VISUALIZATION OF PROTEIN-PROTEIN INTERACTIONS IN THE SECRETORY PATHWAY OF MAMMALIAN CELLS

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von
Beat Nyfeler
aus Gondiswil (BE)

Basel, 2006

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Hans-Peter Hauri und Prof. Dr. Martin Spiess

Basel, den 24. Oktober 2006

Prof. Dr. Hans-Peter Hauri
Dekan

Acknowledgment

I would like to express my gratitude and thanks to...

Hans-Peter Hauri for giving me the opportunity to perform my PhD thesis in his lab and for being a great mentor and motivator always having an open door to answer and discuss my questions.

Stephen Michnick for scientific advice and for giving me the opportunity to spend 6 months in his lab.

Käthy Bucher for her excellent technical assistance, for keeping the Hauri lab going and for her honest opinion and moral support including the uncountable sponsored coffees.

Oliver Nufer for introducing me into the secrets of molecular biology and various biochemical techniques.

Markus Wendeler for being a smart discussion partner, a humorous lab companion and for critically reading my PhD thesis.

Christian Appenzeller, Houchaima Ben-Tekaya, Carine Bonnon, Lionel Breuza, Regula Halbeisen, Sandra Jaggi, Eva Kögler, Sandra Mitrovic, Cécile Vedrenne and Lorenz Waldmeier for sharing thoughts, reagents and for the good atmosphere in the Hauri lab.

Members of the Michnick lab and especially Edi Stefan for scientific support and the cheerful atmosphere in the lab.

Verena Jäggin for expert assistance in FACS sorting and analysis.

Markus Meier, Markus Hämmerle, Angèle Klein, Marianne Liechti and Jny Wittker for keeping the 7th floor of the Biozentrum running.

My family, all my friends and especially Barbara for their moral support and motivation, for their interest in my work and for always believing in me.

Summary

An increasing number of human disorders are being linked to mutations in components of the secretory pathway. One example is combined blood coagulation factor V and VIII deficiency, an autosomal recessive disorder leading to hemophilia due to markedly reduced levels of coagulation factors V and VIII in blood plasma. This disorder is genetically linked to the transmembrane protein ERGIC-53 and its soluble luminal interaction partner MCFD2, both of which reside in the early secretory pathway. ERGIC-53 and MCFD2 constitute a transport receptor complex required for the efficient secretion of blood coagulation factors V and VIII. The strict dependence of some secretory proteins on receptor-mediated transport illustrates the need to understand this process in detail.

The characterization of transport receptors and their cognate cargo proteins is hampered by the weak and transient nature of the underlying protein-protein interactions which take place in the rather specialized luminal environment of the endoplasmic reticulum (ER). As a matter of fact, many luminal protein interactions of secretory and membrane proteins are missed by standard techniques of interaction proteomics such as affinity isolation or the yeast-two hybrid system. To overcome these substantial technical limitations, we tested if a protein fragment complementation assay (PCA) based on the yellow fluorescent protein (YFP) can be applied *in vivo* to capture protein-protein interactions inside the lumen of the secretory pathway. YFP PCA relies on complementing YFP from two non-fluorescent fragments (YFP1 or YFP2) which have been fused to two test proteins. If the two test proteins interact, YFP1 and YFP2 are brought into close proximity which induces the correct folding and reconstitution of fluorescent YFP. By successfully applying a YFP PCA inside the lumen of the ER, we could visualize the interaction between ERGIC-53 and its luminal interaction partners MCFD2, cathepsin Z and cathepsin C in a specific manner. Noteworthy, a direct interaction between cathepsin C and ERGIC-53 has been suspected previously but could not be established by chemical crosslinking and affinity purification-based techniques. Thus, YFP PCA is a powerful tool to capture protein interactions inside the secretory pathway.

To search for additional cargo proteins of ERGIC-53, we developed a novel genomewide fluorescence complementation-based cDNA library screen. For this purpose, we constructed for the first time a cDNA-YFP1 fusion library which meets all

the requirements for probing protein-protein interactions in the lumen of the secretory pathway by YFP PCA. The library was co-transfected with the YFP2-ERGIC-53 bait into mammalian COS-1 cells. Fluorescence activated cell sorting was then used to isolate yellow fluorescent COS-1 cells from which library plasmids were recovered. In a small-scale pilot screen, we identified alpha-1-antitrypsin as potential interaction partner of ERGIC-53 suggesting that ERGIC-53 might bind more cargo proteins than initially assumed. The identification of alpha-1-antitrypsin demonstrates that YFP complementation can be successfully applied to screen a cDNA library for novel protein-protein interactions. This approach should provide a firm basis to map protein interactions inside the secretory pathway in a genomewide setting.

With the ability to visualize and quantify protein interactions between ERGIC-53 and its cargo *in vivo*, YFP PCA is a potent technique to analyze the ERGIC-53/MCFD2 transport receptor complex in more detail. Hence, we used luminal YFP complementation to establish the cargo binding properties of the ERGIC-53/MCFD2 complex and showed that ERGIC-53 can bind cathepsin Z and cathepsin C in a MCFD2-independent manner. This suggests cargo selectivity of the ERGIC-53/MCFD2 complex. While ERGIC-53 can interact with cathepsin Z and cathepsin C in the absence of MCFD2, MCFD2 is selectively required for the recruitment of blood coagulation factors V and VIII. A combination of short interference RNA-mediated ERGIC-53 knockdown, immunofluorescence-based protein localization, and tracking of metabolically labeled MCFD2 revealed a strict dependence of MCFD2 on ERGIC-53 for correct localization and intracellular retention. Our finding that MCFD2 is secreted upon a knockdown of ERGIC-53 explains the lack of MCFD2 that has been reported in ERGIC-53 deficient hemophilic patients suffering from combined blood coagulation factor V and VIII deficiency.

In conclusion, this thesis provides deeper insight into receptor-mediated cargo capture by proposing cargo selectivity of the ERGIC-53/MCFD2 transport complex. Furthermore, the development of the luminal YFP PCA provides attractive and promising perspectives to analyze and screen protein interactions inside the lumen of the secretory pathway.

Table of contents

1. Introduction

Endoplasmic reticulum Golgi intermediate compartment protein-53 (ERGIC-53) is a protein which has been intensively studied during the last 18 years[1]. Although ERGIC-53 has been characterized in detail, knowledge on its biological function as a cargo receptor is still incomplete. The lack of ERGIC-53 in humans causes reduced secretion of blood coagulation factors V and VIII[2, 3]. But why is ERGIC-53 ubiquitously expressed at high levels while blood coagulation factor V and VIII expression is largely restricted to liver cells? Why is ERGIC-53 conserved in a species like *C. elegans* lacking a blood coagulation cascade? Are we missing an essential information about ERGIC-53? To address this point, I decided to search for further ERGIC-53 interaction partners which could give us insights into additional functions of the protein.

Recently, a genetic screen reported the identification of multiple coagulation factor deficiency protein 2 (MCFD2) as luminal ERGIC-53 interacting protein[4]. Like ERGIC-53, MCFD2 is required for the secretion of blood coagulation factors V and VIII. Since the interaction between MCFD2 and ERGIC-53 strictly depends on the presence of calcium, this interaction has not been captured previously by a variety of ERGIC-53 pulldown approaches all of which lacked calcium. This raises the question if there is a better suited method for the identification of ERGIC-53 interaction partners than pulldown experiments.

This thesis combines the area of protein secretion with the method-based field of interaction proteomics. The mammalian secretory pathway is introduced and molecular machineries which act in the lumen of the endoplasmic reticulum (ER) in processes such as protein folding, modification, quality control, degradation and secretion are described in detail. Furthermore, currently available techniques in interaction proteomics are presented and evaluated in regard of their application to capture ERGIC-53 interaction partners or luminal protein-protein interactions, in more general terms. The main focus lies on a novel method called protein fragment complementation assay (PCA)[5].

1.1 The secretory pathway of mammalian cells

1.1.1 Organization of the secretory pathway

The exchange of molecules between the intracellular and extracellular environment is an essential task for eukaryotic and prokaryotic cells. While small molecules such as ions pass cellular membranes through specialized channels, larger molecules such as proteins require more elaborate mechanisms to enter or leave a cell. Eukaryotic cells have developed a complex secretory pathway, composed of diverse membrane-enclosed compartments to regulate the process of secretion (exocytosis) and protein uptake (endocytosis). Figure 1 depicts the mammalian endomembrane system which is responsible for transport and sorting of membrane and soluble proteins. The functional dissection of the secretory pathway began more than 30 years ago in the laboratory of George Palade with the visualization of the secretion process of tritium-labeled proteins by electron microscopy autoradiography and subcellular fractionation. Newly synthesized secretory proteins were shown to be transported in a vectorial process from the ER via the Golgi apparatus to the plasma membrane[6].

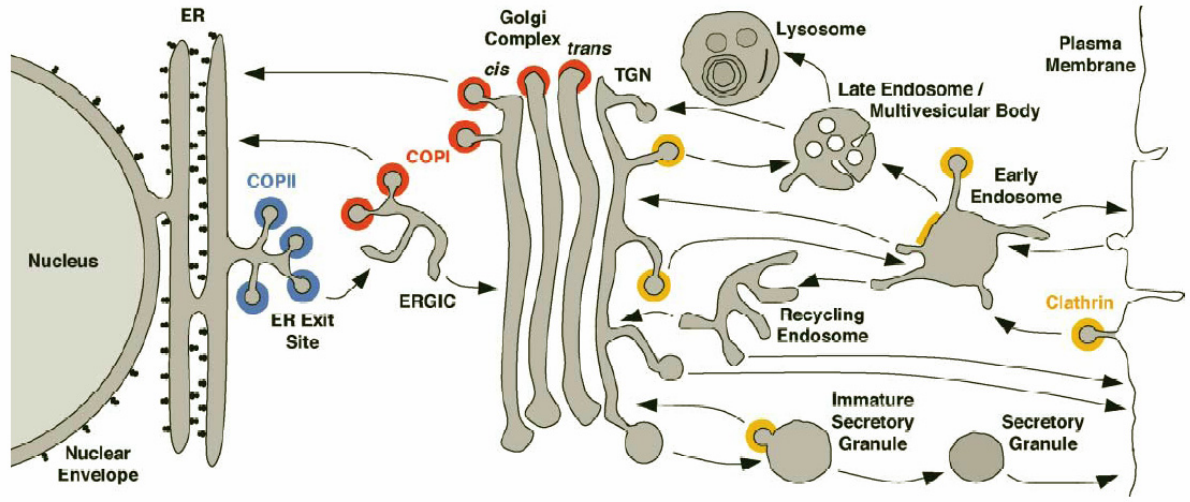


Figure 1: Organization of the mammalian secretory pathway.

The scheme illustrates the different compartments involved in exo- and endocytosis of proteins. The endoplasmic reticulum (ER), the ER Golgi intermediate compartment (ERGIC), the Golgi complex, the endosomal and lysosomal compartments as well as the plasma membrane are interconnected through vesicular transport pathways as indicated by arrows. The localization of coated membranes are depicted in blue for coat protein II (COPII), in red for coat protein I (COPI) and in orange for clathrin (Reproduced from Bonifacino and Glick, 2004[7]).

The early secretory pathway is composed of the ER, the ER Golgi intermediate compartment (ERGIC) and the Golgi apparatus.

ER

The ER can be divided into three morphologically distinct subdomains, namely the nuclear envelope, the ribosome-covered rough ER and the ribosome-free smooth ER[8]. The rough ER is the entry point for secretory and membrane proteins into the secretory pathway and participates, together with the smooth ER, in lipid, cholesterol and steroid biosynthesis[9] as well as cellular detoxification. The lumen of the ER provides a specialized environment assisting folding and modification of most proteins. Since the ER is the main controller of cellular protein folding and degradation, it is an important signaling organelle by regulating various cellular processes through the unfolded protein response (UPR)[10].

ERGIC

The ERGIC consists of tubulovesicular membrane clusters located between the ER and the Golgi. The ERGIC is a stable compartment of long-lived stationary structures which are interconnected by highly mobile, short-lived elements[11]. The ERGIC is the first sorting station for anterograde cargo molecules. Secretory proteins are transported towards the Golgi while ER resident proteins are retrieved back to the ER[12]. Proteomic analysis of the ERGIC identified several chaperones suggesting an additional function of the ERGIC in post-ER protein quality control[13].

Golgi apparatus

The Golgi apparatus consists of a series of flattened cisternal membranes which differ in their protein and lipid composition. During the secretion process, proteins enter the Golgi at the *cis* face and leave at the *trans* face while intra Golgi transport is thought to occur through either vesicular transport or cisternal maturation[14]. The Golgi apparatus is involved in cellular signaling, sphingolipid biosynthesis and is the major processing and sorting compartment[15]. Several Golgi resident enzymes are involved in the processing of N- and O-linked carbohydrate side chains. After completed posttranslational modification, proteins are sorted at the trans Golgi network (TGN) to their final destination such as plasma membrane, secretory granules, endosomes or lysosomes[16].

1.1.2 The concept of vesicular transport

The observation that newly synthesized secretory proteins can be detected during their secretion process within small vesicular carriers led to the fundamental principle of vesicular membrane transport[6]. Vesicles transport proteins and lipids between organelles by budding from a 'donor' compartment and fusing with an 'acceptor' compartment. Figure 2 illustrates the prevailing model of the budding and fusion machinery which ensures exchange of material between different organelles without affecting their homeostasis. To generate a vesicle, cytosolic coat proteins need to be recruited to the membrane and deform the lipid bilayer. To date, three different coatomers have been identified that participate in different intracellular transport pathways. The clathrin coat mediates vesicular budding from the Golgi, the plasma membrane and the endosomal compartments while coat protein I (COPI) and coat protein II (COPII) ensure bidirectional transport between the ER and the Golgi[17].

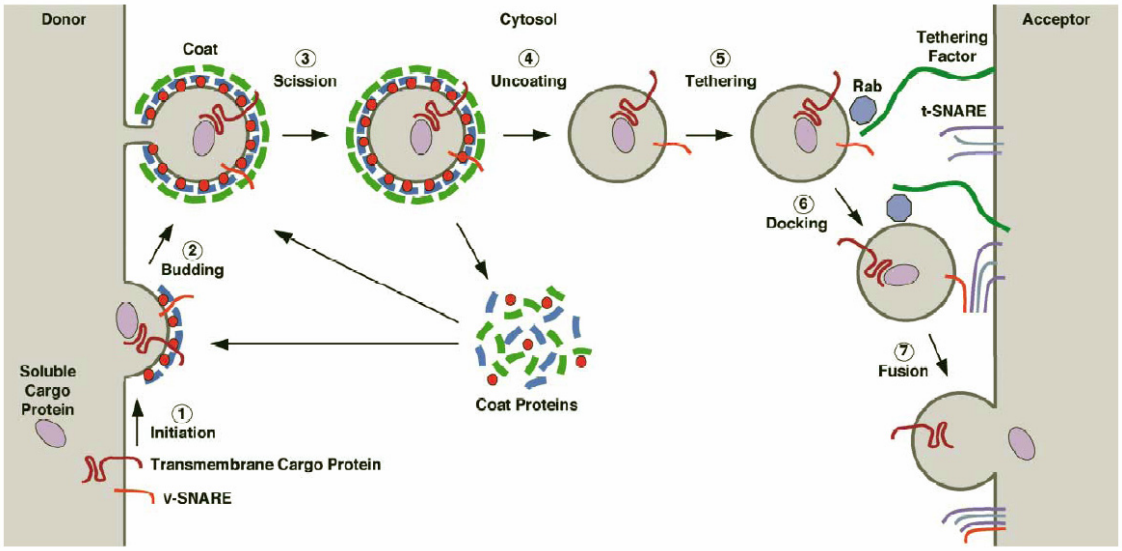


Figure 2: Vesicular budding and fusion machinery.

(1) Coat assembly is initiated by the recruitment of coat proteins (blue) to membrane-associated GTPases (red). Transmembrane cargo and SNARE proteins are recruited into budding structures. (2) Outer coat components (green) bind to the budding complex, polymerize into a mesh-like network and induce vesicular formation. (3) Vesicular scission is caused by the coat itself or by accessory proteins at the neck between the vesicle and the donor compartment. (4) Vesicles are uncoated due to inactivation of membrane-associated GTPases, phosphoinositide hydrolysis or through the action of uncoating enzymes. Cytosolic coatomers are recycled for further rounds of vesicle budding while the uncoated vesicle moves towards the acceptor compartment. (5) Vesicles are tethered to the acceptor membrane through different Rabs and tethering factors. (6) v- and t-SNAREs interact with each other thereby promoting vesicle docking. (7) Assembly of the SNAREs into a four-helix bundle induces membrane fusion (Reproduced from Bonifacino and Glick, 2004[7]).

COPII-dependent trafficking from ER to Golgi is one of the best studied examples of vesicular transport and was primarily unraveled by two independent approaches: yeast genetics and vesicular transport *in vitro*. The laboratory of Randy Scheckman isolated several temperature-sensitive yeast mutants defective in protein secretion (so called *sec* mutants) thereby identifying many important components such as the Sec23, Sec24, Sec31 and Sec13 subunits of the COPII coat[18]. The laboratory of James Rothman set up an *in vitro* transport assay to analyze vesicular trafficking between mammalian Golgi cisternae[19]. Reconstitution of transport with recombinant proteins led to the identification of various proteins involved in vesicular budding and fusion.

The generation of a COPII-coated vesicle starts with the conversion of the small GTPase Sar1 from the GDP- to its GTP-bound state. This GDP/GTP exchange is catalyzed by the ER resident, transmembrane protein Sec12 which acts as guanine exchange factor (GEF)[20]. Active, GTP-bound Sar1 exposes an amphipathic α-helix that facilitates the association with the ER membrane and primes membrane curvature[21]. While membrane-associated Sar1 recruits the Sec23-Sec24 heterodimer as inner layer of the COPII coat by directly interacting with the Sec23 subunit[22], anterograde cargo molecules are recruited into the Sar1-Sec23-Sec24 pre-budding complex through a direct interaction between Sec24 and ER exit signals in the cytosolic tail of transmembrane proteins[23-27]. The pre-budding complex then recruits the Sec13-Sec31 heterotetramer as outer layer of the COPII coat, which polymerizes into a cage-like structure thereby driving vesicle formation and scission[28]. No accessory proteins have been identified for the scission process of COPII-coated vesicles. After scission, free COPII-coated transport carriers, recently visualized by immuno-electron tomography[29], shed their coat and fuse with ERGIC membranes. Subsequently, anterograde vesicular carriers are transported to the Golgi apparatus in a microtubule and dynein/dynactin dependent process[30, 31] and are tethered to the Golgi membrane trough tethering factors such as the oligomeric TRAPP I complex[32] or the coiled-coil protein p115[33]. Together with the ER Golgi SNARE proteins Bet1, Syntaxin5, membrin and Sec22[34, 35], tethering factors guarantee specific fusion with the Golgi membrane. SNARE proteins not only specify but also drive the final membrane fusion step by assembling into a four-helix bundle thereby bringing the opposing membranes close enough to fuse[36]. After membrane fusion, the SNARE complex is untwisted and recycled through the action of α-SNAP and NSF[37].

1.1.3 Protein folding and processing in the secretory pathway

About one third of all newly synthesized proteins enter the secretory pathway for either secretion into extracellular space or sorting to their final intracellular residence. Co-translational translocation into the ER through the Sec61 pore complex represents the general mechanism for entering the secretory pathway[38]. The translocation event is preceded by the recognition of the N-terminal signal sequence in secretory and membrane proteins by the signal recognition particle (SRP)[39]. Binding to the SRP receptor targets the ribosome-nascent polypeptide-SRP complex to the Sec61 pore complex of the ER membrane[40]. Already during the translocation process newly synthesized proteins are modified and start to fold under the assistance of different ER resident enzymes. Disulfide bridge formation, N-glycosylation, signal sequence cleavage and the attachment of glycosyl-phosphatidylinositol (GPI) anchors belong to the major modification events occurring in the ER. Proper protein modifications are often a prerequisite for correct folding as exemplified by the tendency of glycoproteins to misfold and aggregate if glycosylation is inhibited[41]. An elaborate ER quality control system makes sure that only correctly folded proteins can leave the ER. Terminally misfolded proteins are re-translocated into the cytosol for degradation, a process known as ER-associated protein degradation (ERAD)[42]. The molecular machineries participating in folding, modification, ER quality control and ERAD will be described below in detail, and are summarized in Figure 3.

Protein folding

Protein folding in the ER is assisted by several general chaperones. The most abundant and important chaperones are glucose regulated protein 78 (Grp78, also known as BiP) and glucose regulated protein 94 (Grp94) belonging to the Hsp70 and Hsp90 families of classical chaperones. Grp78 binds to hydrophobic sequences in many different proteins while Grp94 seems to limit its interactions to a small subset of substrates[43, 44]. Chaperones assist the folding process most likely by preventing off-pathway folding intermediates thereby keeping the substrate in a folding-competent state. Grp78 and Grp94 can bind the same substrate by sequential interactions[45]. Recently it became evident that the classical ER chaperones form multiprotein complexes with other ER folding enzymes such as protein disulfide isomerase (PDI) which is involved in oxidative protein folding, or Cyclophilin B which belongs to the

cis/trans peptidyl prolyl isomerase (PPIase) family[46, 47]. A heterogeneous multiprotein network seems to cover the distinct requirements of protein folding in the ER[48]. In addition to general chaperones, there exist several protein-specific folding factors[49]. Hsp47 is an example of a collagen-specific binding protein acting as a molecular chaperon in the biosynthesis pathway of collagen[50].

Oxidative protein folding

The formation of correct disulfide bonds is a requirement for many luminal proteins to gain their correct conformation. Inhibition of oxidative protein folding results in an accumulation of misfolded proteins in the ER and an activation of the UPR[10]. Oxidative protein folding is based on the oxidation, reduction and isomerization of disulfide bridges and is catalyzed by a family of proteins called oxidoreductases. Oxidoreductases are composed of thioredoxin-like catalytic domains containing a CXXC active site motif. PDI is the best characterized oxidoreductase and acts as a molecular chaperone as well as an isomerase in the folding process of substrate proteins[51]. The non-catalytic domain of PDI binds unfolded proteins through a hydrophobic binding pocket which allows the catalytic domain to interact with substrate molecules and rearrange disulphide bridges. The human oxidoreductase family consists of 14 PDI-like family members, many of which are still uncharacterized[52].

Signal sequence cleavage

Proteins are targeted to the secretory pathway through a N-terminal signal sequence which is cleaved upon translocation into the ER in most instances. Signal sequence cleavage is performed by a signal peptidase complex composed of the five different membrane proteins SPC12, SPC18, SPC21, SPC22/23 and SPC25. SPC18 and SPC 21, the mammalian homologs of the essential yeast gene Sec11, contain catalytically active serine, histidine and aspartic acid residues and are considered to be proteolytic subunits of the signal peptidase complex[53]. In addition to the signal peptidase complex, a signal peptide peptidase was discovered which liberates and processes signal peptides by intramembrane cleavage[54].

N-linked glycosylation

The addition and processing of N-linked glycans belongs to one of the major tasks of the ER[55, 56]. Glycoproteins contain a variety of structurally distinct N-linked

carbohydrate side chains which arise from the trimming and modification of a 14-saccharide core glycan by diverse ER and Golgi resident glycosidases and glycosyltransferases[57, 58]. The 14-saccharide core glycan is co-translationally transferred from a dolicholpyrophosphate precursor to an asparagine residue in the Asn-X-Ser/Thr consensus sequence of a nascent glycoprotein. This reaction is catalyzed by the oligosaccharyltransferase (OST), an ER protein complex of nine integral membrane proteins, five of which are essential[59]. The different subunits attach OST to the Sec61 translocation pore, recruit dolicholpyrophosphate-oligosaccharide precursors, scan nascent polypeptides for glycosylation motifs and catalyze finally the N-glycosylation reaction[60].

ER quality control

N-glycans play an important role in ER quality control[61]. Depending on the structure of their N-glycans, glycoproteins can be retained in the ER or targeted for degradation. Already before termination of their synthesis, glycoproteins are trimmed by glucosidase I and II which remove the two outermost glucose residues. The resulting mono-glucosylated glycoproteins are bound and retained in the ER by calnexin and calreticulin. These two chaperone-like lectins interact with the thiol oxidoreductase ERp57 which catalyzes the rate-limiting step of protein folding in the ER, namely the formation of intramolecular disulfide bridges[42]. After release from calnexin and calreticulin, the remaining glucose residue is removed by glucosidase II and the folding state of the glycoproteins is monitored by UDP-glucose:glycoprotein glucosyltransferase (UGGT). UGGT recognizing both oligosaccharide and protein moieties[62] and specifically reglucosylates incompletely folded glycoproteins which allows the reassociation with calnexin and calreticulin. Noteworthy, most glycoproteins can acquire their native structure in a single round of association with calnexin and calreticulin[63]. Correctly folded glycoproteins are not reglucosylated and can leave the ER. Apart from calnexin and calreticulin, several other ER enzymes such as Grp78, Grp94 or PDI participate in primary quality control by recognizing exposed hydrophobic regions or unpaired cystein residues and thereby retaining non-native proteins in the ER[42, 49].

ER associated degradation

Prolonged retention of a protein in the ER is an indication of its inability to gain the correct conformation. Terminally misfolded proteins need to be specifically recognized, retranslocated into the cytosol and degraded[64]. In yeast, the slow-acting ER mannosidase I temporally controls the retention time of glycoproteins in the ER and initiates the ERAD process by trimming glycans to their $Man_8GlcNAc_2$ B isoform[65, 66]. This B isoform is specifically recognized by the luminal lectins Htm1p and Yos9p. Interestingly, Yos9p can be found in a luminal complex containing Kar2p, the homologue of mammalian Grp78[67]. Kar2p is thought to prevent aggregation of misfolded proteins by maintaining them in a retranslocation-competent conformation. In mammalian cells, ERAD is also initiated by mannose trimming. Glycoproteins are however trimmed down to $Man_{5-6}GlcNAc_2$[68] and are bound by ER-degradation enhancing α-mannosidase-like proteins 1-3 (EDEM1-3)[69, 70]. Noteworthy, EDEM1 is the mammalian homologue of yeast Htm1p. Binding of Htm1p and Yos9p in yeast or EDEM1-3 in mammalian cells is believed to target misfolded proteins for retranslocation into the cytosol. How exactly and through which channel the process of retranslocation occurs is still a matter of debate[71]. In the cytosol, proteins are completely deglucosylated and targeted via ubiquitination for 26S proteasomal degradation.

GPI anchoring

GPI anchoring is a covalent, post-translational protein modification attaching diverse enzymes, receptors and antigens to the membrane via their carboxy-terminus. GPI is synthesized in a multi-step reaction in the ER and is composed of oligosaccharides as well as inositol phospholipids. A carboxy-terminal, hydrophobic signal marks proteins for GPI anchoring. The addition of GPI is catalyzed by an ER transamidase complex, consisting of at least five protein subunits. The transamidase complex recognizes and cleaves the GPI-anchoring hydrophobic signal before covalently linking GPI[72].

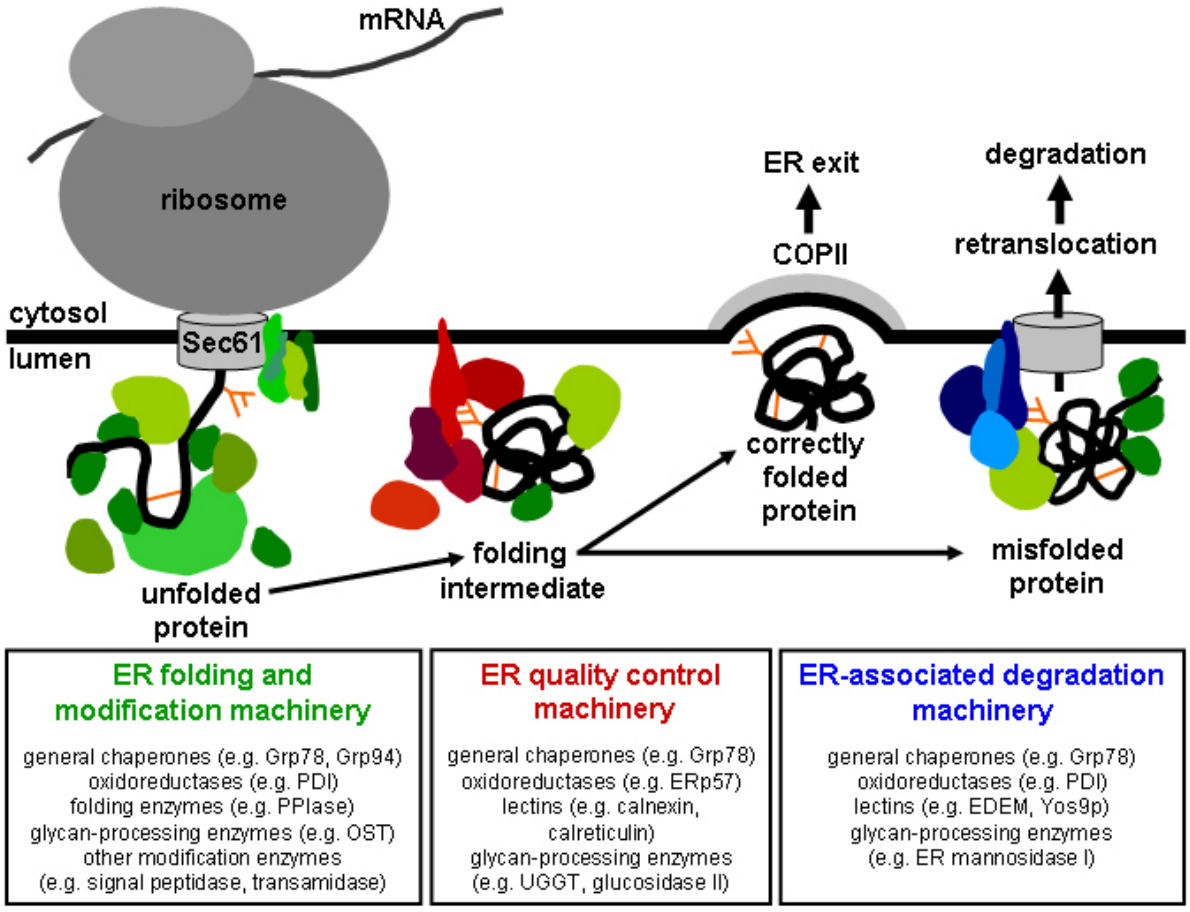


Figure 3: ER folding, modification, quality control and degradation machinery.

Proteins entering the secretory pathway are co-translationally inserted through the Sec61 pore complex into the ER. In the lumen of the ER, proteins are processed by an elaborate folding and modification machinery (green). Oxidoreductases and folding enzymes are thought to participate actively in the folding process while chaperons keep the substrate in a folding-competent conformation by preventing aggregation. ER resident enzymes such as oligosaccharyltransferase (OST), glucosidase I and II, signal peptidase or the transamidase complex modify the folding substrates. The folding and modification process is monitored by an ER quality control machinery (red). General chaperones and oxidoreductases participate in primary quality control by retaining proteins with exposed hydrophobic regions or unpaired cystein residues in the ER. Calnexin, calreticulin, ERp57, glucosidase II and UDP-glucose:glycoprotein glucosyltransferase (UGGT) compose the calnexin-calreticulin cycle retaining glycoproteins in the ER and promote correct folding. While correctly folded proteins can leave the ER in COPII-coated vesicles, terminally misfolded proteins are recognized and degraded by the ER-associated protein degradation (ERAD) machinery (blue). Misfolded luminal proteins are processed by the slow acting ER mannosidase I, recognized by lectins such as Yos9p or ER-degradation enhancing α-mannosidase-like proteins 1-3 (EDEM1-3) and retranslocated into the cytosol for degradation. General chaperons and oxidoreductases participate in the ERAD process as well. The different players of the folding, modification, quality control and ERAD machinery are thought to compose a network of flexible, heterogeneous multiprotein complexes interacting with the substrates to be processed.

1.1.4 Protein exit from the ER

ER exit of transmembrane proteins

Transport competent proteins leave the ER in COPII-coated vesicles. Transmembrane proteins can be selectively incorporated into budding vesicles by direct interaction between their cytosolic tails and the Sec24 subunit of the COPII coat[73]. The interaction with Sec24 is mediated by different conserved amino acid motives, termed ER exits signals. Studies on the vesicular stomatitis virus glycoprotein (VSV-G) led to the identification of the di-acidic (DXE) export motif[74]. ER export of the potassium channel Kir2.1 and the yeast proteins Gap1p and Sys1p also relies on di-acidic export signals[75-77]. The characterization of ERGIC-53 targeting determinants led to the identification of the di-hydrophobic motives as additional class of ER export signals[78]. The di-phenylalanine motif in ERGIC-53 can be substituted by a single phenylalanine or tyrosine at position -2, two leucines or isoleucines at position -1 and -2 or a single valine at position -1. Importantly, the single carboxy-terminal valine can be transplanted to other membrane proteins and accelerate their transport, hence acting as a true ER export signal[23]. The finding that ERGIC-53 hexamerization is required for efficient ER exit suggests that the oligomeric presentation of di-hydrophobic motives enhances the recruitment into COPII-coated vesicles[79, 80]. In addition to ERGIC-53, di-hydrophobic ER exit motives can be found in transmembrane proteins belonging to the p24 family of proteins or the Erv41-Erv46 complex[81, 82]. A third class of ER export signals, the di-basic (R/K-X-R/K) motives, was discovered in Golgi glycosyltransferases. In contrast to the di-acidic and di-hydrophobic signals, di-basic motives interact with the Sar1 component of the COPII coat[27]. How can the COPII coat recruit at least three different ER export signals? The participation of Sec23, Sec24 as well as Sar1 in the cargo binding process is for sure one option. In addition, coat proteins might accommodate different cargo proteins by having multiple binding sites. In Sec24 three different binding sites can bind the di-acidic export motif as well as signals in the ER Golgi SNARE proteins Sed5, Bet1 and Sec22[24, 25]. Another intriguing finding is that the different isoforms of Sec24, four of which can be found in humans, show selective binding preferences for different ER exit signals (unpublished data, M.Wendeler).

ER exit of soluble proteins

In contrast to transmembrane proteins, soluble luminal proteins can not directly interact with the cytosolic COPII coat. Soluble proteins were first proposed to enter budding vesicles at their prevailing luminal concentration, a model based on bulk flow-mediated transport[83]. The bulk flow model was challenged by the finding that certain soluble cargo proteins are enriched in COPII-coated vesicles. Glycosylated pro-α-factor, for example, was enriched about 20 fold in ER-derived vesicles in comparison to the ER. This enrichment is dependent on the membrane protein Erv29p which suggests receptor-mediated cargo capture for ER exit[76]. The concept of receptor-mediated ER export was strengthened by the recent identification of a hydrophobic signal in glycosylated pro-α-factor which is required for binding to Erv29p and for its efficient packaging into COPII-coated vesicles. Importantly, this hydrophobic signal can be transplanted to the ER resident protein Kar2p thereby promoting its incorporation into COPII-coated vesicles[84]. So far, Erv29p is the best characterized cargo receptor in yeast and specifically recruits not only glycosylated pro-α-factor but also vacuolar hydrolases, carboxypeptidase Y and proteinase A[85].

There is also evidence for receptor-mediated ER export of soluble cargo proteins in mammalian cells. The transmembrane protein ERGIC-53 captures soluble proteins such as cathepsin Z, cathepsin C and blood coagulation factors V and VIII in the ER and displays many characteristics of a cargo receptor[3, 86-89]. The di-phenylalanine motif in the cytosolic tail of ERGIC-53 binds to COPII which ensures the recruitment into ER-derived vesicles for anterograde transport[78]. Furthermore, a di-lysine motif that interacts with COPI mediates retrieval back to the ER[90]. Binding to COPII as well as COPI enables ERGIC-53 to cycle between the ER and the ERGIC. ERGIC-53 is a mannose-binding lectin[86, 91] and a histidine residue in its carbohydrate recognition domain modulates the lectin activity in a calcium- and pH-dependent manner. This raises the possibility that calcium and pH control cargo binding and release in different compartments. Loss-of-function mutations in ERGIC-53 lead to combined blood coagulation factor V and VIII deficiency in humans (OMIM # 227'300)[3]. These patients show reduced levels of blood coagulation factors V and VIII in the plasma. A reduced secretion of blood coagulation factors V and VIII was also shown in mammalian cell cultures upon expression of a cycling deficient ERGIC-53 mutant[2]. Interestingly, the efficient secretion of blood coagulation factors V and VIII requires an additional protein,

named MCFD2[4]. MCFD2 is a soluble, luminal protein which interacts with ERGIC-53 in a calcium-dependent manner. The participation of ERGIC-53 as well as MCFD2 in the secretion process of blood coagulation factors V and VIII raises the intriguing possibility of cargo receptor complexes composed of a transmembrane and a soluble protein.

The existence of receptor-mediated ER export of luminal proteins has been documented in yeast as well as mammalian cells[86, 92]. The prevalence of receptor-mediated ER exit, however, remains elusive. Receptor-mediated ER export may account for only a subset of soluble proteins while the majority of proteins might rely on bulk flow-mediated transport (Figure 4). Perhaps only rare soluble proteins require a receptor for efficient recruitment into COPII-coated vesicles. Moreover, cargo receptors might act as secondary quality control by capturing only correctly folded cargo proteins.

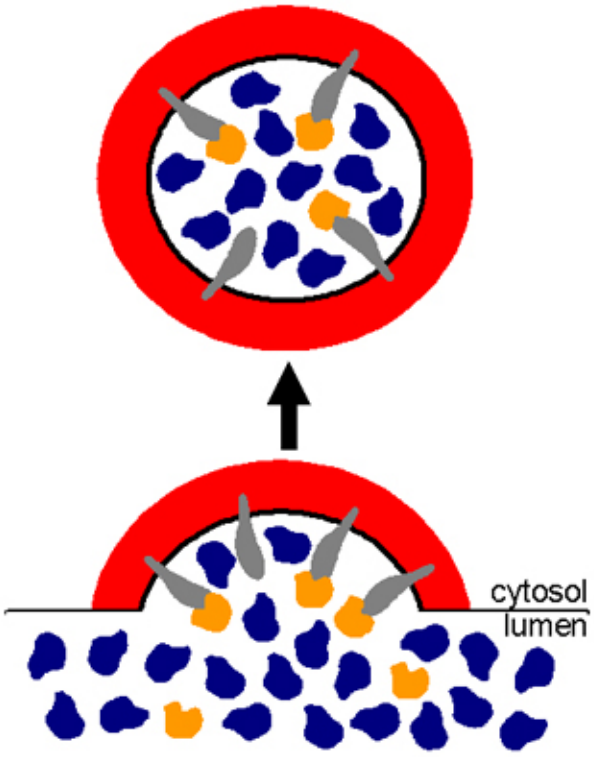


Figure 4: Bulk flow- and receptor-mediated ER export of soluble proteins.

The bulk flow model predicts that abundant soluble proteins (blue) enter COPII-coated vesicles (red) at their prevailing luminal concentration. Transmembrane receptors (gray) are recruited into budding vesicles by direct interaction between their cytosolic tail and the COPII coat. The transmembrane receptors bind luminal cargo proteins (orange) thereby recruiting them into vesicles. Receptor-mediated ER export could enrich rare soluble proteins inside ER-derived vesicles and function as secondary quality control process.

1.2 Interaction proteomics

In 2001, a first draft of the human genome was released[93]. Three years later most gaps in the human genome were closed and the number of protein-coding genes was estimated to be in the range of 20'000-25'000[94]. What is the biological function of all these proteins in a human cell? This challenging question is currently tackled amongst others by interaction proteomics. According to the concept of 'guilt-by-association', proteins that interact with one another fulfill similar biological functions[95, 96]. The analysis of protein interaction networks in yeast demonstrated that most biological processes require protein complexes rather than single proteins[97]. The dissection of multiprotein complexes and the identification of all interacting proteins are therefore crucial to assess the biological function of the many currently uncharacterized proteins. Given that the estimated 5 interaction partners per yeast protein[98] are also true for higher eukaryotes, we are left with at least 100'000 protein interactions per human cell. How can we analyze this plethora of interactions and how can we characterize the different properties of protein interactions as depicted in Figure 5?

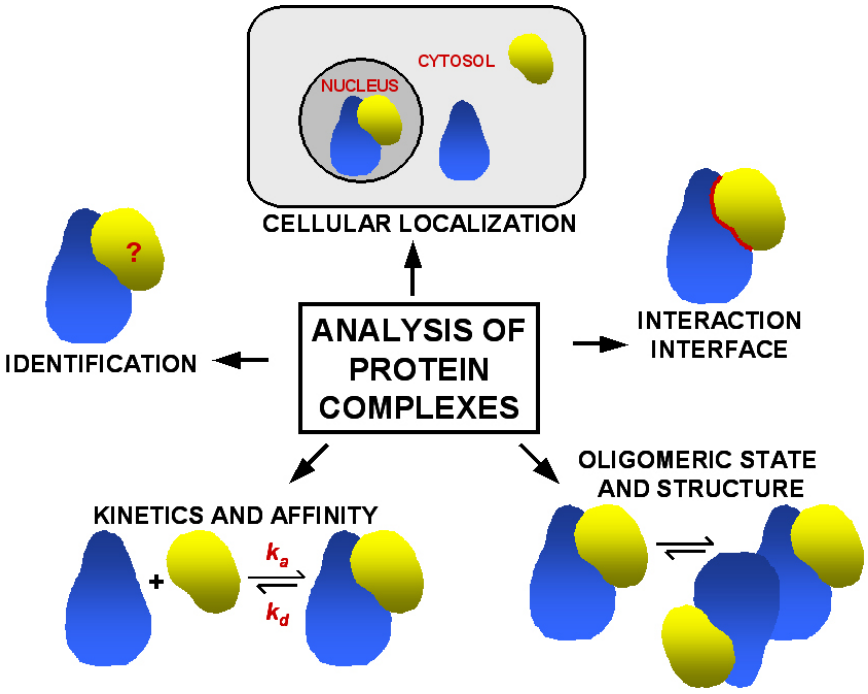


Figure 5: The different levels of characterization of protein complexes.

The detailed analysis of a protein complex requires the identification of interacting proteins as well as the determination of structure, oligomeric state, stoichiometry and cellular localization of the complex. Furthermore, the kinetics and affinities of the individual interactions have to be determined and the amino acid residues making contact at the interaction interface need to be identified.

Interaction proteomics has been greatly facilitated by the availability of genome sequence data and versatile methods for the analysis of protein-protein interactions[99-101]. In order to identify protein interaction partners, methodologies are available for screening large numbers of different protein interactions. Typically, a protein of interest (called 'bait') is used to screen for novel interacting proteins (called 'preys'). Affinity purification of multiprotein complexes, protein microarrays and library-based screening methods such as the yeast-two hybrid system belong to the currently most popular methods in interaction proteomics and are described in more detail on pages 22-28. The emerging field of computational-based prediction of protein-protein interactions and its application to the construction of protein interaction networks is presented on page 29. Once interaction partners are identified, the biochemical and biophysical properties of the underlying protein interactions can be characterized by a number of sophisticated methods (Table 1).

Table 1: Commonly used methods for the characterization of protein complexes

Determination of kinetics and affinities of protein-protein interactions

- surface plasmon resonance
- affinity chromatography
- isothermal titration calorimetry

Analysis of protein-protein interaction interfaces

- nuclear magnetic resonance spectroscopy
- hydroxyl-radical footprinting
- reverse yeast two-hybrid

Determination of the oligomeric state and structure of protein complexes

- gel filtration
- analytical ultracentrifugation
- light scattering
- electron microscopy
- X-ray crystallography
- nuclear magnetic resonance spectroscopy
- atomic force microscopy

Analysis of cellular localization and dynamics of protein complexes

- fluorescence resonance energy transfer
- protein fragment complementation
- single-molecule spectroscopy

1.2.1 Affinity purification-based techniques

One of the most elaborate and broadly applied methods to identify interacting proteins combines affinity-based isolation of protein complexes and mass spectrometry-based protein identification (Figure 6).

Ideally, an endogenous protein is purified with associated proteins from its physiological context using a specific antibody. Since the number of specific antibodies is limited, a more general strategy is to purify a tagged version of the protein. Glutathione S-transferase, small peptide epitopes like FLAG, HA or c-myc as well as poly-His tails can serve as affinity tags for protein isolation using immobilized glutathione, anti-epitope antibodies or nickel ions. Recently, a tandem affinity purification (TAP) tag was introduced which allows two sequential affinity isolation steps thereby increasing the specificity of the isolation procedure under largely physiological conditions[102].

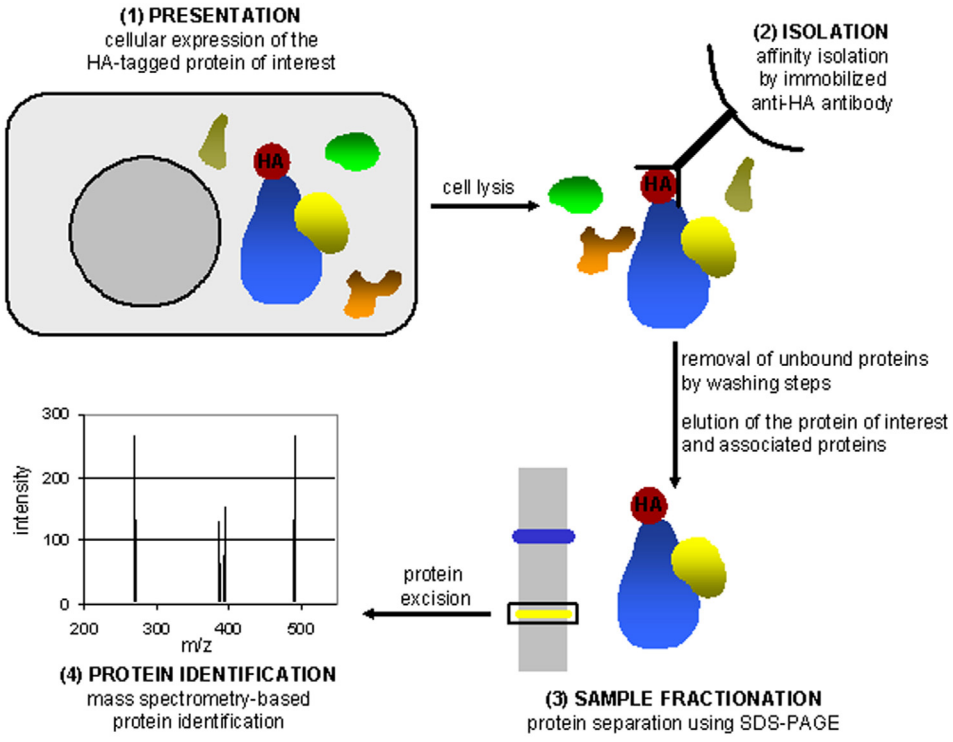


Figure 6: Affinity purification-based identification of protein interaction partners.
A protein of interest is fused to an affinity tag as exemplified by the HA epitope. (1) The HA-tagged protein is expressed and associates with endogenous interaction partners. (2) The protein complex is isolated using anti-HA immunoprecipitation. (3) Unbound proteins are washed away and the protein of interest is eluted with its associated partners and fractionated by SDS-PAGE. (4) Separated proteins are excised, digested with a sequence-specific protease such as trypsin and identified by mass spectrometry.

After immobilizing the endogenous or affinity tagged protein of interest, unbound proteins are removed in distinct washing steps and interacting proteins are eluted, separated by SDS-PAGE and identified by mass spectrometry procedures. Electrospray ionization and matrix-assisted laser desorption/ionization belong to the most commonly used techniques to volatize proteins while ion trap, time-of-flight, quadrupole and Fourier transform ion cyclotrons belong to the most commonly used analyzers for mass measurement[103].

Affinity purification-based methods have the advantage of capturing interaction partners expressed in their native sub-cellular location. Protein complexes composed of multiple interaction partners including soluble and membrane proteins can be dissected. In order to co-isolate interacting proteins, one has to find suitable conditions and buffers allowing the protein interactions to be maintained *in vitro* after cell lysis. Amount and nature of the detergent, salt concentrations as well as the addition of certain cofactors need to be adjusted for different protein interactions. Despite the need of adjusting conditions for different protein interactions, affinity purification-based methods were successfully applied in two large-scale studies. The protein-protein interaction network of *S. cerevisiae* was partially mapped upon affinity isolation of either TAP-tagged[104] or FLAG-tagged[105] yeast proteins followed by the identification of co-isolated proteins by mass-spectrometry. Noteworthy, these two large-scale studies, although internally reproducible, showed little overlap which suggests non-saturating screening conditions in terms of covering only subsets of all possible protein-protein interactions[106].

1.2.2 Protein array-based techniques

Protein array-based techniques probe proteins immobilized in a spatially resolved manner on a solid support, with a molecule of interest (Figure 7). Protein arrays consist of either crude protein extracts or purified, recombinant proteins, spotted on filter membranes or solid surfaces like glass slides.

Far-Western-based screening of high-density protein filters has been successfully applied for the identification of protein interaction partners[107]. High-density protein filters contain arrayed crude protein extracts from individual bacterial clones expressing recombinant proteins from a cDNA library. These filters are probed with the protein of interest which is consequently visualized by Western blotting. The positional information of the Western blot signal allows the identification of the

bacterial clone harboring the cDNA plasmid responsible for the expression of the interaction partner[108]. In addition to Western blot-based detection, high-density protein expression filters can be directly probed with a radiolabeled protein of interest[109].

By developing the immobilization of large number of different proteins on solid surfaces in a functional manner, the first protein microarray experiments were made possible[110, 111]. In 2001, a protein microarray containing the whole yeast proteome was produced by spotting 5800 purified, poly-His tagged yeast proteins onto nickel coated slides. Several calmodulin- and phospholipid-binding proteins were identified by probing the protein microarray with biotinylated calmodulin and liposomes followed by the visualization with fluorescent-labeled streptavidin[112]. The application of microarrays to protein interactions in the human proteome is so far restricted to the analysis of protein interaction domains. An interaction network of the human epidermal growth factor receptors was recently created by measuring the affinities between different SRC homology 2 and phosphotyrosine binding domains[113].

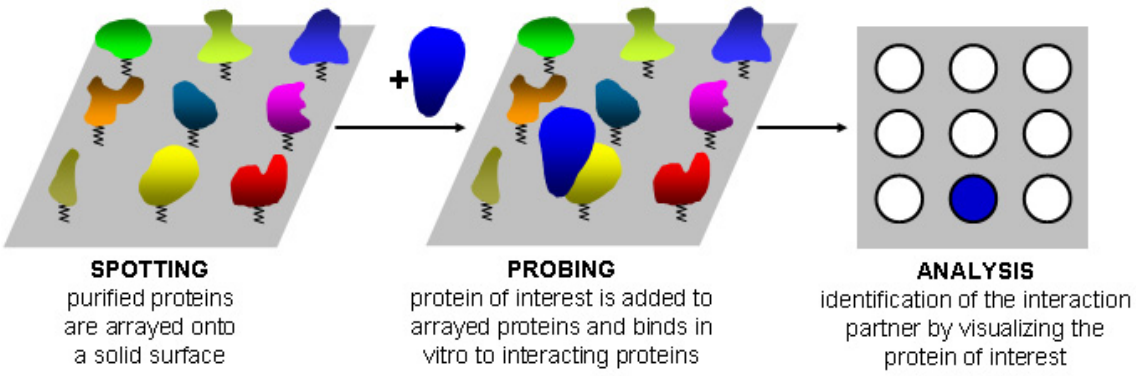


Figure 7: Protein microarray technology.

Individual proteins are purified and immobilized in a functional manner in high density on a solid surface to form a protein microarray. A protein of interest is added to the microarray and interacts *in vitro* with its interaction partners. Visualization of the protein of interest identifies interacting proteins on the microarray.

Array-based techniques have the potential to unravel interaction networks rapidly because once the array is prepared a large number of interactions can be probed at the same time. The possibility to analyze protein-protein, protein-lipid, protein-carbohydrate as well as protein-drug interactions is a key advantage of the technology[114]. The limiting factor of protein microarrays is the need to generate purified, recombinant proteins for spotting. Given the 20'000-25'000 protein coding

human genes, this is associated with a tremendous amount of work. The successful identification of protein interaction partners in microarray-based experiments relies again on finding the right *in vitro* conditions allowing the interactions to take place. Since spotted onto one array, a single condition has to be chosen for all protein interactions leaving no possibility for individual optimization.

1.2.3 Library-based techniques

Library-based methods probe protein interactions between a bait and a library-encoded prey protein. The complexity of the library determines the number of binary protein interactions to be screened. Ideally, a library expresses a representative of each protein in the proteome[99]. Versatile techniques like phage display, the yeast two-hybrid system or protein fragment complementation assays use a library-based approach to identify protein interaction pairs.

Phage Display

The phage display technique is based on the insertion of a DNA sequence library into the genome of a filamentous phage such as M13. Since the DNA sequences are inserted into a coat protein of the phage, fusion proteins are expressed and displayed on the phage surface. An immobilized protein of interest is used to capture phages displaying interaction partners on their surface while non-binding phages are washed away. Protein interactions are detected *in vitro*. In successive rounds of selection (called biopanning) specific phages are enriched and thereafter characterized[115]. The use of complex libraries in combination with an enrichment strategy is a major advantage of the phage display technology. Serious limitations derive from the limited folding and modification capacity of filamentous phages. In addition, analyzable proteins are restricted in size and chemical properties since they need to be able to cross the phage lipid bilayer for being displayed on the extracellular surface. Transmembrane interaction partners and proteins larger than 90 kDa have never been identified by phage display-based screening presumably due to their inability to pass the outer phage membrane[116].

Yeast two-hybrid system

The invention of the yeast two-hybrid system launched the possibility to identify protein-protein interactions *in vivo*[117]. This technique takes advantage of yeast cells and the modular architecture of transcription factors as a screening tool. The DNA-binding domain (DBD) and the transcription activation domain (TAD) of a transcription factor are divided and fused to two separate proteins. If the two proteins interact, DBD and TAD can be brought into close proximity thereby reconstituting a functional transcription factor which activates reporter gene expression in yeast cells (Figure 8). Commonly used reporter genes are *lacZ* for color selection and auxotrophic genes such as *LEU2*, *HIS3* or *ADE2* for growth selection. To identify interaction partners, a protein of interest (bait) is usually fused to the DBD and screened against a cDNA library (prey) fused to the TAD. Under the appropriate selection conditions only yeast cells expressing interacting bait and prey proteins will grow. The yeast two-hybrid technology requires no protein purification steps and detects protein interactions in the living cell. This allows the detection of weak and transient protein interactions with even low stoichiometry interaction partners.

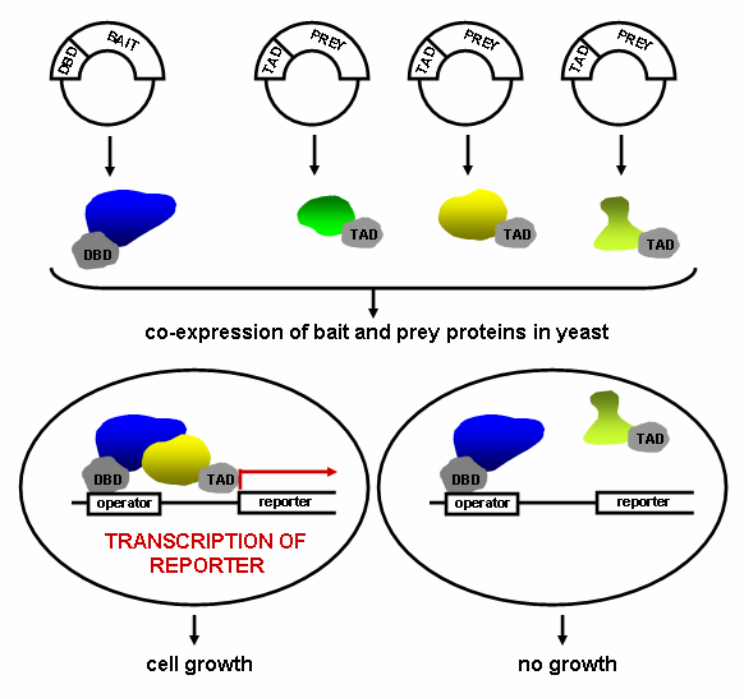


Figure 8: The yeast two-hybrid system. A bait protein, fused to the DBD is co-expressed with a library of prey proteins, fused to the TAD. If the bait interacts with a prey protein, DBD and TAD can reconstitute a functional transcription factor and activate the expression of a reporter gene. Activation of the reporter gene will promote cellular growth under the appropriate selection conditions. If bait and prey protein fail to interact with each other, the reporter gene will not be transcribed and yeast cells will fail to grow in selective medium.

The power of the yeast two-hybrid system has been demonstrated by several genome-wide screens for the proteomes of *S. cerevisiae*[118, 119], *C. elegans*[120] and *Drosophila melanogaster*[121]. A preliminary map of the human proteome was also

published on the basis of the yeast-two hybrid technique[122]. To cover as many human protein interactions as possible, different libraries have to be generated from various tissues and developmental stages. Disadvantages of the yeast two-hybrid system are the identification of many false positive interactions and its restriction to soluble proteins[118, 123]. Furthermore, protein interactions taking place in specialized subcellular compartments might not be revealed by the yeast two-hybrid system since probed in a nuclear environment. During the last decade, many variations of the classical yeast two-hybrid system have been reported[124, 125].

Protein fragment complementation assays

One variation of the classical yeast two-hybrid system takes advantage of the small protein ubiquitin. Test proteins are tagged with either the N-terminal fragment of ubiquitin or the C-terminal fragment of ubiquitin to which a transcription factor is fused. The interaction of the test proteins brings the two fragments of ubiquitin into close proximity reconstituting split-ubiquitin, which is readily recognized and cleaved by ubiquitin-specific proteases in the yeast cytosol[126]. Since ubiquitin-specific proteases cleave at the C-terminus of ubiquitin they release the transcription factor fused to the C-terminal fragment. The released transcription factor can translocate into the nucleus and activate the expression of a reporter gene. The split-ubiquitin technology is particularly useful for the analysis of protein interactions with integral membrane proteins since the protein interaction event and its subsequent detection are spatially separated[127, 128]. The split-ubiquitin system is illustrative for the technological development from the classical yeast two-hybrid system towards more general applicable protein fragment complementation assays (PCA). The detection of protein-protein interactions by PCA relies on the reconstitution of the reporter protein itself from two rationally dissected fragments (Figure 9). This is in contrast to the yeast two-hybrid and split-ubiquitin technology where protein interactions are detected indirectly through the transcriptional activation of reporter genes. The direct detection of the reconstituted reporter allows the PCA to visualize protein interactions in almost every sub-cellular compartment and cell type. Protein fragment complementation has been applied successfully to several reporter proteins such as β-lactamase, dihydrofolate reductase, *Renilla* and firefly luciferase as well as different fluorescent proteins[129-134]. All of these reporter proteins can be reconstituted from two fragments which are brought into close proximity through the interaction of two test proteins.

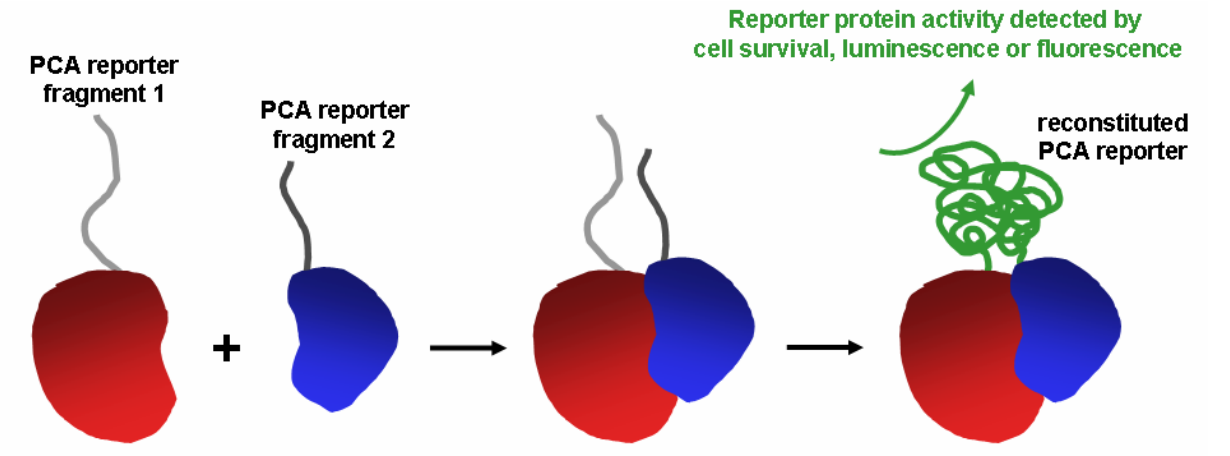


Figure 9: The basic concept of PCA.

A reporter protein is rationally dissected into fragments 1 (light grey) and 2 (dark grey) which are fused to two test proteins (red and blue). The PCA reporter fragments alone show no reporter activity. Upon interaction of the two test proteins, fragments 1 and 2 are brought into close proximity and can reconstitute the PCA reporter by folding into an active conformation. PCA reporter activity can be detected using assays based on cell survival (for DHFR), luminescence (for *Renilla* and firefly luciferase) or fluorescence (for fluorescent proteins).

Protein fragment complementation of fluorescent proteins is also known as bimolecular fluorescence complementation (BiFC) and has gained broad interest in different areas of biology. Fluorescence complementation is an especially attractive approach since fluorescent proteins are known to express, fold and fluoresce in almost every cell type and subcellular compartment[135]. Protein interactions have been visualized between soluble[133] as well as membrane proteins[136] and in different species such as bacteria, fungi, plants and mammalian cells[137]. By testing a library of antiparallel leucine zippers in bacteria, fluorescence complementation was shown to trap even weak and transient interactions with a minimal strength (K_D) of 1mM[135]. Fluorescence complementation is a powerful method to study enzyme-substrate complexes[138], to localize protein complexes[133], to investigate signal transduction cascades[139] and can be used to identify protein interaction partners using a library-based approach[140]. In 2004, the first genomewide fluorescence complementation-based cDNA library screen successfully identified an interaction partner of protein kinase PKB/Akt. In this screen, a human brain cDNA library was tagged with fragment 1 of GFP and PKB/Akt, fused to fragment 2 of GFP, was used as a bait. Bait and prey proteins were co-expressed in mammalian COS cells which were subsequently subjected to fluorescence activated cell sorting (FACS). The Ft1 protein was identified as PKB/Akt interacting protein after isolation and sequencing of prey plasmids from GFP positive COS cells[141].

1.2.4 Computational-based techniques

Apart from experimental techniques various computational-based methods have been developed to analyze protein-protein interactions. Presence or absence of genes in related species, conservation of gene neighborhood, gene fusion events, correlated mutations or similarity of physiological trees were successfully applied in the prediction of protein interactions[142]. With the availability of experimental high-throughput datasets recent attempts in bioinformatics tend to integrate several datasets for the prediction of protein interaction networks[143]. High-throughput interaction datasets from model organisms like *S. cerevisiae*[104, 105, 118, 119], *C. elegans*[120] or *Drosophila melanogaster*[121] are combined with mRNA expression profiles, protein localization data, information regarding protein interaction domains or functional annotations. These often incomplete datasets are integrated using Bayesian statistics[144] leading to quite robust protein interaction networks. By this means, a first map of the human protein-protein interaction network was modeled that predicts about 40'000 protein interactions[145].

To store, categorize and analyze the growing number of protein interactions several databases have been compiled. The Human Protein Interaction Database (HPID, http://www.hpid.org)[146] or the Yeast Protein Database (YPD: http://www.biobase-international.com)[147] are examples of organism specific protein interaction databases. In turn, the Database of Interacting Proteins (DIP: http://dip.doe-mbi.ucla.edu)[148] contains data from different species but is restricted to solely experimentally determined protein-protein interactions. The majority of all protein-protein interactions that can be found in the DIP have been identified by the yeast two-hybrid screening system (52%) and co-immunoprecipitation of interacting proteins (18%)[148].

1.3 Identification of protein-protein interactions in the lumen of the secretory pathway

Which current technique in interaction proteomics is best suited to identify protein-protein interactions in the lumen of the secretory pathway? Luminal protein interactions take place in the rather specialized environment of the ER. The ER milieu differs from the cytosol in respect to redox potential, ion concentrations as well as the presence of a specialized folding machinery. The lumen of the ER is 20 to 100 times more oxidizing than the cytosol[149] and with 20-200μM the concentration of free calcium is at least 200 times higher than in the cytosol[150]. Furthermore, the ER is densely packed with chaperones and modifying enzymes thereby highly outnumbering the substrates to be folded and modified[151]. The folding and modification process in the ER requires many transient protein-protein interactions as exemplified by newly synthesized proteins that bind to various chaperones during their folding process. Many luminal protein interactions are not only transient but also of low-affinity as exemplified by the many carbohydrate-mediated interactions between glycoproteins and ER lectins operating in protein secretion and quality control. Due to their transient and weak nature, luminal protein-protein interactions are difficult to capture. A further challenge is that many luminal protein interactions involve integral membrane proteins. Thus, identification of luminal protein-protein interactions requires a method that captures weak and transient protein interactions between soluble as well as integral membrane proteins, preferentially under oxidizing conditions and high calcium levels.

Affinity purification

Mass spectrometry-based characterization of affinity purified protein complexes can be applied to identify novel, luminal protein-protein interactions as recently demonstrated for the Yos9p-Kar2p-Hrd3p complex involved in ER associated protein degradation[67, 152]. Many luminal protein interactions are not revealed, however, since affinity purification requires a protein isolation step which mainly preserves stable interactions. Chemical crosslinking with thiol-cleavable crosslinkers in intact cells presents a possibility to capture transient protein interactions prior to affinity isolation[153]. By using chemical crosslinking, ER chaperones were shown to form multiprotein networks rather than existing as single, free proteins[46]. Furthermore, the interaction between ERGIC-53 and the lysosomal glycoprotein cathepsin Z was also

identified by chemical crosslinking[86]. Although a powerful technique, chemical crosslinking can capture protein interaction partners only if they display crosslinkable chemical groups in the correct distance and orientation. Thus, not all protein-protein interactions are amenable to chemical crosslinking procedures.

Concerning the detection of protein interactions with transmembrane proteins, affinity purification procedures can be used but require optimization. The analysis of high-throughput protein-protein interaction datasets revealed that affinity isolation-based techniques often fail to detect protein interactions with transmembrane proteins due to the difficulty of their solubilization by standard procedures[123].

Protein microarrays

It is potentially feasible to screen protein interactions by protein microarrays but the technique is still in a developmental stage. The requirement for sophisticated equipment as well as the lack of protein chips covering mammalian proteomes impedes the application of the microarray technology for the identification of novel luminal protein interactions at this stage. The risk of missing weak and transient protein interactions due to their detection *in vitro* is a further disadvantage of the protein microarray technology. Yet, an interesting application of the microarray technology is the analysis of protein-carbohydrate interactions since many important luminal protein interactions are carbohydrate-mediated and different carbohydrate structures can be synthesized on a chip[66, 154].

Yeast two-hybrid system

The yeast-two hybrid system is well established for the identification of cytosolic protein interactions and has been also applied to luminal protein interactions in individual cases[155] but is far from ideal to identify protein interactions in the lumen of the secretory pathway. Many interactions are missed due to misfolding and lack of correct disulfide bridges or post-translational modifications in luminal proteins if expressed in the yeast nucleus. Furthermore, the reducing nature and the low calcium levels of the yeast nucleus constitute a rather suboptimal environment for luminal protein interactions to take place. The restriction of the yeast-two hybrid system to soluble proteins is another disadvantage, but can be overcome by methods such as the split-ubiquitin system. Nevertheless, split-ubiquitin relies on a cleavage step by cytosolic proteases thereby limiting the analysis to cytosolic interactions of integral membrane proteins.

Protein fragment complementation assay

The concept of protein fragment complementation can be applied to protein interactions in virtually every subcellular compartment. Limitations derive from the requisite that the reconstituted reporter protein must be functional and detectable in the particular subcellular compartment. The detection of luminal protein interactions would require a PCA reporter inside the lumen of the secretory pathway. A luminal PCA would monitor luminal protein interactions in their physiological environment, hence under oxidizing conditions and under high calcium levels. Since expressed in their native environment, most luminal bait and prey proteins would be correctly folded and contain the correct post-translational modifications due to access to the ER folding and modification machinery. Importantly, the technique of protein fragment complementation is applicable to protein interactions between soluble as well as membrane proteins but suffers, like the yeast-two hybrid system, from false negative and false positive hits[140, 156]. Of note, PCA-based cDNA library screening has been reported only for the identification of cytosolic protein interactions[140].

Summary

Table 2 summarizes the advantages and disadvantages of the affinity purification-, protein microarray-, yeast two hybrid- and protein fragment complementation-based techniques. Not included are factors such as costs or amount and time of work. In terms of costs, mass spectrometry- and protein microarray-based approaches belong to the more expensive techniques due to the requirement of sophisticated instruments. Concerning amount and time of work, protein microarray- and protein fragment complementation-based screening strategies belong to the more labor-intensive methods since they are still in their developmental stage.

The evaluation of the different methods reveals that the protein microarray technology and the yeast two-hybrid system are not appropriate to identify luminal protein interactions at this stage. In contrast, the mass spectrometry-based characterization of purified protein complexes represents a powerful and broadly applicable strategy to analyze luminal protein interactions and can be combined with *in vivo* crosslinking procedures. A second promising strategy is based on the concept of protein fragment complementation. PCA-based techniques have not been applied to protein interactions in the secretory pathway but have the unique potential to capture luminal protein interactions *in vivo*, analogous to the yeast-two hybrid system for cytosolic protein interactions.

Method	Application	Advantages	Disadvantages	Membrane proteins	Luminal proteins
Affinity purification	*in vitro*	- Identification of multiple interaction partners - Prey proteins are endogenous proteins expressed at physiological conditions - Combination with *in vivo* chemical crosslinking	- Weak and transient protein interactions are often lost during purification - Purification conditions differ for different protein complexes - Low abundant and small proteins are difficult to identify by mass-spectrometry	+/-	+/-
Protein microarray	*in vitro*	- Identification of protein-protein, protein-lipid, protein-carbohydrate as well as protein-drug interactions - Detection of direct protein interactions - High-throughput application	- Detection restricted to binary protein interactions - Weak and transient protein interactions may not be detected - Technique still in developmental stage	+/-	+/-
Yeast two-hybrid	*in vivo*	- Detection of weak and transient protein interactions - Identification of low stoichiometry interaction partners - Protein fragments as well as full-length proteins can be analyzed - Modifiable screening and detection system	- Detection restricted to binary protein interactions - Many false negative hits: nuclear-based detection system limits analysis to cytosolic and nuclear interactions - Many false positive hits: over-expression of bait and prey proteins; some prey proteins show intrinsic transcriptional activation	-	-
Protein fragment complementation	*in vivo*	- All advantages of the yeast-two-hybrid system - Choice of expression system: yeast as well as mammalian cells - Bait and prey proteins are expressed in native environment and interactions are detected at the correct sub-cellular location - Correct post-translational modification of bait and prey proteins - Interactions can be visualized directly by means of fluorescence in living cells	- Detection restricted to binary protein interactions - Screening technique still in developmental stage - False negative hits: sterical hindrance of complementation - Many false positive hits when bait and prey proteins are highly overexpressed	+	+

Table 2: Evaluation of the most popular methods in interaction proteomics.

The different techniques to identify protein-protein interactions are evaluated in terms of their *in vitro/in vivo* application as well as their ability to capture membrane or luminal protein interaction partners. Intrinsic advantages and disadvantages of the methods are summarized.

1.4 References

1. Hauri, H.P., Kappeler, F., Andersson, H. & Appenzeller, C. ERGIC-53 and traffic in the secretory pathway. *J Cell Sci* **113 (Pt 4)**, 587-596 (2000).
2. Moussalli, M. et al. Mannose-dependent endoplasmic reticulum (ER)-Golgi intermediate compartment-53-mediated ER to Golgi trafficking of coagulation factors V and VIII. *J Biol Chem* **274**, 32539-32542 (1999).
3. Nichols, W.C. et al. Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. *Cell* **93**, 61-70 (1998).
4. Zhang, B. et al. Bleeding due to disruption of a cargo-specific ER-to-Golgi transport complex. *Nat Genet* **34**, 220-225 (2003).
5. Michnick, S.W. Protein fragment complementation strategies for biochemical network mapping. *Curr Opin Biotechnol* **14**, 610-617 (2003).
6. Palade, G. Intracellular aspects of the process of protein synthesis. *Science* **189**, 347-358 (1975).
7. Bonifacino, J.S. & Glick, B.S. The mechanisms of vesicle budding and fusion. *Cell* **116**, 153-166 (2004).
8. Pfeffer, S. Membrane domains in the secretory and endocytic pathways. *Cell* **112**, 507-517 (2003).
9. van Meer, G. Transport and sorting of membrane lipids. *Curr Opin Cell Biol* **5**, 661-673 (1993).
10. Schroder, M. & Kaufman, R.J. The mammalian unfolded protein response. *Annu Rev Biochem* **74**, 739-789 (2005).
11. Ben-Tekaya, H., Miura, K., Pepperkok, R. & Hauri, H.P. Live imaging of bidirectional traffic from the ERGIC. *J Cell Sci* **118**, 357-367 (2005).
12. Appenzeller-Herzog, C. & Hauri, H.P. The ER-Golgi intermediate compartment (ERGIC): in search of its identity and function. *J Cell Sci* **119**, 2173-2183 (2006).
13. Breuza, L. et al. Proteomics of endoplasmic reticulum-Golgi intermediate compartment (ERGIC) membranes from brefeldin A-treated HepG2 cells identifies ERGIC-32, a new cycling protein that interacts with human Erv46. *J Biol Chem* **279**, 47242-47253 (2004).
14. Warren, G. & Malhotra, V. The organisation of the Golgi apparatus. *Curr Opin Cell Biol* **10**, 493-498 (1998).
15. Farquhar, M.G. & Palade, G.E. The Golgi apparatus: 100 years of progress and controversy. *Trends Cell Biol* **8**, 2-10 (1998).
16. Rodriguez-Boulan, E. & Musch, A. Protein sorting in the Golgi complex: shifting paradigms. *Biochim Biophys Acta* **1744**, 455-464 (2005).
17. Kirchhausen, T. Three ways to make a vesicle. *Nat Rev Mol Cell Biol* **1**, 187-198 (2000).
18. Novick, P., Field, C. & Schekman, R. Identification of 23 complementation groups required for post-translational events in the yeast secretory pathway. *Cell* **21**, 205-215 (1980).
19. Balch, W.E., Dunphy, W.G., Braell, W.A. & Rothman, J.E. Reconstitution of the transport of protein between successive compartments of the Golgi measured by the coupled incorporation of N-acetylglucosamine. *Cell* **39**, 405-416 (1984).
20. d'Enfert, C., Wuestehube, L.J., Lila, T. & Schekman, R. Sec12p-dependent membrane binding of the small GTP-binding protein Sar1p promotes formation of transport vesicles from the ER. *J Cell Biol* **114**, 663-670 (1991).
21. Lee, M.C. et al. Sar1p N-terminal helix initiates membrane curvature and completes the fission of a COPII vesicle. *Cell* **122**, 605-617 (2005).
22. Bi, X., Corpina, R.A. & Goldberg, J. Structure of the Sec23/24-Sar1 pre-budding complex of the COPII vesicle coat. *Nature* **419**, 271-277 (2002).
23. Nufer, O. et al. Role of cytoplasmic C-terminal amino acids of membrane proteins in ER export. *J Cell Sci* **115**, 619-628 (2002).
24. Mossessova, E., Bickford, L.C. & Goldberg, J. SNARE selectivity of the COPII coat. *Cell* **114**, 483-495 (2003).
25. Miller, E.A. et al. Multiple cargo binding sites on the COPII subunit Sec24p ensure capture of diverse membrane proteins into transport vesicles. *Cell* **114**, 497-509 (2003).
26. Barlowe, C. Signals for COPII-dependent export from the ER: what's the ticket out? *Trends Cell Biol* **13**, 295-300 (2003).
27. Giraudo, C.G. & Maccioni, H.J. Endoplasmic reticulum export of glycosyltransferases depends on interaction of a cytoplasmic dibasic motif with Sar1. *Mol Biol Cell* **14**, 3753-3766 (2003).
28. Stagg, S.M. et al. Structure of the Sec13/31 COPII coat cage. *Nature* **439**, 234-238 (2006).
29. Zeuschner, D. et al. Immuno-electron tomography of ER exit sites reveals the existence of free COPII-coated transport carriers. *Nat Cell Biol* **8**, 377-383 (2006).

30. Presley, J.F. et al. ER-to-Golgi transport visualized in living cells. *Nature* **389**, 81-85 (1997).
31. Watson, P., Forster, R., Palmer, K.J., Pepperkok, R. & Stephens, D.J. Coupling of ER exit to microtubules through direct interaction of COPII with dynactin. *Nat Cell Biol* **7**, 48-55 (2005).
32. Sacher, M. et al. TRAPP I implicated in the specificity of tethering in ER-to-Golgi transport. *Mol Cell* **7**, 433-442 (2001).
33. Allan, B.B., Moyer, B.D. & Balch, W.E. Rab1 recruitment of p115 into a cis-SNARE complex: programming budding COPII vesicles for fusion. *Science* **289**, 444-448 (2000).
34. Hay, J.C. et al. Localization, dynamics, and protein interactions reveal distinct roles for ER and Golgi SNAREs. *J Cell Biol* **141**, 1489-1502 (1998).
35. Ungar, D. & Hughson, F.M. SNARE protein structure and function. *Annu Rev Cell Dev Biol* **19**, 493-517 (2003).
36. Weber, T. et al. SNAREpins: minimal machinery for membrane fusion. *Cell* **92**, 759-772 (1998).
37. May, A.P., Whiteheart, S.W. & Weis, W.I. Unraveling the mechanism of the vesicle transport ATPase NSF, the N-ethylmaleimide-sensitive factor. *J Biol Chem* **276**, 21991-21994 (2001).
38. Osborne, A.R., Rapoport, T.A. & van den Berg, B. Protein translocation by the Sec61/SecY channel. *Annu Rev Cell Dev Biol* **21**, 529-550 (2005).
39. Keenan, R.J., Freymann, D.M., Stroud, R.M. & Walter, P. The signal recognition particle. *Annu Rev Biochem* **70**, 755-775 (2001).
40. Halic, M. et al. Signal recognition particle receptor exposes the ribosomal translocon binding site. *Science* **312**, 745-747 (2006).
41. Helenius, A. How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Mol Biol Cell* **5**, 253-265 (1994).
42. Ellgaard, L., Molinari, M. & Helenius, A. Setting the standards: quality control in the secretory pathway. *Science* **286**, 1882-1888 (1999).
43. Fewell, S.W., Travers, K.J., Weissman, J.S. & Brodsky, J.L. The action of molecular chaperones in the early secretory pathway. *Annu Rev Genet* **35**, 149-191 (2001).
44. Argon, Y. & Simen, B.B. GRP94, an ER chaperone with protein and peptide binding properties. *Semin Cell Dev Biol* **10**, 495-505 (1999).
45. Melnick, J., Dul, J.L. & Argon, Y. Sequential interaction of the chaperones BiP and GRP94 with immunoglobulin chains in the endoplasmic reticulum. *Nature* **370**, 373-375 (1994).
46. Meunier, L., Usherwood, Y.K., Chung, K.T. & Hendershot, L.M. A subset of chaperones and folding enzymes form multiprotein complexes in endoplasmic reticulum to bind nascent proteins. *Mol Biol Cell* **13**, 4456-4469 (2002).
47. Kleizen, B. & Braakman, I. Protein folding and quality control in the endoplasmic reticulum. *Curr Opin Cell Biol* **16**, 343-349 (2004).
48. Sitia, R. & Braakman, I. Quality control in the endoplasmic reticulum protein factory. *Nature* **426**, 891-894 (2003).
49. Ellgaard, L. & Helenius, A. Quality control in the endoplasmic reticulum. *Nat Rev Mol Cell Biol* **4**, 181-191 (2003).
50. Nagata, K. Hsp47: a collagen-specific molecular chaperone. *Trends Biochem Sci* **21**, 22-26 (1996).
51. Yao, Y., Zhou, Y. & Wang, C. Both the isomerase and chaperone activities of protein disulfide isomerase are required for the reactivation of reduced and denatured acidic phospholipase A2. *Embo J* **16**, 651-658 (1997).
52. Ellgaard, L. & Ruddock, L.W. The human protein disulphide isomerase family: substrate interactions and functional properties. *EMBO Rep* **6**, 28-32 (2005).
53. Paetzel, M., Karla, A., Strynadka, N.C. & Dalbey, R.E. Signal peptidases. *Chem Rev* **102**, 4549-4580 (2002).
54. Martoglio, B. Intramembrane proteolysis and post-targeting functions of signal peptides. *Biochem Soc Trans* **31**, 1243-1247 (2003).
55. Helenius, A. & Aebi, M. Intracellular functions of N-linked glycans. *Science* **291**, 2364-2369 (2001).
56. Parodi, A.J. Protein glucosylation and its role in protein folding. *Annu Rev Biochem* **69**, 69-93 (2000).
57. Herscovics, A. Importance of glycosidases in mammalian glycoprotein biosynthesis. *Biochim Biophys Acta* **1473**, 96-107 (1999).
58. Kornfeld, R. & Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem* **54**, 631-664 (1985).

59. Yan, A. & Lennarz, W.J. Unraveling the mechanism of protein N-glycosylation. *J Biol Chem* **280**, 3121-3124 (2005).
60. Chavan, M. & Lennarz, W. The molecular basis of coupling of translocation and N-glycosylation. *Trends Biochem Sci* **31**, 17-20 (2006).
61. Trombetta, E.S. & Parodi, A.J. Quality control and protein folding in the secretory pathway. *Annu Rev Cell Dev Biol* **19**, 649-676 (2003).
62. Sousa, M. & Parodi, A.J. The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. *Embo J* **14**, 4196-4203 (1995).
63. Moremen, K.W. & Molinari, M. N-linked glycan recognition and processing: the molecular basis of endoplasmic reticulum quality control. *Curr Opin Struct Biol* (2006).
64. Romisch, K. Endoplasmic reticulum-associated degradation. *Annu Rev Cell Dev Biol* **21**, 435-456 (2005).
65. Tokunaga, F., Brostrom, C., Koide, T. & Arvan, P. Endoplasmic reticulum (ER)-associated degradation of misfolded N-linked glycoproteins is suppressed upon inhibition of ER mannosidase I. *J Biol Chem* **275**, 40757-40764 (2000).
66. Helenius, A. & Aebi, M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* **73**, 1019-1049 (2004).
67. Denic, V., Quan, E.M. & Weissman, J.S. A Luminal Surveillance Complex that Selects Misfolded Glycoproteins for ER-Associated Degradation. *Cell* **126**, 349-359 (2006).
68. Frenkel, Z., Gregory, W., Kornfeld, S. & Lederkremer, G.Z. Endoplasmic reticulum-associated degradation of mammalian glycoproteins involves sugar chain trimming to Man6-5GlcNAc2. *J Biol Chem* **278**, 34119-34124 (2003).
69. Olivari, S., Galli, C., Alanen, H., Ruddock, L. & Molinari, M. A novel stress-induced EDEM variant regulating endoplasmic reticulum-associated glycoprotein degradation. *J Biol Chem* **280**, 2424-2428 (2005).
70. Mast, S.W. et al. Human EDEM2, a novel homolog of family 47 glycosidases, is involved in ER-associated degradation of glycoproteins. *Glycobiology* **15**, 421-436 (2005).
71. Meusser, B., Hirsch, C., Jarosch, E. & Sommer, T. ERAD: the long road to destruction. *Nat Cell Biol* **7**, 766-772 (2005).
72. Ikezawa, H. Glycosylphosphatidylinositol (GPI)-anchored proteins. *Biol Pharm Bull* **25**, 409-417 (2002).
73. Kuehn, M.J., Herrmann, J.M. & Schekman, R. COPII-cargo interactions direct protein sorting into ER-derived transport vesicles. *Nature* **391**, 187-190 (1998).
74. Nishimura, N. & Balch, W.E. A di-acidic signal required for selective export from the endoplasmic reticulum. *Science* **277**, 556-558 (1997).
75. Ma, D. et al. Role of ER export signals in controlling surface potassium channel numbers. *Science* **291**, 316-319 (2001).
76. Malkus, P., Jiang, F. & Schekman, R. Concentrative sorting of secretory cargo proteins into COPII-coated vesicles. *J Cell Biol* **159**, 915-921 (2002).
77. Votsmeier, C. & Gallwitz, D. An acidic sequence of a putative yeast Golgi membrane protein binds COPII and facilitates ER export. *Embo J* **20**, 6742-6750 (2001).
78. Kappeler, F., Klopfenstein, D.R., Foguet, M., Paccaud, J.P. & Hauri, H.P. The recycling of ERGIC-53 in the early secretory pathway. ERGIC-53 carries a cytosolic endoplasmic reticulum-exit determinant interacting with COPII. *J Biol Chem* **272**, 31801-31808 (1997).
79. Nufer, O., Kappeler, F., Guldbrandsen, S. & Hauri, H.P. ER export of ERGIC-53 is controlled by cooperation of targeting determinants in all three of its domains. *J Cell Sci* **116**, 4429-4440 (2003).
80. Sato, K. & Nakano, A. Oligomerization of a cargo receptor directs protein sorting into COPII-coated transport vesicles. *Mol Biol Cell* **14**, 3055-3063 (2003).
81. Dominguez, M. et al. gp25L/emp24/p24 protein family members of the cis-Golgi network bind both COP I and II coatomer. *J Cell Biol* **140**, 751-765 (1998).
82. Otte, S. & Barlowe, C. The Erv41p-Erv46p complex: multiple export signals are required in trans for COPII-dependent transport from the ER. *Embo J* **21**, 6095-6104 (2002).
83. Wieland, F.T., Gleason, M.L., Serafini, T.A. & Rothman, J.E. The rate of bulk flow from the endoplasmic reticulum to the cell surface. *Cell* **50**, 289-300 (1987).
84. Otte, S. & Barlowe, C. Sorting signals can direct receptor-mediated export of soluble proteins into COPII vesicles. *Nat Cell Biol* **6**, 1189-1194 (2004).
85. Caldwell, S.R., Hill, K.J. & Cooper, A.A. Degradation of endoplasmic reticulum (ER) quality control substrates requires transport between the ER and Golgi. *J Biol Chem* **276**, 23296-23303 (2001).
86. Appenzeller, C., Andersson, H., Kappeler, F. & Hauri, H.P. The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. *Nat Cell Biol* **1**, 330-334 (1999).

87. Cunningham, M.A. et al. LMAN1 is a molecular chaperone for the secretion of coagulation factor VIII. *J Thromb Haemost* **1**, 2360-2367 (2003).
88. Vollenweider, F., Kappeler, F., Itin, C. & Hauri, H.P. Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme. *J Cell Biol* **142**, 377-389 (1998).
89. Zhang, B., Kaufman, R.J. & Ginsburg, D. LMAN1 and MCFD2 form a cargo receptor complex and interact with coagulation factor VIII in the early secretory pathway. *J Biol Chem* **280**, 25881-25886 (2005).
90. Kappeler, F., Itin, C., Schindler, R. & Hauri, H.P. A dual role for COOH-terminal lysine residues in pre-Golgi retention and endocytosis of ERGIC-53. *J Biol Chem* **269**, 6279-6281 (1994).
91. Itin, C., Roche, A.C., Monsigny, M. & Hauri, H.P. ERGIC-53 is a functional mannose-selective and calcium-dependent human homologue of leguminous lectins. *Mol Biol Cell* **7**, 483-493 (1996).
92. Belden, W.J. & Barlowe, C. Role of Erv29p in collecting soluble secretory proteins into ER-derived transport vesicles. *Science* **294**, 1528-1531 (2001).
93. Lander, E.S. et al. Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).
94. ConsortiumInternational, H.G.S. Finishing the euchromatic sequence of the human genome. **431**, 931-945 (2004).
95. Auerbach, D., Thaminy, S., Hottiger, M.O. & Stagljar, I. The post-genomic era of interactive proteomics: facts and perspectives. *Proteomics* **2**, 611-623 (2002).
96. Oliver, S. Guilt-by-association goes global. *Nature* **403**, 601-603 (2000).
97. Gavin, A.C. et al. Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**, 631-636 (2006).
98. Krogan, N.J. et al. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* **440**, 637-643 (2006).
99. Phizicky, E., Bastiaens, P.I., Zhu, H., Snyder, M. & Fields, S. Protein analysis on a proteomic scale. *Nature* **422**, 208-215 (2003).
100. Piehler, J. New methodologies for measuring protein interactions in vivo and in vitro. *Curr Opin Struct Biol* **15**, 4-14 (2005).
101. Monti, M., Orru, S., Pagnozzi, D. & Pucci, P. Interaction proteomics. *Biosci Rep* **25**, 45-56 (2005).
102. Rigaut, G. et al. A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**, 1030-1032 (1999).
103. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198-207 (2003).
104. Gavin, A.C. et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141-147 (2002).
105. Ho, Y. et al. Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* **415**, 180-183 (2002).
106. Bader, G.D. & Hogue, C.W. Analyzing yeast protein-protein interaction data obtained from different sources. *Nat Biotechnol* **20**, 991-997 (2002).
107. Grelle, G. et al. Identification of VCP/p97, carboxyl terminus of Hsp70-interacting protein (CHIP), and amphiphysin II interaction partners using membrane-based human proteome arrays. *Mol Cell Proteomics* **5**, 234-244 (2006).
108. Mahlknecht, U., Ottmann, O.G. & Hoelzer, D. Far-Western based protein-protein interaction screening of high-density protein filter arrays. *J Biotechnol* **88**, 89-94 (2001).
109. de Graaf, K. et al. Characterization of cyclin L2, a novel cyclin with an arginine/serine-rich domain: phosphorylation by DYRK1A and colocalization with splicing factors. *J Biol Chem* **279**, 4612-4624 (2004).
110. Zhu, H. et al. Analysis of yeast protein kinases using protein chips. *Nat Genet* **26**, 283-289 (2000).
111. MacBeath, G. & Schreiber, S.L. Printing proteins as microarrays for high-throughput function determination. *Science* **289**, 1760-1763 (2000).
112. Zhu, H. et al. Global analysis of protein activities using proteome chips. *Science* **293**, 2101-2105 (2001).
113. Jones, R.B., Gordus, A., Krall, J.A. & MacBeath, G. A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature* **439**, 168-174 (2006).
114. Zhu, H. & Snyder, M. Protein chip technology. *Curr Opin Chem Biol* **7**, 55-63 (2003).

115. Smothers, J.F., Henikoff, S. & Carter, P. Phage display. Affinity selection from biological libraries. *Science* **298**, 621-622 (2002).
116. Rhyner, C., Kodzius, R. & Crameri, R. Direct selection of cDNAs from filamentous phage surface display libraries: potential and limitations. *Curr Pharm Biotechnol* **3**, 13-21 (2002).
117. Fields, S. & Song, O. A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-246 (1989).
118. Ito, T. et al. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A* **98**, 4569-4574 (2001).
119. Uetz, P. et al. A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* **403**, 623-627 (2000).
120. Li, S. et al. A map of the interactome network of the metazoan C. elegans. *Science* **303**, 540-543 (2004).
121. Giot, L. et al. A protein interaction map of Drosophila melanogaster. *Science* **302**, 1727-1736 (2003).
122. Stelzl, U. et al. A human protein-protein interaction network: a resource for annotating the proteome. *Cell* **122**, 957-968 (2005).
123. von Mering, C. et al. Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* **417**, 399-403 (2002).
124. Fashena, S.J., Serebriiskii, I. & Golemis, E.A. The continued evolution of two-hybrid screening approaches in yeast: how to outwit different preys with different baits. *Gene* **250**, 1-14 (2000).
125. Lee, J.W. & Lee, S.K. Mammalian two-hybrid assay for detecting protein-protein interactions in vivo. *Methods Mol Biol* **261**, 327-336 (2004).
126. Johnsson, N. & Varshavsky, A. Split ubiquitin as a sensor of protein interactions in vivo. *Proc Natl Acad Sci U S A* **91**, 10340-10344 (1994).
127. Miller, J.P. et al. Large-scale identification of yeast integral membrane protein interactions. *Proc Natl Acad Sci U S A* **102**, 12123-12128 (2005).
128. Stagljar, I., Korostensky, C., Johnsson, N. & te Heesen, S. A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Proc Natl Acad Sci U S A* **95**, 5187-5192 (1998).
129. Galarneau, A., Primeau, M., Trudeau, L.E. & Michnick, S.W. Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. *Nat Biotechnol* **20**, 619-622 (2002).
130. Remy, I. & Michnick, S.W. Visualization of biochemical networks in living cells. *Proc Natl Acad Sci U S A* **98**, 7678-7683 (2001).
131. Paulmurugan, R. & Gambhir, S.S. Monitoring protein-protein interactions using split synthetic renilla luciferase protein-fragment-assisted complementation. *Anal Chem* **75**, 1584-1589 (2003).
132. Luker, K.E. et al. Kinetics of regulated protein-protein interactions revealed with firefly luciferase complementation imaging in cells and living animals. *Proc Natl Acad Sci U S A* **101**, 12288-12293 (2004).
133. Hu, C.D., Chinenov, Y. & Kerppola, T.K. Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol Cell* **9**, 789-798 (2002).
134. Jach, G., Pesch, M., Richter, K., Frings, S. & Uhrig, J.F. An improved mRFP1 adds red to bimolecular fluorescence complementation. *Nat Methods* **3**, 597-600 (2006).
135. Tsien, R.Y. The green fluorescent protein. *Annu Rev Biochem* **67**, 509-544 (1998).
136. Chen, C.D., Oh, S.Y., Hinman, J.D. & Abraham, C.R. Visualization of APP dimerization and APP-Notch2 heterodimerization in living cells using bimolecular fluorescence complementation. *J Neurochem* **97**, 30-43 (2006).
137. Kerppola, T.K. Visualization of molecular interactions by fluorescence complementation. *Nat Rev Mol Cell Biol* **7**, 449-456 (2006).
138. von der Lehr, N. et al. The F-box protein Skp2 participates in c-Myc proteosomal degradation and acts as a cofactor for c-Myc-regulated transcription. *Mol Cell* **11**, 1189-1200 (2003).
139. Remy, I., Montmarquette, A. & Michnick, S.W. PKB/Akt modulates TGF-beta signalling through a direct interaction with Smad3. *Nat Cell Biol* **6**, 358-365 (2004).
140. Remy, I. & Michnick, S.W. A cDNA library functional screening strategy based on fluorescent protein complementation assays to identify novel components of signaling pathways. *Methods* **32**, 381-388 (2004).
141. Remy, I. & Michnick, S.W. Regulation of apoptosis by the Ft1 protein, a new modulator of protein kinase B/Akt. *Mol Cell Biol* **24**, 1493-1504 (2004).
142. Valencia, A. & Pazos, F. Computational methods for the prediction of protein interactions. *Curr Opin Struct Biol* **12**, 368-373 (2002).

143. Jansen, R. et al. A Bayesian networks approach for predicting protein-protein interactions from genomic data. *Science* **302**, 449-453 (2003).
144. Eddy, S.R. What is Bayesian statistics? *Nat Biotechnol* **22**, 1177-1178 (2004).
145. Rhodes, D.R. et al. Probabilistic model of the human protein-protein interaction network. *Nat Biotechnol* **23**, 951-959 (2005).
146. Han, K., Park, B., Kim, H., Hong, J. & Park, J. HPID: the Human Protein Interaction Database. *Bioinformatics* **20**, 2466-2470 (2004).
147. Csank, C. et al. Three yeast proteome databases: YPD, PombePD, and CalPD (MycoPathPD). *Methods Enzymol* **350**, 347-373 (2002).
148. Xenarios, I. et al. DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. *Nucleic Acids Res* **30**, 303-305 (2002).
149. Hwang, C., Sinskey, A.J. & Lodish, H.F. Oxidized redox state of glutathione in the endoplasmic reticulum. *Science* **257**, 1496-1502 (1992).
150. Bygrave, F.L. & Benedetti, A. What is the concentration of calcium ions in the endoplasmic reticulum? *Cell Calcium* **19**, 547-551 (1996).
151. Marquardt, T., Hebert, D.N. & Helenius, A. Post-translational folding of influenza hemagglutinin in isolated endoplasmic reticulum-derived microsomes. *J Biol Chem* **268**, 19618-19625 (1993).
152. Carvalho, P., Goder, V. & Rapoport, T.A. Distinct Ubiquitin-Ligase Complexes Define Convergent Pathways for the Degradation of ER Proteins. *Cell* **126**, 361-373 (2006).
153. Sinz, A. Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *J Mass Spectrom* **38**, 1225-1237 (2003).
154. Ratner, D.M. et al. Probing protein-carbohydrate interactions with microarrays of synthetic oligosaccharides. *Chembiochem* **5**, 379-382 (2004).
155. Chung, K.T., Shen, Y. & Hendershot, L.M. BAP, a mammalian BiP-associated protein, is a nucleotide exchange factor that regulates the ATPase activity of BiP. *J Biol Chem* **277**, 47557-47563 (2002).
156. Ozalp, C., Szczesna-Skorupa, E. & Kemper, B. Bimolecular fluorescence complementation analysis of cytochrome p450 2c2, 2e1, and NADPH-cytochrome p450 reductase molecular interactions in living cells. *Drug Metab Dispos* **33**, 1382-1390 (2005).

2. Aim of the thesis

2.1 PCA-based analysis of luminal protein interactions

The analysis of luminal protein interactions by protein fragment complementation holds the promise to uncover novel interactions missed by other methods. So far, β-lactamase, dihydrofolate reductase, *Renilla* and firefly luciferase as well as fluorescent proteins were successfully applied as PCA reporters for the detection of cytosolic protein interactions. Can any of these PCA strategies be adopted to probe for protein interactions in the lumen of the secretory pathway? Since the yellow fluorescent protein (YFP) can be readily detected in the luminal environment of the ER, a YFP-based PCA approach was examined in this thesis regarding its potential to detect luminal protein interactions. ERGIC-53 and its interacting proteins MCFD2, cathepsin Z and cathepsin C were used as a proof of concept to test the feasibility and specificity of the method. Furthermore, a YFP PCA-based cDNA library screening strategy was established to identify interaction partners of ERGIC-53 in a genomewide setting.

2.2 Characterization of the ERGIC-53/MCFD2 complex

The identification of MCFD2 as a luminal interaction partner of ERGIC-53 led to the notion of cargo receptor complexes in the field of receptor-mediated ER export. Why do we need a cargo receptor complex? What is the function of MCFD2? Genetic and biochemical data show that MCFD2 is required for the secretion of blood coagulation factors V and VIII. It remains unknown, however, if MCFD2 functions as a specific factor for these proteins or has a more general role in the binding of cargo glycoproteins to ERGIC-53. Work in this thesis addressed the question of cargo selectivity of the ERGIC-53/MCFD2 receptor complex by combining cellular depletion of MCFD2 with YFP PCA-based quantification of the cargo binding properties of ERGIC-53. Moreover, the ERGIC-53/MCFD2 protein complex was dissected in regard of its localization and subunit interdependence. For this purpose, immunofluorescence-based protein localization studies were combined with short interference RNA (siRNA)-mediated protein knockdowns.

3. Results

3.1 Capturing protein interactions in the secretory pathway of living cells

Manuscript published in Proceedings of the National Academy of Sciences of the United States of America. 2005 May 3; 102(18): 6350–6355.

An online version is available at:
http://www.pnas.org/cgi/content/full/102/18/6350

Capturing protein interactions in the secretory pathway of living cells

Beat Nyfeler*, Stephen W. Michnick†, and Hans-Peter Hauri*‡

*Department of Pharmacology and Neurobiology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland; and †Département de Biochimie, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, PQ, Canada H3C 3J7

Communicated by Marilyn Gist Farquhar, University of California at San Diego, La Jolla, CA, March 10, 2005 (received for review January 12, 2005)

The secretory pathway is composed of membrane compartments specialized in protein folding, modification, transport, and sorting. Numerous transient protein–protein interactions guide the transport-competent proteins through the secretory pathway. Here we have adapted the yellow fluorescent protein (YFP)-based protein fragment complementation assay (PCA) to detect protein–protein interactions in the secretory pathway of living cells. Fragments of YFP were fused to the homooligomeric cargo-receptor lectin endoplasmic reticulum Golgi intermediate compartment (ERGIC)-53, to the ERGIC-53-interacting multicoagulation factor deficiency protein MCFD2, and to ERGIC-53's cargo glycoprotein cathepsin Z. YFP PCA analysis revealed the oligomerization of ERGIC-53 and its interaction with MCFD2, as well as its lectin-mediated interaction with cathepsin Z. Mutation of the lectin domain of ERGIC-53 selectively decreased YFP complementation with cathepsin Z. Using YFP PCA, we discovered a carbohydrate-mediated interaction between ERGIC-53 and cathepsin C. We conclude that YFP PCA can detect weak and transient protein interactions in the secretory pathway and hence is a powerful approach to study luminal processes involved in protein secretion. The study extends the application of PCA to carbohydrate-mediated protein–protein interactions of low affinity.

ERGIC-53 | lectin cargo receptor | protein fragment complementation assay | protein–protein interaction

Eukaryotic cells have evolved a secretory pathway that is composed of characteristic membrane compartments, including the endoplasmic reticulum (ER), the ER-Golgi intermediate compartment (ERGIC), and the Golgi apparatus. Approximately one-third of all cellular proteins are translocated into the lumen of the ER, where modification, folding, and oligomerization occur, before proteins are further transported along the secretory pathway. The folding and modification processes involve numerous ER resident proteins that are believed to operate as a quality control machinery that surveys correct folding in the ER (1, 2). After acquisition of transport competence, the secretory proteins exit the ER by a receptor-mediated mechanism (3, 4) or by bulk flow. The interactions between proteins of the ER quality control machinery and their substrates, as well as between cargo receptors and their cargo, are often of a weak and transient nature and therefore difficult to study. Traditional techniques for studying protein–protein interactions, such as yeast two-hybrid assays, may not be adequate to reveal interactions among these proteins, given that the yeast two-hybrid approach identifies interactions in a reducing (cytoplasm and nucleus) rather than oxidizing (ER) environment. We therefore explored the possibility of adapting the protein fragment complementation assay (PCA) (5) to studying protein interactions in the secretory pathway.

The basic concept of PCA relies on engineering reporter protein fragments that exhibit no functional activity by themselves and do not spontaneously fold. The fragments are fused to two interacting proteins. The interaction of the hybrid proteins brings the two reporter fragments into proximity, where they fold into the active 3D structure of the complete reporter protein. PCA has been described by using a variety of proteins, including β-lactamase, dihydrofolate reductase, *Renilla* and firefly luciferases, and GFP and yellow fluorescent protein (YFP) as reporters (6–13). The GFP and YFP PCAs have proven particularly simple for detection and library screening of cytosolic, membrane, and nuclear protein–protein and protein–RNA interactions (13–16).

In the present study, we have explored the suitability of the YFP-based PCA technique for visualizing protein–protein interactions in the lumen of the early secretory pathway by using ERGIC-53 as model protein. ERGIC-53 is a homooligomeric nonglycosylated type I transmembrane protein cycling in the early secretory pathway (17, 18). It contains a functional lectin domain (19) and acts as a cargo receptor for a subset of glycoproteins, including blood coagulation factors V and VIII (20), cathepsin Z (catZ) (3), and presumably cathepsin C (catC) (21). Chemical crosslinking revealed a direct and carbohydrate-dependent interaction of the lysosomal protein catZ with ERGIC-53 (3). catC is also a lysosomal protein, related to catZ. The secretion of catC is reduced if a dominant-negative ER-locked form of ERGIC-53 is expressed, but all attempts to show a direct interaction between catC and ERGIC-53 have failed (21). Recently, MCFD2 was identified as an ERGIC-53 interacting protein (22). MCFD2 is a nonglycosylated luminal protein, which contains two EF-hands and binds ERGIC-53 in a calcium-dependent manner, but its precise role in cargo transport is unknown.

In the current study, we have adopted the YFP PCA to the secretory pathway and visualized the oligomerization of ERGIC-53 as well as its interaction with catZ and MCFD2. The power of YFP PCA is demonstrated by the detection of the lectin-mediated interaction between ERGIC-53 and catC. We anticipate that YFP PCA will be a valuable approach for studying the different processes involved in protein secretion and the basis for the first robust screens of protein–protein interactions in the secretory pathway.

Methods

Antibodies. The following antibodies were used: mouse mAb G1/93 against human ERGIC-53 (17), mouse mAb A1/182 against BAP31 (18), rabbit pAb against the KDEL receptor (23) (kind gift of H.-D. Söling, Max-Planck-Institut für Biophysikalische Chemie, Göttingen, Germany), mouse mAb G1/133 against giantin (24), and mouse mAb against GFP (Roche Applied Science). Note that anti-GFP recognizes YFP fragment 2 but not fragment 1.

DNA Constructs. The construction of pcDNA3 vectors (Invitrogen) containing the sequences of YFP fragments 1 (YFP1; amino acids 1–158) and 2 (YFP2; amino acids 159–239) was

Freely available online through the PNAS open access option.

Abbreviations: catC, cathepsin C; catZ, cathepsin Z; ER, endoplasmic reticulum; ERGIC, ER-Golgi intermediate compartment; PCA, protein fragment complementation assay; YFP, yellow fluorescent protein.

‡To whom correspondence should be addressed. E-mail: hans-peter.hauri@unibas.ch.

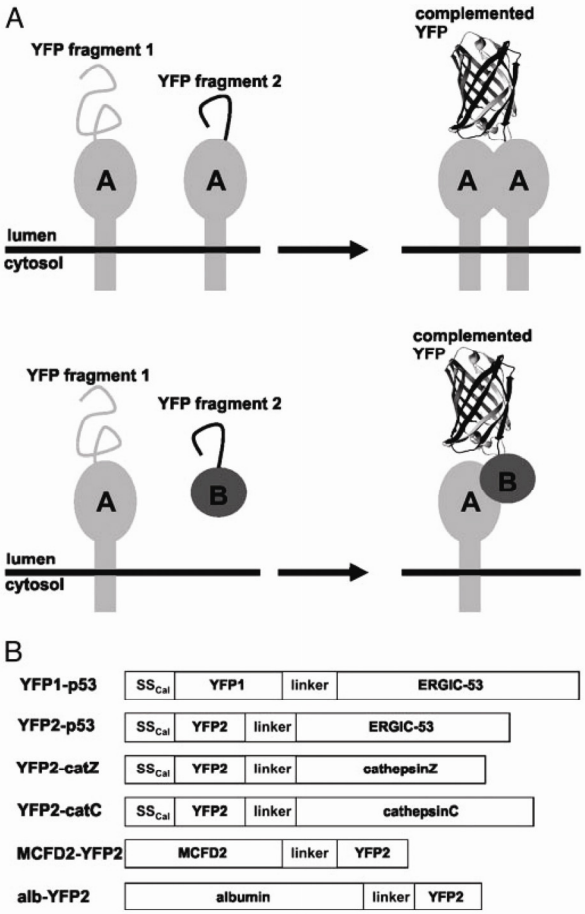


Fig. 1. YFP fragment complementation in the lumen of the secretory pathway. (*A*) Principle of the YFP PCA used in the current study. YFP fragments 1 (amino acids 1–158) and 2 (amino acids 159–239) are fused to the luminal part of a given transmembrane protein A or to a soluble luminal protein B. Homooligomerization of the membrane protein A or its interaction with B brings the two fragments of YFP into close proximity and leads to complementation into functional fluorescent YFP by folding into an active 3D structure. (*B*) Modular composition of the fusion proteins used in this study. ERGIC-53 (p53) served as an oligomeric transmembrane cargo receptor. catZ, catC, MCFD2, and albumin were used as soluble luminal proteins. For the N-terminal fusion of the YFP fragments, the endogenous signal sequence of the proteins was replaced by the signal sequence of calreticulin (SS$_{cal}$).

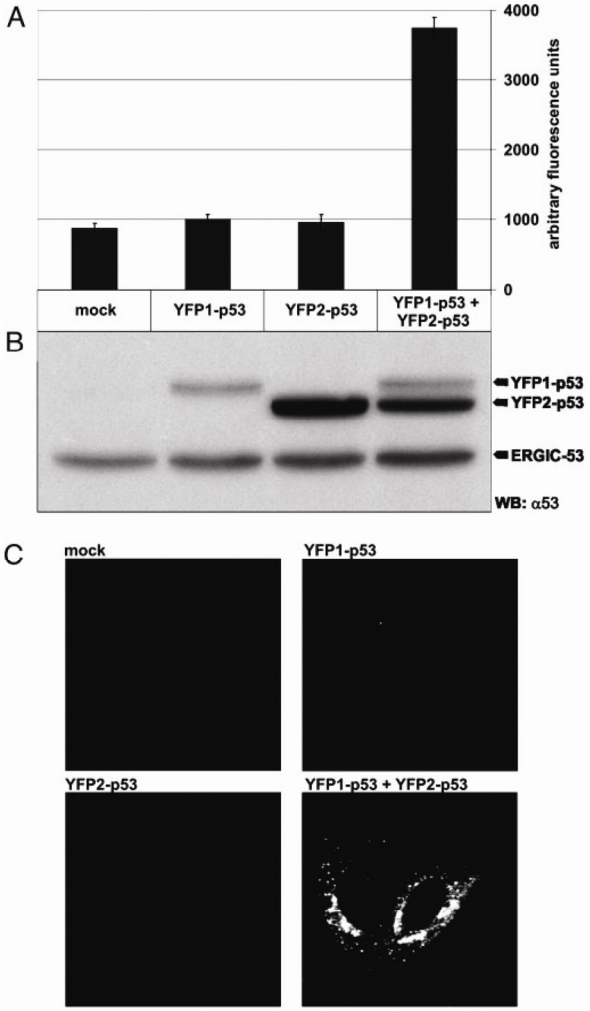


Fig. 2. Visualization of ERGIC-53 oligomerization by YFP PCA. (*A*) YFP fragment complementation was detected by fluorometric analysis by using microtiter plates with cell suspensions of HeLa cells expressing the indicated fusion proteins. Coexpression of YFP1-p53 and YFP2-p53 resulted in YFP fluorescence, whereas the expression of the single fusion proteins led to no detectable YFP fluorescence. (*B*) Immunoblot analysis using anti-ERGIC-53 verifies expression of the fusion proteins YFP1-p53 and YFP2-p53. Equal protein amounts are present in all lanes, as revealed by similar levels of endogenous ERGIC-53. (*C*) Visualization of the oligomerization of YFP1-p53 and YFP2-p53 in the early secretory pathway of HeLa cells by fluorescence microscopy of live cells. Note that the YFP fragments show no fluorescence signal when transfected individually.

described previously (13). The sequence coding for the 10-aa $(GGGGS)_2$ linker was introduced in 5′ or 3′ of the YFP fragments (13). cDNAs encoding MCFD2 and albumin were amplified by PCR and subcloned 5′ of the linker-YFP2 sequence (resulting in fusion proteins with YFP fragments at the C terminus). ERGIC-53 cDNA had to be subcloned 3′ of linker-YFP fragments, because its C terminus is cytosolic. For 3′ subcloning, the DNA sequence coding for the signal sequence of calreticulin (SS$_{cal}$) was introduced into pcDNA3 by ligation of annealed phosphorylated oligonucleotides. The efficient signal sequence of calreticulin was used, because the endogenous signal sequence of ERGIC-53 was inefficient in ER translocation of recombinant ERGIC-53 fusion proteins. DNA sequences of YFP1- and YFP2-linker were amplified by PCR and subcloned 3′ of SS$_{cal}$. cDNAs encoding ERGIC-53, catZ, and catC lacking the signal sequence were amplified by PCR and subcloned 3′ of the SS$_{cal}$-YFP fragment linker. The composition of all recombinant fusion proteins used in this study is shown in Fig. 1*B*. YFP1 containing the Q69M mutation (25), as well as ERGIC-53 containing the N156A mutation, was generated by introducing point mutations with the QuikChange site-directed mutagenesis kit (Stratagene).

Cell Culture and DNA Transfection. HeLa cells were grown in DMEM, supplemented with 10% FBS, 1× nonessential amino acids and antibiotics. For fluorometric analysis and metabolic labeling, HeLa cells were grown in six-well plates. For fluorescence microscopy, HeLa cells were grown on poly(L-lysine)-

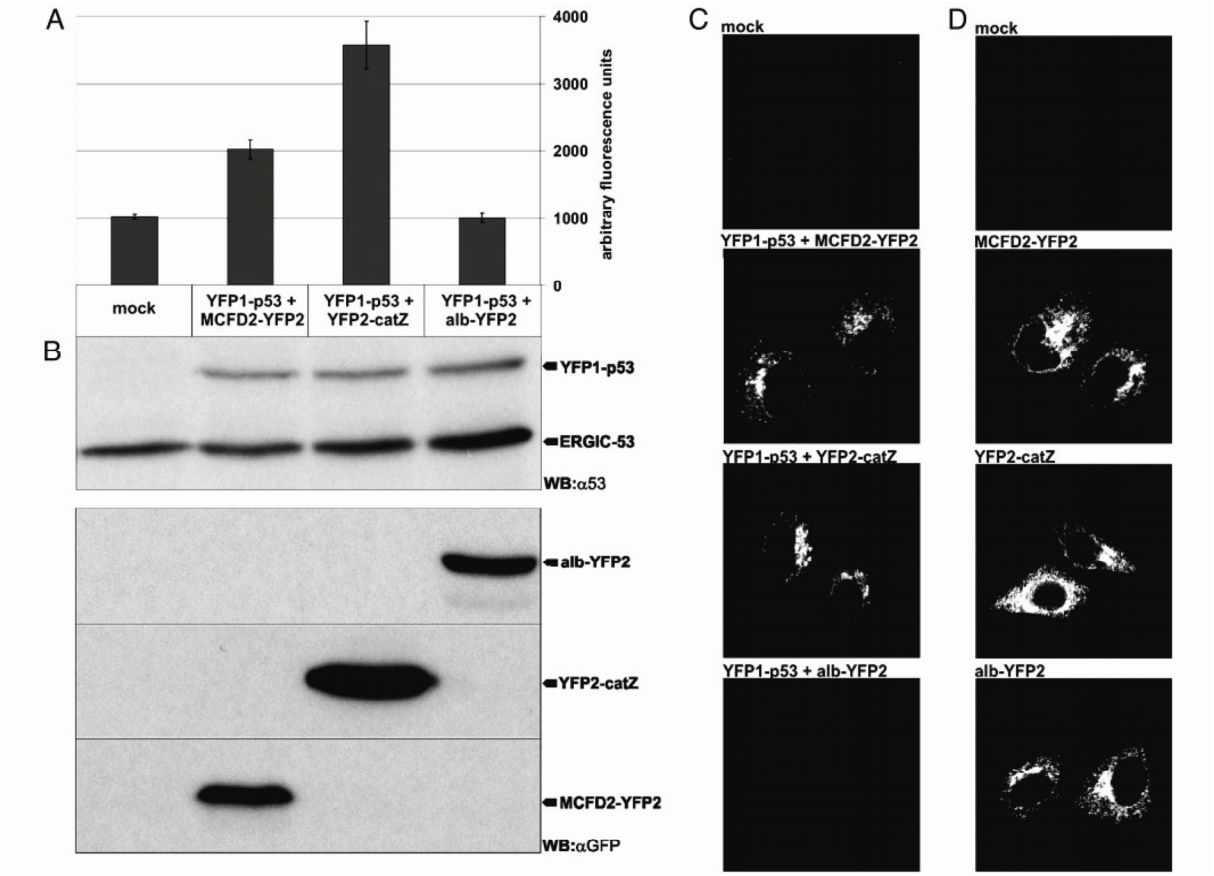


Fig. 3. YFP PCA can specifically detect interactions of ERGIC-53 with MCFD2 and catZ. (*A*) YFP fragment complementation indicates interactions of ERGIC-53 with MCFD2 and catZ but not with albumin. Fluorescence was measured by fluorometry of cell suspensions in microtiter plates. (*B*) Expression of the fusion proteins was visualized by immunoblotting with anti-ERGIC-53 and anti-GFP. (*C*) Fluorescence microscopy of live cells expressing the indicated fusion proteins. (*D*) Immunofluorescence using anti-GFP confirmed the expression of MCFD2-YFP2, YFP2-catZ, and alb-YFP2 in the early secretory pathway.

coated glass slides. Cells were transfected at ≈60% confluence by using FuGENE6 (Roche Applied Science) according to the manufacturer's instructions.

YFP Fluorometric Analysis. Twenty-four hours after transfection, cells were washed with PBS, harvested by trypsinization, and resuspended in 1 ml of PBS. Cells were then pelleted by centrifugation, resuspended in 200 μl of PBS, transferred to black 96-well microtiter plates (Nunc), and subjected to fluorometric analysis by using a Victor2 fluorometer (PerkinElmer). Excitation and emission wavelengths of 485 and 535 nm, respectively, were used. Data from three independent experiments were averaged, and error bars indicate standard deviations. For fluorescence microscopy, cells were washed twice with PBS and mounted under a glass coverslip. Live cells were analyzed by laser-scanning confocal microscopy (TCS NT, Leica, Deerfield, IL). Representative images of single optical sections are shown.

Immunofluorescence. Twenty-four hours after transfection, HeLa cells were fixed in 3% paraformaldehyde, permeabilized in PBS containing 0.1% saponin, and incubated with the primary antibody. For anti-GFP immunofluorescence, cells were stained with Alexa Fluor 488-conjugated goat-anti-mouse IgG (Molecular Probes). For colocalization studies, cells were stained with Alexa Fluor 568-conjugated goat-anti mouse IgG or goat-anti rabbit IgG (Molecular Probes) and analyzed by using confocal microscopy.

Immunoblot Analysis. Protein samples were prepared from cells used for fluorometric analysis in microtiter plates. Protein samples were separated by SDS/PAGE, transferred to nitrocellulose membranes, immunoblotted with anti-ERGIC-53 and anti-GFP, and visualized by enhanced chemiluminescence (Amersham Pharmacia Biosciences).

Metabolic Labeling. Twenty-four hours after transfection, cells were pulsed for 15 min with 100 μCi (1 Ci = 37 GBq) of [^{35}S]methionine (PerkinElmer) and chased for the indicated times in HeLa culture medium containing 10 mM L-methionine. Cells were lysed in 1% Nonidet P-40/50 mM Tris·HCl, pH 7.5/150 mM NaCl/PMSF, and the lysate was cleared by centrifugation at 100,000 × *g* for 1 h. The chase medium was cleared from cell debris by centrifugation at 20,000 × *g* for 10 min. Cleared samples were immunoprecipitated with anti-ERGIC-53 and -GFP. Immunoprecipitates were separated by SDS/PAGE and proteins visualized by fluorography.

Results

The principle of YFP PCA used in the current study is illustrated in Fig. 1*A*. To optimize YFP as a reporter for studying protein–

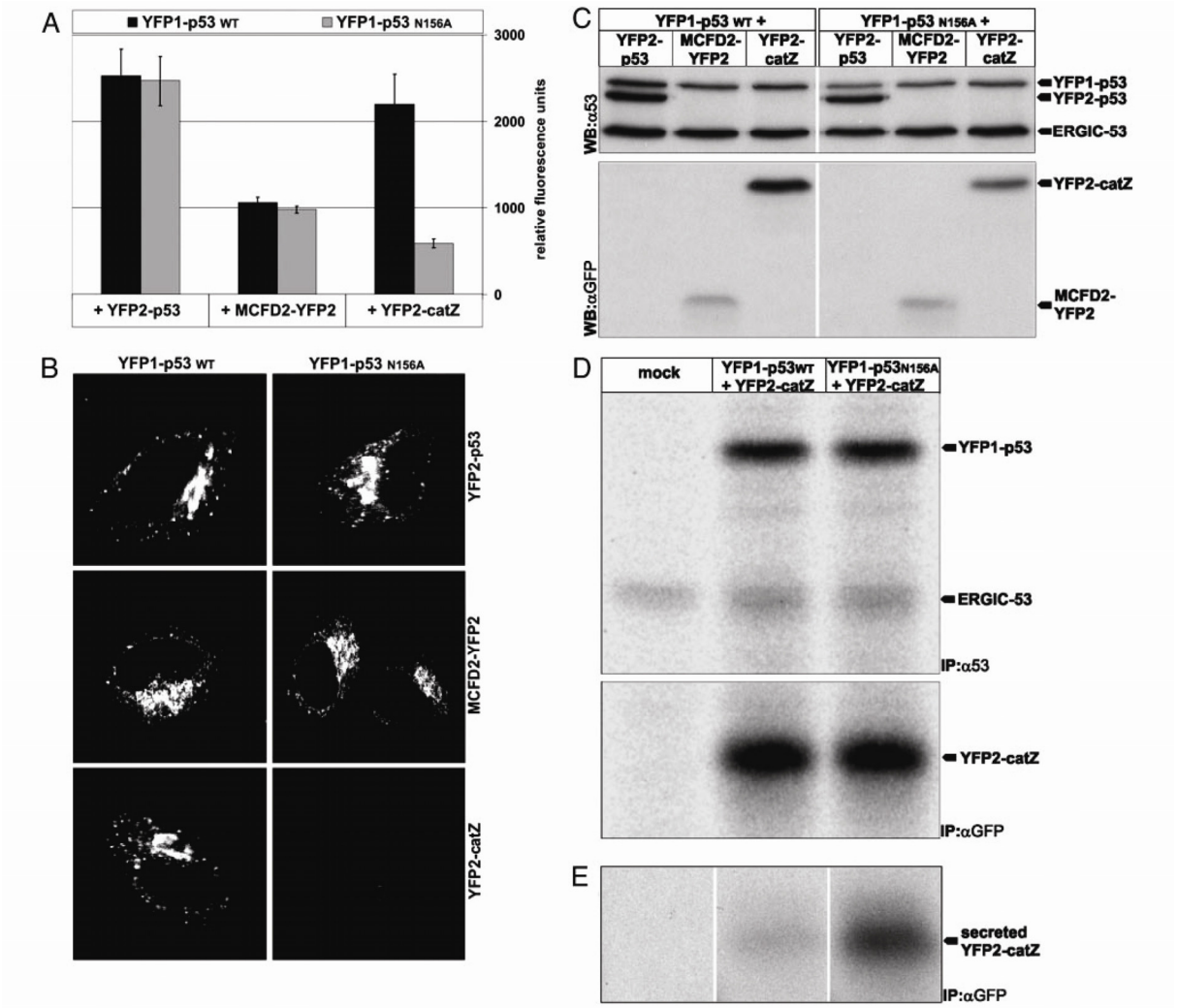


Fig. 4. Inactivation of the lectin activity of ERGIC-53 selectively impairs lectin-mediated interaction with catZ. (*A*) Fluorometric analysis of cell suspensions in microtiter plates shows that inactivation of the carbohydrate recognition domain of ERGIC-53 by the N156A mutation selectively impaired the lectin-mediated interaction with catZ. ERGIC-53 oligomerization and its interaction with MCFD2 are not affected. Relative fluorescence values are shown in which the background fluorescence of mock transfected cells was subtracted. (*B*) Fluorescence microscopy of living HeLa cells expressing the indicated fusion proteins. (*C*) Expression of all fusion proteins was analyzed by immunoblotting with anti-ERGIC-53 and anti-GFP. (*D*) A 15-min [^{35}S]methionine pulse followed by immunoprecipitation with anti-ERGIC-53 and anti-GFP shows equal rates of synthesis of YFP2-catZ cotransfected with either YFP1-p53$_{WT}$ or YFP1-p53$_{N156A}$ (fluorogram). (*E*) Transfected HeLa cells were pulsed for 15 min with [^{35}S]methionine and chased for 1 h. Anti-GFP immunoprecipitation of the chase medium reveals reduced secretion of YFP2-catZ cotransfected with YFP1-p53$_{WT}$ as compared with YFP-p53$_{N156A}$ (fluorogram).

protein interactions in the secretory pathway, its glutamine residue at position 69 was replaced by a methionine. YFP Q69M, known as citrine, shows improved photostability and expression in different cellular compartments, including the secretory pathway (25). To test whether the YFP PCA reveals protein–protein interactions in the lumen of the secretory pathway, we first studied the homooligomerization of the type I transmembrane protein ERGIC-53. YFP fragments 1 and 2 were fused to the luminal N terminus of ERGIC-53, resulting in YFP1-p53 and YFP2-p53 (Fig. 1*B*). Individual expression of either YFP1-p53 or YFP2-p53 alone gave no detectable YFP fluorescence in living cells, indicating that the YFP fragments *per se* have no intrinsic fluorescence if expressed in the secretory pathway. Strong YFP fluorescence was observed, however, when YFP1-p53 and YFP2-p53 were coexpressed, which allowed for YFP fragment complementation triggered by ERGIC-53 oligomerization (Fig. 2*A*). The fluorescence pattern (Fig. 2*C*) observed for complemented YFP in live cells is typical for the early secretory pathway (26). Double-labeling experiments showed overlap with the ER marker BAP31, the ERGIC- and cis-Golgi-localized KDEL-receptor, and the cis-medial Golgi marker giantin (Fig. 6, which is published as supporting information on the PNAS web site). We conclude that YFP PCA can detect homooligomerization of membrane proteins in the secretory pathway.

We next determined whether YFP PCA can detect the protein-protein-mediated interaction between ERGIC-53 and MCFD2 as well as the protein-carbohydrate-mediated interaction between ERGIC-53 and its glycoprotein cargo catZ. As a

negative control, we included the nonglycosylated protein albumin, the secretion of which does not depend on ERGIC-53. YFP fragment 2 was fused to the C terminus of MCFD2 and albumin as well as to the N terminus of catZ (Fig. 1*B*). C-terminal tagging of catZ was not possible, because it interferes with ERGIC-53 binding (C. Appenzeller and H.-P.H., unpublished data). Fluorometric analysis detected YFP fragment complementation if YFP1-p53 was coexpressed with MCFD2-YFP2 or YFP2-catZ (Fig. 3*A*). In both cases, the YFP fluorescence was localized to the early secretory pathway (Fig. 3*C*). Notably, albumin-YFP2 did not induce YFP fragment complementation with YFP1-p53, although it was correctly expressed at similar levels as MCFD2-YFP2 (Fig. 3 *B* and *D*). The results demonstrate the selectivity and specificity of YFP PCA to detect luminal protein–protein interactions in the secretory pathway. Moreover, coexpression of YFP1-p53 with MCFD2-YFP2 showed similar fluorescence as coexpression of YFP2-p53 with MCFD2-YFP1 (data not shown), indicating exchangeability of the YFP fragments.

To further investigate the sensitivity and specificity of YFP PCA, we studied the effect of inactivation of the lectin domain of ERGIC-53 by the N156A mutation. This point mutation abolishes binding of ERGIC-53 to mannose beads (19) as well as to catZ (3). Fluorometric analysis revealed that the N156A mutation specifically decreased the lectin-dependent interaction of ERGIC-53 with catZ, whereas ERGIC-53 oligomerization, as well as its interaction with MCFD2, was not affected (Fig. 4*A*). The fluorometric results were confirmed by fluorescence microscopy of live cells (Fig. 4*B*). Immunoblot analysis showed higher steady-state levels of YFP2-catZ when coexpressed with YFP1-p53$_{WT}$ than with YFP1-p53$_{N156A}$ (Fig. 4*C*). A possible explanation for this difference may be decreased dissociation of YFP2-catZ from YFP1-p53$_{WT}$ after complementation of the YFP fragments, which would result in intracellular accumulation of YFP2-catZ. Stabilization of the interaction between fusion proteins by YFP fragment complementation has indeed been noticed (12). To test this possibility, we studied synthesis and secretion of YFP2-catZ in pulse–chase experiments using [^{35}S]methionine. Fig. 4*D* shows that YFP2-catZ was synthesized at equal rates, irrespective of whether it was coexpressed with WT or N156A ERGIC-53. In contrast, YFP2-catZ secretion was less efficient when coexpressed with YFP1-p53$_{WT}$ than with YFP1-p53$_{N156A}$ (Fig. 4*E*). These findings explain the differences in the steady-state protein amount of YFP2-catZ (Fig. 4*C*) and support the notion that catZ is retained intracellularly by WT ERGIC-53 because of decreased dissociation after YFP fragment complementation.

So far, we have provided the proof of concept of YFP PCA for detecting luminal protein interactions in the secretory pathway that had previously been established by alternative techniques. To search for a novel interaction, we applied YFP PCA to catC. The secretion of catC was shown to be delayed in cells overexpressing dominant-negative ER-retained ERGIC-53 (21), but all previous attempts to demonstrate a direct interaction by pull-down or crosslinking experiments have failed. In contrast, YFP PCA can detect a direct interaction between ERGIC-53 and catC and shows that this interaction depends on a functional lectin domain of ERGIC-53 (Fig. 5). This result conclusively establishes that catC is a cargo glycoprotein of ERGIC-53 and demonstrates the power of YFP PCA to visualize novel protein–protein interactions of weak, transient, and glycan-mediated nature in the secretory pathway that escape detection by coimmunoprecipitation and chemical crosslinking.

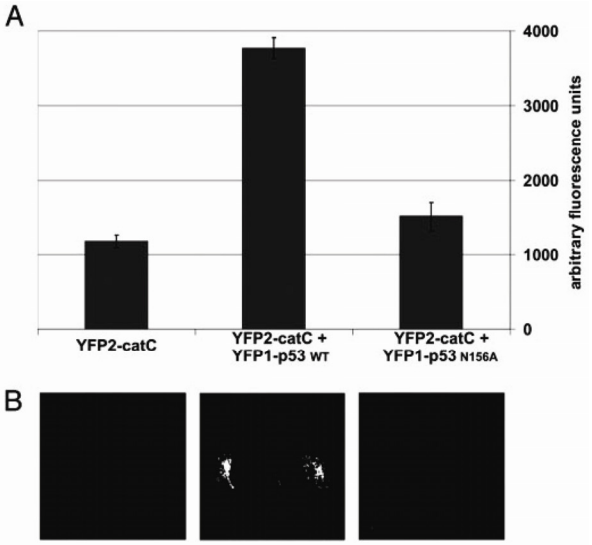


Fig. 5. ERGIC-53 interacts with catC. (*A*) Fluorometric analysis of YFP fragment complementation of ERGIC-53 and catC. The interaction of ERGIC-53 with catC depends on a functional lectin domain of ERGIC-53. (*B*) Fluorescence microscopy of live cells expressing the indicated fusion proteins.

Discussion

The control of secretion requires a plethora of protein interactions on both the cytoplasmic and luminal sides of the secretory pathway. A particular difficulty has been to detect luminal interactions, many of which are of low affinity and occur in an oxidizing environment and hence are often not amenable to traditional approaches. The YFP PCA described here can overcome these limitations. It allows the visualization, in live cells, of luminal protein–protein interactions of different natures: (*i*) interactions among transmembrane proteins; (*ii*) protein-protein-based interactions between a transmembrane and a luminal protein; and (*iii*) glycan-mediated interactions between a membrane lectin and its cargo glycoprotein. Although not tested here, we anticipate that specific interactions between two soluble proteins in the lumen of the secretory pathway will also be detectable by YFP PCA. Further improvements to the citrine YFP PCA, such as increased solubility of fragments, could be achieved through protein engineering strategies (27), although solubility of YFP was not a problem in our experiments.

Most notably, YFP PCA can detect protein–protein interactions that are transient and of low affinity, as illustrated for the lectin interaction of ERGIC-53 with catZ. The detection of such low-affinity interactions may be facilitated by the stabilization of the complex by the reconstituted YFP. This also means that caution should be taken in applying the YFP PCA to quantitative kinetic analysis of protein complex dissociation in limited cases, for which dissociation is rapid compared with other processes like protein degradation. YFP PCA exhibits several major advantages over other methods for the investigation of protein–protein interactions, including: (*i*) its simplicity and reliability, (*ii*) its high sensitivity using YFP citrine that enables analysis of interactions among proteins expressed at levels comparable to many endogenous proteins, and (*iii*) direct visualization of protein–protein interactions in their normal compartmental environment of living cells.

Fluorometric analysis of complemented YFP results in a quantitative read-out. YFP complementation monitored for the interaction between ERGIC-53 and MCFD2 shows about half the fluorescence intensity compared with ERGIC-53 oligomerization and its interaction with catZ or catC. N-terminal tagging of MCFD2 did not significantly increase YFP fragment complementation for the interaction between ERGIC-53 and MCFD2 (data not shown). One possible reason for lower YFP fragment

complementation may be a lower expression level of MCFD2-YFP2 compared with YFP2-catZ (Fig. 3*B*). ERGIC-53 binding of catZ and catC is lectin-dependent, whereas binding of MCFD2 is lectin-independent (Figs. 4*A* and 5*A*). The position of this lectin-independent binding site on ERGIC-53 could also account for lower YFP fragment complementation by imposing spatial constraints.

In this study, we not only provide the proof of concept of the YFP PCA approach but also apply it to detect a direct protein interaction of low affinity between ERGIC-53 and catC that has been surmised but was experimentally unproven. The catC finding is important for several reasons. It validates a previous proposal based on catZ that catC carries an ER-export signal (28). Moreover, it paves the way for the identification of additional ligands of ERGIC-53. Such ligands have been proposed to exist on the basis of secretion assays performed in HeLa cells expressing dominant-negative ERGIC-53 (21). In more general terms, the ability of the YFP PCA to detect transient protein interactions of low affinity will greatly facilitate the analysis of protein complexes involved in virtually all luminal functions of the secretory pathway. Of particular interest is that YFP PCA can detect carbohydrate-mediated interactions between a lectin and its ligands. Numerous lectins in the secretory pathway are involved in protein folding (2), degradation (29–31), transport (32, 33), and sorting (34). For some of these lectins, the ligands are unknown or insufficiently characterized and hence the YFP PCA may aid in the elucidation of lectin functions. Last, the feasibility of screening cDNA-fusion libraries for interacting partners identified by high-throughput YFP PCA analysis (35) should prove a useful approach for revealing new components involved in luminal processes controlling secretion.

We thank K. Bucher for expert technical assistance. This work was supported by the University of Basel and the Swiss National Science Foundation.

1. Gething, M. J. & Sambrook, J. (1992) *Nature* **355,** 33–45.
2. Ellgaard, L., Molinari, M. & Helenius, A. (1999) *Science* **286,** 1882–1888.
3. Appenzeller, C., Andersson, H., Kappeler, F. & Hauri, H. P. (1999) *Nat. Cell. Biol.* **1,** 330–334.
4. Belden, W. J. & Barlowe, C. (2001) *Science* **294,** 1528–1531.
5. Michnick, S. W., Remy, I., Campbell-Valois, F. X., Vallee-Belisle, A. & Pelletier, J. N. (2000) *Methods Enzymol.* **328,** 208–230.
6. Galarneau, A., Primeau, M., Trudeau, L. E. & Michnick, S. W. (2002) *Nat. Biotechnol.* **20,** 619–622.
7. Spotts, J. M., Dolmetsch, R. E. & Greenberg, M. E. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 15142–15147.
8. Remy, I. & Michnick, S. W. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 7678–7683.
9. Paulmurugan, R., Umezawa, Y. & Gambhir, S. S. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 15608–15613.
10. Luker, K. E., Smith, M. C., Luker, G. D., Gammon, S. T., Piwnica-Worms, H. & Piwnica-Worms, D. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 12288–12293.
11. Ghosh, I., Hamilton, A. D. & Regan, L. (2000) *J. Am. Chem. Soc.* **122,** 5658–5659.
12. Hu, C. D., Chinenov, Y. & Kerppola, T. K. (2002) *Mol. Cell* **9,** 789–798.
13. Remy, I., Montmarquette, A. & Michnick, S. W. (2004) *Nat. Cell. Biol.* **6,** 358–365.
14. Remy, I. & Michnick, S. W. (2004) *Mol. Cell. Biol.* **24,** 1493–1504.
15. de Virgilio, M., Kiosses, W. B. & Shattil, S. J. (2004) *J. Cell Biol.* **165,** 305–311.
16. Rackham, O. & Brown, C. M. (2004) *EMBO J.* **23,** 3346–3355.
17. Schweizer, A., Fransen, J. A., Bachi, T., Ginsel, L. & Hauri, H. P. (1988) *J. Cell Biol.* **107,** 1643–1653.
18. Klumperman, J., Schweizer, A., Clausen, H., Tang, B. L., Hong, W., Oorschot, V. & Hauri, H. P. (1998) *J. Cell Sci.* **111,** 3411–3425.
19. Itin, C., Roche, A. C., Monsigny, M. & Hauri, H. P. (1996) *Mol. Biol. Cell* **7,** 483–493.
20. Moussalli, M., Pipe, S. W., Hauri, H. P., Nichols, W. C., Ginsburg, D. & Kaufman, R. J. (1999) *J. Biol. Chem.* **274,** 32539–32542.
21. Vollenweider, F., Kappeler, F., Itin, C. & Hauri, H. P. (1998) *J. Cell Biol.* **142,** 377–389.
22. Zhang, B., Cunningham, M. A., Nichols, W. C., Bernat, J. A., Seligsohn, U., Pipe, S. W., McVey, J. H., Schulte-Overberg, U., de Bosch, N. B., Ruiz-Saez, A., *et al.* (2003) *Nat. Genet.* **34,** 220–225.
23. Majoul, I., Sohn, K., Wieland, F. T., Pepperkok, R., Pizza, M., Hillemann, J. & Soling, H. D. (1998) *J. Cell Biol.* **143,** 601–612.
24. Linstedt, A. D. & Hauri, H. P. (1993) *Mol. Biol. Cell* **4,** 679–693.
25. Griesbeck, O., Baird, G. S., Campbell, R. E., Zacharias, D. A. & Tsien, R. Y. (2001) *J. Biol. Chem.* **276,** 29188–29194.
26. Ben-Tekaya, H., Miura, K., Pepperkok, R. & Hauri, H. P. (2005) *J. Cell Sci.* **118,** 357–367.
27. Cabantous, S., Terwilliger, T. C. & Waldo, G. S. (2005) *Nat. Biotechnol.* **23,** 102–107.
28. Appenzeller-Herzog, C., Nyfeler, B., Burkhard, P., Santamaria, I., Lopez-Otin, C. & Hauri, H. P. (2005) *Mol. Biol. Cell* **16,** 1258–1267.
29. Nakatsukasa, K., Nishikawa, S., Hosokawa, N., Nagata, K. & Endo, T. (2001) *J. Biol. Chem.* **276,** 8635–8638.
30. Jakob, C. A., Bodmer, D., Spirig, U., Battig, P., Marcil, A., Dignard, D., Bergeron, J. J., Thomas, D. Y. & Aebi, M. (2001) *EMBO Rep.* **2,** 423–430.
31. Hosokawa, N., Wada, I., Hasegawa, K., Yorihuzi, T., Tremblay, L. O., Herscovics, A. & Nagata, K. (2001) *EMBO Rep.* **2,** 415–422.
32. Nufer, O., Mitrovic, S. & Hauri, H. P. (2003) *J. Biol. Chem.* **278,** 15886–15896.
33. Neve, E. P., Svensson, K., Fuxe, J. & Pettersson, R. F. (2003) *Exp. Cell Res.* **288,** 70–83.
34. Dahms, N. M. & Hancock, M. K. (2002) *Biochim. Biophys. Acta* **1572,** 317–340.
35. Remy, I. & Michnick, S. W. (2004) *Methods* **32,** 381–388.

Supporting information

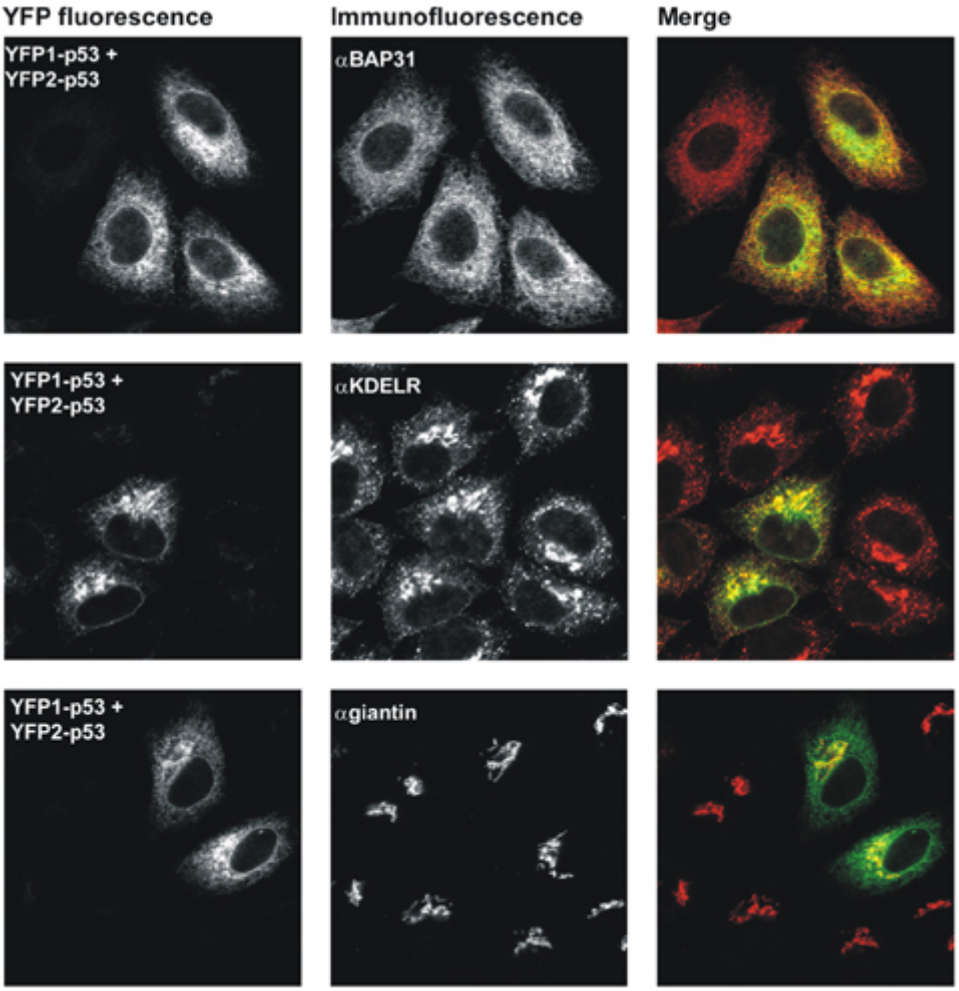


Fig. 6. Reconstituted yellow fluorescent protein (YFP) fluorescence after YFP1-p53 and YFP2-p53 oligomerization localizes to the early secretory pathway. HeLa cells coexpressing YFP1-p53 and YFP2-p53 were analyzed 24 h after transfection. Colocalization of the reconstituted YFP fluorescence with the endoplasmic reticulum marker BAP31, the ERGIC/cis-Golgi marker KDEL receptor, and the cis/medial-Golgi marker giantin demonstrates its localization to the early secretory pathway.

3.2 Development of a genomewide screening procedure to identify cargo proteins of ERGIC-53

Manuscript in preparation.

Beat Nyfeler, Eduard Stefan, Stephen Michnick, Hans-Peter Hauri

Abstract

ERGIC-53 is a transmembrane lectin which cycles in the early secretory pathway and acts as a cargo receptor for a subset of glycoproteins including cathepsin Z, cathepsin C and blood coagulation factors V and VIII. To address the question of whether ERGIC-53 captures additional cargo proteins, we developed a genomewide screening procedure based on the complementation of the yellow fluorescent protein (YFP). To this end, a cDNA-YFP1 fusion library was constructed, expressed in mammalian COS-1 cells and screened using YFP2-tagged ERGIC-53 as bait protein. COS-1 cells that contained complemented YFP were isolated by fluorescence activated cell sorting (FACS) and library plasmid were recovered and analyzed. Here we describe the evaluation of the different steps of the screening procedure and provide results of a first small-scale pilot screen which identified alpha-1-antitrypsin as a potential cargo protein of ERGIC-53. This finding suggests that ERGIC-53 might bind more glycoproteins than initially assumed. Furthermore, our results clearly demonstrate the feasibility to screen a cDNA library for luminal interaction partners of ERGIC-53 by means of fluorescence complementation. Thus, this study provides a firm basis for genomewide screening of protein-protein interactions inside the secretory pathway of mammalian cells.

Introduction

The lumen of the endoplasmic reticulum (ER) provides a unique environment for proteins to be folded and modified[1]. One third of all newly synthesized proteins is co-translationally inserted into the ER and processed in its oxidizing milieu by multiple ER resident enzymes. An elaborate ER quality control system monitors the conformational state of the folding substrates and retains them in the ER during the folding process[2]. While terminally misfolded proteins are re-translocated to the cytosol for degradation[3], correctly folded proteins can be transported to the ER Golgi intermediate compartment (ERGIC) in coat protein II (COPII)-coated vesicles[4]. The Sec24 subunit of the COPII coat selectively recruits transmembrane proteins into anterograde vesicles by directly interacting with ER exit motifs in their cytosolic tails[5]. In turn, transmembrane proteins can capture soluble cargo proteins and incorporate them into COPII-coated vesicles[6-9] according to the model of receptor-mediated ER export of soluble proteins. Although attractive, this model suffers from the low number of transmembrane receptors and cognate soluble cargo proteins which have been identified so far. The best characterized mammalian cargo receptor is the ER Golgi intermediate compartment protein-53 (ERGIC-53)[10, 11]. This type I membrane protein captures with its luminal domain soluble cargo proteins such as cathepsin Z, cathepsin C and blood coagulation factors V and VIII[9, 12-14], while a di-hydrophobic ER exit motif in its cytosolic tail ensures packaging into COPII-coated vesicles[15]. Is ERGIC-53 a cargo receptor for only a small subset of glycoproteins? Are cathepsin Z, cathepsin C and blood coagulation factors V and VIII the only cargo proteins of ERGIC-53? Recent findings that ERGIC-53 also binds mutant immunoglobulin chains[16] and nicastrin[17] indicate the possibility of further, yet unidentified cargo proteins.

For this reason, we set out to screen for ERGIC-53 cargo proteins in a genomewide setting. As a method of choice, a cDNA library screening strategy based on yellow fluorescence complementation was developed. The concept of yellow fluorescence complementation relies on fusing two non-fluorescent fragments of the yellow fluorescent protein (YFP) to two separate proteins which are then expressed in mammalian cells[18]. If the proteins interact, the two fragments of YFP are brought into close proximity which induces the correct folding and reconstitution of fluorescent YFP *in vivo*. In a previous study, we have provided the proof of concept for YFP complementation-based visualization of protein-protein interactions

inside the lumen of the secretory pathway. Using luminal YFP complementation, the oligomerization of ERGIC-53 as well as its interactions with cathepsin Z, cathepsin C and multiple coagulation factor deficiency protein 2 (MCFD2) was readily visualized[19]. With the identification of Ft1 as novel interaction partner of protein kinase B, fluorescence complementation-based cDNA library screening has been established for cytosolic protein interactions[20, 21]. The present study describes now the development of a screening procedure that is applicable inside the lumen of the secretory pathway. In order to screen for luminal protein-protein interactions, a human adult liver cDNA-YFP1 fusion library was constructed and assessed. The different steps in the screening procedure such as co-expression of bait and prey proteins, fluorescent-activated cell sorting (FACS), prey plasmid recovery and hit validation were evaluated. Furthermore, a small-scale pilot screen for ERGIC-53 cargo proteins was undertaken which resulted in the identification of alpha-1-antitrypsin as potential interaction partner of ERGIC-53. The development of this yellow fluorescence complementation-based cDNA library screening approach should be generally applicable for genomewide screening of protein-protein interactions in the lumen of the secretory pathway.

Results

In order to screen for novel ERGIC-53 cargo proteins, a yellow fluorescence complementation-based cDNA library screening procedure was established as illustrated in Figure 1. ERGIC-53 was tagged with YFP fragment 2 (YFP2-ERGIC-53) and used as a bait while prey proteins were expressed from a cDNA-YFP1 fusion library (Figure 1A). Of note, YFP fragment 2 of the ERGIC-53 bait protein is localized to the lumen of the secretory pathway which restricts the screen to luminal protein-protein interactions (Figure 1B). If mammalian cells co-express YFP2-ERGIC-53 with an interacting YFP1-tagged prey protein, the two non-fluorescent YFP fragments are brought into close proximity and can reconstitute fluorescent YFP. Yellow fluorescent cells harboring reconstituted YFP can be readily collected by FACS and used for the isolation and validation of prey plasmids (Figure 1C)[21].

Construction of the cDNA-YFP1 library

YFP complementation-based screening of protein interactions inside the secretory pathway requires the translocation of both bait and prey proteins into the ER. Thus, in a cDNA library screening-based approach, it has to be ensured that library-encoded prey proteins can reach the lumen of the ER. For this purpose, cDNA inserts were tagged at their C-terminus with YFP fragment 1 to preserve endogenous N-terminal signal sequences in membrane and secretory proteins. As a consequence, cDNA inserts have to contain their 5' ends (to ensure the expression of the signal sequence) but lack their 3' ends (to prevent the stop codon from terminating translation prior to YFP1). These requirements are difficult to meet and we first tried to construct a cDNA fusion library from scratch. For this purpose, liver mRNA was reverse transcribed using random primers, second strands were synthesized, and generated cDNAs were ligated, after linker addition and size fractionation, into a YFP1 containing mammalian expression vector. With around 10^5 independent clones that harbored mainly small inserts deriving from mitochondrial DNA (data not shown), the generated library was not applicable to YFP complementation-based protein interaction screening. Therefore, we decided to switch to a cDNA library subcloning approach as depicted in Figure 2. cDNA inserts were subcloned from a human adult liver X-NubG library (Figure 2A) which has been initially constructed to screen protein-protein interactions using the split-ubiquitin system[22]. This library was chosen for the following reasons: (i) secretory proteins are

highly represented since adult liver, a secretory tissue, was used as source for cDNA library construction, (ii) N-terminal signal sequences of membrane and secretory proteins are preserved since random primers were utilized for reverse transcription which enriches the 5' ends of cDNA inserts, (iii) the cDNA library has a high complexity containing plasmids from $1.3x10^6$ independent clones, (iiii) cDNA inserts can be reliably subcloned via two *SfiI* restriction sites. Since cDNA inserts are flanked with two GGCCNNNNNGGCC *SfiI* recognition sites that differ in their 5 middle base pairs, subcloning can be performed without loss of orientation. In addition, the 14 base pair long recognition site of *SfiI* is rare and cDNA inserts are unlikely to be fragmented during subcloning.

To subclone the library, inserts were excised from the X-NubG library, separated by agarose gel-electrophoresis, purified according to their size and inserted into mammalian pcDNA3 expression vectors harboring YFP fragment 1 in all three reading frames (Figure 2C and D). The subcloning strategy resulted in a cDNA-YFP1 library that contained plasmids from approximately 10^6 independent clones with cDNA inserts ranging from 1 to 2.5kb in size (Figure 3A).

Assessment of the quality of the cDNA-YFP1 library

To evaluate the cDNA-YFP1 library, 16 randomly picked plasmids were analyzed for insert size and insert properties. For this purpose, inserts were excised by *SfiI* and sequenced. The digestion with *SfiI* demonstrated that all 16 clones contained inserts of a 1 to 2.5 kb size range (Figure 3B). Furthermore, sequence analysis revealed that 14 of 16 clones contained cDNA inserts coding for proteins of already known function while one clone contained an uncharacterized open reading frame and one clone could not be sequenced (Figure 3C). In order to express a functional YFP1 fusion protein, a cDNA insert should contain its start codon to initiate translation but should lack a stop codon prior to the YFP fragment 1. Hence, the 14 clones that contained cDNA inserts of known proteins were examined for containing their start codon but lacking a stop codon. While all 14 cDNA inserts contained their start codon, 6 lacked their stop codon which indicates that around 40% of the cDNA inserts (6 out of 14) are present in a functional manner. In addition, cDNA inserts were analyzed for having an N-terminal signal sequence. 7 out of the 14 cDNA inserts indeed contained an N-terminal signal sequence suggesting that membrane and secretory proteins are highly represented in the generated cDNA-YFP1 library. To complete the evaluation of the cDNA-YFP1 library, we tested if its co-expression with the YFP2-ERGIC-53

bait protein was able to reconstitute YFP molecules. As shown in Figure 3D, co-expression of YFP2-ERGIC-53 with the cDNA-YFP1 library in HeLa cells led to the detection of around 0.1% yellow fluorescent cells which were absent if YFP2-ERGIC-53 or the cDNA-YFP1-library were expressed alone. This indicates that the cDNA-YFP1 library expresses functional YFP1 fusion proteins which can interact with YFP2-ERGIC-53 and can reconstitute fluorescent YFP.

Plasmid recovery from FACS-sorted cells

A crucial point in the YFP complementation-based cDNA library screen represents the expression and recovery of library plasmids. Standard transfection reagents, including FuGENE6, transfect several plasmids per cell and thereby allow high protein expression levels. Since libraries contain many different plasmids, one can not rely on the uptake of multiple copies of the same plasmid per cell. Therefore, we performed the library screen in COS-1 cells that express the large T antigen and can replicate transfected plasmids containing the SV40 eukaryotic origin of replication[23]. Upon replication by COS-1 cells, each cDNA-YFP1 library plasmid is present in multiple copies per transfected cell which guarantees sufficient prey protein expression and enhances the chances for the subsequent recovery of the library plasmid. As depicted in Figure 4A, YFP2-ERGIC-53 was co-transfected with the cDNA-YFP1 library or two control plasmids. As negative and positive control plasmids, HA- and YFP1-tagged MCFD2 were used, respectively. To limit the transfection of YFP1-MCFD2 in the positive control sample to maximally one plasmid copy per cell, YFP1-MCFD2 was diluted 1:1000 in an excess of MCFD2-HA negative control plasmid. Furthermore, bait and prey plasmids were transfected at a ratio of 10:1 to ensure that all transfected cells take up the bait plasmid but only a minimal number of prey plasmids. FACS-based analysis revealed no yellow fluorescent cells for the negative control, while 0.23% and 0.10% yellow fluorescent cells were detectable for the positive control and the library sample, respectively (Figure 4A and B). This result clearly confirms that the cDNA-YFP1 library indeed expresses interaction partners of YFP2-ERGIC-53 and can reconstitute yellow fluorescence. The analysis of yellow fluorescence intensities indicates that the library sample shows two to three times lower signals if compared to the positive control.

To evaluate plasmid recovery from FACS-sorted cells, 500 yellow fluorescent cells were collected from the positive control sample. Total DNA was isolated and transformed into bacteria which were then grown on ampicillin-containing plates to

recover only clones harboring prey plasmids. This procedure resulted in 600 transformants which corresponds to approximately one recovered plasmid per sorted cell. Since YFP1-MCFD2 was diluted 1:1000 in the MCFD2-HA negative control, the amount of MCFD2-HA in the recovered prey plasmids is indicative for the number of plasmids which are co-transfected together with YFP1-MCFD2 but do not contribute to fluorescence complementation. In order to determine the ratio of YFP1-MCFD2 and MCFD2-HA, individual transformants were picked and analyzed by PCR screen. From 64 checked transformants, 9 contained YFP1-MCFD2 while 53 contained MCFD2-HA (Figure 4C). The recovery of YFP1-MCFD2 in 9 cases demonstrates the feasibility to recover and enrich prey plasmids which are 'positive' for fluorescence complementation. Nevertheless, the detection of MCFD2-HA in 54 cases indicates that the majority of recovered prey plasmids represent co-transfected plasmids which have not contributed to fluorescence complementation. The possibility that missorted cells interfere with this analysis is unlikely, since FACS-sorting resulted in a 95% pure population of yellow fluorescent cells (data not shown). The finding that only every seventh recovered prey plasmids is 'positive' and has contributed to fluorescence complementation implies that a rather high number of prey plasmids need to be validated in the library screen.

Screening for ERGIC-53 interaction partners

To test if ERGIC-53 interaction partners can be identified by screening the cDNA-YFP1 library, a small scale pilot screen was performed. To this end, yellow fluorescent cells were collected by FACS from the library sample (Figure 4A and 4B). From a total of $2x10^6$ cells, 500 fluorescent cells were sorted and used to isolate total DNA. Transformation of the isolated DNA into bacteria resulted in about 500 transformants which corresponds again to one recovered prey plasmid per sorted cell. Subsequently, cDNA-YFP1 library plasmids were isolated from 24 individual transformants, co-expressed with YFP2-ERGIC-53 in COS-1 cells and validated by fluorometric analysis (Figure 5A). From the 24 screened library plasmids, plasmid number 17 reconstituted significant yellow fluorescence upon co-transfection with YFP2-ERGIC-53 (Figure 5B). The analysis of library plasmid number 17 by *SfiI* digestion revealed an insert size of approximately 1.3 kb (Figure 5C) and sequencing identified the insert as alpha-1-antitrypsin. It is of note, that the start codon, the signal sequence, and all amino acids of alpha-1-antitrypsin except the last two are present in library plasmid 17 (Figure 5D). Since fusion to YFP fragment 1 is in frame, library

plasmid 17 expresses YFP1 tagged alpha-1-antitrypsin in a functional manner. The identification of alpha-1-antitrypsin in the small-scale screen for ERGIC-53 cargo proteins clearly demonstrates that yellow fluorescence complementation-based cDNA library screening is a powerful approach to identify potential interaction partners of ERGIC-53.

Discussion

The identification of ERGIC-53 cargo proteins is hampered by the weak and transient nature of the underlying protein-protein interactions. While standard affinity purification procedures fail to isolate cargo proteins of ERGIC-53, chemical crosslinking procedures have been successfully applied to detect the interactions between ERGIC-53 and its cargo proteins cathepsin Z[9] and blood coagulation factor VIII[14]. However, not all protein-protein interactions are amenable to crosslinking as exemplified by the inability to chemically crosslink cathepsin C to ERGIC-53[13]. With the application of the YFP protein fragment complementation assay (PCA) to interactions in the lumen of the secretory pathway, a technique has been developed for capturing cargo proteins of ERGIC-53 in a sensitive and specific manner[19]. While YFP PCA-based visualization of the interactions between ERGIC-53 and cathepsin Z, cathepsin C or MCFD2 was based on already known or suspected protein interactions, the present study extends the application of luminal YFP complementation to the identification of novel protein interactions. The feasibility of fluorescence complementation-based cDNA library screening is demonstrated by pulling out YFP1-tagged alpha-1-antitrypsin from a cDNA-YFP1 library as interaction partner of YFP2-ERGIC-53. Alpha-1-antitrypsin was identified in a small scale screen of 24 recovered library plasmids. An up-scaled version of this screen holds the promise to identify further interaction partners of ERGIC-53. Our small scale screen also reveals that the majority of recovered prey plasmids (23 out of 24) are not able to significantly reconstitute yellow fluorescence if individually expressed with YFP2-ERGIC-53. This finding is in agreement with the recovery of the many MCFD2-HA plasmids (53 out of 64) from the positive control sample. Thus, several prey plasmids are taken up per cell during transfection while only very few really contribute to fluorescence complementation. In order to obtain a reasonable number of validated hits, either a high number of library plasmids need to be screened, or transfection conditions have to be optimized. By minimizing the amount of transfected DNA and transfection reagent and by using ten times less prey than bait plasmids, we have already minimized to a certain extent the number of library plasmids taken up per cell. Nevertheless, other transfection methods such as electroporation should be tested for their ability to transfect even fewer plasmid copies per cell. In the case of electroporation, the YFP2-ERGIC-53 bait protein should be stably transfected into COS-1 cells to ensure its expression.

Our pilot screen identified an interaction of alpha-1-antitrypsin and ERGIC-53. Several considerations support the specificity of this interaction. First, YFP PCA has been shown to capture protein interactions between ERGIC-53 and its cargo proteins in a specific manner[19]. Second, the detection of only 0.1% yellow fluorescent cells upon co-transfection of YFP2-ERGIC-53 and the cDNA-YFP1 library argues against an unspecific protein-protein interaction. Third, an almost full-length version of alpha-1-antitrypsin was expressed since library plasmid number 17 contains the complete coding sequence of alpha-1-antitrypsin except the last two amino acids. This increases the likelihood that alpha-1-antitrypsin is present in its native conformation and does not unspecifically stick to ERGIC-53. Fourth, since ERGIC-53 is a mannose-binding lectin and alpha-1-antitrypsin is triply N-glycosylated, the detected interaction might be carbohydrate-mediated. The carbohydrate-dependence of the interaction between alpha-1-antitrypsin and ERGIC-53, can be probed by YFP PCA using the lectin deficient ERGIC-53 N156A mutant.

The interaction between alpha-1-antitrypsin and ERGIC-53 needs to be verified by a YFP PCA independent approach. One possibility is to probe the interaction by chemical crosslinking in hepatoma cell lines such as HepG2 which express both alpha-1-antitrypsin and ERGIC-53 endogenously.

In conclusion, this study has two major implications. First, YFP complementation can be successfully applied to screen cDNA libraries for novel protein-protein interactions in the lumen of the secretory pathway. Second, ERGIC-53 might bind more glycoproteins than initially assumed. An up-scaled version of this screen will allow the search for ERGIC-53 cargo proteins in a genomewide setting and should be generally applicable to luminal protein-protein interactions in the secretory pathway.

Material and Methods

Construction of pcDNA3[SfiI-linker-YFP1] vectors for library subcloning

Linker-YFP1 fragments covering all three reading frames of YFP fragment 1 were PCR amplified from pcDNA3[MCFD2-YFP1][19] using 5'-CTA AAG CTT GGT GGC GGT GGC TCT GGA GG-3' (reading frame 1), 5'-CTA AAG CTT AGG TGG CGG TGG CTC TGG AGG-3' (reading frame 2) or 5'-CTA AAG CTT CAG GTG GCG GTG GCT CTG GAG G -3' (reading frame 3) as forward and 5'-TAC GGC TCG AGT TAC TTG TAC AGC TCG TCC ATG C -3' as reverse primers. PCR fragments were cloned into the pCMV-Script vector (Stratagene) via the *HindIII* and *XhoI* sites. It is of note that YFP1 contains the citrine mutation (Q69M)[19, 25]. Two *SfiI* restriction sites were introduced into the pCMV[linker-YFP1] vector by inserting annealed and phosphorylated 5'- AAT TCG GCC ATT ACG GCC AGG CCT TTA ATT AAG GCC GCC TCG GCC-3' and 5'- AGC TGG CCG AGG CGG CCT TAA TTA AAG GCC TGG CCG TAA TGG CCG-3' oligos via the *EcoRI* and *HindIII* sites. Finally, the SfiI-linker-YFP1 fragments were subcloned from the pCMV vector into the pcDNA3 vector (Invitrogen) via *EcoRI* and *XhoI*. The pcDNA3[YFP1] vectors covering all three reading frames are illustrated in Figure 2 C and D and were verified by sequencing.

Generation of cDNA-YFP1 library

A human adult liver cDNA-NubG library (Dualsystems) was digested for 3h with *SfiI* (NEB) at 50°C and the product was separated by agarose gel-electrophoresis. DNA was visualized by ethidium bromide, and inserts were excised from the gel in four fractions according to size (fraction A: 4-2.5kb, fraction B: 2.5-1.5kb, fraction C: 1.5-1kb, fraction D: 1-0.3kb). The four insert fractions were purified using the QIAquick gel extraction kit (QIAGEN) and ligated into *SfiI* digested, dephosphorylated and gel purified pcDNA3[YFP1] vectors. Ligation products were ethanol precipitated and transformed into competent MC1061 bacteria[26] by electroporation. Transformed bacteria were plated onto LB-plates containing 100μg/ml ampicillin and grown over night at 30°C. For each of the four cDNA fractions, eight transformants were picked, amplified in liquid culture before plasmids were isolated using the GenElute Plasmid Miniprep kit (SIGMA). Isolated plasmids were then analyzed by *SfiI* digest and sequenced using 5'-CAA ATG GGC GGT AGG CGT GTA CG-3' as forward and 5'-CGT CCA GCT CGA CCA GGA TGG-3' as reverse primer. Transformants deriving

from insert fraction A contained mainly a contaminating, 2.5 kb large, ampicillin resistant vector harboring no cDNA insert. Transformants deriving from insert fractions B and C contained various cDNA inserts differing in size and nature (Figure 3C). Transformants deriving from insert fraction D contained mainly products from mitochondrial genome. Consequently, insert fractions B and C were chosen to generate the final cDNA-YFP1 library. Ligation, transformation and platting were scaled up until a total of around $1x10^6$ transformants was obtained. All transformants were scraped into LB medium, pooled and library plasmids were isolated using the Plasmid Maxi kit (QIAGEN).

Construction of the YFP2-ERGIC-53 bait plasmid

The YFP2-ERGIC-53 bait was constructed in the pCMV-Script vector by two subsequent subcloning steps. First, the SS-YFP2 fragment was subcloned from pcDNA3[SS-YFP2-catZ][19] into the pCMV vector using *EcoRI*. ERGIC-53 lacking its endogenous signal sequence was then subcloned from pcDNA3[SS-YFP2-ERGIC-53][19] using *XhoI*. The pCMV[SS-YFP2-ERGIC-53] construct was verified by sequencing.

Cell Culture and Transfection

HeLa cells (ATCC, CCL-2) were grown in DMEM, supplemented with 10% fetal bovine serum, 1x non-essential amino acids and antibiotics. COS-1 cells (ATCC, CRL-1650) were grown in DMEM, supplemented with 10% fetal bovine serum and antibiotics. For FACS analysis and fluorometric analysis cells were grown in 35mm dishes while cells for FACS sorting were grown 100mm dishes. Cells were transfected 24h after plating using FuGENE6 (Roche Applied Science). Cells in 35mm dishes were transfected with 1.5μl FuGENE6 and 500ng DNA and cells in 100mm dishes were transfected with 7.5μl FuGENE6 and 2.5μg DNA. Bait and prey plasmids were co-transfected at a ratio of 10:1.

FACS Analysis and Sorting

48h after transfection, cells were washed with PBS, trypsinized and resuspended in PBS containing 0.1% bovine serum albumin and 5mM EDTA. FACS analysis was performed on a CYAN ADP analyzer (Dako) using the Summit v4.3 software (Dako). Cell sorting was performed on a FACS Vantage SE (Becton Dickinson). An excitation

wavelength of 488nm was used and emission was recorded at 530nm (green fluorescence) and 680nm (red fluorescence). FuGENE6 transfected COS-1 cells showed significant autofluorescence. Plotting of green versus red fluorescence resulted in autofluorescent cells lying in the diagonal. Consequently, YFP expressing cells could be exclusively detected in the green fluorescent channel which allowed accurate analysis and sorting.

Recovery of prey plasmids

Genomic DNA was isolated from FACS-sorted cells using the DNeasy Tissue kit (QIAGEN) according to the manufacturer's instructions. The isolated DNA was ethanol precipitated, resuspended in 10μl water and transformed into XL-10 gold ultracompetent bacteria (Stratagene). Transformed bacteria were grown over night at 37°C on LB-plates containing 100μg/ml ampicillin. Ampicillin selects transformants that harbor prey plasmids since only prey plasmids contain the ampicillin resistance marker. Transformants were either probed by PCR screen or amplified in liquid culture before plasmids were isolated using the GenElute Plasmid Miniprep kit (SIGMA).

PCR Screen

Individual transformants were picked, transferred into 20μl water, boiled and probed by PCR using the FastStart PCR Master (Roche Applied Science) and 5'-GAA CCC ACT GCT TAC TGG CTT ATC G-3' as forward and 5'-TGG CAA CTA GAA GGC ACA GTC GAG G-3' as reverse primer. The two primers anneal to the pcDNA3 vector backbone and amplify 1149bp and 666bp fragments for pcDNA3[SS-YFP1-MCFD2][19] and pcDNA3[SS-HA-MCFD2], respectively. PCR products were separated by agarose gel-electrophoresis and visualized by ethidium bromide staining.

Fluorometric analysis

Fluorometric analysis was performed as described previously[19].

Acknowledgement

We thank Käthy Bucher for expert technical assistance and Verena Jäggin for the professional assistance in FACS analysis and sorting. This work was supported by the University of Basel and the Swiss National Science Foundation.

References

1. Gething, M.J. & Sambrook, J. Protein folding in the cell. *Nature* **355**, 33-45 (1992).
2. Ellgaard, L., Molinari, M. & Helenius, A. Setting the standards: quality control in the secretory pathway. *Science* **286**, 1882-1888 (1999).
3. Romisch, K. Endoplasmic reticulum-associated degradation. *Annu Rev Cell Dev Biol* **21**, 435-456 (2005).
4. Lee, M.C., Miller, E.A., Goldberg, J., Orci, L. & Schekman, R. Bi-directional protein transport between the ER and Golgi. *Annu Rev Cell Dev Biol* **20**, 87-123 (2004).
5. Barlowe, C. Signals for COPII-dependent export from the ER: what's the ticket out? *Trends Cell Biol* **13**, 295-300 (2003).
6. Belden, W.J. & Barlowe, C. Role of Erv29p in collecting soluble secretory proteins into ER-derived transport vesicles. *Science* **294**, 1528-1531 (2001).
7. Otte, S. & Barlowe, C. Sorting signals can direct receptor-mediated export of soluble proteins into COPII vesicles. *Nat Cell Biol* **6**, 1189-1194 (2004).
8. Appenzeller-Herzog, C. et al. Carbohydrate- and conformation-dependent cargo capture for ER-exit. *Mol Biol Cell* **16**, 1258-1267 (2005).
9. Appenzeller, C., Andersson, H., Kappeler, F. & Hauri, H.P. The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. *Nat Cell Biol* **1**, 330-334 (1999).
10. Hauri, H.P., Kappeler, F., Andersson, H. & Appenzeller, C. ERGIC-53 and traffic in the secretory pathway. *J Cell Sci* **113 (Pt 4)**, 587-596 (2000).
11. Schweizer, A., Fransen, J.A., Bachi, T., Ginsel, L. & Hauri, H.P. Identification, by a monoclonal antibody, of a 53-kD protein associated with a tubulo-vesicular compartment at the cis-side of the Golgi apparatus. *J Cell Biol* **107**, 1643-1653 (1988).
12. Nichols, W.C. et al. Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. *Cell* **93**, 61-70 (1998).
13. Vollenweider, F., Kappeler, F., Itin, C. & Hauri, H.P. Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme. *J Cell Biol* **142**, 377-389 (1998).
14. Zhang, B., Kaufman, R.J. & Ginsburg, D. LMAN1 and MCFD2 form a cargo receptor complex and interact with coagulation factor VIII in the early secretory pathway. *J Biol Chem* **280**, 25881-25886 (2005).
15. Kappeler, F., Klopfenstein, D.R., Foguet, M., Paccaud, J.P. & Hauri, H.P. The recycling of ERGIC-53 in the early secretory pathway. ERGIC-53 carries a cytosolic endoplasmic reticulum-exit determinant interacting with COPII. *J Biol Chem* **272**, 31801-31808 (1997).
16. Mattioli, L. et al. ER storage diseases: a role for ERGIC-53 in controlling the formation and shape of Russell bodies. *J Cell Sci* **119**, 2532-2541 (2006).
17. Morais, V.A. et al. N-glycosylation of human nicastrin is required for interaction with the lectins from the secretory pathway calnexin and ERGIC-53. *Biochim Biophys Acta* (2006).
18. Kerppola, T.K. Visualization of molecular interactions by fluorescence complementation. *Nat Rev Mol Cell Biol* **7**, 449-456 (2006).
19. Nyfeler, B., Michnick, S.W. & Hauri, H.P. Capturing protein interactions in the secretory pathway of living cells. *Proc Natl Acad Sci U S A* **102**, 6350-6355 (2005).
20. Remy, I. & Michnick, S.W. Regulation of apoptosis by the Ft1 protein, a new modulator of protein kinase B/Akt. *Mol Cell Biol* **24**, 1493-1504 (2004).
21. Remy, I. & Michnick, S.W. A cDNA library functional screening strategy based on fluorescent protein complementation assays to identify novel components of signaling pathways. *Methods* **32**, 381-388 (2004).
22. Stagljar, I., Korostensky, C., Johnsson, N. & te Heesen, S. A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Proc Natl Acad Sci U S A* **95**, 5187-5192 (1998).
23. Gluzman, Y. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23**, 175-182 (1981).
24. Carrell, R.W. alpha 1-Antitrypsin: molecular pathology, leukocytes, and tissue damage. *J Clin Invest* **78**, 1427-1431 (1986).
25. Griesbeck, O., Baird, G.S., Campbell, R.E., Zacharias, D.A. & Tsien, R.Y. Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *J Biol Chem* **276**, 29188-29194 (2001).
26. Casadaban, M.J. & Cohen, S.N. Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. *J Mol Biol* **138**, 179-207 (1980).

Figure Legend

Figure 1: Yellow fluorescence complementation-based cDNA screening for ERGIC-53 interaction partners.

Fluorescence complementation was applied to screen a cDNA library for ERGIC-53 interaction partners. To this end, a cDNA-YFP1 library was generated to express YFP1-tagged prey proteins while YFP2-ERGIC-53 was used as bait. The cDNA-YFP1 library and YFP2-ERGIC-53 were constructed in mammalian expression vectors harboring the bacterial markers for ampicillin and kanamycin resistance, respectively (A). To probe protein-protein interactions, YFP2-ERGIC-53 and prey-YFP1 fusion proteins were co-expressed in mammalian COS-1 cells. Yellow fluorescence is only detectable if YFP2-ERGIC-53 is co-expressed with an interacting YFP1 fusion protein which brings YFP fragments 1 and 2 into close proximity and allows reconstitution of YFP (B). Yellow fluorescent cells were collected by fluorescence-activated cell sorting (FACS) and DNA was isolated and transformed into bacteria which were then grown on ampicillin-containing plates to selectively recover prey plasmids. In order to validate the interaction with ERGIC-53, prey plasmids were isolated and individually analyzed for YFP complementation upon co-expression with YFP2-ERGIC-53 (C).

Figure 2: Construction of the cDNA-YFP1 library

The cDNA-YFP1 library applied in this study was generated by subcloning cDNA inserts from a cDNA-NubG library. Information regarding the source, construction method, library vector, complexity, average insert size as well as insert size range of the cDNA-NubG library is indicated (A). cDNA inserts were excised from the cDNA-NubG library by *SfiI* digest, fractionated according to their size, purified and subcloned into the mammalian pcDNA3 expression vector containing YFP fragment 1 as schematically depicted in (B). A detailed map of the pcDNA3[YFP1] vector is illustrated (C). cDNA inserts were ligated into the vector via the two indicated *SfiI* restriction sites just in front of the linker-YFP1 fragment. The ten amino acid GGGGSGGGGS linker precedes YFP1 to ensure flexibility of YFP fragment 1 for fluorescence complementation. While the SV40 origin of replication allows plasmid amplification in mammalian cells harboring the large T-antigen, expression of YFP1 fusion proteins is guaranteed by the CMV promoter in the pcDNA3 vector. To cover all possible reading frames in pcDNA3[YFP1], inserts of 0, 1 or 2 base pairs (orange)

were introduced between the *SfiI* sites (blue) and the linker region (green). The translation products of the three reading frames are illustrated (D).

Figure 3: Assessment of the quality of the cDNA-YFP1 library

The subcloning strategy depicted in Figure 2 resulted in the construction of a cDNA-YFP1 library with a complexity of about $1x10^6$ clones. To assess the quality of the cDNA-YFP1 library, inserts were excised by *SfiI* digestion using either the total library (A) or 16 randomly picked clones (B). *SfiI* digestion of the total library reveals a smear of cDNA inserts ranging from around 1 to 2.5 kb in size while the analysis of the 16 individual clones demonstrates that DNA inserts are present in all plasmids. To characterize the inserts of the 16 picked clones, plasmids were sequenced. Sequence analysis was used to identify the inserts and to test if the inserts contained their start and stop codon and a N-terminal signal sequence (C). Note that the cDNA-YFP1 library is enriched in secretory and membrane proteins harboring N-terminal signal sequences and includes many functional cDNA inserts containing their start codon but lacking their stop codon. In order to evaluate if the cDNA-YFP1 library can express prey proteins that interact with YFP2-ERGIC-53 and thereby reconstitute yellow fluorescence, HeLa cells were transfected and subjected to FACS-based analysis (D). While no yellow fluorescent cells were detectable for cells transfected with either YFP2-ERGIC-53 or the cDNA-YFP1 library alone, 0.1% yellow fluorescent cells appeared in the case of co-transfected cells.

Figure 4: Plasmid recovery from FACS-sorted cells

YFP2-ERGIC-53 was co-expressed in COS-1 cells with either MCFD2-HA as negative control, 1:1000 diluted YFP1-MCFD2 as positive control or the cDNA-YFP1 library (A). 48h after transfection, yellow fluorescent cells were analyzed by FACS (B). Note that yellow fluorescent cells were absent in the negative control sample while 0.23% and 0.10% yellow fluorescent cells were detectable in the positive control and library sample, respectively. In order to evaluate plasmid recovery, yellow fluorescent cells were FACS-sorted from the positive control sample, total DNA was isolated and transformed into bacteria which were then grown on ampicillin-containing plates to select transformants that contained prey plasmids. From the 600 obtained transformants, 64 were probed by PCR screen for containing either YFP1-MCFD2 or MCFD2-HA (C). The recovery of pcDNA3[YFP1-MCFD2] in just 9 cases

indicates that approximately only every seventh prey plasmid is ‘positive’ in the sense of having contributed to yellow fluorescence complementation.

Figure 5: Screening for ERGIC-53 interaction partners

To screen for ERGIC-53 interaction partners, COS-1 cells were co-transfected with YFP2-ERGIC-53 and the cDNA-YFP1 library, yellow fluorescent cells were collected by FACS and library plasmids were recovered as described in *Material and Methods*. 24 isolated library plasmids were individually expressed with YFP2-ERGIC-53 in COS-1 cells which were then subjected to fluorometric analysis (A). MCFD2-HA and YFP1-MCFD2 were used as negative and positive controls, respectively. Note that library plasmid 17 induced significant yellow fluorescence upon co-transfection with YFP2-ERGIC-53. This finding was confirmed by re-transfecting COS-1 cells and re-measuring yellow fluorescence by fluorometric analysis (B). *SfiI* digestion of library plasmid 17 excised a 1.3 kb insert (C) which was identified as alpha-1-antitrypsin by sequencing (D). The red bar marks the region of the alpha-1-antitrypsin mRNA and protein which is represented in library plasmid 17.

Figure 1: Yellow fluorescence complementation-based cDNA screening for ERGIC-53 interaction partners.

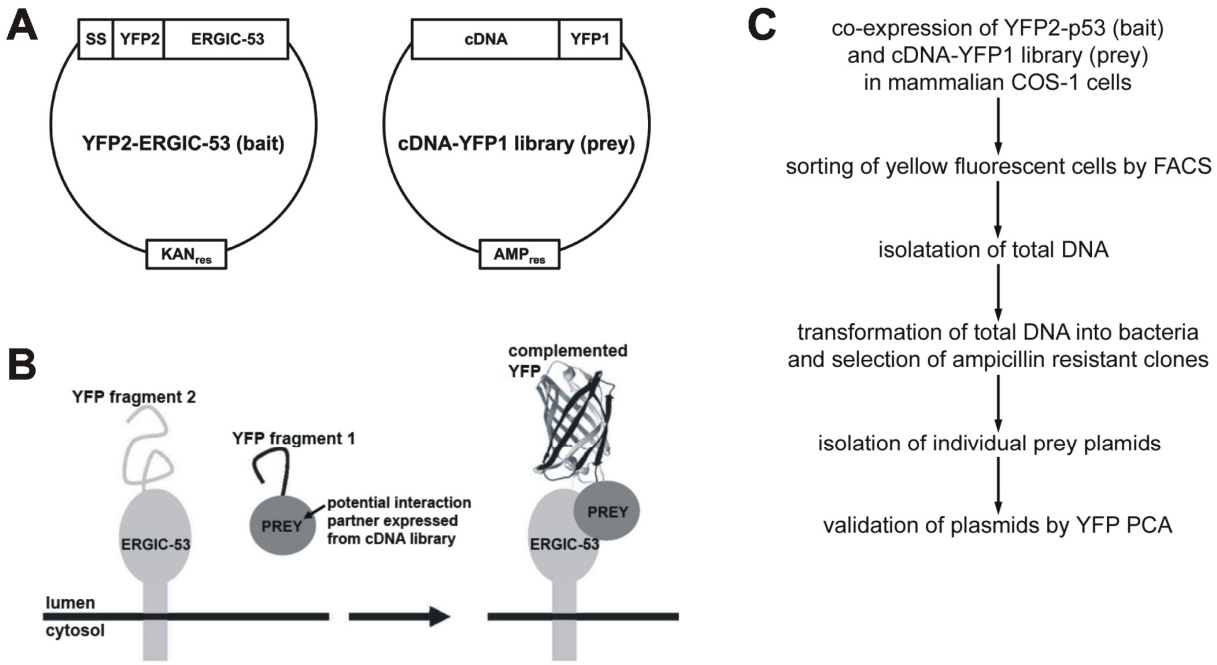


Figure 2: Construction of the cDNA-YFP1 library.

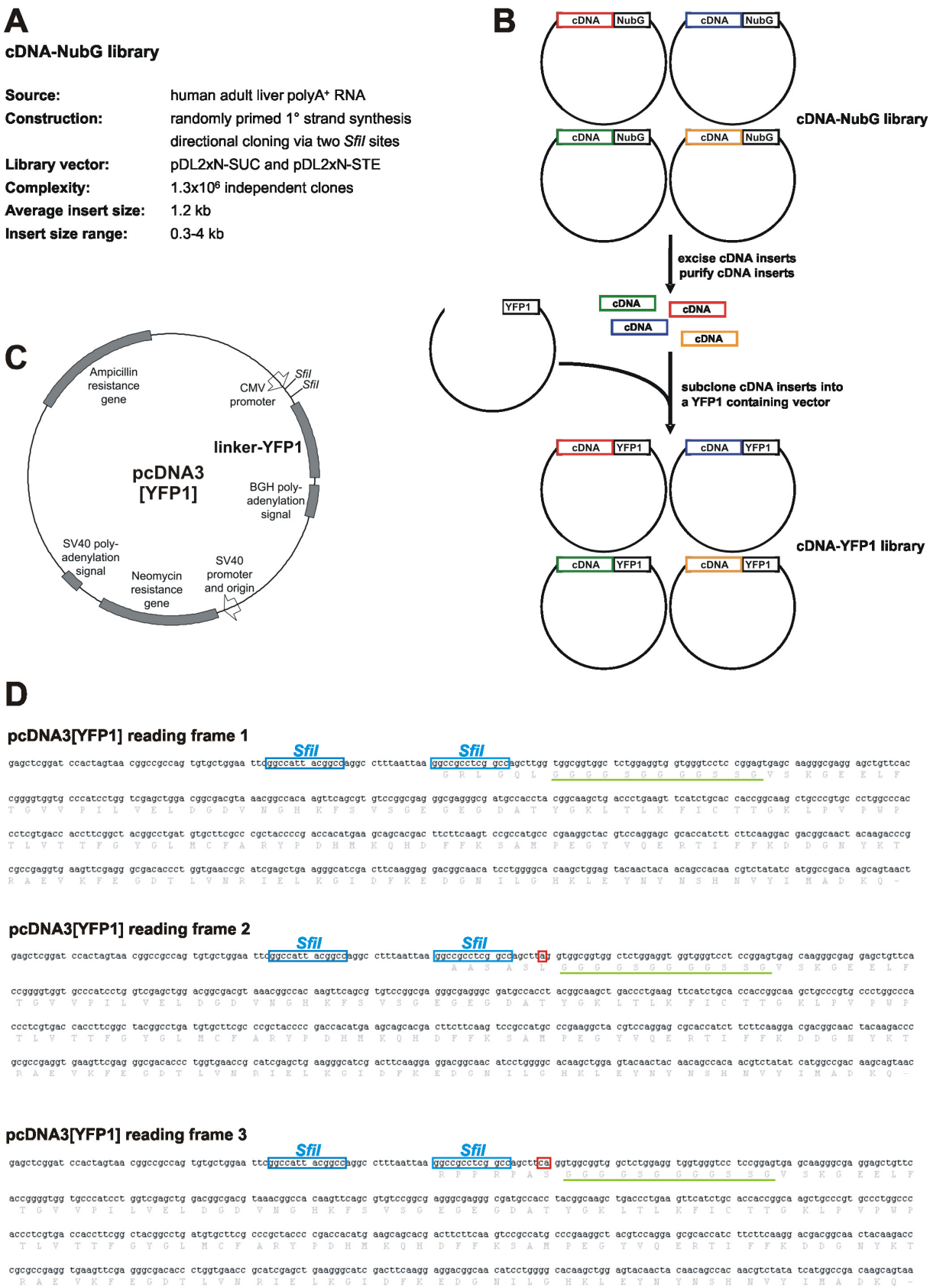
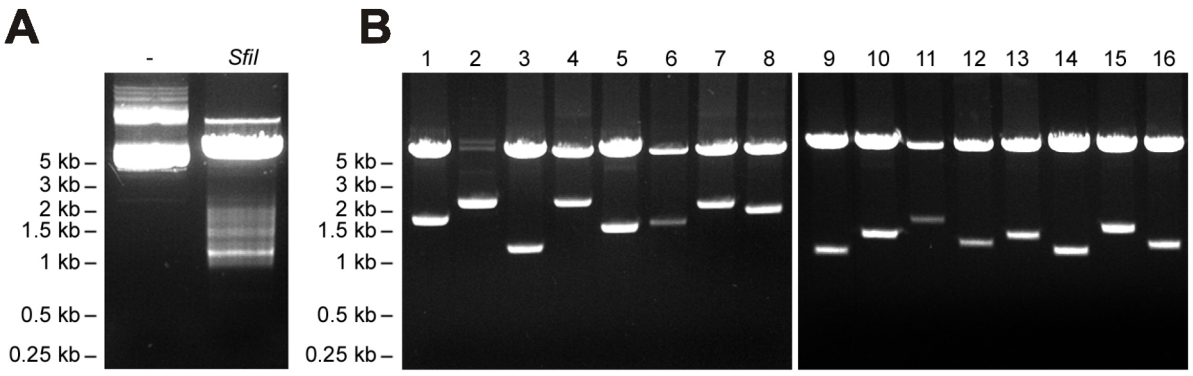


Figure 3: Assessment of the quality of the cDNA-YFP1 library.



C

	vector	insert size	insert	start codon	signal sequence	stop codon
1	pcDNA3	1.9 kb	hydroxyacyl-CoA dehydrogenase	+	-	+
2	nd.	2.2 kb	nd.	nd.	nd.	nd.
3	pcDNA3	1.3 kb	alpha1-antitrypsin	+	+	-
4	pcDNA3	2.2 kb	hexokinase1	+	-	-
5	pcDNA3	1.5 kb	vitronectin	+	+	+
6	pcDNA3	1.5 kb	v-myc viral oncogene homolg	+	-	-
7	pcDNA3	2 kb	nucleobindin1	+	+	+
8	pcDNA3	1.9kb	tyrosine aminotransferase	+	-	+
9	pcDNA3	1.1 kb	open reading frame 68, chromosome 16	nd.	nd.	nd.
10	pcDNA3	1.4 kb	hemopexin	+	+	+
11	pcDNA3	1.6 kb	cytochrome P450 CYP2E1	+	-	+
12	pcDNA3	1.3 kb	sirtuin	+	-	-
13	pcDNA3	1.4 kb	alpha1-antitrypsin	+	+	+
14	pcDNA3	1.1 kb	AMBP protein	+	+	-
15	pcDNA3	1.5 kb	fibrinogen gamma chain	+	+	+
16	pcDNA3	1.3 kb	cytochrome P450 CYP2D6	+	-	-

D

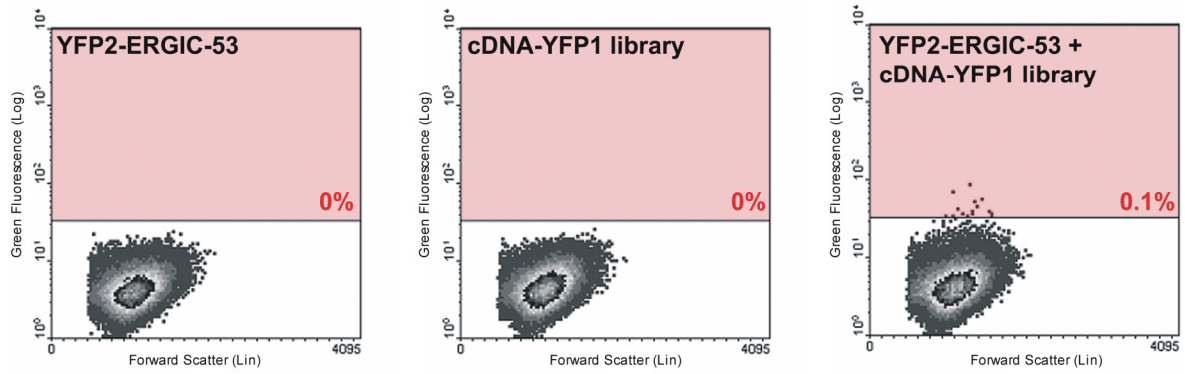
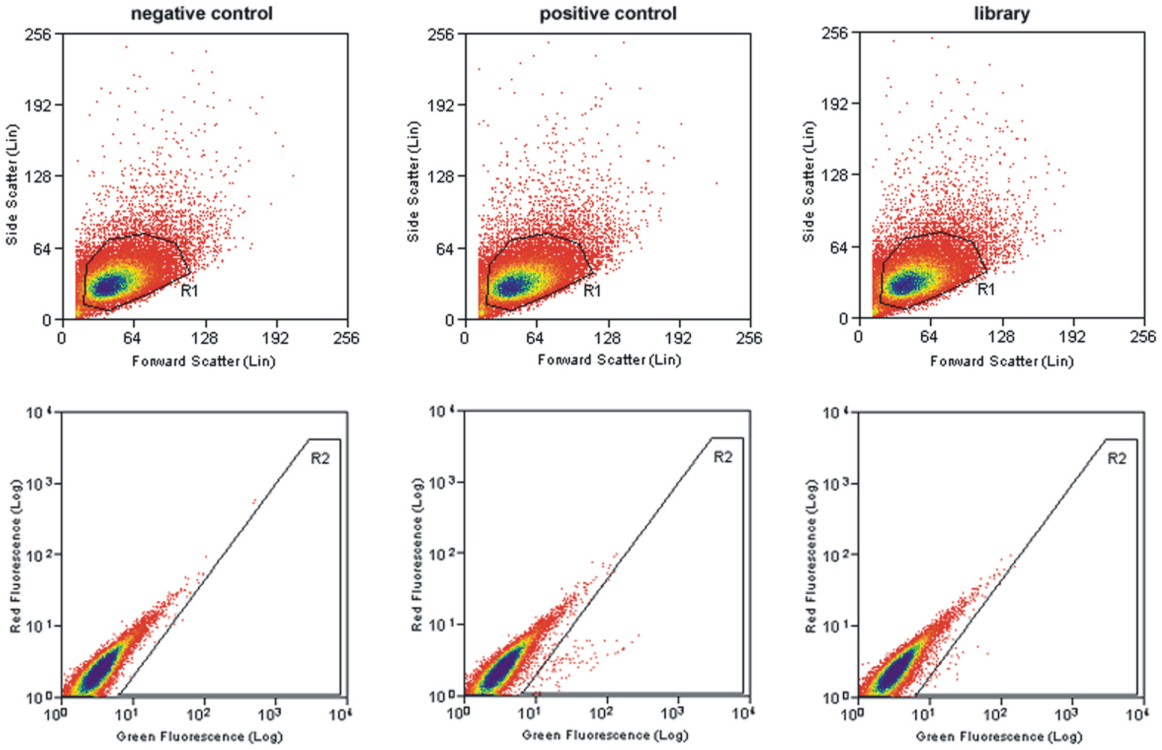


Figure 4: Plasmid recovery from FACS-sorted cells

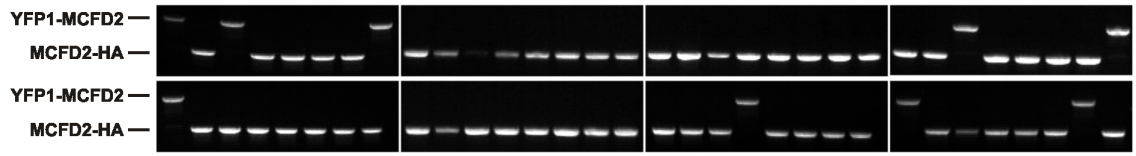
A

Sample	Bait	Prey	total cells analyzed	cells in R1	cells in R2
negative control	YFP2-ERGIC-53	MCFD2-HA	54877	49494 (90.2%)	**2 (0.00%)**
positive control	YFP2-ERGIC-53	YFP1-MCFD2 + MCFD2-HA (1:1000)	55158	49287 (89.4%)	**127 (0.23%)**
library	YFP2-ERGIC-53	cDNA-YFP1 library	55036	49409 (89.8%)	**51 (0.10%)**

B

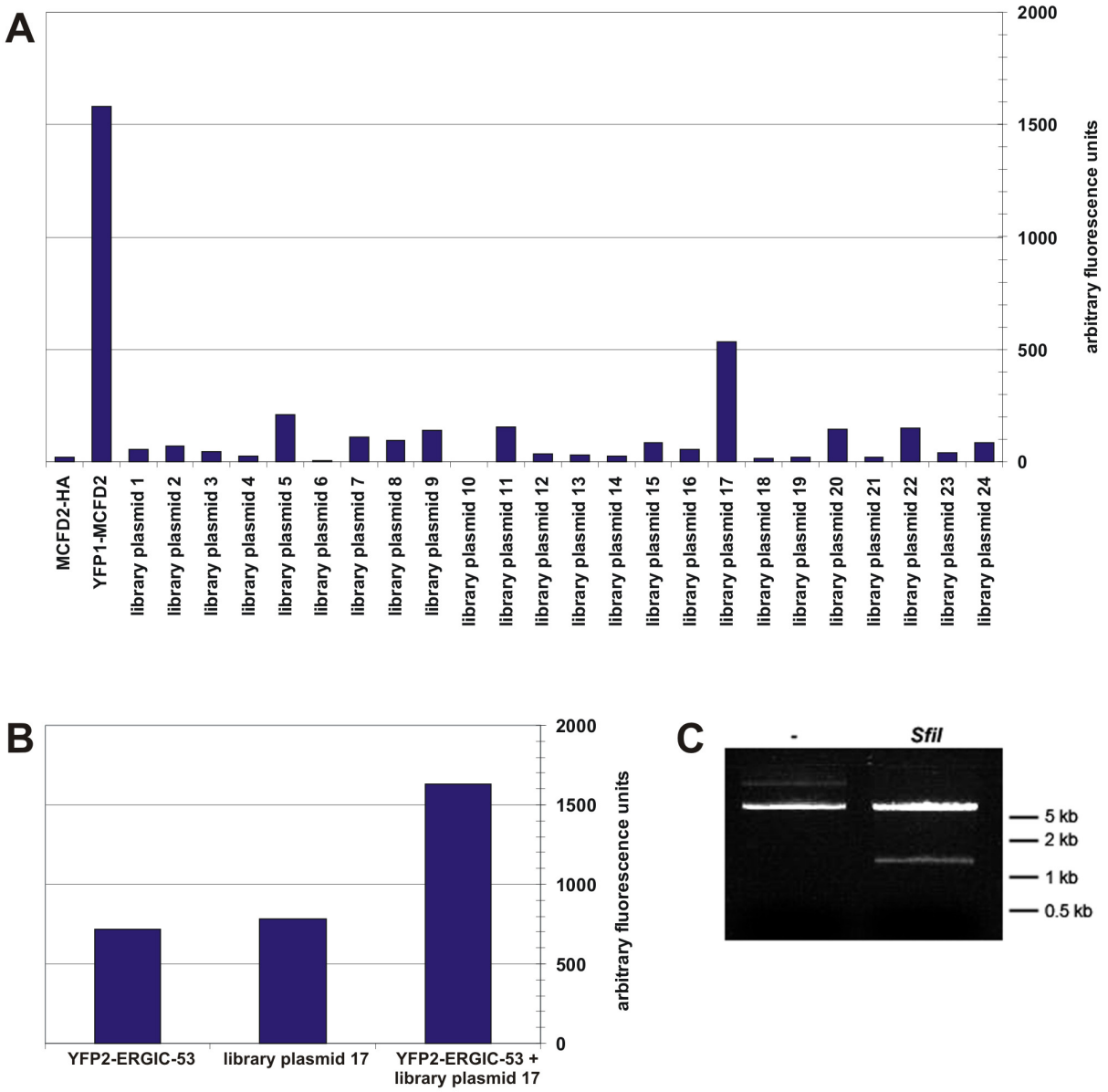


C



Transformants containing YFP1-MCFD2: 9
Transformants containing MCFD2-HA: 53

Figure 5: Screening for ERGIC-53 interaction partners



D alpha-1-antitrypsin:

```
ccactgctta aatacggacg aggacagggc cctgtctcct cagcttcagg caccaccact gacctgggac agtgaatcga caatgccgtc ttctgtctcg tggggcatcc tcctgctggc aggcctgtgc tgcctggtcc ctgtctccct   mRNA
                                                                                M  P  S  S  V  S  W  G  I  L  L  L  A  G  L  C  C  L  V  P  V  S   protein

ggctgaggat ccccagggag atgctgccca gaagacagat acatcccacc atgatcagga tcacccaacc ttcaacaaga tcacccccaa cctggctgag ttcgccttca gcctataccg ccagctggca caccagtcca acagcaccaa
L  A  E  D  P  Q  G  D  A  A  Q  K  T  D  T  S  H  H  D  Q  D  H  P  T  F  N  K  I  T  P  N  L  A  E  F  A  F  S  L  Y  R  Q  L  A  H  Q  S  N  S  T

tatcttcttc tccccagtga gcatcgctac agcctttgca atgctctccc tggggaccaa ggctgacact cacgatgaaa tcctggaggg cctgaatttc aacctcacgg agattccgga ggctcagatc catgaaggct tccaggaact
N  I  F  F  S  P  V  S  I  A  T  A  F  A  M  L  S  L  G  T  K  A  D  T  H  D  E  I  L  E  G  L  N  F  N  L  T  E  I  P  E  A  Q  I  H  E  G  F  Q  E

cctccgtacc ctcaaccagc cagacagcca gctccagctg accaccggca atggcttgtt cctcagcgag ggcctgaagc tagtggataa gtttttggag gatgttaaaa agttgtacca ctcagaagcc ttcactgtca acttcgggga
L  L  R  T  L  N  Q  P  D  S  Q  L  Q  L  T  T  G  N  G  L  F  L  S  E  G  L  K  L  V  D  K  F  L  E  D  V  K  K  L  Y  H  S  E  A  F  T  V  N  F  G

caccgaagag gccaagaaac agatcaacga ttacgtggag aagggtactc aagggaaaat tgtggatttg gtcaaggagc ttgacagaga cacagttttt gctctggtga attacatctt ctttaaaggc aaatgggaga gaccctttga
D  T  E  E  A  K  K  Q  I  N  D  Y  V  E  K  G  T  Q  G  K  I  V  D  L  V  K  E  L  D  R  D  T  V  F  A  L  V  N  Y  I  F  F  K  G  K  W  E  R  P  F

agtcaaggac accgaggaag aggacttcca cgtggaccag gtgaccaccg tgaaggtgcc tatgatgaag cgtttaggca tgtttaacat ccagcactgt aagaagctgt ccagctgggt gctgctgatg aaatacctgg gcaatgccac
E  V  K  D  T  E  E  E  D  F  H  V  D  Q  V  T  T  V  K  V  P  M  M  K  R  L  G  M  F  N  I  Q  H  C  K  K  L  S  S  W  V  L  L  M  K  Y  L  G  N  A

cgccatcttc ttcctgcctg atgaggggaa actacagcac ctggaaaatg aactcaccca cgatatcatc accaagttcc tggaaaatga agacagaagg tctgccagct tacatttacc caaactgtcc attactggaa cctatgatct
T  A  I  F  F  L  P  D  E  G  K  L  Q  H  L  E  N  E  L  T  H  D  I  I  T  K  F  L  E  N  E  D  R  R  S  A  S  L  H  L  P  K  L  S  I  T  G  T  Y  D

gaagagcgtc ctgggtcaac tgggcatcac taaggtcttc agcaatgggg ctgacctctc cggggtcaca gaggaggcac ccctgaagct ctccaaggcc gtgcataagg ctgtgctgac catcgacgag aaagggactg aagctgctgg
L  K  S  V  L  G  Q  L  G  I  T  K  V  F  S  N  G  A  D  L  S  G  V  T  E  E  A  P  L  K  L  S  K  A  V  H  K  A  V  L  T  I  D  E  K  G  T  E  A  A

ggccatgttt ttagaggcca tacccatgtc tatccccccc gaggtcaagt tcaacaaacc ctttgtcttc ttaatgattg accaaaatac caagtctccc ctctttcatgg gaaaagtggt gaatcccacc caaaaataac tgcctctcgc
G  A  M  F  L  E  A  I  P  M  S  I  P  P  E  V  K  F  N  K  P  F  V  F  L  M  I  D  Q  N  T  K  S  P  L  F  M  G  K  V  V  N  P  T  Q  K  -

tcctcaaccc ctcccctcca tccctggccc cctccctgga tgacattaaa gaagggttga gctggaaaaa aaaaaaaaaa aaa
```

3.3 Cargo selectivity of the ERGIC-53/MCFD2 transport receptor complex

Manuscript published in Traffic. 2006 November; 7(11): 1473-1481.

An online version is available at:
http://www.blackwell-synergy.com/doi/full/10.1111/j.1600-0854.2006.00483.x

Cargo Selectivity of the ERGIC-53/MCFD2 Transport Receptor Complex

Beat Nyfeler[1], Bin Zhang[2], David Ginsburg[2], Randal J. Kaufman[2] and Hans-Peter Hauri[1,*]

[1]Biozentrum, University of Basel, CH-4056 Basel, Switzerland
[2]Life Sciences Institute, the Departments of Biological Chemistry, Internal Medicine, Human Genetics Howard Hughes Medical Institute, University of Michigan, Ann Arbor, Mi 48109, USA
**Corresponding author: Hans-Peter Hauri, hans-peter.hauri@unibas.ch*

Exit of soluble secretory proteins from the endoplasmic reticulum (ER) can occur by receptor-mediated export as exemplified by blood coagulation factors V and VIII. Their efficient secretion requires the membrane lectin ER Golgi intermediate compartment protein-53 (ERGIC-53) and its soluble luminal interaction partner multiple coagulation factor deficiency protein 2 (MCFD2), which form a cargo receptor complex in the early secretory pathway. ERGIC-53 also interacts with the two lysosomal glycoproteins cathepsin Z and cathepsin C. Here, we tested the subunit interdependence and cargo selectivity of ERGIC-53 and MCFD2 by short interference RNA-based knockdown. In the absence of ERGIC-53, MCFD2 was secreted, whereas knocking down MCFD2 had no effect on the localization of ERGIC-53. Cargo binding properties of the ERGIC-53/MCFD2 complex were analyzed *in vivo* using yellow fluorescent protein fragment complementation. We found that MCFD2 is dispensable for the binding of cathepsin Z and cathepsin C to ERGIC-53. The results indicate that ERGIC-53 can bind cargo glycoproteins in an MCFD2-independent fashion and suggest that MCFD2 is a recruitment factor for blood coagulation factors V and VIII.

Key words: cargo receptor, endoplasmic reticulum, ER–Golgi intermediate compartment, lectin, protein fragment complementation, protein retention, protein secretion

Received 14 June 2006, revised and accepted for publication 4 August 2006

After folding, N-gylcosylation and oligomerization, newly synthesized secretory proteins leave the endoplasmic reticulum (ER) in coat protein II (COP II)-coated vesicles (1,2). Membrane proteins can be recruited into COP II-coated vesicles by interaction of their cytosolic tails with the Sec23/24 complex of the COP II coat (3–6). In contrast, soluble secretory proteins cannot rely on a direct interaction with the COP II coat for topological reasons. Their export is believed to occur by either bulk flow or receptor-mediated export (7). According to the bulk flow model, proteins enter COP II-coated vesicles by default due to their high concentration in the ER (8). In the receptor-mediated export model, membrane receptors bind to soluble cargo proteins, thereby recruiting them into COP II-coated vesicles (9).

Interactions of cargo receptors with ER export signals of soluble secretory proteins have been characterized only recently. Two such cargo receptors have been studied in detail. The yeast membrane protein Erv29p binds a hydrophobic ER export signal in its cargo protein glycosylated pro-alpha-factor (10,11). The mammalian membrane protein ER Golgi intermediate compartment protein-53 (ERGIC-53) recognizes an ER export signal in the cargo protein cathepsin Z that is composed of a combined oligosaccharide/peptide structure (12,13). ERGIC-53 is a 53-kDa type 1 membrane protein that operates as a mannose lectin cycling between the ER and the ERGIC (14–17). The cytosolic diphenylalanine motif in ERGIC-53 interacts with the COP II coat, thereby recruiting ERGIC-53 and its bound cargo to anterograde vesicles (6). A dilysine motif in the cytosolic tail of ERGIC-53 mediates retrieval back to the ER by interacting with the coat protein I (COP I) coat (18). ERGIC-53 acts as a cargo receptor for two lysosomal glycoproteins cathepsin Z and cathepsin C (12,19,20). Mutations in ERGIC-53 can lead to combined factor V and factor VIII deficiency in humans (OMIM #227300). Patients with loss of function mutations in ERGIC-53 show reduced levels of blood coagulation factors V and VIII in their plasma (21). Biochemical studies established a role of ERGIC-53 as a cargo receptor required for efficient transport of factors V and VIII in cultured mammalian cells (22). Recently, the multiple coagulation factor deficiency 2 gene (MCFD2) was identified as a second locus responsible for blood coagulation factor V and VIII deficiency (23). The MCFD2 gene encodes a soluble 16-kDa protein in the lumen of the ER. The protein possesses two EF-hand domains and interacts with ERGIC-53 in a calcium-dependent manner. Chemical cross-linking showed an interaction of factor VIII with both MCFD2 and ERGIC-53, suggesting that ERGIC-53 and MCFD2 operate as a cargo receptor complex (23,24).

Here, we have investigated the interdependence and cargo selectivity of the two subunits of the ERGIC-53/MCFD2 receptor complex by a knockdown approach using short interference RNA (siRNA). We show that ERGIC-53 is strictly required for the retention of MCFD2 in the early secretory pathway. Depletion of MCFD2 by siRNA did not affect binding of cathepsin Z and cathepsin C to ERGIC-53. The results suggest that MCFD2 is a protein specifically required for factors V and VIII but not for the lectin function of ERGIC-53.

Results

Localization of endogenous MCFD2

Overexpressed MCFD2 was previously shown to co-localize with ERGIC-53 to the ERGIC (23). To study the localization of endogenous MCFD2, immunofluorescence microscopy experiments were performed in HeLa cells. Figure 1 shows that endogenous MCFD2 co-localized with ERGIC-53 but only minimally with the two ER markers B-cell receptor associated protein 31 (BAP31) and cytoskeleton-linking membrane protein 63 (CLIMP-63). Next, we examined if MCFD2 cycles in the early secretory pathway. Cycling proteins are known to accumulate in ERGIC clusters in response to brefeldin A (BFA) (25). Brefeldin A treatment indeed led to the accumulation of MCFD2 in ERGIC-53 clusters, while the localization of the two ER resident proteins BAP31 and CLIMP-63 remained unchanged (Figure 1). These data suggest that endogenous MCFD2 localizes to the ERGIC and cycles in the early secretory pathway.

ERGIC-53 retains MCFD2 in the early secretory pathway

Do MCFD2 and ERGIC-53 form a stable complex during their entire cycling journey? There are indications that ER and ERGIC may differ in their luminal pH and ionic properties (26). It is possible, therefore, that the calcium-dependent interaction of MCFD2 and ERGIC-53 does not persist throughout the entire cycling process. Co-immunoprecipitation of MCFD2 and ERGIC-53 from BFA-treated cells (not shown) suggests that MCFD2 and ERGIC-53 interact in ERGIC clusters. Moreover, ERGIC-53-deficient lymphoblasts contain only trace amounts of intracellular MCFD2, indicating that MCFD2 requires ERGIC-53 either for stabilization or for intracellular retention (23). Experimental demonstration of a requirement for intracellular retention of MCFD2 by ERGIC-53 would directly prove a physiologically relevant post-ER interaction of the two cycling proteins. To test this, we studied the effect of siRNA-based ERGIC-53 silencing on the localization of MCFD2. Three siRNA duplexes designed against different ERGIC-53 target sequences were tested for knockdown efficiencies. The most efficient siRNA duplex reduced ERGIC-53 levels in HeLa cells to less than 10% within 3 days. In ERGIC-53-depleted cells, only trace amounts of intracellular MCFD2 could be detected by immunoblotting. Immunofluorescence microscopy confirmed that MCFD2 was no longer detectable in these cells. The effect on MCFD2 is specific because the localization of BAP31 and giantin was not affected (Figure 2A) and total secretion of ^{35}S-methionine-labeled proteins was unchanged (not shown). Although ERGIC-53 is a major protein of the ERGIC, its depletion does not seem to impair the morphology of the early secretory pathway as indicated by the normal localization of organelle markers of the early secretory pathway including numerous ER, ERGIC and Golgi proteins (not shown). An intact early secretory pathway after ERGIC-53 depletion is in line with previous observations showing that mistargeting of ERGIC-53 to the ER does not result in morphological changes of the early secretory pathway (19).

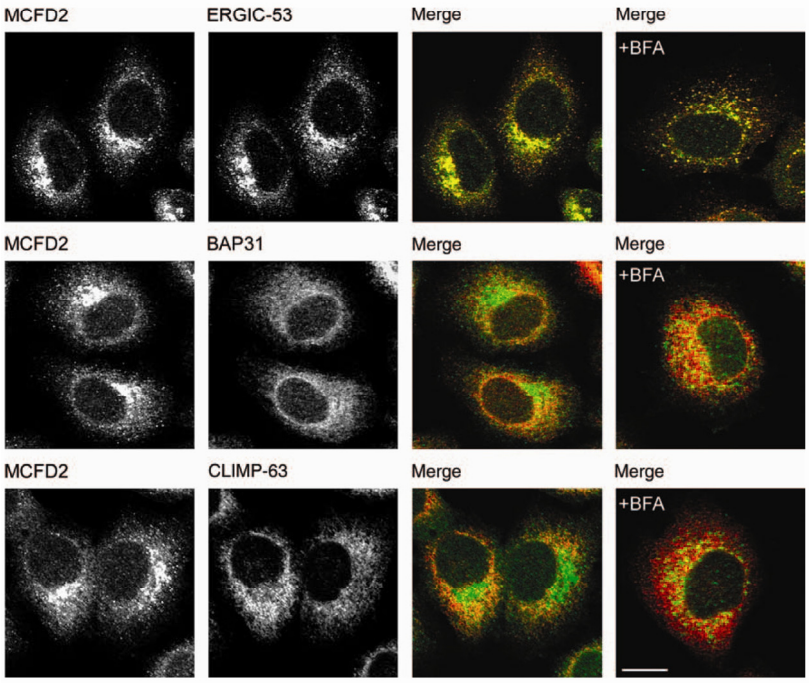


Figure 1: MCFD2 co-localizes with ERGIC-53. Localization of MCFD2 in HeLa cells visualized by double immunofluorescence microscopy using organelle marker antibodies for ERGIC (ERGIC-53, affinity-purified polyclonal antibody) and ER (BAP31 and CLIMP-63; mAbs). The cells were left untreated or incubated with brefeldin A (+BFA) for 90 min at a concentration of 10 μg/mL. The MCFD2 is shown in green, while BAP31, CLIMP-63 and ERGIC-53 are shown in red. Bar = 10 μm.

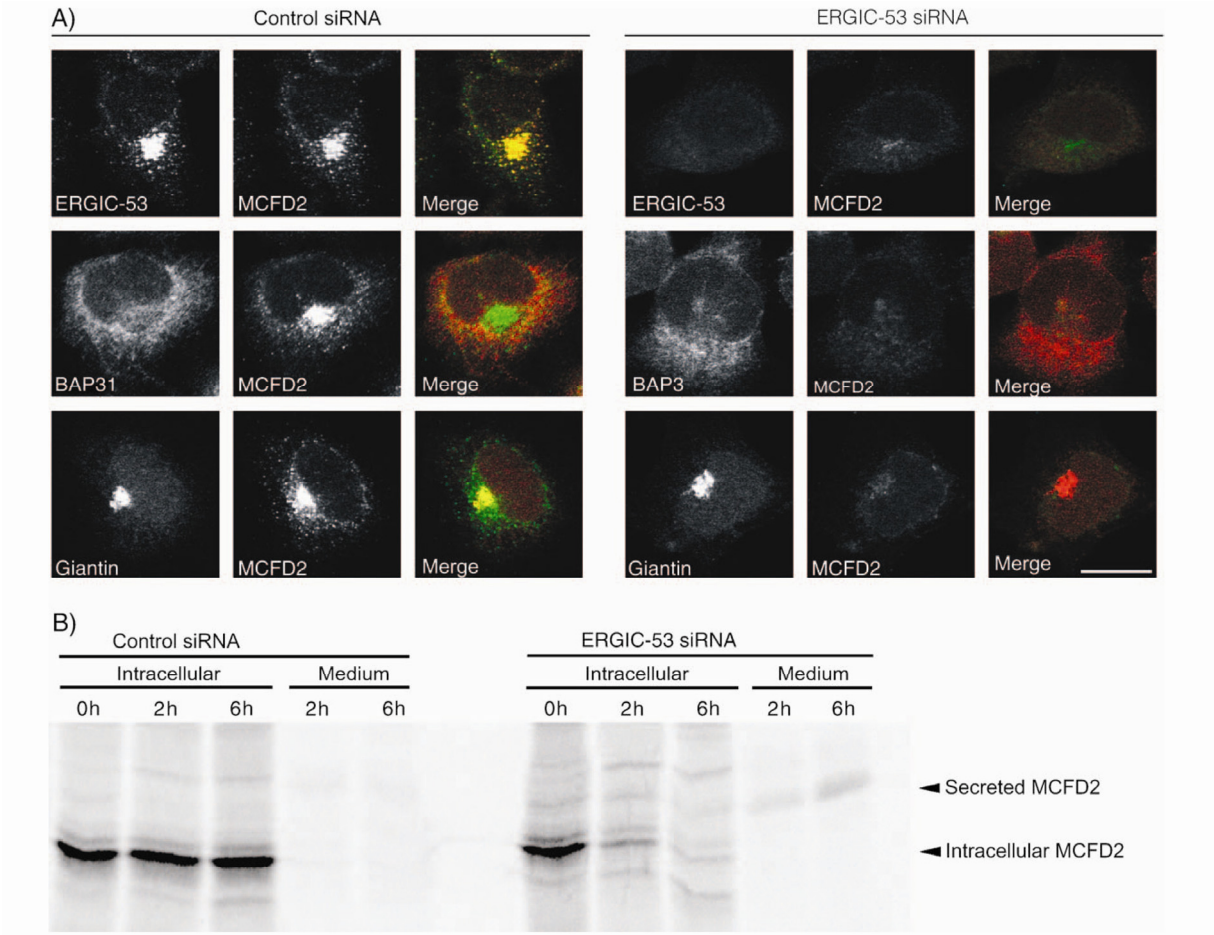


Figure 2: ERGIC-53 is required for the intracellular retention of MCFD2. A) HeLa cells were transiently transfected with control or ERGIC-53 siRNA duplexes. Seventy-two hours after transfection, MCFD2 was visualized by double immunofluorescence microscopy together with ERGIC-53, BAP31 or giantin (Golgi). In ERGIC-53-depleted HeLa cells, only trace amounts of MCFD2 are detectable. The effect on MCFD2 is specific because BAP31 and giantin are unaffected. Bar = 10 μm. B) Control and ERGIC-53 siRNA-transfected HeLa cells were pulsed with ^{35}S-methionine for 1 h and chased for the indicated times. The MCFD2 was immunoprecipitated from cell lysates and from conditioned medium. Note that depletion of ERGIC-53 leads to secretion of MCFD2.

To investigate the disappearance of MCFD2 after ERGIC-53 knockdown, pulse–chase experiments with ^{35}S-methionine were performed. Cell lysate and conditioned medium of control and ERGIC-53 siRNA-transfected HeLa cells were probed for ^{35}S-methionine-labeled MCFD2 after 2 and 6-h chase periods. In ERGIC-53-depleted cells, intracellular MCFD2 disappeared and was detected as secreted protein in the conditioned medium (Figure 2B). Secreted MCFD2 showed a higher apparent M_r than the initially synthesized protein due to *O*-glycosylation (24). *O*-glycosylation may render MCFD2 less accessible to the antibody, accounting for the only partial recovery of MCFD2 in the conditioned medium. Alternatively, a fraction of MCFD2 may be degraded rather than secreted. Secreted, *O*-glycosylated MCFD2 can also be recovered from conditioned medium on overexpression of the protein (Figure 3). The results shown in Figure 2 show that ERGIC-53 is strictly required for intracellular retention of MCFD2, indicating that MCFD2 and ERGIC-53 interact in post-ER compartments of unperturbed cells. In HeLa, COS, HepG2 and several glioblastoma cell lines, endogenous ERGIC-53 retains all MCFD2. No secreted MCFD2 could be detected in the conditioned medium (data not shown).

MCFD2 interacts with ERGIC-53 in the ER

Next, we examined if MCFD2 and ERGIC-53 can also interact in the ER. To this end, MCFD2 was localized in HeLa KKAA cells that express a dominant-negative form of ERGIC-53 that is retained in the ER by a C-terminal KKAA retention signal (19,27). In the presence of tetracycline, the expression of ERGIC-53-KKAA is repressed. Under these conditions, ERGIC-53 and MCFD2 co-localized to the ERGIC and were both BFA sensitive (Figure 4A) like in

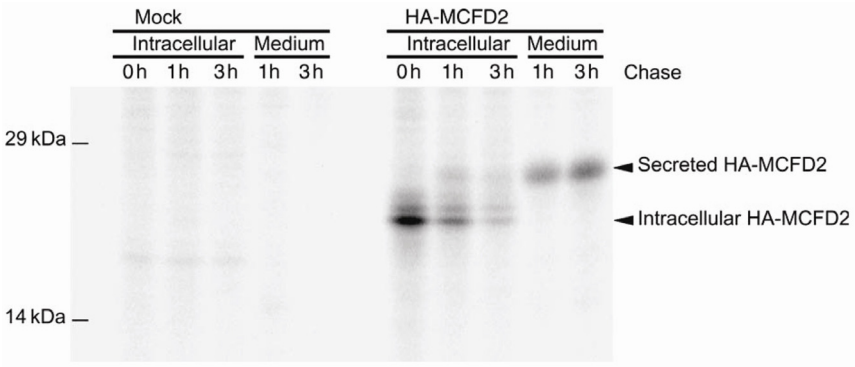


Figure 3: Overexpressed MCFD2 is secreted. HA-MCFD2-transfected and mock-transfected HeLa cells were labeled with ^{35}S-methionine for 30 min and chased for the indicated times. MCFD2 was immunoprecipitated from cell lysates and conditioned medium using polyclonal anti-MCFD2. Intracellular HA-MCFD2 disappears during the 3-h chase period and can be recovered from conditioned medium. Secreted HA-MCFD2 shows a shift in M_r due to *O*-glycosylation (24).

wild-type HeLa cells. Removal of tetracycline, which induces the expression of ERGIC-53-KKAA, mislocalized MCFD2 to the ER (Figure 4B). ER localization of MCFD2 is indicated by co-localization with BAP31 as well as resistance to BFA. The mislocalization of MCFD2 by ER-retained ERGIC-53 shows that the two proteins can also interact in the ER.

ERGIC-53 binds certain cargo glycoproteins in an MCFD2-independent fashion

Previous genetic and biochemical data showed that ERGIC-53 and MCFD2 form a cargo receptor complex, recognizing factors V and VIII (23,24). Factor VIII can be cross-linked to both ERGIC-53 and MCFD2 with similar efficiency, indicating the existence of a triple complex.

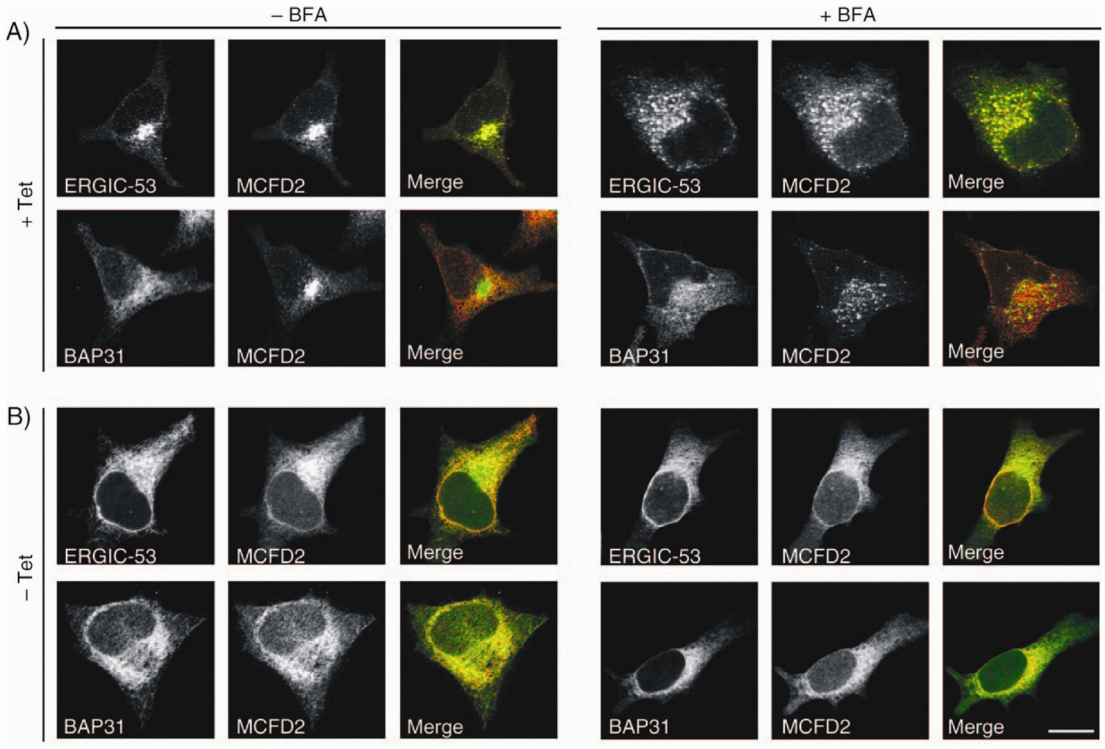


Figure 4: ERGIC-53-KKAA retains MCFD2 in the ER. MCFD2, BAP31 and ERGIC-53 were visualized by immunofluorescence microscopy in HeLa cells expressing ER-retained ERGIC-53-KKAA in a tetracycline-dependent inducible manner. Prior to fixation, the cells were treated with BFA (+BFA) (10 μg/mL for 90 min) or left untreated (−BFA). A) In the presence of tetracycline (+Tet), the expression of ERGIC-53-KKAA is repressed and the cells express only endogenous ERGIC-53. MCFD2 co-localizes with ERGIC-53 and is BFA sensitive. B) Withdrawal of tetracycline (−Tet) for 48 h induces ERGIC-53-KKAA that retains all ERGIC-53 in the ER due to a dominant-negative effect. Expression of ERGIC-53-KKAA relocalizes MCFD2 to the ER. Co-localization with BAP31 and BFA resistance show the ER retention of MCFD2. Bar = 10 μm.

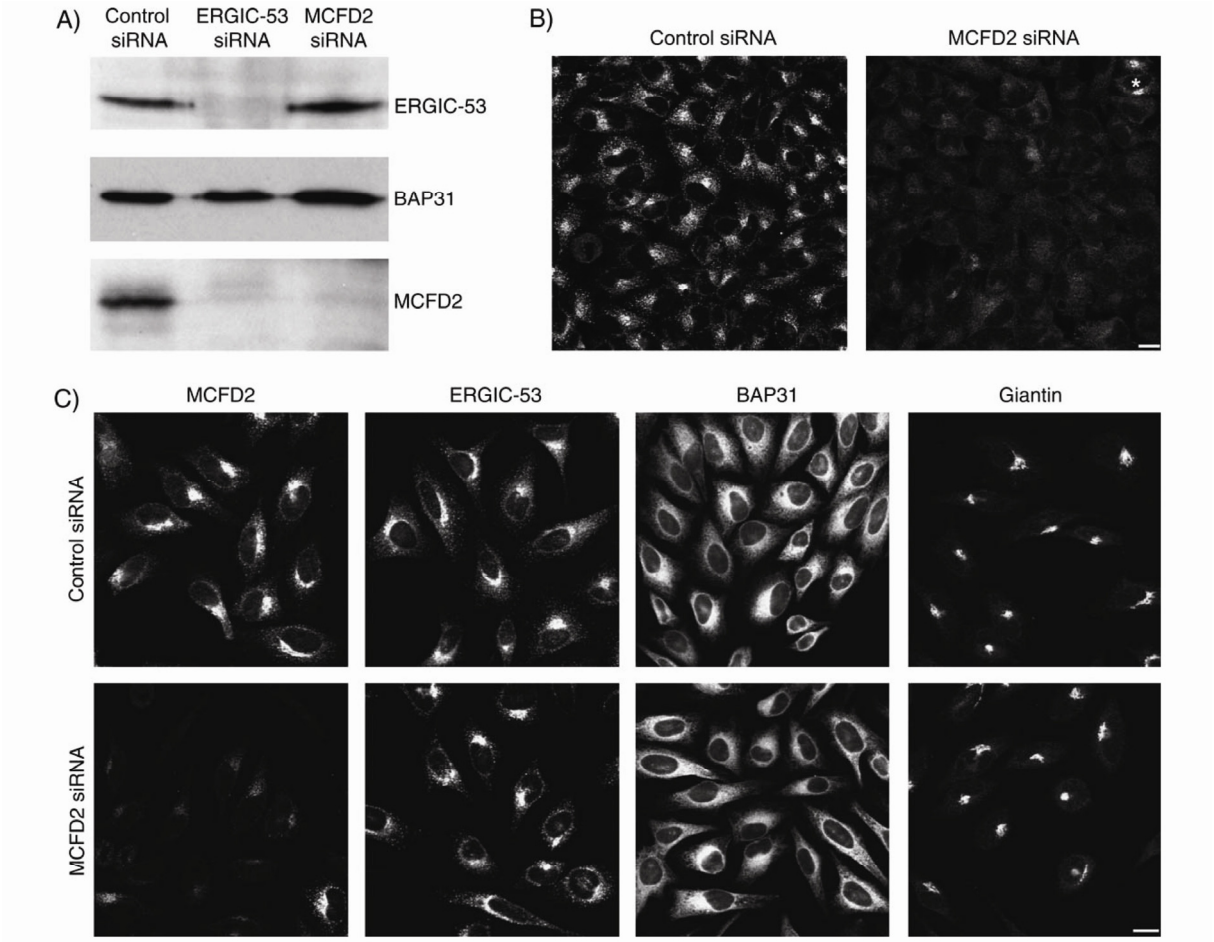


Figure 5: Depletion of MCFD2 affects neither localization nor expression of ERGIC-53. A) HeLa cells were transiently transfected with control, ERGIC-53 and MCFD2 siRNA duplexes. Seventy-two hours after transfection, ERGIC-53, BAP31 and MCFD2 were visualized by immunoblotting. Transfection of ERGIC-53 siRNA depletes both ERGIC-53 and MCFD2, whereas transfection of MCFD2 siRNA only depletes MCFD2. BAP31 levels are unaffected by ERGIC-53- and MCFD2-specific siRNA transfection. B) The very high (>95%) transfection efficiency of MCFD2 siRNA duplexes was visualized by immunofluorescence microscopy in control and MCFD2 siRNA-transfected HeLa cells 72 h after transfection. A nontransfected cell is marked (*). C) MCFD2, ERGIC-53, BAP31 and giantin were visualized by immunofluorescence microscopy in control and MCFD2 siRNA-transfected cells. Depletion of MCFD2 from HeLa cells does not change the localization pattern of ERGIC-53, BAP31 or giantin. Bar = 10 μm.

ERGIC-53 also interacts with the two lysosomal glycoproteins cathepsin Z and cathepsin C (12,19,20). To test if the MCFD2 subunit of the cargo receptor complex is required more generally for cargo binding, an siRNA-based knockdown approach was taken. Six siRNA duplexes, designed against different MCFD2 target sequences, were probed for knockdown efficiencies. The most efficient siRNA duplex reduced MCFD2 in HeLa cells to less than 10% within 3 days. Transfection efficiency of siRNA was very high as only very few cells stained positive for MCFD2 72 h after transfection (Figure 5B). Depletion of MCFD2 in HeLa cells affected neither the localization nor the protein level of ERGIC-53 (Figure 5A,C), which is consistent with the data from MCFD2-deficient lymphoblasts (23). Likewise, the localization of BAP31 and giantin was unaffected (Figure 5C).

To analyze whether cargo binding of ERGIC-53 requires MCFD2, we combined siRNA-mediated MCFD2 depletion with a recently established yellow fluorescent protein (YFP)-based protein fragment complementation assay (PCA) (20). The basic concept of the YFP PCA relies on splitting YFP into two fragments (YFP1 and YFP2) that exhibit no fluorescence by themselves. When fused to two interacting proteins, the two YFP fragments can be brought into close proximity where they can complement to functional, fluorescent YFP by folding into an active 3D structure (28). Using YFP PCA, we have previously visualized the oligomerization of ERGIC-53, its interaction with MCFD2 and its lectin-mediated interactions with cathepsin Z and cathepsin C (20). Mutagenesis of the lectin domain of ERGIC-53 selectively abolished YFP complementation

with cathepsin Z and cathepsin C (20), demonstrating that YFP PCA is a powerful technique to study specific, carbohydrate-mediated interactions between ERGIC-53 and cargo proteins *in vivo*.

Here, we applied the YFP PCA in MCFD2-depleted HeLa cells to investigate the role of MCFD2 on the oligomerization and glycoprotein-binding properties of ERGIC-53. Figure 6B shows that ERGIC-53 oligomerization and its interaction with cathepsin Z and cathepsin C are independent on the presence of MCFD2. Immunoblotting revealed equal expression of the YFP PCA constructs in control and MCFD2 siRNA-transfected cells, and silencing of MCFD2 was efficient (Figure 6C). As an internal control, the interaction between ERGIC-53 and MCFD2 was analyzed. After transfection of MCFD2-specific siRNA, no YFP complementation of ERGIC-53 and MCFD2 was observed because the expression of MCFD2-YFP2 was silenced. The results shown in Figure 6 clearly show that ERGIC-53 does not require MCFD2 to bind cathepsin Z and cathepsin C *in vivo*. Furthermore, MCFD2 is not required for ERGIC-53 oligomerization, which is consistent with data derived from MCFD2-deficient lymphoblasts (24). MCFD2-independent binding of cathepsin Z to ERGIC-53 is further supported by the fact that neither chemical cross-linking nor YFP PCA-based experiments showed a direct interaction of MCFD2 and cathepsin Z (not shown).

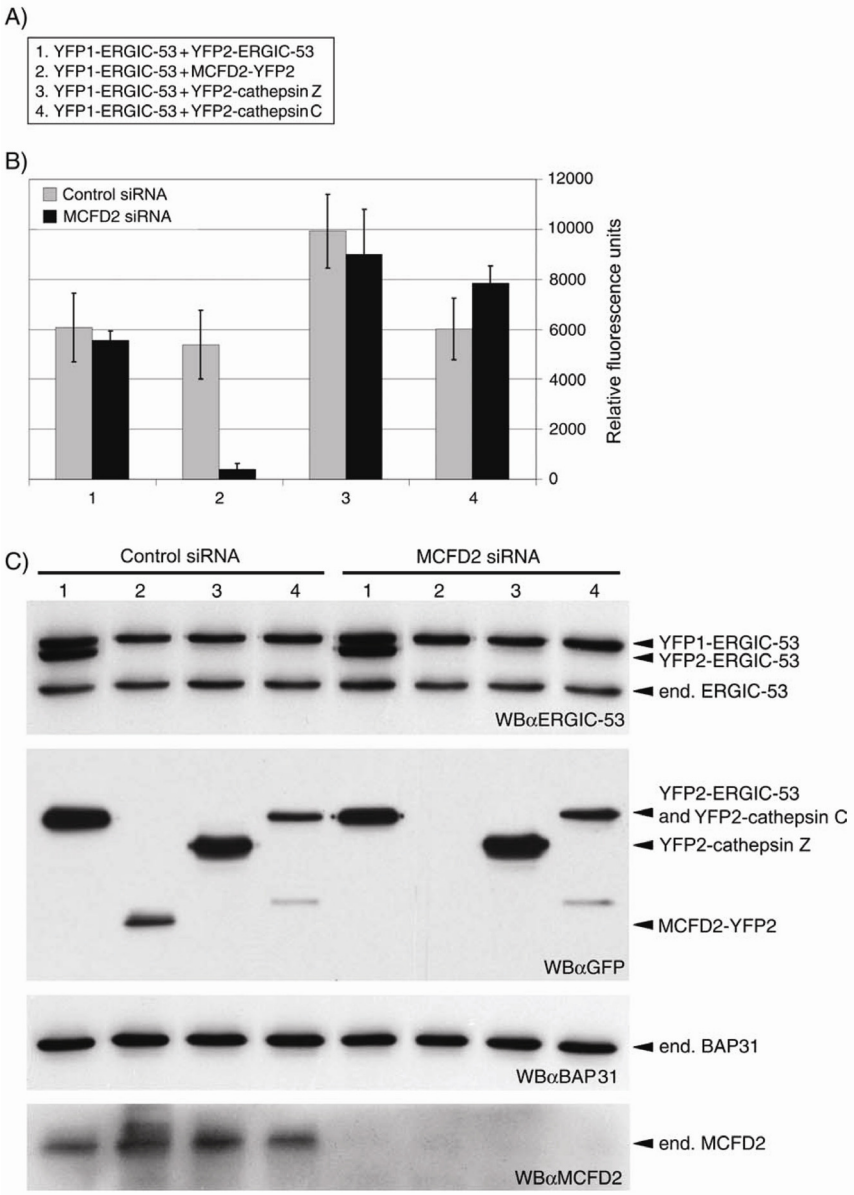


Figure 6: Yellow fluorescent protein PCA-based analysis of ERGIC-53-mediated protein interactions. A) The indicated pairs of YFP PCA constructs were expressed in control and MCFD2 siRNA-transfected HeLa cells. B) YFP fragment complementation was detected by fluorometric analysis using microtiter plates with cell suspensions of transfected HeLa cells (20). Background fluorescence of mock-transfected cells was subtracted, and the relative fluorescence units are shown. Data from three independent experiments were averaged and the error bars represent the standard deviation. MCFD2 siRNA affects neither the ERGIC-53 oligomerization nor the interactions between ERGIC-53 and cathepsin Z or cathepsin C. C) Lysates of fluorometrically analyzed cells were probed by Western blotting (WB) using mAbs against ERGIC-53, GFP, BAP31 and MCFD2. Equal expression of the YFP PCA constructs in control and MCFD2 siRNA-transfected cells is visualized by immunoblotting using anti-ERGIC-53 and anti-GFP antibodies. Equal protein amounts are present in all lanes as revealed by similar levels of endogenous ERGIC-53 and BAP31. MCFD2 siRNA depletes endogenous MCFD2 and silences the expression of the MCFD2-YFP2 construct.

Discussion

Receptor-mediated export of secretory proteins from the ER is a universal feature of eukaryotic cells; yet, the diversity of this process is still largely unknown. Conceivably, receptor-mediated ER export is required to increase transport efficiency of less abundant secretory proteins or for secretory cargo undergoing a second quality control step (2). Receptor-mediated ER export can also account for temporal and spatial co-ordination of the secretion of specific signaling molecules (29). The molecular basis underlying cargo receptor interactions is beginning to emerge although only few cases have been studied (11,13). ERGIC-53 and MCFD2 are particularly interesting because they are both required for the efficient secretion of blood coagulation factors V and VIII and constitute a cargo receptor complex composed of two subunits (23,24,30).

Here, we have described a detailed characterization of the ERGIC-53/MCFD2 receptor complex using a combination of immunofluorescence, gene silencing and PCA approaches to provide insight into the interdependence and cargo selectivity of the two subunits. Endogenous MCFD2 co-localized with ERGIC-53 in untreated and BFA-treated HeLa cells and thus shows characteristic features of a protein cycling between the ERGIC and the ER. Remarkably, both cycling and intracellular retention of MCFD2 are strictly dependent on ERGIC-53. ER-retained ERGIC-53-KKAA mislocalized MCFD2 to the ER, indicating that MCFD2 can bind to ERGIC-53 in the ER and requires ERGIC-53 for anterograde transport. Depletion of ERGIC-53 by siRNA resulted in the secretion of MCFD2; hence, MCFD2 requires ERGIC-53 also for retrograde transport back to the ER. We conclude that ERGIC-53 and MCFD2 form a stable complex and cycle together in the early secretory pathway due to the cytosolic diphenylalanine and dilysine motifs in ERGIC-53. In COS cells, endogenous ERGIC-53 and MCFD2 have similarly long half-lives (>26 h) (24), which is in line with a stable complex of the two proteins.

ERGIC-53-deficient patients have only trace amounts of intracellular MCFD2 (23). Our results provide now an explanation for this lack of MCFD2 by showng MCFD2 secretion on ERGIC-53 depletion in cell culture. Most soluble secretory proteins that localize to the early secretory pathway of mammalian cells carry a C-terminal KDEL tetrapeptide motif and are retained by binding to the KDEL receptor (31). This receptor captures the KDEL proteins in ERGIC and *cis* Golgi and recycles them back to the ER, providing a general mechanism for protein retention early in the secretory pathway (32). In contrast, MCFD2 is retained by a specific interaction with ERGIC-53, exemplifying a KDEL-receptor-independent retention mechanism in the early secretory pathway.

Overexpression of MCFD2 also results in its secretion, suggesting that the protein is secreted as soon as the retention capacity of ERGIC-53 is saturated. This finding may have physiological implications in view of a recent report proposing that rat MCFD2 (formerly termed stem cell-derived neuronal stem cell supporting factor (SDNSF)) can act as an autocrine or paracrine factor in maintaining stem cell potential and neurogenesis in the adult central nervous system (33). In HeLa, COS, HepG2 and several glioblastoma cell lines, however, we could not detect any secreted endogenous MCFD2, suggesting that endogenous ERGIC-53 is capable of retaining all MCFD2 in these cells, although secretion of minute amounts below the detection limit cannot be excluded. ERGIC-53 and MCFD2 levels are induced in response to cellular stress (34,35). This raises the intriguing possibility that abundant MCFD2 can be secreted and act in some signaling events under certain conditions.

Factor VIII can be cross-linked to both MCFD2 and ERGIC-53, arguing for a triple complex early in the secretory pathway (24). Interestingly, factor VIII can also be cross-linked to the MCFD2 D129E mutant that is unable to interact with ERGIC-53 (24). Thus, MCFD2 interacts with factor VIII in an ERGIC-53-independent manner. To determine whether MCFD2 is a general cargo recruitment factor for ERGIC-53, we tested a putative interaction of MCFD2 with cathepsin Z or cathepsin C. Although serving as a cargo for ERGIC-53, cathepsin Z could not be directly cross-linked to MCFD2. Likewise, protein interaction studies using YFP PCA could not reveal an interaction between MCFD2 and cathepsin Z or cathepsin C, arguing against a role of MCFD2 in the recruitment of these cargo proteins. To confirm these findings, the interaction between ERGIC-53 and cathepsin Z or cathepsin C was studied after MCFD2 knockdown. YFP PCA showed that the interaction between ERGIC-53 and cathepsin Z or cathepsin C in living cells is not affected by the depletion of MCFD2. Binding of cathepsin Z and cathepsin C to ERGIC-53 is unlikely to be mediated by residual MCFD2. First, siRNA-mediated knockdown is very efficient, and hardly any residual MCFD2 is detectable by immunoblot (Figures 5A and 6C). Second, immunofluorescence-based analysis of siRNA transfection efficiency (Figure 5B) shows that residual MCFD2 derives from only very few apparently nontransfected cells, while the big majority of cells is entirely depleted of MCFD2. Third, in the YFP PCA-based protein interaction analysis, endogenous ERGIC-53 and YFP1-ERGIC-53 compete for residual MCFD2, which amplifies the MCFD2 depletion effect. In support of these *in vivo* data, ERGIC-53 can bind to immobilized mannose *in vitro* when purified without MCFD2 in the absence of calcium (26).

The results of our study indicate that MCFD2 is not required for the lectin-mediated binding of certain glycoproteins to ERGIC-53 in intact cells. We propose that the MCFD2/ERGIC-53 receptor complex possesses dual-binding properties. ERGIC-53 binds glycoproteins, such as

cathepsin Z or cathepsin C, in an MCFD2-independent fashion. Conversely, MCFD2 interacts with factors V and VIII in an ERGIC-53-independent manner. Binding of MCFD2 to ERGIC-53 recruits factors V and VIII to ERGIC-53, thereby ensuring efficient ER export. In support of this model, unglycosylated factor VIII but not unglycosylated cathepsin Z can be cross-linked to the MCFD2/ERGIC-53 complex (12,24). Recruitment of specific cargo molecules by a luminal subunit of a receptor complex adds another layer of complexity to receptor-mediated ER export. It will be interesting in the future to elucidate how MCFD2 recruits factors V and VIII to ERGIC-53. This will require new methodology as we were unable to visualize this process by YFP PCA. Mechanistic insight into this recruitment process will be important to understand how soluble secretory proteins can be captured for ER export.

Materials and Methods

Antibodies

The following antibodies were used: mouse monoclonal antibody (mAb) G1/93 against human ERGIC-53 (14) (ALX-804-602; Alexis, Lausen, Switzerland), mouse mAb against human MCFD2 (23), goat polyclonal antibody (pAb) against human MCFD2 (R&D Systems, Minneapolis, MN, USA), rabbit pAb against human MCFD2 (23), mouse mAb A1/182 against BAP31 (16) (ALX-804-601; Alexis), mouse mAb G1/133 against giantin (36) (ALX-804-600; Alexis), mouse mAb against hemagglutinin (HA; Covance, Princeton, NJ, USA) and mouse mAb against green fluorescent protein (GFP; Roche Applied Science, Basel, Switzerland).

Cell culture

HeLa cells (CCL-2; ATCC, Manassas, VA, USA) were grown in DMEM, supplemented with 10% fetal bovine serum, 1× nonessential amino acids and antibiotics. HeLa KKAA cells were cultured as described previously (19). For fluorometric analysis and metabolic labeling, cells were grown in six-well plates. For fluorescence microscopy cells, were grown on poly-L-lysine-coated glass slides.

siRNA transfection

siRNA oligos were purchased from Qiagen (Venlo, The Netherlands) and Eurogentec (Seraing, Belgium). Three and six siRNA oligos were designed against ERGIC-53 and MCFD2, respectively. The most efficient siRNA oligo was determined by immunoblotting and chosen for all further experiments. ERGIC-53 was knocked down using 5′-GGACAGAAUCGUAUUCAUCdTdT-3′ as sense and 5′-GAUGAAUACGAUUCUGUCCdTdT-3′ as antisense oligo. The MCFD2 was knocked down using 5′-AGAAGGUGUCAUCAACAAAdTdT-3′ as sense and 5′-UUUGUUGAUGACACCUUCUdAdG-3′ as antisense oligo. Nonsilencing control siRNA was purchased from Qiagen. A final siRNA concentration of 5 nM was used for transfection directly after cell plating using HiPerFect (Qiagen) according to the manufacturer's instructions.

DNA transfection

Cloning of pcDNA3[YFP1-p53], pcDNA3[YFP2-p53], pcDNA3[MCFD2-YFP2], pcDNA3[YFP2-catZ] and pcDNA3[YFP2-catC] was described previously (13,20). pcDNA3[HA-MCFD2] was generated by inserting HA-MCFD2 without its signal sequence (generated by polymerase chain reaction amplification) into the pcDNA3 vector containing the artificial signal sequence of calreticulin (20). DNA constructs were transfected with FuGENE6 (Roche Applied Science) according to the manufacturer's instructions. For siRNA and DNA co-transfection, DNA constructs were transfected 24 h after cell plating and siRNA transfection. Cells were analyzed 48 h after DNA transfection, which corresponds to 72 h after siRNA transfection.

Immunofluorescence microscopy

Brefeldin A (10 μg/mL; Epicentre, Madison, WI, USA) was added to the cells 90 min prior to fixation. Cells were fixed in 3% para-formaldehyde and permeabilized for 5 min in PBS containing 3% BSA and 0.2% Triton-X-100. Primary antibodies were added for 30 min in PBS containing 3% BSA. After rinsing, the secondary antibodies conjugated with Alexa Fluor 488 or Alexa Fluor 568 (Molecular Probes, Leiden, The Netherlands) were added for 30 min. Cells were washed in PBS, embedded and analyzed by laser scanning confocal microscopy (TCS NT; Leica, Wetzlar, Germany). For indirect immunofluorescence using mouse mAb against human MCFD2, 0.5% SDS and 5% β-mercaptoethanol was added during permeabilization step (37). The use of mouse immunoglobulin G (IgG) 1 (for MCFD2) and mouse IgG2a (for BAP31 and CLIMP-63) specific secondary antibodies allowed double staining of MCFD2 with BAP31 or CLIMP-63.

Immunoblotting

Protein samples were prepared as described previously (20), separated by SDS–PAGE; transferred to nitrocellulose membranes; immunoblotted with anti-ERGIC-53, anti-BAP31, anti-MCFD2 and anti-GFP and visualized by enhanced chemiluminescence (Amersham Bioscience, Uppsala, Sweden).

Metabolic labeling

Cells were deprived of L-methionine for 20 min, pulsed for 60 min with 100 μCi ^{35}S-methionine (Perkin Elmer, Wellesley, MA, USA) and chased for the indicated times in HeLa culture medium containing 10 mM L-methionine. Cells were lysed in 1% Triton-X-100, 50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 2 mM $CaCl_2$ and phenylmethylsulfonylfluoride (PMSF), and the lysate was cleared by centrifugation at 100 000 × ***g*** for 1 h. The chase medium was cleared from cell debris by centrifugation at 20 000 × ***g*** for 5 min. Cleared samples were immunoprecipitated with anti-MCFD2. Immunoprecipitates were separated by SDS–PAGE, and radiolabeled bands were imaged using a phosphorimager (Molecular Dynamics, Sunnyvale, CA, USA).

Yellow fluorescent protein fluorometric analysis

Fluorometric analysis was performed as described previously (20). Data from three independent experiments were averaged.

Acknowledgments

We thank K. Bucher for expert technical assistance. B. Z. is a recipient of a Career Development Award from the National Hemophilia Foundation, USA. D. G. and R. J. K. are Howard Hughes Medical Institute Investigators. This work was supported by the University of Basel, the Swiss National Science Foundation (H.-P. H) and the National Institutes of Health grants PO1 HL057346 (D. G.) and HL052173 (R. J. K.).

References

1. Lee MC, Miller EA, Goldberg J, Orci L, Schekman R. Bi-directional protein transport between the ER and Golgi. Annu Rev Cell Dev Biol 2004;20:87–123.
2. Ellgaard L, Molinari M, Helenius A. Setting the standards: quality control in the secretory pathway. Science 1999;286:1882–1888.
3. Miller EA, Beilharz TH, Malkus PN, Lee MC, Hamamoto S, Orci L, Schekman R. Multiple cargo binding sites on the COPII subunit Sec24p ensure capture of diverse membrane proteins into transport vesicles. Cell 2003;114:497–509.
4. Mancias JD, Goldberg J. Exiting the endoplasmic reticulum. Traffic 2005;6:278–285.

5. Kappeler F, Klopfenstein DR, Foguet M, Paccaud JP, Hauri HP. The recycling of ERGIC-53 in the early secretory pathway. ERGIC-53 carries a cytosolic endoplasmic reticulum-exit determinant interacting with COPII. J Biol Chem 1997;272:31801–31808.
6. Nufer O, Guldbrandsen S, Degen M, Kappeler F, Paccaud JP, Tani K, Hauri HP. Role of cytoplasmic C-terminal amino acids of membrane proteins in ER export. J Cell Sci 2002;115:619–628.
7. Warren G, Mellman I. Bulk flow redux? Cell 1999;98:125–127.
8. Wieland FT, Gleason ML, Serafini TA, Rothman JE. The rate of bulk flow from the endoplasmic reticulum to the cell surface. Cell 1987; 50:289–300.
9. Kuehn MJ, Herrmann JM, Schekman R. COPII-cargo interactions direct protein sorting into ER-derived transport vesicles. Nature 1998;391: 187–190.
10. Belden WJ, Barlowe C. Role of Erv29p in collecting soluble secretory proteins into ER-derived transport vesicles. Science 2001;294: 1528–1531.
11. Otte S, Barlowe C. Sorting signals can direct receptor-mediated export of soluble proteins into COPII vesicles. Nat Cell Biol 2004;6:1189–1194.
12. Appenzeller C, Andersson H, Kappeler F, Hauri HP. The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. Nat Cell Biol 1999; 1:330–334.
13. Appenzeller-Herzog C, Nyfeler B, Burkhard P, Santamaria I, Lopez-Otin C, Hauri HP. Carbohydrate- and conformation-dependent cargo capture for ER-exit. Mol Biol Cell 2005;16:1258–67.
14. Schweizer A, Fransen JA, Bachi T, Ginsel L, Hauri HP. Identification, by a monoclonal antibody, of a 53-kD protein associated with a tubulo-vesicular compartment at the cis-side of the Golgi apparatus. J Cell Biol 1988;107:1643–1653.
15. Itin C, Roche AC, Monsigny M, Hauri HP. ERGIC-53 is a functional mannose-selective and calcium-dependent human homologue of leguminous lectins. Mol Biol Cell 1996;7:483–493.
16. Klumperman J, Schweizer A, Clausen H, Tang BL, Hong W, Oorschot V, Hauri HP. The recycling pathway of protein ERGIC-53 and dynamics of the ER-Golgi intermediate compartment. J Cell Sci 1998;111: 3411–3425.
17. Ben-Tekaya H, Miura K, Pepperkok R, Hauri HP. Live imaging of bidirectional traffic from the ERGIC. J Cell Sci 2005;118:357–367.
18. Itin C, Schindler R, Hauri HP. Targeting of protein ERGIC-53 to the ER/ERGIC/cis-Golgi recycling pathway. J Cell Biol 1995;131:57–67.
19. Vollenweider F, Kappeler F, Itin C, Hauri HP. Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme. J Cell Biol 1998;142:377–389.
20. Nyfeler B, Michnick SW, Hauri HP. Capturing protein interactions in the secretory pathway of living cells. Proc Natl Acad Sci U S A 2005; 102:6350–6355.
21. Nichols WC, Seligsohn U, Zivelin A, Terry VH, Hertel CE, Wheatley MA, Moussalli MJ, Hauri HP, Ciavarella N, Kaufman RJ, Ginsburg D. Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. Cell 1998;93:61–70.
22. Moussalli M, Pipe SW, Hauri HP, Nichols WC, Ginsburg D, Kaufman RJ. Mannose-dependent endoplasmic reticulum (ER)-Golgi intermediate compartment-53-mediated ER to Golgi trafficking of coagulation factors V and VIII. J Biol Chem 1999;274:32539–32542.
23. Zhang B, Cunningham MA, Nichols WC, Bernat JA, Seligsohn U, Pipe SW, McVey JH, Schulte-Overberg U, de Bosch NB, Ruiz-Saez A, White GC, Tuddenham EG, Kaufman RJ, Ginsburg D. Bleeding due to disruption of a cargo-specific ER-to-Golgi transport complex. Nat Genet 2003;34:220–225.
24. Zhang B, Kaufman RJ, Ginsburg D. LMAN1 and MCFD2 form a cargo receptor complex and interact with coagulation factor VIII in the early secretory pathway. J Biol Chem 2005;280:25881–25886.
25. Breuza L, Halbeisen R, Jeno P, Otte S, Barlowe C, Hong W, Hauri HP. Proteomics of endoplasmic reticulum-Golgi intermediate compartment (ERGIC) membranes from brefeldin A-treated HepG2 cells identifies ERGIC-32, a new cycling protein that interacts with human Erv46. J Biol Chem 2004;279:47242–47253.
26. Appenzeller-Herzog C, Roche AC, Nufer O, Hauri HP. pH-induced conversion of the transport lectin ERGIC-53 triggers glycoprotein release. J Biol Chem 2004;279:12943–12950.
27. Andersson H, Kappeler F, Hauri HP. Protein targeting to endoplasmic reticulum by dilysine signals involves direct retention in addition to retrieval. J Biol Chem 1999;274:15080–15084.
28. Kerppola TK. Visualization of molecular interactions by fluorescence complementation. Nat Rev Mol Cell Biol 2006;7:449–56.
29. Bokel C, Dass S, Wilsch-Brauninger M, Roth S. Drosophila Cornichon acts as cargo receptor for ER export of the TGFalpha-like growth factor Gurken. Development 2006;133:459–470.
30. Zhang B, McGee B, Yamaoka JS, Guglielmone H, Downes KA, Minoldo S, Jarchum G, Peyvandi F, de Bosch NB, Ruiz-Saez A, Chatelain B, Olpinski M, Bockenstedt P, Sperl W, Kaufman RJ et al. Combined deficiency of factor V and factor VIII is due to mutations in either LMAN1 or MCFD2. Blood 2006;107:1903–1907.
31. Lewis MJ, Pelham HR. A human homologue of the yeast HDEL receptor. Nature 1990;348:162–163.
32. Lewis MJ, Pelham HR. Ligand-induced redistribution of a human KDEL receptor from the Golgi complex to the endoplasmic reticulum. Cell 1992;68:353–364.
33. Toda H, Tsuji M, Nakano I, Kobuke K, Hayashi T, Kasahara H, Takahashi J, Mizoguchi A, Houtani T, Sugimoto T, Hashimoto N, Palmer TD, Honjo T, Tashiro K. Stem cell-derived neural stem/progenitor cell supporting factor is an autocrine/paracrine survival factor for adult neural stem/progenitor cells. J Biol Chem 2003;278:35491–35500.
34. Nyfeler B, Nufer O, Matsui T, Mori K, Hauri HP. The cargo receptor ERGIC-53 is a target of the unfolded protein response. Biochem Biophys Res Commun 2003;304:599–604.
35. Spatuzza C, Renna M, Faraonio R, Cardinali G, Martire G, Bonatti S, Remondelli P. Heat shock induces preferential translation of ERGIC-53 and affects its recycling pathway. J Biol Chem 2004;279:42535–42544.
36. Linstedt AD, Hauri HP. Giantin, a novel conserved Golgi membrane protein containing a cytoplasmic domain of at least 350 kDa. Mol Biol Cell 1993;4:679–693.
37. Blum R, Pfeiffer F, Feick P, Nastainczyk W, Kohler B, Schafer KH, Schulz I. Intracellular localization and in vivo trafficking of p24A and p23. J Cell Sci 1999;112:537–548.

4. Conclusions and Perspectives

4.1 YFP PCA-based analysis of protein interaction

4.1.1 Proof of concept of luminal YFP PCA

In this thesis, I addressed the question of whether YFP protein fragment complementation is a suitable tool to analyze protein-protein interactions in the lumen of the secretory pathway by studying ERGIC-53 and its luminal interaction partners[1]. The results clearly demonstrate the feasibility and simplicity of the luminal YFP PCA in visualizing the homo-oligomerization of ERGIC-53 as well as its interaction with cathepsin Z, cathepsin C and MCFD2. Thus, the luminal YFP PCA is applicable to protein interactions involving both integral membrane and soluble proteins. Additionally, the specificity of the luminal YFP PCA was demonstrated by showing that a mutation of the ERGIC-53 lectin domain selectively decreased the carbohydrate-mediated interactions with cathepsin Z and cathepsin C. Furthermore, the luminal YFP PCA was successfully applied to establish the cargo selectivity of the ERGIC-53/MCFD2 receptor complex *in vivo*. By using a YFP PCA-based quantification in MCFD2-depleted cells, ERGIC-53 was shown to bind glycoproteins in a MCFD2-independent manner. Finally, the potential of the luminal YFP PCA to identify novel protein-protein interactions was demonstrated in a small scale cDNA library screen. In a pilot screen, alpha-1-antitrypsin was identified from a cDNA-YFP1 fusion library as potential interaction partner of ERGIC-53. With the development of this YFP complementation-based cDNA library screening procedure, a firm basis is provided to search for luminal protein interaction partners in a genomewide setting. Collectively, these results clearly demonstrate that YFP PCA is an attractive tool to study and screen protein-protein interactions inside the secretory pathway.

4.1.2 Evaluation of fluorescence complementation

Fluorescent protein fragment complementation has been applied to many different protein interactions since its discovery[2, 3]. A key attraction of a fluorescent protein as PCA reporter is the possibility to visualize protein interactions in living cells and to determine the intracellular localization of functional protein complexes. This is of importance since a protein might be distributed throughout the cell while performing its biological function only in a specific subcellular compartment due to restrictive localization of interacting proteins. While immunofluorescence-based approaches

stain the total pool of a protein, YFP PCA can be used to selectively visualize the fraction of functional protein found in a complex. This is relevant for many luminal protein interactions occurring in the process of protein folding and modification. Grp78, for example, binds just the non-native fraction of a protein which constitutes only a small subfraction of the total protein and is difficult to analyze by conventional methods.

By combining fragments of green, yellow, cyan and blue fluorescent proteins, multicolor fluorescence complementation assays were developed that can visualize multiple protein interactions in a single cell[4]. Importantly, multicolor fluorescence complementation can be used to examine competitive protein interactions in complexes where several proteins compete for the same interaction partner[5].

Like every method, YFP PCA harbors not only advantages. The properties of fluorescence complementation are determined by the folding kinetics and the stability of GFP and its colour variants. GFP folds into an 11-stranded β-barrel containing a co-axial α-helix[6]. Once assembled, this structure is very stable making fluorescence complementation an essentially irreversible process[7]. Irreversibility facilitates the capture of transient protein interactions but complicates the analysis of dissociation processes in protein complexes. For example, one can think of two transcription factors which interact in the cytosol and translocate together into the nucleus where they dissociate again. In a fluorescence complementation-based analysis, the stabilization of the interaction between the two transcription factors could result in a false positive fluorescence signal in the nucleus. Therefore, the consequences of trapping interaction partners by association of the fluorescent protein fragments have to be analyzed in detail. In the case of ERGIC-53, the dissociation of cathepsin Z was prevented after YFP fragment complementation[1] and the process of cargo protein release could not be studied by YFP PCA. For releasing cathepsin Z, ERGIC-53 was proposed to undergo a pH-induced conformational change[8] which might not be powerful enough to dissociate the YFP fragments after complementation. To address the dynamic features of protein interactions, a fluorescence resonance energy transfer (FRET)-based approach should be used[9, 10]. FRET relies on the transfer of energy, in form of excitation energy, between fluorescent donor and acceptor molecules. To study protein-protein interaction by FRET, a pair of interacting proteins is tagged with two spectrally distinct full-length fluorescent proteins whose fluorescence intensities and lifetimes are measured. Since FRET strictly depends on

the distance between the donor and acceptor fluorescent protein and since full-length fluorescent proteins interact only minimally with each other[11], a FRET-based approach can readily determine the dynamic behavior of protein complexes in real-time[12].

Protein fragment complementation probes binary interactions between recombinant fusion proteins. The need to express fusion proteins has two major drawbacks, namely false negative hits caused by the steric hindrance of the fusion proteins and false positive hits due to overexpression of the interaction partners[13]. Depending on the localization of the tag, complementation may be hindered due to distance or orientation constraints. Peptide linkers between the proteins and the reporter fragments minimize but do not completely avoid the problem of steric hindrance. Thus, a 'non-interaction' prediction for a protein pair based on protein fragment complementation has to be further analyzed in terms of steric constraints. A possible solution to the problem of steric hindrance might be a transposon-based random tagging strategy. Random tagging of the glutamate receptor subunit GluR1 with a transposon that contained GFP was reported to be an efficient way to select functional GFP insertion sites[14]. By taking advantage of *in vitro* transposition, a YFP fragment 1 or 2 containing transposon could be used to generate a whole population of randomly tagged proteins. After co-expression of this tagged protein population with its potential interaction partner, a FACS-based analysis may reveal if the population contains complementation-competent fusion proteins. Since *in vitro* transposition generates a huge diversity of different fusion proteins[14], sterical constraints should be overcome.

The problem of false positive hits due to the overexpression of interaction partners is a common problem of many methods in interaction proteomics. For fluorescence complementation, the problem of overexpression artefacts is further enhanced by the intrinsic tendency of the fluorescent protein fragments to self-associate[3]. In yeast, recombinant DNA can be introduced into the genome by homologous recombination[15]. This allows the expression of recombinant proteins from their endogenous promoter at physiological levels. Mammalian cells are not easily amenable to homologous recombination and require viral- or plasmid-based expression of the recombinant proteins which often results in expression levels much higher than those of the corresponding endogenous proteins. Nevertheless, overexpression artefacts can be minimized by using expression vectors with weak promoters, by optimizing the transfection conditions in terms of transfection reagent

and transfected DNA amount and by controlling the expression time. An expression level comparable to the endogenous protein is desirable and can be assessed by Western blotting. In the case of ERGIC-53, the luminal YFP PCA visualized interaction partners of ERGIC-53 in a specific manner since the YFP1-p53 and YFP2-p53 bait proteins were expressed at a moderate level similar to that of endogenous ERGIC-53.

4.1.3 Future prospects

What are the perspectives for the luminal YFP PCA? It will be exciting to search for ERGIC-53 interacting proteins in a genomewide setting by scaling up the fluorescence complementation-based cDNA library screen using YFP2-ERGIC-53 as bait protein. Also other luminal bait proteins may be used since fluorescence complementation-based cDNA library screening can be generally applied to luminal proteins. The animal lectin vesicular integral-membrane protein 36 (VIP-36) is an attractive candidate. Like ERGIC-53, VIP-36 is proposed to act as a cargo receptor but not a single interaction partner has been identified so far[16-18].

In another promising approach, one could try to map the multiprotein network of ER resident proteins[19, 20]. For this purpose, important players in the ER folding and modification process would be tagged with YFP fragments 1 or 2 and probed against each others. To avoid overexpression artefacts, this approach should be performed in yeast cells taking advantage of homologous recombination. A detailed mapping of the ER multiprotein network might result in a more system-based view of protein folding and modification in the ER.

Since fluorescence complementation is an irreversible process, dynamic properties of protein interactions can not be reliably studied. To analyze the numerous transient protein interactions in the lumen of the secretory pathway in real time, a reversible PCA would be of exceptional importance. By destabilizing the protein interaction interface between the two PCA reporter fragments by systematic amino acid mutations, a reversible PCA might be developed. Since fluorescent proteins fold into a very stable β-barrel structure, other PCA reporter proteins such as *Renilla* luciferase should be preferentially used for the development of a reversible PCA. Once established, a reversible PCA would also allow the screening of compound libraries for their potential to modulate protein interactions. With its increasing implications in diabetes, neurodegeneration and cancer[21], UPR signaling

displays an attractive target for a compound library screen. Bortezomib is an example of a chemical compound that links ER stress signaling with cancer therapy[22]. Bortezomib is a proteasome inhibitor, which enhances ER stress but at the same time seems to inhibit PERK-mediated UPR signaling and render cancer cells more sensitive to chemotherapy[23, 24]. The mammalian UPR is induced by ER stress-mediated dimerization of the transmembrane stress sensors IRE1 and PERK and by ER stress-regulated ER to Golgi trafficking of ATF6[25]. A reversible PCA that monitors the dimerization of IRE1 or PERK might provide a basis to screen for additional chemical compounds which modulate UPR signaling thereby identifying novel potential drug candidates.

4.2 Receptor-mediated ER export

4.2.1 Characterization of the ERGIC-53/MCFD2 protein complex

ERGIC-53 and MCFD2 are both required for the efficient secretion of blood coagulation factors V and VIII[26, 27]. Since ERGIC-53 and MCFD2 interact not only with each other but also with blood coagulation factor VIII, the existence of a cargo receptor complex composed of a transmembrane and a soluble protein was proposed[28]. This raises the intriguing possibility that a luminal protein interacts with cargo proteins and recruits them to a transmembrane receptor for ER export. To gain further insight into cargo receptor complexes, the ERGIC-53/MCFD2 protein complex was characterized in this thesis with regard to subunit interdependence, localization, and cargo binding properties. The data strongly suggest that ERGIC-53 and MCFD2 form a stable protein complex since the endogenous proteins cycle in the early secretory pathway and interact in both ER and ERGIC. Noteworthy, MCFD2 depends on ERGIC-53 for intracellular retention. How does MCFD2 bind to ERGIC-53? Pulldown experiments revealed that MCFD2 requires the presence of calcium and two intact EF-hand motifs for its interaction with ERGIC-53. EF-hand motifs are calcium binding helix-loop-helix protein structures which can be found in more than 200 different calcium binding proteins[29]. Most of these proteins are cytosolic albeit some EF-hand containing luminal proteins are known. The luminal proteins Cab45, reticulocalbin, Erc-55 and calumenin contain multiple EF-hand domains and constitute a protein family termed CREC with yet unidentified function[30]. A further EF-hand containing protein in the secretory pathway is calnuc which localizes to the lumen of the Golgi as well as to the cytosol. Interestingly, calnuc shows some

similarities with MCFD2 by displaying two EF-hand motifs and being O-glycosylated during the secretion process[31, 32].

Although a $^{45}Ca^{2+}$-binding assay could not determine the calcium binding affinity of MCFD2[33], both EF-hand motifs in MCFD2 are likely to be occupied due to the high calcium levels in the lumen of the ER. Calcium-bound MCFD2 might display an ERGIC-53 binding-competent conformation while deprivation of calcium may induce a conformational change in MCFD2 releasing it from ERGIC-53. This mechanism is in agreement with findings that *in vivo* depletion of luminal calcium by thapsigargin treatment disrupts the interaction between MCFD2 and ERGIC-53. Thapsigargin treatment resulted in the appearance of O-glycosylated MCFD2 which can no longer bind to ERGIC-53[28]. Since thapsigargin is a strong inducer of ER stress by depleting luminal calcium levels and since ERGIC-53 and MCFD2 are both target genes of the UPR[34, 35], it is tempting to speculate that ER stress can modulate the interaction between MCFD2 and ERGIC-53. In future experiments, it will be interesting to examine if there is a functional relationship between ERGIC-53, MCFD2 and ER stress. A further unanswered question concerns the MCFD2 binding site on ERGIC-53. The crystal structure of ERGIC-53 revealed a large surface patch of conserved residues opposite to the lectin binding site[36]. This surface patch could act as a potential MCFD2 binding site. Nuclear magnetic resonance (NMR) spectroscopy would be a powerful method to map the interactions interface between ERGIC-53 and MCFD2[37]. In chemical shift perturbation experiments using purified ERGIC-53 and MCFD2, the amino acid residues involved in binding could be determined. To verify the NMR data, one would then mutagenize the assigned amino acid residues in ERGIC-53 and MCFD2 and quantify the interaction by YFP PCA *in vivo*.

4.2.2 ERGIC-53/MCFD2 receptor-mediated cargo transport

Why does secretion of blood coagulation factors V and VIII require both ERGIC-53 and MCFD2? Depletion of MCFD2 by siRNA turned out to be a powerful approach to analyze the function of ERGIC-53 independent of MCFD2. Conversely, an ERGIC-53 depletion strategy could not be used to study MCFD2 individually since MCFD2 requires ERGIC-53 for its intracellular retention.

As shown in this thesis, ERGIC-53 can bind cathepsin Z and cathepsin C in a MCFD2-independent manner. Together with the finding that blood coagulation factor

VIII can be crosslinked to both ERGIC-53 and MCFD2[28], these results lead to the following model of cargo selectivity of the ERGIC-53/MCFD2 complex. ERGIC-53 can interact with cathepsin Z and cathepsin C by itself and requires MCFD2 only for the recruitment of blood coagulation factor V and VIII. The definitive proof of cargo selectivity of the ERGIC-53/MCFD2 receptor complex awaits further confirmation since we were unable to examine a direct interaction between ERGIC-53/MCFD2 and blood coagulation factors V and VIII by YFP PCA. Our model derives from the analysis of the cargo binding properties of ERGIC-53 and needs to be complemented with cargo transport studies. Although ERGIC-53 binds soluble glycoproteins beyond any doubts, its function as transport receptor has not been proven. Evidence for ERGIC-53-mediated transport is strong for blood coagulation factors V and VIII and supported by data deriving from genetic and biochemical approaches[26, 28, 38, 39]. For cathepsin Z and cathepsin C, a delayed secretion was shown upon expression of the dominant-negative ER retained KKAA mutant of ERGIC-53 in cell culture[40, 41]. Evidence of a cathepsin Z and cathepsin C transport defect in ERGIC-53-deficient cells has not been reported yet. By analyzing the secretion of endogenous cathepsin C in HeLa cells after siRNA-mediated ERGIC-53 depletion, no transport defect could be detected (unpublished data). Since only a small fraction of cathepsin C is secreted while the majority is targeted to lysosomes, quantification of secreted cathepsin C may not be adequate to draw a reliable conclusion concerning ERGIC-53-mediated transport. Additionally, it can not be completely excluded that siRNA-mediated depletion of ERGIC-53 leaves sufficient residual protein to ensure proper cargo transport. This is rather unlikely, however, since the ERGIC-53 knockdown was very efficient and an effect on MCFD2 was readily detectable. To what extent anterograde transport of cathepsin Z and cathepsin C indeed depends on ERGIC-53 should be critically addressed in future studies. The availability of ERGIC-53 knockout mice (Bin Zhang, personal communication) offers new ways to address this question. It will be possible to determine the levels of cathepsin Z and cathepsin C in purified lysosomes from these mice. To address transport kinetics, cathepsin C and cathepsin Z transport may be analyzed in primary cell cultures derived from ERGIC-53 knockout mice. As control, the concentration of blood coagulation factors V and VIII can be readily quantified in the blood plasma of these animals and the secretion of blood coagulation factors can be studied in primary hepatocyte cultures[42-44]. Such experiments should definitively clarify the role of ERGIC-53 in anterograde cargo transport.

The fluorescence complementation-based cDNA library screen in this thesis identified alpha-1-antitrypsin as novel potential interaction partner of ERGIC-53 suggesting that ERGIC-53 might capture more glycoproteins than initially assumed. It is possible that ERGIC-53 has a rather broad affinity towards glycoproteins which might facilitate glycoprotein secretion in general. For this purpose, it would be interesting to determine the carbohydrate specificity of ERGIC-53, for instance by frontal affinity chromatography[16] or carbohydrate microarrays[45]. Since other intracellular lectins such as VIP-36[17, 18] or VIPL[46, 47] might share ERGIC-53 redundant functions, they may compensate for a loss of ERGIC-53. The situation is different for blood coagulation factors V and VIII, the recruitment of which depends on the specific protein-protein interaction between ERGIC-53 and MCFD2. Therefore, other intracellular lectins might be largely unable to compensate for their secretion. This MCFD2 dependence would explain why the loss of ERGIC-53 has a specific effect on the secretion of blood coagulation factors V and VIII without affecting the secretion of other glycoproteins.

How can we envisage the cargo receptor function of ERGIC-53 based on the currently available data? One possibility is that ERGIC-53-mediated cargo capture reflects a secondary quality control step. By capturing only the correctly folded fraction of its cargo molecules, ERGIC-53 may control the incorporation of native proteins into COPII-coated vesicles. Time-resolved analysis of ERGIC-53 cargo binding endorses a secondary quality control mechanism by showing that the interaction between ERGIC-53 and cathepsin Z is maximal after an initial lag-period which is paralleled by folding of cathepsin Z[41]. To interact with ERGIC-53, cathepsin Z requires a folded β-hairpin/oligosaccharide structure including an intracellular disulfide bridge. This indicates that ERGIC-53 binds a native conformation of cathepsin Z. As a further possibility, ERGIC-53 might capture cargo proteins to facilitate their oligomerization. Recently, ERGIC-53 was reported to be involved in the formation of Russell bodies which are dilated ER membranes containing aggregated immunoglobin heavy chains[48, 49]. In the absence of light chains, mutant Ig-μ heavy chains were shown to interact with ERGIC-53 and condense into aggregates. Hexamerization of ERGIC-53 was speculated to seed this condensation and aggregation process. In more general terms, ERGIC-53 monomers may capture cargo molecules and by hexamerization promote the oligomerization of bound cargo proteins. While the oligomeric state of cathepsin Z is unknown, cathepsin C exists as

a tetramer and oligomerization was suggested to occur before its transport into lysosomes[50]. The analysis of the oligomerization process of cathepsin C in ERGIC-53-depleted cells might reveal if ERGIC-53 is required for cargo protein oligomerization.

In summary, the function of ERGIC-53 is still not entirely known. Nevertheless, the characterization of the cargo selectivity of the ERGIC-53/MCFD2 protein complex and the development of the luminal YFP PCA offer promising perspectives to address this issue. By analyzing ERGIC-53 knockout mice, by determining the carbohydrate-specificity of ERGIC-53, and by identifying additional ERGIC-53 interaction partners in the genomewide YFP PCA-based protein interaction screen, novel functional insights may arise.

4.3 References

1. Nyfeler, B., Michnick, S.W. & Hauri, H.P. Capturing protein interactions in the secretory pathway of living cells. *Proc Natl Acad Sci U S A* **102**, 6350-6355 (2005).
2. Ghosh, I., Hamilton, A.D. & Regan, L. Antiparallel leucine zipper-directed protein reassembly: Application to the green fluorescent protein. *Journal of the American Chemical Society* **122**, 5658-5659 (2000).
3. Kerppola, T.K. Visualization of molecular interactions by fluorescence complementation. *Nat Rev Mol Cell Biol* **7**, 449-456 (2006).
4. Hu, C.D. & Kerppola, T.K. Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat Biotechnol* **21**, 539-545 (2003).
5. Grinberg, A.V., Hu, C.D. & Kerppola, T.K. Visualization of Myc/Max/Mad family dimers and the competition for dimerization in living cells. *Mol Cell Biol* **24**, 4294-4308 (2004).
6. Ormo, M. et al. Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **273**, 1392-1395 (1996).
7. Magliery, T.J. et al. Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. *J Am Chem Soc* **127**, 146-157 (2005).
8. Appenzeller-Herzog, C., Roche, A.C., Nufer, O. & Hauri, H.P. pH-induced conversion of the transport lectin ERGIC-53 triggers glycoprotein release. *J Biol Chem* **279**, 12943-12950 (2004).
9. Zhang, J., Campbell, R.E., Ting, A.Y. & Tsien, R.Y. Creating new fluorescent probes for cell biology. *Nat Rev Mol Cell Biol* **3**, 906-918 (2002).
10. Jares-Erijman, E.A. & Jovin, T.M. FRET imaging. *Nat Biotechnol* **21**, 1387-1395 (2003).
11. Zacharias, D.A., Violin, J.D., Newton, A.C. & Tsien, R.Y. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **296**, 913-916 (2002).
12. Miyawaki, A. Fluorescence imaging of physiological activity in complex systems using GFP-based probes. *Curr Opin Neurobiol* **13**, 591-596 (2003).
13. Ito, T. et al. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A* **98**, 4569-4574 (2001).
14. Sheridan, D.L. et al. A new way to rapidly create functional, fluorescent fusion proteins: random insertion of GFP with an in vitro transposition reaction. *BMC Neurosci* **3**, 7 (2002).
15. Rothstein, R. Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. *Methods Enzymol* **194**, 281-301 (1991).
16. Kamiya, Y. et al. Sugar-binding properties of VIP36, an intracellular animal lectin operating as a cargo receptor. *J Biol Chem* **280**, 37178-37182 (2005).
17. Fiedler, K., Parton, R.G., Kellner, R., Etzold, T. & Simons, K. VIP36, a novel component of glycolipid rafts and exocytic carrier vesicles in epithelial cells. *Embo J* **13**, 1729-1740 (1994).
18. Fiedler, K. & Simons, K. Characterization of VIP36, an animal lectin homologous to leguminous lectins. *J Cell Sci* **109 (Pt 1)**, 271-276 (1996).
19. Kleizen, B. & Braakman, I. Protein folding and quality control in the endoplasmic reticulum. *Curr Opin Cell Biol* **16**, 343-349 (2004).
20. Meunier, L., Usherwood, Y.K., Chung, K.T. & Hendershot, L.M. A subset of chaperones and folding enzymes form multiprotein complexes in endoplasmic reticulum to bind nascent proteins. *Mol Biol Cell* **13**, 4456-4469 (2002).
21. Zhao, L. & Ackerman, S.L. Endoplasmic reticulum stress in health and disease. *Curr Opin Cell Biol* **18**, 444-452 (2006).
22. Garber, K. Researchers target unfolded protein response in cancerous tumor growth. *J Natl Cancer Inst* **98**, 512-514 (2006).
23. Nawrocki, S.T. et al. Bortezomib sensitizes pancreatic cancer cells to endoplasmic reticulum stress-mediated apoptosis. *Cancer Res* **65**, 11658-11666 (2005).
24. Nawrocki, S.T. et al. Bortezomib inhibits PKR-like endoplasmic reticulum (ER) kinase and induces apoptosis via ER stress in human pancreatic cancer cells. *Cancer Res* **65**, 11510-11519 (2005).
25. Bernales, S., Papa, F.R. & Walter, P. Intracellular Signaling by the Unfolded Protein Response. *Annu Rev Cell Dev Biol* (2006).
26. Nichols, W.C. et al. Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. *Cell* **93**, 61-70 (1998).

27. Zhang, B. et al. Bleeding due to disruption of a cargo-specific ER-to-Golgi transport complex. *Nat Genet* **34**, 220-225 (2003).
28. Zhang, B., Kaufman, R.J. & Ginsburg, D. LMAN1 and MCFD2 form a cargo receptor complex and interact with coagulation factor VIII in the early secretory pathway. *J Biol Chem* **280**, 25881-25886 (2005).
29. Kawasaki, H., Nakayama, S. & Kretsinger, R.H. Classification and evolution of EF-hand proteins. *Biometals* **11**, 277-295 (1998).
30. Honore, B. & Vorum, H. The CREC family, a novel family of multiple EF-hand, low-affinity Ca(2+)-binding proteins localised to the secretory pathway of mammalian cells. *FEBS Lett* **466**, 11-18 (2000).
31. Lin, P., Fischer, T., Weiss, T. & Farquhar, M.G. Calnuc, an EF-hand Ca(2+) binding protein, specifically interacts with the C-terminal alpha5-helix of G(alpha)i3. *Proc Natl Acad Sci U S A* **97**, 674-679 (2000).
32. Lavoie, C., Meerloo, T., Lin, P. & Farquhar, M.G. Calnuc, an EF-hand Ca(2+)-binding protein, is stored and processed in the Golgi and secreted by the constitutive-like pathway in AtT20 cells. *Mol Endocrinol* **16**, 2462-2474 (2002).
33. Toda, H. et al. Stem cell-derived neural stem/progenitor cell supporting factor is an autocrine/paracrine survival factor for adult neural stem/progenitor cells. *J Biol Chem* **278**, 35491-35500 (2003).
34. Spatuzza, C. et al. Heat shock induces preferential translation of ERGIC-53 and affects its recycling pathway. *J Biol Chem* **279**, 42535-42544 (2004).
35. Nyfeler, B., Nufer, O., Matsui, T., Mori, K. & Hauri, H.P. The cargo receptor ERGIC-53 is a target of the unfolded protein response. *Biochem Biophys Res Commun* **304**, 599-604 (2003).
36. Velloso, L.M., Svensson, K., Schneider, G., Pettersson, R.F. & Lindqvist, Y. Crystal structure of the carbohydrate recognition domain of p58/ERGIC-53, a protein involved in glycoprotein export from the endoplasmic reticulum. *J Biol Chem* **277**, 15979-15984 (2002).
37. Zuiderweg, E.R. Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry* **41**, 1-7 (2002).
38. Cunningham, M.A. et al. LMAN1 is a molecular chaperone for the secretion of coagulation factor VIII. *J Thromb Haemost* **1**, 2360-2367 (2003).
39. Moussalli, M. et al. Mannose-dependent endoplasmic reticulum (ER)-Golgi intermediate compartment-53-mediated ER to Golgi trafficking of coagulation factors V and VIII. *J Biol Chem* **274**, 32539-32542 (1999).
40. Vollenweider, F., Kappeler, F., Itin, C. & Hauri, H.P. Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme. *J Cell Biol* **142**, 377-389 (1998).
41. Appenzeller-Herzog, C. et al. Carbohydrate- and conformation-dependent cargo capture for ER-exit. *Mol Biol Cell* **16**, 1258-1267 (2005).
42. Baumann, H., Jahreis, G.P. & Gaines, K.C. Synthesis and regulation of acute phase plasma proteins in primary cultures of mouse hepatocytes. *J Cell Biol* **97**, 866-876 (1983).
43. Wilson, D.B., Salem, H.H., Mruk, J.S., Maruyama, I. & Majerus, P.W. Biosynthesis of coagulation Factor V by a human hepatocellular carcinoma cell line. *J Clin Invest* **73**, 654-658 (1984).
44. Biron-Andreani, C. et al. Secretion of functional plasma haemostasis proteins in long-term primary cultures of human hepatocytes. *Br J Haematol* **125**, 638-646 (2004).
45. Ratner, D.M. et al. Probing protein-carbohydrate interactions with microarrays of synthetic oligosaccharides. *Chembiochem* **5**, 379-382 (2004).
46. Neve, E.P., Svensson, K., Fuxe, J. & Pettersson, R.F. VIPL, a VIP36-like membrane protein with a putative function in the export of glycoproteins from the endoplasmic reticulum. *Exp Cell Res* **288**, 70-83 (2003).
47. Nufer, O., Mitrovic, S. & Hauri, H.P. Profile-based data base scanning for animal L-type lectins and characterization of VIPL, a novel VIP36-like endoplasmic reticulum protein. *J Biol Chem* **278**, 15886-15896 (2003).
48. Mattioli, L. et al. ER storage diseases: a role for ERGIC-53 in controlling the formation and shape of Russell bodies. *J Cell Sci* **119**, 2532-2541 (2006).
49. Valetti, C., Grossi, C.E., Milstein, C. & Sitia, R. Russell bodies: a general response of secretory cells to synthesis of a mutant immunoglobulin which can neither exit from, nor be degraded in, the endoplasmic reticulum. *J Cell Biol* **115**, 983-994 (1991).
50. Muno, D., Ishidoh, K., Ueno, T. & Kominami, E. Processing and transport of the precursor of cathepsin C during its transfer into lysosomes. *Arch Biochem Biophys* **306**, 103-110 (1993).

Curriculum Vitae

Name: **Beat Nyfeler**

Date of birth: **20.07.1978**

Place of origin: **Gondiswil (BE)**

Nationality: **Swiss**

Marital status: **Single**

Present address: **Department of Pharmacology/Neurobiology**
Biozentrum
University of Basel
Klingelbergstr.70
CH-4056 Basel
Switzerland

E-mail address: **beat.nyfeler@unibas.ch**

EDUCATION

PhD thesis at the Biozentrum, University of Basel, Switzerland. **2003-2006**
PhD thesis performed under the guidance of Professor Hans-Peter Hauri.
Title of PhD thesis: 'Visualization of protein-protein interactions in the secretory pathway of mammalian cells'.

Research collaboration at the Université de Montréal, Canada. **2005**
Research visit (6 months) in the group of Professor Stephen Michnick.

Study of Biology II at the University of Basel, Switzerland. **1998-2002**
Diploma degree in Biochemistry, Microbiology, Cell Biology, Biophysics and Structural Biology.
Diploma thesis performed under the guidance of Professor Hans-Peter Hauri.
Title of Master thesis: 'ERGIC-53 is a target of the unfolded protein response'.

Internship at Novartis Crop Protection, Research Triangle Parc, NC, USA. **1998**
Training (6 months) in the Herbicid Division under the guidance of Dr. Joshua Levin.

Highschool at the Gymnasium Bäumlihof, Basel, Switzerland. **1989-1997**

PUBLICATIONS

Cargo selectivity of the ERGIC-53/MCFD2 cargo receptor complex.
Nyfeler B, Zhang B, Ginsburg D, Kaufman RJ, Hauri HP. Accepted in Traffic, 2006.

Capturing protein interactions in the secretory pathway of living cells.
Nyfeler B, Michnick SW, Hauri HP. Proc Natl Acad Sci U S A. 2005 May 3; 102(18): 6350-6355.

Carbohydrate- and Conformation-dependent Cargo Capture for ER-Exit.
Appenzeller-Herzog C, Nyfeler B, Burkhard P, Santamaria I, Lopez-Otin C, Hauri HP. Mol Biol Cell. 2005 Mar; 16(3): 1258-1267.

The cargo receptor ERGIC-53 is a target of the unfolded protein response.
Nyfeler B, Nufer O, Matsui T, Mori K, Hauri HP. Biochem Biophys Res Commun. 2003 May 16;304(4):599-604.

Arabidopsis genes essential for seedling viability: isolation of insertional mutants and molecular cloning.
Budziszewski GJ, Lewis SP, Glover LW, Reineke J, Jones G, Ziemnik LS, Lonowski J, Nyfeler B, Aux G, Zhou Q, McElver J, Patton DA, Martienssen R, Grossniklaus U, Ma H, Law M, and Levin JZ. Genetics. 2001 December; 159(4): 1765–1778.