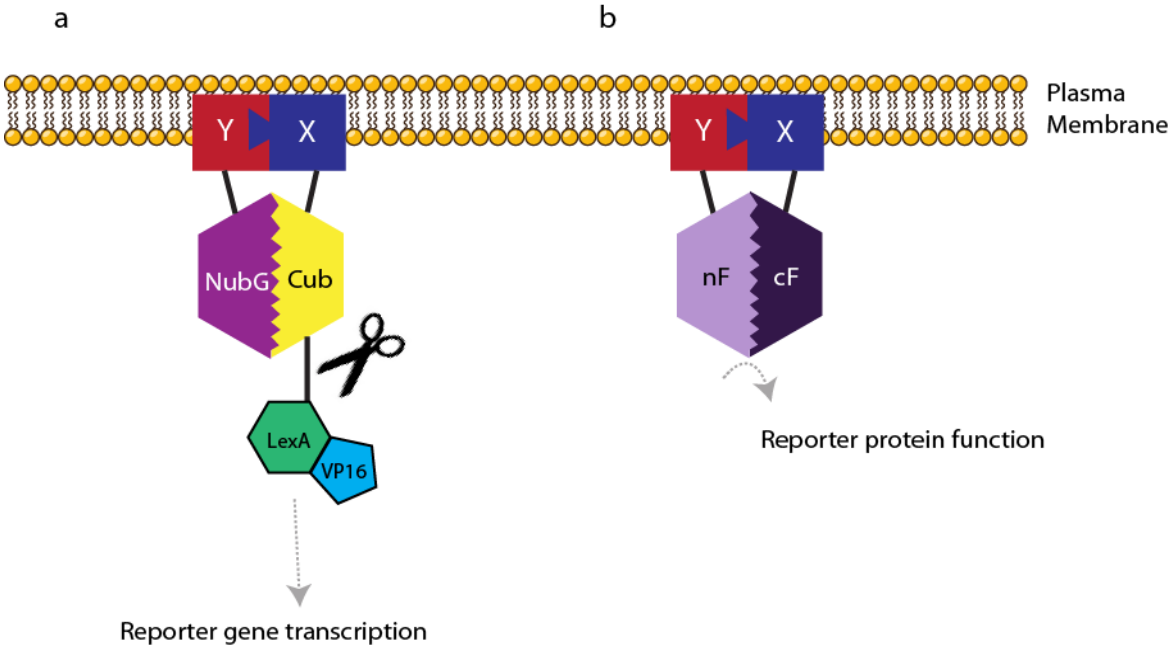

Figure 2.4: a. Split-ubiquitin system (or the membrane yeast two-hybrid system). Reassembly of native ubiquitin is provided by the bait-prey (X and Y) interaction. Thus, UBPs can cleave ubiquitin and the transcription factor of the reporter gene is released. **b.**Protein complementation assay. Two fragments of a protein are brought together by an interaction between bait (X) and prey (Y). Reconstitution results in direct reporter protein function.

For systematic identification of interactions of membrane proteins with both membrane and soluble proteins, the mating-based split ubiquitin system was developed [40], [41]. This method has a similar readout as the split-ubiquitin system, and additionally the split-ubiquitin system was further improved to develop a high-throughput method. In this improved approach, the bait and prey proteins are fused to the two halves of ubiquitin and expressed in yeast cells of opposite mating type (a and α). Upon mating, the diploid cell co-expresses the proteins. An interaction brings the two fragments of ubiquitin into close proximity resulting in the cleavage and release of transcription factor to activate an auxotrophic marker (Figure 2.5) [41].The advantage of this technique is that it is time efficient, i.e. instead of transforming the gene of interest through a complete cDNA library, only the opposite mating type of the library-containing yeast is transformed. Then an array screen is applied. Using the mating-based split ubiquitin system they developed, Fields and colleagues identified 1985 putative interactions among 705 integral membrane proteins [41].

Recently, Deribe *et al*. [42] identified novel interaction partners of the human epidermal growth factor receptor (EGFR). EGFR is a single-span cell surface receptor implicated in the regulation of crucial cellular functions including cell growth, proliferation and cell survival [43], [44]. In

their work, they fused a human fetal brain cDNA library C-terminally to NubG and transformed it into yeast containing a EGFR-Cub-transcription factor construct. After isolation and identification of plasmids from positive yeasts, they identified 87 proteins that bound to the receptor and validated the complete set of these proteins with bioinformatics and a subset by immunoprecipitation (including proteins like HSP70 and α-adaptin). Among these proteins, they also showed that activity of histone deacetylase 6 (HDAC6), a cytoplasmic lysine deacetylase, is regulated by EGFR upon interaction. Furthermore, HDAC6 deacetylates α-tubulin in stimulated cells and a feedback mechanism was identified in which HDAC6 is inactivated by EGFR through phosphorylation.

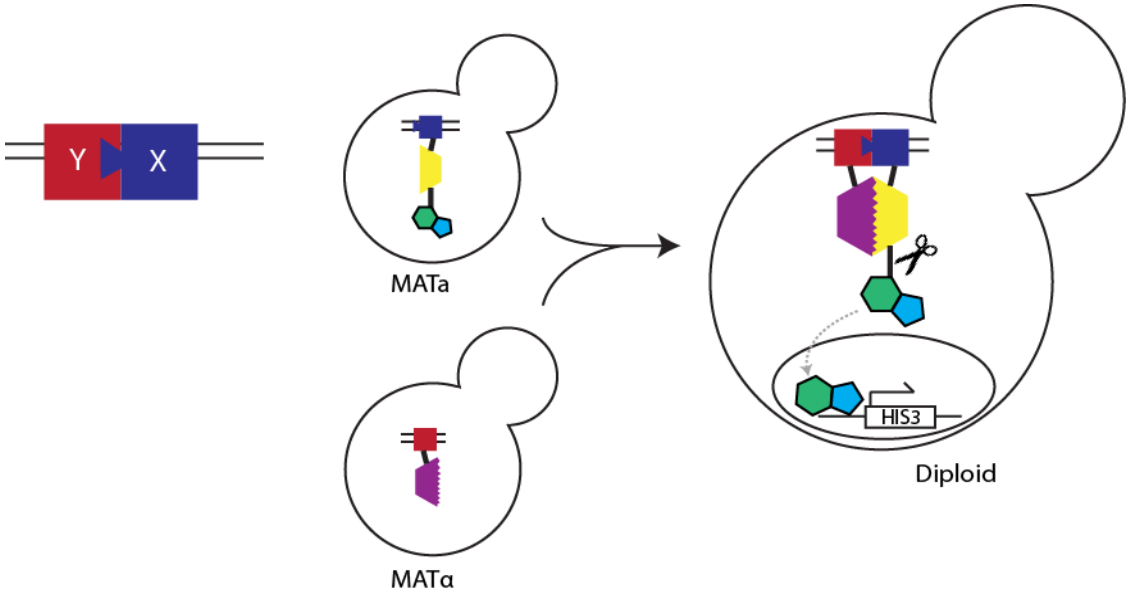

Figure 2.5: Two integral membrane proteins (X and Y) are fused to the two halves of ubiquitin and expressed in cells of opposite mating type. Upon interaction in diploid cells, which expresses both proteins, ubiquitin is reconstituted and cleaved by UBPs releasing the transcription factor to activate auxotrophic gene transcription. (This figure is adapted from the original study [41])

Split-ubiquitin can also be applied to multi-pass membrane proteins such as G-protein-coupled receptors (GPCRs). In one study, a GPCR superfamily member human mu-opioid receptor (MOR) was used as the bait protein [45]. Mu-opioid receptors bind to opioid agonist drugs, such as morphine and heroin, and chronic exposure may lead to opioid dependence causing neuronal alterations [46–48]. In this study, a MOR-Cub-transcription factor construct and NubG fusions of a fetal brain cDNA library were transformed into *S.cerevisiae* reporter strain THY.AP4.Out of 6 million colonies, 104 positive clones represented 10 distinct human proteins. Moreover, two of them were found to encode the C-terminus of GPR177, an orphan GPCR protein identified as a MOR interacting protein [45]. GPR117 is the mammalian ortholog of *Drosophila* Wntless/Evi/Sprinter, an evolutionary conserved protein that plays an important role in secretion of Wnt proteins from Wnt-producing cells [49–51]. In the follow up experiments they showed that morphine treatment promotes formation of MOR/GPR117 complexes followed by inhibition of Wnt secretion. In this study, the previously

unknown role for GPR177 in regulating the response to opioid drugs has been discovered using split-ubiquitin system.

Another recent example of successful application of split-ubiquitin system is the work by Paumi *et al.* who investigated interaction partners of a yeast ABC transporter Ycf1p [52]. ABC transporter superfamily of membrane transporters is ubiquitous in all organisms from microbes to mammals and they actively transport a broad range of substrates across cell membranes [53–55]. In their previous study ([41] also mentioned above), out of selected 705 annotated integral membrane proteins, only 365 were found to be competent for split-ubiquitin screening. Many of the remaining 340 membrane proteins were untestable due to their high expression levels resulting in mislocalization and self-activation of the reporter gene in the absence of prey. Paumi and his co-workers added an additional YFP protein to Cub, resulting in a C terminal Cub-YFP-TF tag for the bait protein (named integrated MYTH). Thus, they could verify the localization of expressed protein by fluorescence microscopy. Using their approach, they identified six Ycf1p-interacting proteins and among these proteins Tus1p, a cytosolic GEF for Rho1p, has an activating function on Ycf1p via the small GTPase Rho1p [52].

The split-ubiquitin system is the most employed yeast two-hybrid approach for membrane protein interactions. It has been shown that it can be applied to yeast [52], plant [56] and mammalian[42], [45], [57] membrane proteins. Moreover, this approach can be used to identify interactions of plasma membrane proteins, ER proteins [58] and outer nuclear proteins [59]. One of the most advantageous features of this system is having small ubiquitin fragments in hybrid proteins, which minimizes the possibility of steric hindrance. Additionally, even though the bait protein is a membrane protein, prey can be either a (trans)membrane protein like the bait or can be a soluble protein. However, it is strictly required that the bait protein fused to a transcription factor must be anchored to a membrane or at least have strong affinity for a membrane. Since soluble bait proteins can diffuse to the nucleus without interacting Nub and independently start transcription, soluble membrane-associated proteins cannot be used as baits for split-ubiquitin system with a transcription factor as reporter. Moreover, false positives may arise from proteolysis of the fusion protein and release of the transcription factor by unknown process (e.g. quality control mechanisms in the ER). However, all possibilities mentioned for false positive readout can be eliminated by proper control experiments. Moreover, split-ubiquitin based cDNA libraries for yeast, *C. elegans*, Drosophila and several mammalian tissues are available commercially, which also makes the split-ubiquitin system cost-effective.

2.4 G-protein fusion based yeast two-hybrid system

In another variation of the yeast two-hybrid system, working dynamics of heterotrimeric G-proteins, consisting of G_{α}, G_{β} and G_{γ} subunits are used. Upon GPCR activation, a conformational change in the G-protein α-subunit occurs and triggers exchange of GDP for GTP. This exchange, results in dissociation of the $G_{\beta/\gamma}$ subunits. Then the G_{α} and $G_{\beta/\gamma}$ subunits activate different downstream pathways until GTP is hydrolyzed and the G-protein reverts to its initial inactive conformation [16]. In

the G-protein fusion based yeast two-hybrid approach, the protein of interest, i.e. the bait is an integral membrane protein and the prey is expressed as a soluble fusion to a G-protein γ-subunit in G_γ-deficient yeast mutant [60]. The interaction between bait and prey recruits the G_γ fusion protein to the membrane where they sequester the G_β subunit, thus disrupting the formation of the heterotrimeric G-protein complex and the following downstream signaling in yeast [60]. In this assay, yeast survival and growth is only feasible when the bait protein does not interact with the prey (Figure 2.6).

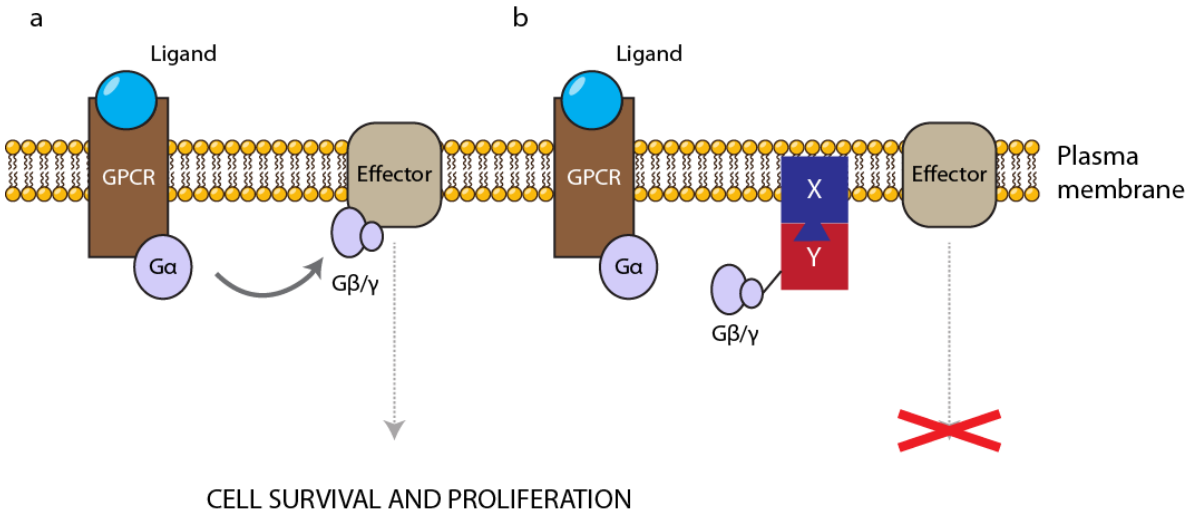

Figure 2.6: G-protein fusion based yeast two-hybrid system **a.** Dynamics of a heterotrimeric G-protein. The activation of a GPCR leads to dissociation of Gβ/γ from Gα and activation of downstream effectors. **b.** In the Gγ-deficient yeast strain, a hybrid of the cytosolic prey (Y) protein with Gγ interacts with Gβ and membrane-bound bait (X) protein. Then, the interaction of Gβ to activated Gα is prevented, thus inhibiting downstream signaling and gene expression.

Subsequent to the discovery of their approach, the authors tested two well-known interactions by this method. First, they tested the interaction of syntaxin 1a with nSec1. Syntaxin 1a has a function in synaptic vesicle fusion and it is a proven drug target. nSec1 binding to syntaxin 1a prevents its assembly with the SNARE core complex, a protein complex required for bilayer fusion [61]. For their second screen, they also tested interaction between fibroblast growth factor-receptor 3 and SNT-1 [60]. Both interactions validated that the G-protein fusion based system can be employed for membrane protein interactions.

Even though the G-protein fusion based yeast two-hybrid system has a major advantage of having only one of the two binding partners as a fusion protein, this technique is not employed as often as others. The main disadvantage of the approach is that the interaction must be measured as reduction in growth. In fact this method was developed to discover drugs which can disturb the interactions of membrane proteins, i.e. when growth of yeast is inhibited in the case of interaction between bait and prey, a drug, which abolishes the interaction, can regain growth. The authors tested the interaction between syntaxin 1a and nSec1 since inhibitors of syntaxin 1a are used to treat muscle spasms that result from improper neurotransmitter release [60]. Despite the fact that this method was developed to screen drugs on known membrane protein interactions, it is still an applicable approach

to identify membrane protein-protein interaction. However, a big limitation of G-protein fusion based yeast two-hybrid system is that the prey proteins can only be soluble proteins.

2.5 SCINEX-P system

The SCINEX-P (screening for interaction between extracellular proteins) system allows the analysis of protein-protein interactions that occur in the oxidizing environment of the ER [62]. In eukaryotic cells, proteins that are destined for secretion and cell surface localization as transmembrane proteins are processed in the ER. The ER offers a unique oxidizing environment, in which ER-resident chaperones enable proper folding and the formation of disulfide bonds[63]. Moreover, N-linked glycosylation is important for proper folding of proteins in the ER [64]. Considering the fact that proteins which are required to be processed in the ER in order to gain functionality or conformation, cannot be subjected to any of the yeast two-hybrid methods mentioned previously, the SCINEX-P system was developed, which is based on the properties of Ire1p, a type I transmembrane protein residing in the yeast ER that is responsible for the so-called unfolded protein response (UPR) [65]. When proper protein folding is abolished, unfolded or incorrectly folded proteins accumulate in the ER. This kind of stress is responded to by transcription of genes encoding ER-resident chaperones and enzymes that assist protein folding in the ER lumen. In yeast, the stress signal from the ER to the nucleus is transmitted by Ire1p. The N-terminal luminal domain of Ire1p functions as a sensor for the folding of ER proteins and controls the dimerization of the transmembrane protein. The C-terminal cytosolic part is a kinase and activated upon dimerization to start the stress response. As it responds to stress, activated Ire1p induces the production of the transcriptional activator Hac1p which later binds to the unfolded protein response elements in the promoter region of the genes encoding ER-resident proteins [66–68].

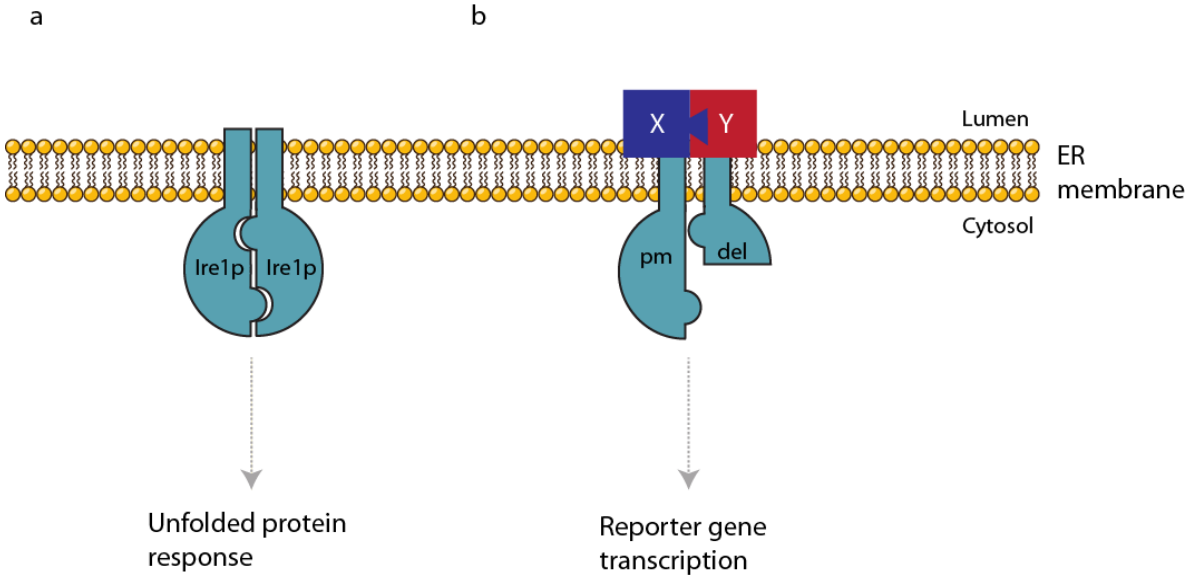

Figure 2.7: The unfolded protein response and its application for the detection of protein interactions. **a.** The ER resident transmembrane protein Ire1p functions as a sensor for unfolded protein, which induces Ire1p dimerization and eventually transcription from the unfolded response elements. **b.** In the SCINEX-P system, the components of the unfolded protein response pathway have been modified. One of the proteins of interest is fused to Ire1p carrying a point mutation (pm) and the other protein is fused to a C-terminal deletion mutant of Ire1p (del).

In the SCINEX-P system, proteins of interest are fused to the luminal N-termini of mutated Ire1p proteins. These mutations of Ire1p (one has a point mutation and the other contains a C-terminal deletion) block autonomous dimerization of the protein unless they are brought to close proximity by the interaction of the two proteins of interest. Interaction between proteins can be measured by the activity of the reporter gene which is inserted upstream of unfolded protein response elements which are activated by Hac1p [62]. This system is more favorable to be employed when at least one of the proteins is not a transmembrane protein on the ER, where previously mentioned systems, such as split-ubiquitin, cannot be applied.

To test their approach, Urech *et al*. tested the interaction between two different single chain antibodies (anti-GCN4wt and λ-Graft) and their common epitope the leucine-zipper of the yeast transcription factor Gcn4p. Moreover, in another study, Pollock *et al*. identified the interacting residues of calnexin with ERp57. Calnexin is a single span type I transmembrane protein which is known to be working cooperatively with ERp57,a soluble ER resident protein, for the quality control system of the ER [69].

The SCINEX-P system can detect interactions between extracellular and/or ER-derived membrane proteins. With this approach, the authors were able to detect an interaction that could not be observed using the conventional yeast two-hybrid system [62], [70]. On the other hand, this method is not applied as often as other yeast-two hybrid systems because of the technique being limited to the ER lumen. Moreover, even though most of the membrane proteins are derived from the ER, they might not be completely functional there, since they can be modified in the Golgi or even on the plasma membrane.

2.6 Protein Complementation Assays

Despite the fact that the protein complementation assay (PCA) is not a yeast two-hybrid method, it is noteworthy to mention it under yeast two-hybrid systems. It was well-known over 60 years that fragments of proteins can reassemble into functional complexes as in ribonuclease and β-galactosidase [71], [72]. However, the concept of split-protein reassembly to test protein-protein interactions was introduced in 1994 by Johnsson and Varshavsky, who demonstrated that when they are in close proximity two fragments of a ubiquitin can still function like a native ubiquitin [34]. As described above, many protein-protein interactions, including membrane proteins, were identified using this feature of ubiquitin. By using the same principle, many proteins are dissected into such fragments to test protein-protein interaction. The only difference, however, is instead of expressing a reporter gene as it is in split-ubiquitin system, split proteins function as reporters themselves (Figure 2.4b). For instance, interaction of bait and prey proteins fused to two halves of β-galactosidase protein can be observed by the β-galactosidase activity, rather than introducing a β-galactosidase gene as a reporter in a split-ubiquitin system. In principle, most of the proteins can be split into such fragments; however, there are a few criteria that must be met for the successful design of a split-

protein reporter. Each protein fragment by itself should not demonstrate any activity; affinity of the fragments in the absence of the interaction between bait and prey should be negligible; and the reassembled protein must provide a readily measurable read out. Up to date many reporter proteins were used to identify or test protein-protein interactions. As mentioned, β-galactosidase is one of the earliest ones to be used [73]. In a setup with split β-galactosidase proteins, interaction of candidate proteins can be measured by using a colorimetric assay. Other reporter proteins such as GFP [74], YFP [75] and luciferase [76–78] whose readouts can be detected by fluorescence techniques are also employed to test/identify interactions. Moreover, the auxotrophic marker Trp1p, a yeast tryptophan synthesizing enzyme, is also used as reporter protein in PCA [79]. Besides those mentioned, dihydrofolatereductase [80], [81], β-lactamase [82], [83], cytosine deaminase [84], TEV protease [85], thymidine kinase [86] and chorismatemutase[87] are among the proteins used as reporter genes.

Tarassov *et al*. performed a systemic binary screen for protein-protein interactions at a genome wide scale in *S. cerevisiae* using a PCA based on a dihydrofolate reductase assay [88]. They included 1124 endogenously expressed proteins and identified 2770 interactions. Among these interactions membrane proteins were also included, such as V-ATPase complexes and ESCRTs (endosomal sorting complexes required for transport) [88].

PCA is a useful and often preferred method to study membrane protein interactions. Up to date about 10 percent of yeast protein interactions were identified by PCA (Figure 2.1). The only difference between PCA and the yeast two-hybrid system is that there is no transcriptional step in PCA. However, this can be an advantage, since interaction of proteins can give direct readouts. However, it must be noted that the reassembly of some of the split-protein reporters (e.g. GFP) is irreversible. This feature can limit the ability to gain temporal information on interactions, although they are useful to trap low affinity interactions [89].

2.7 Discussion

The budding yeast *S. cerevisiae* has several advantages such as short life cycle, possession of eukaryotic secretory machinery and post-translational modifications. Another powerful feature of yeast is easier manipulation of the genome compared to mammalian cells. Thus it is also used to establish new technologies to study interactions of eukaryotic proteins.

The yeast two-hybrid technologies presented here are effective tools to fish for membrane protein interacting partners. As summarized in Table 2.1, not all of these methods can be used to test an interaction between two membrane proteins. Even though the bait protein is a membrane protein, some of these methods can only be used to fish for soluble proteins (the rRRS, G-protein fusion based system). However, methods like the split-ubiquitin and the protein complementation assay have minor limitations making these methods the most often employed.

Even though yeast two-hybrid approaches are cost-effective and readily applied, they can be problematic due to some features of the host organism. First of all, it is a known fact that expression of some membrane proteins might be toxic for yeast. If an auxotrophic marker is used to observe an

interaction, toxicity of a membrane protein may give a false negative. Second, expression levels of the proteins of interest are also challenging. For instance, in a split-ubiquitin system, a highly expressed bait protein can mislocalize to the nucleus and initiate reporter gene transcription on its own. Or else, in an rRRS setup, mRas-fused prey protein with a low expression level may not be able to find its partner on plasma membrane. Moreover, if a non-yeast membrane protein is tested, it must always be taken into consideration that missing or additional cofactors and proteins in the yeast milieu may give false readouts. Another possible drawback is that if protein-protein interactions depend on free amino or carboxyl termini, steric hindrance due to the fusion can give a negative readout. That is why most interactions identified by the yeast two-hybrid are also tested by biochemical assays.

Table 2.1: Summary of the described yeast two-hybrid methods

Method	Possible baits	Possible preys	Cellular compartment
Classical yeast two-hybrid system	Soluble proteins	Soluble proteins	Nucleus
Sos-recruitment system	Soluble proteins	Membrane localized proteins	Plasma membrane periphery
Ras-recruitment system	Soluble proteins	Membrane localized proteins	Plasma membrane periphery
Reverse Ras-recruitment system	Membrane proteins	Soluble proteins	Plasma membrane periphery
Split-ubiquitin system	Membrane proteins	Soluble or membrane proteins	Cytosolic side of any membrane
G-protein fusion system	Membrane proteins	Soluble or membrane proteins	Plasma membrane periphery
SCINEX-P	Extracellular or TM proteins	Extracellular or TM proteins	Endoplasmic reticulum
Protein complementation assay	Membrane proteins	Soluble or membrane proteins	Cytosolic side of any membrane

Despite the fact that the original yeast two-hybrid system has limitations for membrane proteins, the improved two-hybrid systems listed here made it possible to study many membrane proteins, including extracellular, outer nuclear and other organellar membrane proteins. Moreover, with the help of available cDNA libraries from many tissues of several organisms, yeast two-hybrid systems might be one of the most effective tools to identify membrane protein interactions.

Chapter 3

INCORPORATION OF PHOTO-CROSSLINKING AMINO ACIDS INTO PROTEINS IN A SITE SPECIFIC MANNER

3.1 Introduction

Crosslinking is a way to confirm or identify protein-protein interactions. Especially for studying transient interactions, linking two interacting proteins by a covalent bond is a convenient approach. In chemical crosslinking, when exposed to a common crosslinker, most of the functional groups react quickly enough to permit a 'freezing' in place of interactions, thus fixing the interacting molecules in a complex sufficiently stable for isolation and characterization [90]. Following the crosslinking, the protein of interest is purified with the crosslinked proteins and interacting partners are identified by either mass spectrometry techniques or using known antibodies. Formaldehyde, an amine-reactive chemical crosslinker, due to its small size has been employed numerous times to crosslink proteins through primary amines. However, it can spontaneously form polymers in solution and is not protein-specific, thus it can crosslink other macromolecules such as lipids and DNA. Moreover, the complex chemistry of formaldehyde can result in heterogeneous products making analysis by mass spectrometry difficult [91]. NHS-esters have relatively less complex chemistry and are more specific for proteins. However, they have limitations for crosslinking and the cell membrane is not permeable for these chemicals. [90], [91].

A direct way to study protein-protein interactions in living cells is crosslinking proteins by using unnatural photo-reactive amino acids. In 1986, Kauer *et al*. described the photo-crosslinking amino acid p-benzoyl-L-phenylalanine (pBpa, Figure 3.2), and small peptides containing photo-crosslinking amino acids were obtained by chemical synthesis [92]. Moreover, it has been demonstrated that photo-crosslinking amino acids can be incorporated into proteins by an *in vitro* assay using chemically misacylated tRNAs [93], [94]. However, the yields in chemical misacylation are low and the proteins are generated *in vitro*. In another study, unnatural amino acids have been used for acetylcholine receptors in *Xenopus* oocytes by microinjection of a chemically misacylated tRNA and the relevant mRNA [95]. Although methods using chemically misacylated tRNAs are quite useful, they are limited to the production of small quantities of protein due to their stoichiometric nature.

About a decade ago, the ability to incorporate unnatural amino acids into proteins at defined sites *in vivo* was introduced by the Schultz Lab [96]. Here, we report the technology of *in vivo* incorporation of photo-crosslinking amino acids developed by the same group to study protein-protein interactions. Moreover, we report studies in which this new technology was employed in membrane protein interactions.

3.2 Expanding the Genetic Code with Photo-Crosslinking Amino Acids

Addition of a new amino acid to the genetic repertoire *in vivo* requires new components for the translational machinery [96]. First of all, a photo-crosslinking amino acid is required, which cannot be a substrate of any of the endogenous aminoacyl-tRNA synthetases. The amino acid must be efficiently transported into the cytoplasm when it is added to the growth medium. Secondly, a new tRNA is needed that is not recognized by the existing host aminoacyl-tRNA synthetases, but functions efficiently in translation (i.e. an orthogonal [meaning that it functions independently of the endogenous synthetases and tRNAs] tRNA). This orthogonal tRNA must be able to deliver the novel amino acid in response to a codon that does not encode any of the common 20 amino acids (nonsense or four-base codon). Finally, a new aminoacyl-tRNA synthetase (an orthogonal synthetase) is also required. This synthetase aminoacylates the orthogonal tRNA, but does not recognize any of the endogenous tRNAs of the host organism. Moreover, this synthetase must use only the desired photo-crosslinking amino acid but not any of the other 20 amino acids as substrate (Figure 3.1) [96].

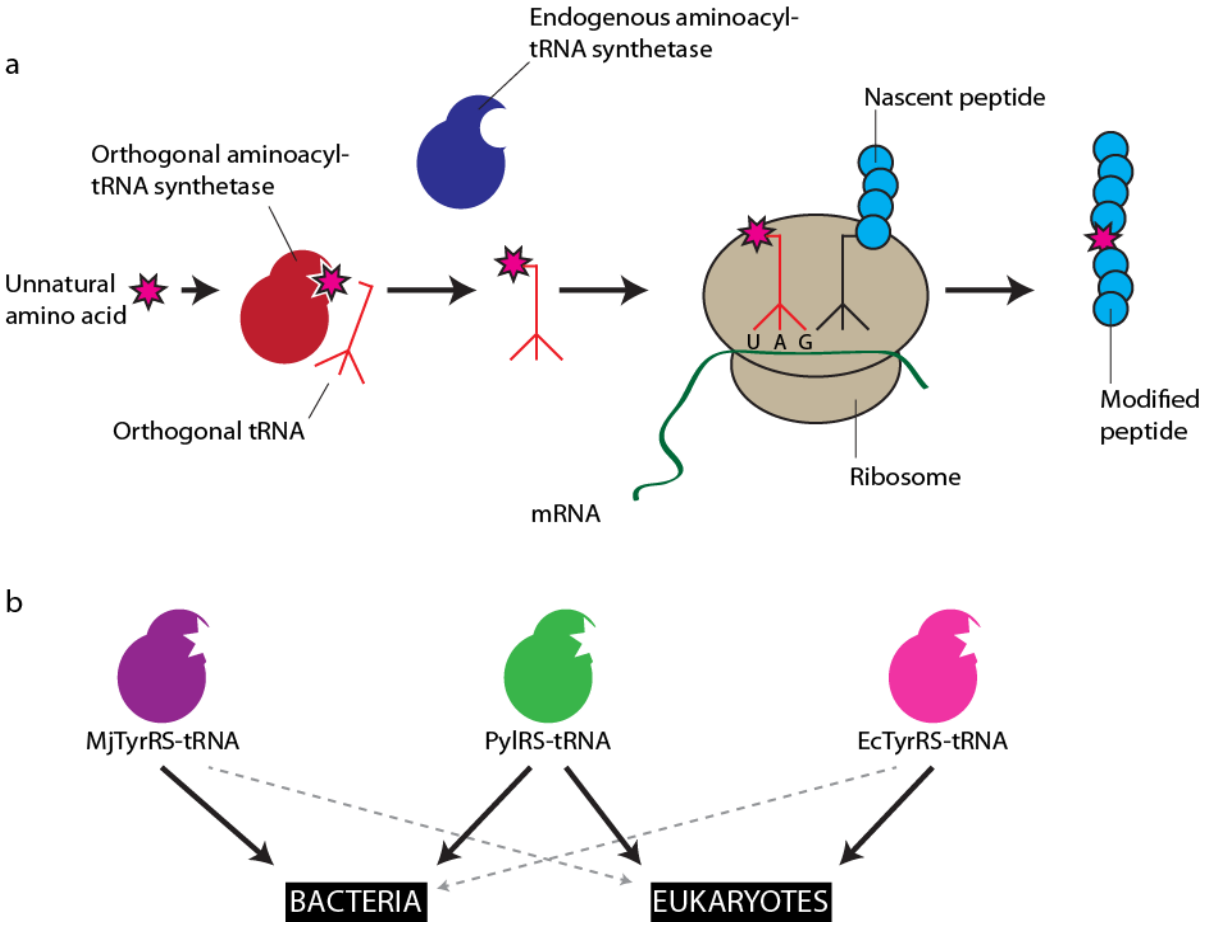

Figure 3.1: a. To introduce an unnatural photo-crosslinking amino acid into a protein, the amino acid is added to the growth medium and taken up by the cell. The photo-crosslinking amino acid is recognized by an orthogonal aminoacyl-tRNA synthetase and attached to an orthogonal amber codon suppressor tRNA, allowing its incorporation into the peptide in response to the amber stop codon UAG. **b.** The pyrrolysyl-tRNA synthetase (PylRS)-tRNA pair from *Methanosarcina barkeri* or *Methanosarcina mazei* is orthogonal for both bacteria and eukaryotes. On the other hand, *M. jannaschii* tyrosyl-tRNA synthetase (MjTyrRS)-tRNA pair is orthogonal in bacteria; and *E. coli* tyrosyl-tRNA synthetase (EcTyrRS)-tRNA is orthogonal

(Figure 3.1 continued) only in eukaryotes. The dashed arrows connect synthetase-tRNA pairs to organisms which they cross-react with their endogenous translational components. (This figure is adapted from Davis and Chin [97])

Initial efforts to encode unnatural amino acids in *E. coli* were aimed at engineering an endogenous aminoacyl-tRNA synthetase / tRNA pair to be orthogonal in *E. coli*. However, misaminoacylation of native tRNAs by engineered aminoacyl-tRNA synthetases resulted in false incorporation of unnatural amino acids. In order to overcome this problem, *in vitro* studies identified orthogonal pairs from other organisms which do not interact with *E. coli* tRNAs and aminoacyl-tRNA synthetases. One candidate pair is the tyrosyl tRNA (Mj-tRNA^{Tyr}) and its synthetase (MjTyrRS) from *Methanococcus jannaschii*, an archaebacterium. Tyrosyl tRNA of *M. jannaschii* has a different identity element from *E.coli* tyrosyl tRNA (First two base pair of the acceptor stem is CG in *M. janaschii* and GC in *E. coli*) [96]. Moreover, MjTyrRS does not have an editing mechanism and therefore, should not proofread if an unnatural amino acid is charged to the tRNA. Since the amber stop codon is the least used among the three stop codons (7%) in *E. coli* and rarely terminates any essential genes [98], Mj-tRNA^{Tyr} was mutated to an amber suppressor tRNA, Mj-tRNA$^{Tyr}_{CUA}$, which can be efficiently aminoacylated by MjTyrRS and is a poor substrate for the *E. coli* synthetases. In order to reduce the interaction of Mj-tRNA$^{Tyr}_{CUA}$ with *E. coli* synthetases, 11 nucleotides of the tRNA that are known not to interact with the MjTyrRS were randomly mutated to generate a suppressor tRNA library. This library was passed through a negative selection which removes tRNAs that are aminoacylated by *E. coli*synthetases and then through a positive selection for tRNAs that can be efficiently aminoacylated by MjTyrRS ([96] and references within)

In order to alter the amino acid specificity of MjTyrRS, in order for it to charge the Mj-tRNA$^{Tyr}_{CUA}$ with an unnatural amino acid, five residues in the active site of MjTyrRS were mutated according to the crystal structure of the homologous tyrosyl-tRNA synthetase from *Bacillus stearothermophilus*. A large library of MjTyrRS mutants consisting of more than 10^8 mutants was developed and subjected to a double-sieve selection. In the positive selection, suppression of an amber stop codon introduced to a nonessential position in the chloramphenicol resistance gene, chloramphenicol acetyl transferase, was aimed for. Cells transformed with the mutant MjTyrRS library and Mj-tRNA$^{Tyr}_{CUA}$gene were selected for their survival in media containing the unnatural amino acid and chloramphenicol. Since suppression efficiency of Mj-tRNA$^{Tyr}_{CUA}$ was already known, cells can survive chloramphenicol treatment if mutant MjTyrRS can charge the orthogonal tRNA with a natural or unnatural amino acid. The surviving cells were then subjected to a negative selection, in which the surviving cells were grown in the presence of chloramphenicol and in the absence of unnatural amino acids. Thus, cells that can introduce only unnatural amino acids are not able to grow in the presence of chloramphenicol. Those cells that did not survive were isolated in a replica plate and identified as orthogonal aminoacyl-tRNAsynthetases ([96] and references within).

Subsequently to the development of the method to incorporate unnatural amino acids into *E. coli* proteins, the first photo-crosslinking amino acid introduced into *E. coli* translational machinery in response to the amber stop codon TAG was pBpa [99].It was demonstrated that mutant aminoacyl-tRNA synthetase MjBpaRS is able to charge mutant tRNA with pBpa and pBpa can be incorporated to amber sites with high-fidelity [99]. Subsequently to the introduction of the first photo-crosslinking

amino acid, other photo-crosslinking amino acids p-azido-L-phenylalanine (pAzpa) [100], benzofuranylalanine[101], 4'-[3-(trifluoromethyl)-3H-diazirin-3-yl]-L-phenylalanine (TfmdPhe) [102] and 3'-azylbutyl-N-carbamoyl-lysine (AbK, Figure 3.2) [103] were also incorporated into the*E. coli* translational machinery.

Table 3.1: List of the most common employed photo-crosslinking amino acids and host organisms in which they are incorporated into proteins

Host organism	Origin of orthogonal aminoacyltRNA synthetase	Photo-crosslinking amino acid
E. coli	*M. jannaschii*	pBpa[99]
E. coli	*M. jannaschii*	pAzpa[100]
E. coli	*E. coli*	Benzofuranylalanine[101]
E. coli	*M. jannaschii*	TfmdPhe[102]
E. coli	*Methanosarcina barkeri / M. mazei*	AbK[103]
S. cerevisiae	*E. coli*	pAzpa[104]
S. cerevisiae	*E. coli*	pBpa[105]
S. cerevisiae	*Methanosarcina barkeri / M. mazei*	AbK[106]
Mammalian	*E. coli*	pBpa[107]
Mammalian	*E. coli*	pAzpa[108]
Mammalian	*E. coli*	TfmdPhe[109]
Mammalian	*Methanosarcina barkeri / M. mazei*	AbK[103]
Mammalian	*Methanosarcina mazei*	TmdZLys[110]

Following these developments in bacteria, manipulating the translational machinery of more complex organisms, such as yeast has occurred. Similar steps as described for *E. coli* are to be used for eukaryotes. However, the translational machinery is not well conserved between prokaryotes and eukaryotes. Thus, components that were employed for *E. coli* (tRNAs and aminoacyl-tRNA synthetases) may not be used for eukaryotes, since they may cross-react with endogenous tRNAs and synthetases. Thus, instead of *M. jannaschii*, translational components (tyrosyl-tRNA synthetase (EcTyr-RS) and tRNA) originating from *E. coli* were employed to incorporate amino acids pAzpa, pBpa and AbK into proteins in *S.cerevisiae*[104–106]. After incorporation of the first unnatural amino acid into mammalian cells using *E.coli* aminoacyl-tRNA synthetase[111], it is followed by photo-crosslinking amino acids, such as pBpa[107] and pAzpa[108], as listed in Table 3.1.

3.3 In vivo Photo-Crosslinking of Membrane Proteins

In a typical experiment, the unnatural photo-crosslinking amino acid is genetically installed at a single site in a protein as described above. Then the protein is irradiated with ultraviolet light. Then, a

binding partner in the close proximity of the photo-crosslinking amino acid becomes covalently trapped, which allows the subsequent identification steps such as by mass spectrometry.

Crosslinking experiments using genetically encoded pBpa allowed identification of many protein-protein interactions among cytoplasmic or membrane proteins [107], [109], [112–116]. Besides the fact that pBpa is one of the earliest available amino acids used in crosslinking studies [92], it is also chemically stable and reacts with C-H bonds upon excitation at 350-360 nm, which is a wavelength that avoids protein and nucleic acid damage [117].

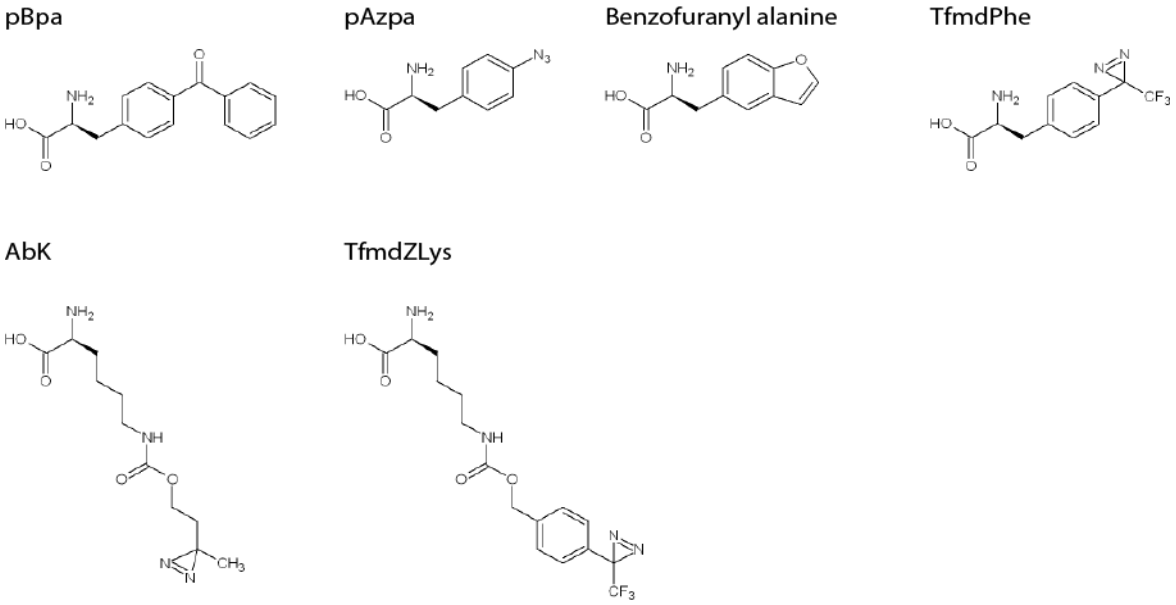

Figure 3.2: Chemical structures of photo-crosslinking amino acids mentioned in this study.

In their work, Mori and Ito studied the interaction of SecY and SecA at a detailed level in *E. coli* using the technique described above. In bacteria SecYEG (SecY-SecE-SecG) is a transmembrane complex that participates in cotranslational integration of membrane proteins and the SecA- and SecB-dependent post-translational translocation of periplasmic and outer membrane proteins. Molecular details of SecA-SecYEG interaction remain poorly understood, including where the binding interface between SecA and SecYEG is. To investigate this interface between SecA and SecY, 53 different amino acid positions of the six cytoplasmic domains of SecY were mutated to pBpa by introducing amber mutations in the secY gene. Subsequent to photo-crosslinking and purification of the SecY protein, immunoblotting against SecA protein has resulted in identification of SecA interacting residues of SecY protein [118].

In another study in yeast, the direct interaction of Tom7, a subunit of the mitochondrial outer membrane TOM40 complex responsible for β-barrel membrane proteins to enter mitochondria, with Tom40 protein and Mdm10 protein was identified. Tom40 protein is another component of the TOM40 complex, which is mainly formed by the beta-barrel pore of the Tom40 protein. Mdm10 occupies a site of the TOB^{core} complex, which is the complex employed by the outer membrane proteins to be

inserted in the outer membrane after they passed through TOM40 complex. To test whether Tom40 and Mdm10 interact with Tom7, the amber codon was introduced at various positions in the Tom7 gene to express pBpa in yeast cells. Subsequent to expression and UV radiation of cells, crosslinked partner proteins of Tom7, which are Tom40 and Mdm10, were identified by immunoblotting with antibodies against these two proteins [116]. Moreover, it was shown in this study that Tom7 interacts with both Tom40 and Mdm10 through the same side of its transmembrane helix.

Although pBpa is the most frequently used photo-crosslinking amino acid, it is not the best one available. The benzophenone group of pBpa is sterically very demanding, and when placed at a protein-protein interface, it may destroy the interaction of interest [103]. Due to its small size and flexible nature, AbK is an alternative to pBpa[103][119]. Moreover, improved efficiency of AbK incorporation is an advantage of this amino acid [119]. Another photo-crosslinking amino acid TfmdPhe has also several advantages over pBpa incorporation in mammalian cells [102][109]. TfmdPhe has double the yield of pBpa incorporation in CHO cells [109]. Moreover, the crosslinked product was detected for TfmdPhe after 1 minute of illumination, and maximum yield was reached after 5 minutes. On the other hand, maximum yield of pBpa was only reached after 10-15 minutes [109]. This indicates that when pBpa is used the possibility of non-specific crosslinking is higher. Additionally, the newly discovered photo-crosslinking amino acid, TmdZLys presents a long-range photo-crosslinking alternative in mammalian cells with its relatively long side chain moiety [110].

3.4 Discussion

Even though the incorporation of photo-crosslinking amino acids is a novel technology it became a preferred method to study protein-protein interactions. Besides the fact that it is *in vivo*, this technology does not require any fusion or truncation process to be applied, thus causing minimum chance of steric hindrance or conformational perturbation. Because of these advantages, this technique became the system of choice to study membrane proteins in bacteria [118], [120], [121] or in the budding yeast [116], [122]. Even though no membrane studies using this novel method was done in mammalian cells, incorporation of photo-crosslinking amino acids into mammalian proteins was achieved using the same principle[107], [108], [123]. Moreover, it was used to find binding partners of a soluble protein in mammalian cells [109].

Another advantage of this method is that it is site-specific. Since photo-crosslinking is not for long-range, only interacting amino acids or residues close to the interaction surface can be crosslinked. Using this advantage, interaction sites of known interacting proteins were also identified [118]. On the other hand, this can also be a drawback. For instance, to identify binding partners of a bait protein, substituting only one amino acid residue with a photo-crosslinking amino acid will result in the proteins which interact with the substituted amino acid residue. Thus, in order to find all the binding partners of a protein, several amino acids in the protein must be substituted one-by-one, since altering many amino acid residues at once may cause conformational changes. Fortunately for

transmembrane proteins, the cytoplasmic portion of the protein, in general, contains the interacting residues of these proteins, which may reduce to amount of work needed.

Some new developments have been reported for this method. For instance, it is reported that a new strain of bacteria was produced which contains the required genes to synthesize the unnatural amino acid of interest from simple carbon sources [124]. The next step could be a strain which can synthesize the required photo-crosslinking amino acids. Moreover, over-expressed suppressor $tRNA_{CUA}$ using polycistronic constructs [125], [126] can increase the efficiency of incorporation of unnatural amino acids. Since suppressor tRNA is in a competition with translational termination factors to response the amber codon, over-expression of tRNA can reduce the chance of termination factors stopping translation.

Despite the fact that the technology of *in vivo* incorporation of photo-crosslinking amino acids might be time-costly due to the required design of several cDNA sequences, it is still an efficient technique to identify binding or interacting partners of membrane proteins.

Chapter 4

SUMMARIZING DISCUSSION

Membrane proteins are essential for many cellular processes including reaction to surrounding environment, multicellular structure, protein folding and energy production. In order to understand molecular mechanisms of membrane proteins, knowledge of membrane protein interactions is as significant as functional and structural information. Various biochemical techniques are available for analysis of protein-protein interactions. However, such techniques are *in vitro* and the outcome may be affected by purification steps and not sensitive to transient and weak interactions. Moreover, studying membrane proteins with many biochemical techniques is difficult, since their recombinant expression may be toxic for the host organism. Subsequent to the solubilization of these proteins with detergents, preserving the functional conformation is challenging. Here I described two methods to study interactions of membrane proteins *in vivo*.

In advanced yeast two-hybrid techniques, the original yeast-two hybrid method, which is to study proteins in the nucleus, was improved to study membrane proteins. Using the budding yeast *S.cerevisiae* to test membrane protein interactions has several advantages, such as the possession of an eukaryotic secretory machinery including post-translational modifications. Moreover, the short life cycle and easy manipulation of the genome are among the advantages of using yeast. In **the reverse Ras-recruitment system**, the inability of the cdc25-2 yeast strain to grow at high temperatures was taken advantage of. In the rRRS setup, membrane-localized bait mediates the Ras protein fused cytosolic prey, thus growth can be regained. The **Split-ubiquitin system** is the most often employed yeast two-hybrid system to study membrane proteins. In this system, two interacting proteins can bring two halves of a ubiquitin moiety together and the reporter protein, which is fused two one of these halves, can be released by ubiquitin-specific proteases. Reporter gene can be a transcription factor and start the transcription of a reporter gene or it can be another protein whose release can be detected by SDS-PAGE. In this system, while bait can be a membrane protein on the cytoplasmic site of any membrane, prey can be either a membrane protein or a soluble protein. The **Protein Complementation Assay** uses the same principle as the split-ubiquitin system that reconstitutes two split halves of a protein by a protein-protein interaction, however; the only difference is that in PCA, the reconstituted protein functions as reporter protein. The **SCINEX-P** system is to detect protein-protein interactions which occur in the ER, and takes advantage of the unfolded protein response mechanism. Two interacting proteins dimerize the unfolded protein sensor protein and initiate the response pathway, in which unfolded protein response elements are attached to reporter genes. This method is generally employed to detect interactions between extracellular proteins. In **G-protein fusion based yeast two-hybrid system**, the prey is fused to G_γ subunit of heterotrimeric G-protein and upon interaction between membrane protein bait and fusion protein prey, G_γ subunit cannot function its role in GPCR downstream pathway. Therefore, growth of yeast stops when interaction between bait and prey occurs.

Up to date, over 40% of known protein-protein interactions were identified by yeast-two hybrid system and PCA [14]. Moreover, a project so-called Membrane-protein Interaction Network Database (MIND) is on its way to identify interactions of over 4000 membrane protein in *Arabidopsis thaliana* using yeast two-hybrid and PCA (http://www.associomics.org/Associomics/Home.html). Even though it is readily and often employed, results obtained from yeast two-hybrid system experiments might be biased since these experiments are conducted under non-physiological expression conditions. That is why information gained from yeast two-hybrid systems is verified with a different assay, such as affinity purification, using endogenous expression levels of proteins.

Besides the two-hybrid methods, introduction of unnatural photo-crosslinking amino acids into proteins by manipulating the translational machinery is a novel method to study interactions of membrane proteins. These amino acids can be incorporated into proteins in a site specific manner by manipulating the translational machinery in bacteria and eukaryotes. In order to achieve this, a tRNA which can suppress a codon in the endogenous translational system, an aminoacyl-tRNA synthetase which can aminoacylate exclusively the suppressor tRNA with exclusively the photo-crosslinking amino acid is required. Using this system, bacterial and eukaryotic membrane protein interaction studies have been performed. A major advantage of using this system is there is no truncation or fusion of the protein of interest. However, to identify an interaction using this method, several constructs of protein of interest must be prepared, since the physical range of crosslinking is limited.

Here we listed the best two available *in vivo* technologies to study interactions of membrane proteins with other proteins. With the help of available cDNA libraries, the yeast two-hybrid system is the most often employed method to identify protein-protein interactions on membranes since it is cost-effective, time-efficient and readily applied. On the other hand, site specific incorporation of photo-crosslinking amino acids may not be among the most often used techniques, since it is a novel approach. However, it can still be considered among the best available *in vivo* methods, since it is site specific and cross-linking occurs at physiological expression levels without any truncation or fusion process. In comparison to soluble proteins, membrane proteins are more challenging to study *in vitro*. It must be always considered that no technology can identify all interactions, and all the technologies have a certain fraction of false positives and negatives. The systems listed here are the most readily applied, high throughput or specific methods currently available.

REFERENCES:

[1] E. S. Lander, “Initial impact of the sequencing of the human genome.,” *Nature*, vol. 470, no. 7333, pp. 187-97, Feb. 2011.

[2] M. Clamp et al., “Distinguishing protein-coding and noncoding genes in the human genome ,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104 , no. 49 , pp. 19428-19433, Dec. 2007.

[3] M. S. Almén, K. J. V. Nordström, R. Fredriksson, and H. B. Schiöth, “Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin.,” *BMC biology*, vol. 7, p. 50, Jan. 2009.

[4] E. Wallin and G. von Heijne, “Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms.,” *Science*, vol. 7, no. 4, pp. 1029-38, Apr. 1998.

[5] T. Reiss, “Drug discovery of the future: the implications of the human genome project,” *Trends in biotechnology*, vol. 19, no. 12. Elsevier Science Publishers, pp. 496-499, 01-Dec-2001.

[6] A. Brückner, C. Polge, N. Lentze, D. Auerbach, and U. Schlattner, “Yeast two-hybrid, a powerful tool for systems biology.,” *International journal of molecular sciences*, vol. 10, no. 6, pp. 2763-88, Jun. 2009.

[7] W. B. Frommer and O. Ninnemann, “Heterologous Expression of Genes in Bacterial, Fungal, Animal, and Plant Cells,” *Annual Review of Plant Physiology and Plant Molecular Biology*, vol. 46, no. 1, pp. 419-444, Jun. 1995.

[8] G. Karp, *Cell and Molecular Biology: Concepts and Experiments*. 2009, pp. 127-128.

[9] L. Keegan, G. Gill, and M. Ptashne, “Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein ,” *Science*, vol. 231 , no. 4739 , pp. 699-704, Feb. 1986.

[10] S. Fields and O.-kyu Song, “A novel genetic system to detect protein-protein interactions,” *Nature*, vol. 340, no. 6230, pp. 245-246, Jul. 1989.

[11] C. T. Chien, P. Bartel, R. Sternglanz, and S. Fields, “The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 21, pp. 9578-82, Nov. 1991.

[12] F. Tirode, C. Malaguti, F. Romero, R. Attar, J. H. Camonis, and J. M. Egly, “A Conditionally Expressed Third Partner Stabilizes or Prevents the Formation of a Transcriptional Activator in a Three-hybrid System,” *Journal of Biological Chemistry*, vol. 272, no. 37, pp. 22995-22999, Sep. 1997.

[13] D. J. SenGupta, B. Zhang, B. Kraemer, P. Pochart, S. Fields, and M. Wickens, “A three-hybrid system to detect RNA-protein interactions in vivo ,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93 , no. 16 , pp. 8496-8501, Aug. 1996.

[14] S. Kerrien et al., “The IntAct molecular interaction database in 2012 ,” *Nucleic Acids Research*, vol. 40 , no. D1 , p. D841-D846, Jan. 2012.

[15] Y. Wang and H. G. Dohlman, “Regulation of G protein and mitogen-activated protein kinase signaling by ubiquitination: insights from model organisms.,” *Circulation research*, vol. 99, no. 12, pp. 1305-14, Dec. 2006.

[16] O. Vögler, J. M. Barceló, C. Ribas, and P. V. Escribá, “Membrane interactions of G proteins and other related proteins,” *Biochimica et Biophysica Acta (BBA) - Membranes*, vol. 1778, pp. 1640-1652, 2008.

[17] R. Dechant and M. Peter, “Nutrient signals driving cell growth,” *Current opinion in cell biology*, vol. 20, pp. 678-687, 2008.

[18] S. Zaman, S. I. Lippman, X. Zhao, and J. R. Broach, “How Saccharomyces Responds to Nutrients,” *Annual Review of Genetics*, vol. 42, pp. 27-81, 2008.

[19] R. Tisi, F. Belotti, C. Paiardi, F. Brunetti, and E. Martegani, “The budding yeast RasGEF Cdc25 reveals an unexpected nuclear localization.,” *Biochimica et biophysica acta*, vol. 1783, no. 12, pp. 2363-74, Dec. 2008.

[20] A. Aronheim, D. Engelberg, N. Li, N. Al-Alawi, J. Schlessinger, and M. Karin, “Membrane targeting of the nucleotide exchange factor Sos is sufficient for activating the Ras signaling pathway.,” *Cell*, vol. 78, no. 6, pp. 949-61, Sep. 1994.

[21] A. Aronheim, E. Zandi, and H. Hennemann, “Isolation of an AP-1 repressor by a novel method for detecting protein-protein interactions . Isolation of an AP-1 Repressor by a Novel Method for Detecting Protein-Protein Interactions,” *Molecular Cell Biology*, vol. 17, no. 6, pp. 3094-3102, 1997.

[22] Y. C. Broder, S. Katz, and A. Aronheim, “The ras recruitment system, a novel approach to the study of protein-protein interactions.,” *Current biology□: CB*, vol. 8, no. 20, pp. 1121-4, Oct. 1998.

[23] M. Hubsman, G. Yudkovsky, and A. Aronheim, “A novel approach for the identification of protein-protein interaction with integral membrane proteins.,” *Nucleic Acids Research*, vol. 29, no. 4, p. E18, Feb. 2001.

[24] P. Frankel et al., “RalA interacts with ZONAB in a cell density-dependent manner and regulates its transcriptional activity.,” *The EMBO journal*, vol. 24, no. 1, pp. 54-62, Jan. 2005.

[25] P. Frankel et al., “Ral and Rho-dependent activation of phospholipase D in v-Raf-transformed cells.,” *Biochemical and biophysical research communications*, vol. 255, no. 2, pp. 502-7, Feb. 1999.

[26] D. O. Henry et al., “Ral GTPases contribute to regulation of cyclin D1 through activation of NF-kappaB.,” *Molecular and cellular biology*, vol. 20, no. 21, pp. 8084-92, Nov. 2000.

[27] N. D. de Ruiter, R. M. Wolthuis, H. van Dam, B. M. Burgering, and J. L. Bos, “Ras-dependent regulation of c-Jun phosphorylation is mediated by the Ral guanine nucleotide exchange factor-Ral pathway.,” *Molecular and cellular biology*, vol. 20, no. 22, pp. 8480-8, Nov. 2000.

[28] S. Moskalenko, D. O. Henry, C. Rosse, G. Mirey, J. H. Camonis, and M. a White, “The exocyst is a Ral effector complex.,” *Nature cell biology*, vol. 4, no. 1, pp. 66-72, Jan. 2002.

[29] W. Volknandt, J. Pevsner, L. a Elferink, and R. H. Scheller, “Association of three small GTP-binding proteins with cholinergic synaptic vesicles.,” *FEBS Letters*, vol. 317, no. 1–2, pp. 53-6, Feb. 1993.

[30] B. T. Kinsella, R. a Erdman, and W. a Maltese, “Carboxyl-terminal isoprenylation of ras-related GTP-binding proteins encoded by rac1, rac2, and ralA.,” *The Journal of biological chemistry*, vol. 266, no. 15, pp. 9786-94, May 1991.

[31] S. C. Falsetti et al., “Geranylgeranyltransferase I inhibitors target RalB to inhibit anchorage-dependent growth and induce apoptosis and RalA to inhibit anchorage-independent growth.,” *Molecular and cellular biology*, vol. 27, no. 22, pp. 8003-14, Nov. 2007.

[32] T. Kokkola et al., “Somatostatin receptor 5 is palmitoylated by the interacting ZDHHC5 palmitoyltransferase.,” *FEBS Letters*, vol. 585, no. 17, pp. 2665-70, Sep. 2011.

[33] G. Cao et al., “Methionine sulfoxide reductase B1 (MsrB1) recovers TRPM6 channel activity during oxidative stress.,” *The Journal of biological chemistry*, vol. 285, no. 34, pp. 26081-7, Aug. 2010.

[34] N. Johnsson and A. Varshavsky, “Split ubiquitin as a sensor of protein interactions in vivo.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, no. 22, pp. 10340-4, Oct. 1994.

[35] A. Hershko and A. Ciechanover, “The Ubiquitin System for Protein Degradation,” *Annual Review of Biochemistry*, vol. 61, no. 1, pp. 761-807, 1992.

[36] I. Stagljar, C. Korostensky, N. Johnsson, and S. te Heesen, “A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 9, pp. 5187-92, Apr. 1998.

[37] S. Wittke, N. Lewke, S. Müller, and N. Johnsson, “Probing the molecular environment of membrane proteins in vivo.,” *Molecular biology of the cell*, vol. 10, no. 8, pp. 2519-30, Aug. 1999.

[38] A. Varshavsky, “The N-end rule pathway of protein degradation.,” *Genes to cells□: devoted to molecular & cellular mechanisms*, vol. 2, no. 1, pp. 13-28, Jan. 1997.

[39] B. S. Hooker, D. J. Bigelow, and C.-T. Lin, “Methods for mapping of interaction networks involving membrane proteins,” *Biochemical and Biophysical Research Communications*, vol. 363, no. 3, pp. 457-461, Nov. 2007.

[40] P. Obrdlik et al., “K+ channel interactions detected by a genetic system optimized for systematic studies of membrane,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 33, pp. 12242-12247, 2004.

[41] J. P. Miller et al., “Large-scale identification of yeast integral membrane protein interactions.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 34, pp. 12123-8, Aug. 2005.

[42] Y. L. Deribe et al., “Regulation of epidermal growth factor receptor trafficking by lysine deacetylase HDAC6.,” *Science*, vol. 2, no. 102, p. ra84, Jan. 2009.

[43] Y. Yarden and M. X. Sliwkowski, “Untangling the ErbB signalling network,” *Nature reviews. Molecular cell biology*, vol. 2, no. 2, pp. 127-137, Feb. 2001.

[44] D. Hoeller, S. Volarevic, and I. Dikic, “Compartmentalization of growth factor receptor signalling,” *Current Opinion in Cell Biology*, vol. 17, no. 2, pp. 107-111, Apr. 2005.

[45] J. Jin et al., “Interaction of the mu-opioid receptor with GPR177 (Wntless) inhibits Wnt secretion: potential implications for opioid dependence.,” *BMC neuroscience*, vol. 11, p. 33, Jan. 2010.

[46] A. J. Eisch, M. Barrot, C. A. Schad, D. W. Self, and E. J. Nestler, “Opiates inhibit neurogenesis in the adult rat hippocampus ,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97 , no. 13 , pp. 7579-7584, Jun. 2000.

[47] Y. Li, H. Wang, L. Niu, and Y. Zhou, “Chronic morphine exposure alters the dendritic morphology of pyramidal neurons in visual cortex of rats,” *Neuroscience Letters*, vol. 418, no. 3, pp. 227-231, May 2007.

[48] K. F. Hauser, P. J. McLaughlin, and I. S. Zagon, “Endogenous opioids regulate dendritic growth and spine formation in developing rat brain,” *Brain Research*, vol. 416, no. 1, pp. 157-161, Jul. 1987.

[49] R. M. Goodman et al., “Sprinter: a novel transmembrane protein required for Wg secretion and signaling ,” *Development* , vol. 133 , no. 24 , pp. 4901-4911, Dec. 2006.

[50] C. Banziger, D. Soldini, C. Schutt, P. Zipperlen, G. Hausmann, and K. Basler, “Wntless, a Conserved Membrane Protein Dedicated to the Secretion of Wnt Proteins from Signaling Cells,” *Cell*, vol. 125, no. 3. Cell Press, pp. 509-522, 05-May-2006.

[51] K. Bartscherer, N. Pelte, D. Ingelfinger, and M. Boutros, “Secretion of Wnt Ligands Requires Evi, a Conserved Transmembrane Protein,” *Cell*, vol. 125, no. 3. Cell Press, pp. 523-533, 05-May-2006.

[52] C. M. Paumi et al., “Mapping protein-protein interactions for the yeast ABC transporter Ycf1p by integrated split-ubiquitin membrane yeast two-hybrid analysis.,” *Molecular cell*, vol. 26, no. 1, pp. 15-25, Apr. 2007.

[53] M. Dean, “The Genetics of ATP Binding Cassette Transporters,” in *Phase II Conjugation Enzymes and Transport Systems*, vol. Volume 400, and L. P. B. T.-M. in E. Helmut Sies, Ed. Academic Press, 2005, pp. 409-429.

[54] D. Taglicht and S. Michaelis, “Saccharomyces cerevisiae ABC proteins and their relevance to human health and disease,” in *ABC Transporters: Biochemical, Cellular, and Molecular Aspects*, vol. Volume 292, M. M. G. B. T.-M. in E. Suresh V. Ambudkar, Ed. Academic Press, 1998, pp. 130-162.

[55] J. Sheps, S. Ralph, Z. Zhao, D. Baillie, and V. Ling, “The ABC transporter gene family of Caenorhabditis elegans has implications for the evolutionary dynamics of multidrug resistance in eukaryotes,” *Genome Biology*, vol. 5, no. 3, p. R15, 2004.

[56] A. Reinders, W. Schulze, S. Thaminy, I. Stagljar, W. B. Frommer, and J. M. Ward, “Intra- and intermolecular interactions in sucrose transporters at the plasma membrane detected by the split-ubiquitin system and functional assays.,” *Structure (London, England□: 1993)*, vol. 10, no. 6, pp. 763-72, Jun. 2002.

[57] S. M. Gisler et al., “Monitoring Protein-Protein Interactions between the Mammalian Integral Membrane Transporters and PDZ-interacting Partners Using a Modified Split-ubiquitin Membrane Yeast Two-hybrid System,” *Molecular & Cellular Proteomics*, vol. 7, no. 7, pp. 1362-1377, Mar. 2008.

[58] M. J. Massaad and A. Herscovics, “Interaction of the endoplasmic reticulum alpha 1,2-mannosidase Mns1p with Rer1p using the split-ubiquitin system.,” *Journal of cell science*, vol. 114, no. Pt 24, pp. 4629-35, Dec. 2001.

[59] M. D. McGee, I. Stagljar, and D. a Starr, “KDP-1 is a nuclear envelope KASH protein required for cell-cycle progression.,” *Journal of cell science*, vol. 122, no. Pt 16, pp. 2895-905, Aug. 2009.

[60] K. N. Ehrhard, J. J. Jacoby, X. Y. Fu, R. Jahn, and H. G. Dohlman, “Use of G-protein fusions to monitor integral membrane protein-protein interactions in yeast.,” *Nature biotechnology*, vol. 18, no. 10, pp. 1075-9, Oct. 2000.

[61] K. M. S. Misura, R. H. Scheller, and W. I. Weis, “Three-dimensional structure of the neuronal-Sec1-syntaxin 1a complex ,” *Nature*, vol. 404, no. 6776, pp. 355-362, Mar. 2000.

[62] D. M. Urech, P. Lichtlen, and A. Barberis, “Cell growth selection system to detect extracellular and transmembrane protein interactions,” *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1622, no. 2, pp. 117-127, Jul. 2003.

[63] L. Ellgaard, M. Molinari, and A. Helenius, “Setting the Standards: Quality Control in the Secretory Pathway ,” *Science*, vol. 286 , no. 5446 , pp. 1882-1888, Dec. 1999.

[64] P. M. Rudd et al., “Roles for glycosylation of cell surface receptors involved in cellular immune recognition,” *Journal of Molecular Biology*, vol. 293, no. 2, pp. 351-366, Oct. 1999.

[65] J. S. Cox, C. E. Shamu, and P. Walter, “Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase,” *Cell*, vol. 73, no. 6. Cell Press, pp. 1197-1206, 18-Jun-1993.

[66] C. Y. Liu, H. N. Wong, J. A. Schauerte, and R. J. Kaufman, “The Protein Kinase/Endoribonuclease IRE1α That Signals the Unfolded Protein Response Has a Luminal N-terminal Ligand-independent Dimerization Domain ,” *Journal of Biological Chemistry*, vol. 277 , no. 21 , pp. 18346-18356, May 2002.

[67] C. Y. Liu, M. Schröder, and R. J. Kaufman, “Ligand-independent Dimerization Activates the Stress Response Kinases IRE1 and PERK in the Lumen of the Endoplasmic Reticulum ,” *Journal of Biological Chemistry*, vol. 275 , no. 32 , pp. 24881-24885, Aug. 2000.

[68] Y. Kozutsumi, M. Segal, K. Normington, M.-J. Gething, and J. Sambrook, “The presence of malfolded proteins in the endoplasmic reticulum signals the induction of glucose-regulated proteins,” *Nature*, vol. 332, no. 6163, pp. 462-464, Mar. 1988.

[69] S. Pollock et al., “Specific interaction of ERp57 and calnexin determined by NMR spectroscopy and an ER two-hybrid system.,” *The EMBO journal*, vol. 23, no. 5, pp. 1020-9, Mar. 2004.

[70] A. A. der Maur et al., “Direct in Vivo Screening of Intrabody Libraries Constructed on a Highly Stable Single-chain Framework ,” *Journal of Biological Chemistry*, vol. 277 , no. 47 , pp. 45075-45085, Nov. 2002.

[71] A. Ullmann, F. Jacob, and J. Monod, “Characterization by in vitro complementation of a peptide corresponding to an operator-proximal segment of the β-galactosidase structural gene of Escherichia coli,” *Journal of Molecular Biology*, vol. 24, no. 2, pp. 339-343, Mar. 1967.

[72] F. M. RICHARDS, “ON THE ENZYMIC ACTIVITY OF SUBTILISIN-MODIFIED RIBONCLEASE,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 44, no. 2, pp. 162-166, 1958.

[73] F. Rossi, C. a Charlton, and H. M. Blau, “Monitoring protein-protein interactions in intact eukaryotic cells by beta-galactosidase complementation.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 16, pp. 8405-10, Aug. 1997.

[74] I. Ghosh and A. Hamilton, “Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein,” *Journal of the American Chemical Society*, vol. 122, no. 11, pp. 5658-5659, 2000.

[75] C.-D. Hu, Y. Chinenov, and T. K. Kerppola, “Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation.,” *Molecular cell*, vol. 9, no. 4, pp. 789-98, Apr. 2002.

[76] R. Paulmurugan and S. Gambhir, “Monitoring protein-protein interactions using split synthetic renilla luciferase protein-fragment-assisted complementation,” *Analytical chemistry*, vol. 75, no. 7, pp. 15608-15613, 2003.

[77] K. E. Luker, M. C. P. Smith, G. D. Luker, S. T. Gammon, H. Piwnica-Worms, and D. Piwnica-Worms, “Kinetics of regulated protein-protein interactions revealed with firefly luciferase complementation imaging in cells and living animals.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 33, pp. 12288-93, Aug. 2004.

[78] I. Remy and S. W. Michnick, “A highly sensitive protein-protein interaction assay based on Gaussia luciferase,” *Nature methods*, vol. 3, no. 12, pp. 977-979, 2006.

[79] P. Tafelmeyer, N. Johnsson, and K. Johnsson, “Transforming a (beta/alfa) 8 -Barrel Enzyme into a Split-Protein Sensor through Directed Evolution,” *Chemistry & Biology*, vol. 11, pp. 681-689, 2004.

[80] J. N. Pelletier, K. M. Arndt, A. Plückthun, and S. W. Michnick, “An in vivo library-versus-library selection of optimized protein-protein interactions.,” *Nature biotechnology*, vol. 17, no. 7, pp. 683-90, Jul. 1999.

[81] I. Remy and S. W. Michnick, “Clonal selection and in vivo quantitation of protein interactions with protein-fragment complementation assays.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 10, pp. 5394-9, May 1999.

[82] T. Wehrman, B. Kleaveland, J.-H. Her, R. F. Balint, and H. M. Blau, “Protein-protein interactions monitored in mammalian cells via complementation of beta -lactamase enzyme fragments.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 6, pp. 3469-74, Mar. 2002.

[83] A. Galarneau, M. Primeau, L.-E. Trudeau, and S. W. Michnick, “Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions.,” *Nature biotechnology*, vol. 20, no. 6, pp. 619-22, Jun. 2002.

[84] P. H. Ear and S. W. Michnick, “A general life-death selection strategy for dissecting protein functions.,” *Nature methods*, vol. 6, no. 11, pp. 813-6, Nov. 2009.

[85] M. Wehr, R. Laage, U. Bolz, and T. Fischer, “Monitoring regulated protein-protein interactions using split TEV,” *Nature methods*, vol. 3, no. 12, 2006.

[86] T. F. Massoud, R. Paulmurugan, and S. Gambhir, “A molecularly engineered split reporter for imaging protein-protein interactions with positron emission tomography.,” *Nature medicine*, vol. 16, no. 8, pp. 921-6, Aug. 2010.

[87] M. M. Müller, H. Kries, E. Csuhai, P. Kast, and D. Hilvert, “Design, selection, and characterization of a split chorismate mutase.,” *Protein science*, vol. 19, no. 5, pp. 1000-10, May 2010.

[88] K. Tarassov et al., “An in vivo map of the yeast protein interactome.,” *Science*, vol. 320, no. 5882, pp. 1465-70, Jun. 2008.

[89] B. Nyfeler, S. W. Michnick, and H.-P. Hauri, “Capturing protein interactions in the secretory pathway of living cells ,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102 , no. 18 , pp. 6350-6355, May 2005.

[90] O. Crociani, “Biochemical methods to study the interactions between integrins and ion channels.,” *Advances in experimental medicine and biology*, vol. 674, pp. 23-32, Jan. 2010.

[91] A. Sinz, “Investigation of protein-protein interactions in living cells by chemical crosslinking and mass spectrometry.,” *Analytical and bioanalytical chemistry*, vol. 397, no. 8, pp. 3433-40, Aug. 2010.

[92] J. Kauer, S. Erickson-Viitanen, H. J. Wolfe, and W. DeGrado, “p-Benzoyl-L-phenylalanine, a new photoreactive amino acid. Photolabeling of calmodulin with a synthetic calmodulin-binding peptide,” *Journal of Biological Chemistry*, vol. 261, no. 15, pp. 10695-10700, 1986.

[93] V. W. Cornish, D. R. Benson, C. A. Altenbach, K. Hideg, W. L. Hubbell, and P. G. Schultz, “Site-specific incorporation of biophysical probes into proteins ,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91 , no. 8 , pp. 2910-2914, Apr. 1994.

[94] T. Kanamori, S.-I. Nishikawa, I. Shin, P. G. Schultz, and T. Endo, “Probing the environment along the protein import pathways in yeast mitochondria by site-specific photocrosslinking ,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94 , no. 2 , pp. 485-490, Jan. 1997.

[95] M. W. Nowak, J. P. Gallivan, S. K. Silverman, C. G. Labarca, D. A. Dougherty, and H. A. Lester, “[28] In vivo incorporation of unnatural amino acids into ion channels in Xenopus oocyte expression system,” in *Ion Channels Part B*, vol. Volume 293, P. M. C. B. T.-M. in Enzymology, Ed. Academic Press, 1998, pp. 504-529.

[96] L. Wang, A. Brock, B. Herberich, and P. G. Schultz, “Expanding the genetic code of Escherichia coli.,” *Science*, vol. 292, no. 5516, pp. 498-500, Apr. 2001.

[97] L. Davis and J. W. Chin, “Designer proteins: applications of genetic code expansion in cell biology.,” *Nature reviews. Molecular cell biology*, vol. 13, no. 3, pp. 168-82, Jan. 2012.

[98] J. Xie and P. G. Schultz, “Adding amino acids to the genetic repertoire.,” *Current opinion in chemical biology*, vol. 9, no. 6, pp. 548-54, Dec. 2005.

[99] J. Chin, A. Martin, and D. King, “Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 17, pp. 11020-7, 2002.

[100] J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang, and P. G. Schultz, “Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli.,” *Journal of the American Chemical Society*, vol. 124, no. 31, pp. 9026-7, Aug. 2002.

[101] T. Bentin, R. Hamzavi, J. Salomonsson, H. Roy, M. Ibba, and P. E. Nielsen, “Photoreactive bicyclic amino acids as substrates for mutant Escherichia coli phenylalanyl-tRNA synthetases.,” *The Journal of biological chemistry*, vol. 279, no. 19, pp. 19839-45, May 2004.

[102] E. M. Tippmann, W. Liu, D. Summerer, A. V. Mack, and P. G. Schultz, “A genetically encoded diazirine photocrosslinker in Escherichia coli.,” *Chembiochem□: a European journal of chemical biology*, vol. 8, no. 18, pp. 2210-4, Dec. 2007.

[103] C. Chou, R. Uprety, L. Davis, J. W. Chin, and A. Deiters, “Genetically encoding an aliphatic diazirine for protein photocrosslinking,” *Chemical Science*, vol. 2, no. 3, p. 480, 2011.

[104] A. Deiters, T. A. Cropp, M. Mukherji, J. W. Chin, J. C. Anderson, and P. G. Schultz, “Adding amino acids with novel reactivity to the genetic code of Saccharomyces cerevisiae.,” *Journal of the American Chemical Society*, vol. 125, no. 39, pp. 11782-3, Oct. 2003.

[105] J. W. Chin, T. A. Cropp, J. C. Anderson, M. Mukherji, Z. Zhang, and P. G. Schultz, “An expanded eukaryotic genetic code.,” *Science*, vol. 301, no. 5635, pp. 964-7, Aug. 2003.

[106] S. M. Hancock, R. Uprety, A. Deiters, and J. W. Chin, “Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair.,” *Journal of the American Chemical Society*, vol. 132, no. 42, pp. 14819-24, Oct. 2010.

[107] N. Hino, Y. Okazaki, T. Kobayashi, and A. Hayashi, “Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid,” *Nature methods*, vol. 2, no. 3, pp. 3-8, 2005.

[108] W. Liu, A. Brock, S. Chen, and P. G. Schultz, “Genetic incorporation of unnatural amino acids into proteins in mammalian cells,” *Nature methods*, vol. 4, no. 3, pp. 239-244, 2007.

[109] N. Hino et al., “Genetic incorporation of a photo-crosslinkable amino acid reveals novel protein complexes with GRB2 in mammalian cells.,” *Journal of molecular biology*, vol. 406, no. 2, pp. 343-53, Feb. 2011.

[110] T. Yanagisawa, N. Hino, F. Iraha, T. Mukai, K. Sakamoto, and S. Yokoyama, “Wide-range protein photo-crosslinking achieved by a genetically encoded N(ε)-(benzyloxycarbonyl)lysine derivative with a diazirinyl moiety.,” *Molecular bioSystems*, vol. 8, no. 4, pp. 1131-5, Apr. 2012.

[111] K. Sakamoto et al., “Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells ,” *Nucleic Acids Research* , vol. 30 , no. 21 , pp. 4692-4699, Nov. 2002.

[112] S. K. Lakshmipathy et al., “Identification of nascent chain interaction sites on trigger factor.,” *The Journal of biological chemistry*, vol. 282, no. 16, pp. 12186-93, Apr. 2007.

[113] D. Boos, C. Kuffer, R. Lenobel, R. Körner, and O. Stemmann, “Phosphorylation-dependent binding of cyclin B1 to a Cdc6-like domain of human separase.,” *The Journal of biological chemistry*, vol. 283, no. 2, pp. 816-23, Jan. 2008.

[114] N. Mohibullah, “Site-specific cross-linking of TBP in vivo and in vitro reveals a direct functional interaction with the SAGA subunit Spt3,” *Genes & development*, pp. 2994-3006, 2008.

[115] S. Tagami et al., “Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein.,” *Nature*, vol. 468, no. 7326, pp. 978-82, Dec. 2010.

[116] K. Yamano, S. Tanaka-Yamano, and T. Endo, “Tom7 regulates Mdm10-mediated assembly of the mitochondrial import channel protein Tom40.,” *The Journal of biological chemistry*, vol. 285, no. 53, pp. 41222-31, Dec. 2010.

[117] J. W. Chin and P. G. Schultz, “In vivo photocrosslinking with unnatural amino Acid mutagenesis.,” *Chembiochem□: a European journal of chemical biology*, vol. 3, no. 11, pp. 1135-7, Nov. 2002.

[118] H. Mori and K. Ito, “Different modes of SecY-SecA interactions revealed by site-directed in vivo photo-cross-linking.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 44, pp. 16159-64, Oct. 2006.

[119] H.-wang Ai, W. Shen, A. Sagi, P. R. Chen, and P. G. Schultz, “Probing protein-protein interactions with a genetically encoded photo-crosslinking amino acid.,” *Chembiochem□: a European journal of chemical biology*, vol. 12, no. 12, pp. 1854-7, Aug. 2011.

[120] S. Okuda and H. Tokuda, “Model of mouth-to-mouth transfer of bacterial lipoproteins through inner membrane LolC, periplasmic LolA, and outer membrane LolB.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 14, pp. 5877-82, Apr. 2009.

[121] E. Freinkman, S.-sin Chng, and D. Kahne, “The complex that inserts lipopolysaccharide into the bacterial outer membrane forms a two-protein plug-and-barrel,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 6, pp. 2486-2491, 2010.

[122] P. Carvalho and A. Stanley, “Retrotranslocation of a misfolded luminal ER protein by the ubiquitin-ligase Hrd1p,” *Cell*, vol. 143, no. 4, pp. 579-591, 2010.

[123] T. Mukai, T. Kobayashi, N. Hino, T. Yanagisawa, K. Sakamoto, and S. Yokoyama, “Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases.,” *Biochemical and biophysical research communications*, vol. 371, no. 4, pp. 818-22, Jul. 2008.

[124] R. a Mehl et al., “Generation of a bacterium with a 21 amino acid genetic code.,” *Journal of the American Chemical Society*, vol. 125, no. 4, pp. 935-9, Jan. 2003.

[125] Y. Ryu and P. G. Schultz, “Efficient incorporation of unnatural amino acids into proteins in Escherichia coli,” *Nature Methods*, vol. 3, no. 4, pp. 263-265, Apr. 2006.

[126] S. Chen, P. G. Schultz, and A. Brock, “An Improved System for the Generation and Analysis of Mutant Proteins Containing Unnatural Amino Acids in Saccharomyces cerevisiae,” *Journal of Molecular Biology*, vol. 371, no. 1, pp. 112-122, Aug. 2007.