IDENTIFICATION OF COMPONENTS OF THE INTRACELLULAR TRANSPORT MACHINERY OF ACYLATED PROTEINS BY A GENOME-WIDE RNAI SCREEN

DISSERTATION

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto Carola University of Heidelberg, Germany
for the degree of Doctor of Natural Sciences

Julia Ritzerfeld
2009

“Ich weiß, dass ich nichts weiß.”
Sokrates

DISSERTATION

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto Carola University of Heidelberg, Germany
for the degree of Doctor of Natural Sciences

presented by
Master of Science in Molecular Bioengineering
Julia Ritzerfeld
born in Berlin

oral examination:

IDENTIFICATION OF COMPONENTS OF THE INTRACELLULAR TRANSPORT MACHINERY OF ACYLATED PROTEINS BY A GENOME-WIDE RNAI SCREEN

Referees:

Prof. Dr. Walter Nickel

Prof. Dr. Thomas Söllner

LIST OF PUBLICATIONS

Tournaviti, S., Hannemann, S., Terjung, S., Kitzing, T.M., Stegmayer, C., Ritzerfeld, J., Walther, P., Grosse, R., Nickel, W., and Fackler, O.T. (2007). SH4-domain-induced plasma membrane dynamization promotes bleb-associated cell motility. J Cell Sci *120*, 3820-3829.

Remmele, S., Ritzerfeld, J., Nickel, W. and Hesser, J. (2008): Automated cell analysis tool for a genome-wide RNAi screen. MIAAB (attached in appendix)

Tournaviti, S., Pietro, E.S., Terjung, S., Schafmeier, T., Wegehingel, S., Ritzerfeld, J., Schulz, J., Smith, D.F., Pepperkok, R., and Nickel, W. (2009). Reversible phosphorylation as a molecular switch to regulate plasma membrane targeting of acylated SH4 domain proteins. Traffic *10*, 1047-1060.

SUMMARY

Targeting of peripheral membrane proteins to different cellular compartments is often mediated by post-translational fatty acylation. For example, N-terminal SH4-domains containing dual lipid modifications mediate reversible attachment to intracellular membranes of a variety of proteins such as the Src family of kinases. Myristoylation and subsequent palmitoylation of the SH4-domain are not only required for stable membrane anchoring, but are also essential for targeting and transport to the plasma membrane. Their association with membrane microdomains, which are enriched in cholesterol and sphingolipids and moreover contain a specific set of proteins, is crucial for functionality and provides a means of spatio-temporal regulation of acylated proteins. Even though many SH4-domain-containing proteins are functionally quite well characterized, little is known about the intracellular machinery mediating transport to the plasma membrane. It has been hypothesized that their association with membrane microdomains already occurs at the level of intracellular membranes and is a prerequisite for targeting and transport. It was the aim of this study to identify gene products involved in intracellular transport of acylated proteins to the plasma membrane employing a genome-wide RNAi screening approach.

For this purpose, we established a stable human model cell line, which simultaneously expresses two distinct plasma-membrane-associated acylated reporter proteins. This cell line was adapted to a high-content screening platform, which is based on reverse transfection of a genome-wide siRNA library and data acquisition by automated widefield microscopy. To analyze imaging data in an unbiased and quantitative manner, an automated single-cell image analysis tool was developed. This tool identifies and compartmentalizes individual cells and determines intensity distributions of the two fluorescent reporter proteins, thus identifying experimental conditions under which intracellular retention of one or both reporter proteins occurs. Primary screening followed by subsequent validation with independent siRNAs resulted in the identification of 60 gene products, which caused intracellular retention of one or both acylated reporter proteins. Interestingly, we were able to identify enzymes involved in lipid homeostasis and microdomain-associated proteins. These findings corroborate the hypothesis that partitioning into membrane microdomains is a crucial step in targeting and transport of acylated proteins to the plasma membrane. Moreover, we were able to identify kinases, phosphatases and other proteins, which may exert regulatory functions in this process. The exact role of these factors in transport of SH4-domain-containing proteins to the plasma membrane will be elucidated in further studies.

ZUSAMMENFASSUNG

Membranassoziation und intrazelluläre Sortierung peripherer Membranproteine wird häufig durch die post-translationale Anheftung von Fettsäuren vermittelt. Ein Beispiel sind N-terminale SH4-Domänen, die duale Lipidmodifikationen enthalten und die reversible Membranassoziation, zum Beispiel von Mitgliedern der Familie der Src Kinasen, vermitteln. Eine Myristoylierung und darauffolgende Palmitoylierung der SH4-Domäne ist dabei nicht nur für eine stabile Membranverankerung notwendig, sondern ist zudem essentiell für den intrazellulären Transport zur Plasmamembran. Ihre Assoziation mit Membranmikrodomänen, die Cholesterin, Sphingomyelin und spezifische Proteine selektiv anreichern, ist für die Funktion von Src Kinasen essentiell. Während die Funktionen der Src Kinasen detailliert erforscht wurden, ist über die intrazelluläre Maschinerie, die ihren Transport zur Plasmamembran vermittelt, nur wenig bekannt. Es wurde die Hypothese aufgestellt, dass die Assoziation mit Mikrodomänen bereits an intrazellulären Membranen erfolgt und eine Voraussetzung für den Transport zur Plasmamembran ist. Das Ziel der vorliegenden Arbeit war es, durch eine auf RNA Interferenz basierten, genomweiten Analyse Genprodukte zu identifizieren, die am intrazellulären Transport von acylierten Proteinen zur Plasmamembran beteiligt sind.

Zu diesem Zweck wurde ein stabiles humanes Zellsystem etabliert, in dem simultan zwei unterschiedliche plasmamembranassoziierte acylierte Reporterproteine exprimiert werden können. Diese Zelllinie wurde an eine *High Content Screening* Plattform adaptiert, die auf reverser Transfektion mit einer genomweiten siRNA-Bibliothek und anschließender automatischer Weitfeldmikroskopie beruht. Um Bilddaten objektiv und quantitativ auswerten zu können, wurde ein automatisiertes Bildanalyseprogramm entwickelt, welches einzelne Zellen erkennen und kompartimentalisieren kann. Über eine Bestimmung der intrazellulären Intensitätsverteilung konnten experimentelle Bedingungen identifiziert werden, die durch die intrazelluläre Retention eines oder beider Reporterproteine charakterisiert waren. Die Primäranalyse und eine anschließende Validierung führten zur Identifizierung von insgesamt 60 Genprodukten, deren RNAi-vermittelte Expressionsunterdrückung eine intrazelluläre Retention eines oder beider Reporterproteine hervorrief. Interessanterweise konnten wir dabei Enzyme identifizieren, die an der zellulären Lipidhomeostase beteiligt sind. Diese Ergebnisse untermauern die Hypothese, dass eine Assoziation mit Membranmikrodomänen ein essentieller Schritt im Transport von acylierten Proteinen zur Plasmamembran ist. Darüber hinaus haben wir verschiedene Kinasen, Phosphatasen und andere Proteine identifiziert, die eine regulatorische Funktion in diesem Prozess spielen könnten. Die genaue Rolle dieser Faktoren im Transport von SH4-domänenhaltigen Proteinen zur Plasmamembran kann nun in weiterführenden Arbeiten untersucht werden.

1 INTRODUCTION

A characteristic hallmark of eukaryotic cells is the existence of distinct membrane-bound compartments called organelles. These organelles fulfil different functions and are therefore characterized by a specific composition of functional and structural components. Molecular constituents of organelles are specifically targeted to their destination by different sorting determinants and pass other compartments on their way to their final destination. This process requires a highly regulated way of targeting and transport between different compartments and a clear distinction between organelle-resident components and transient by-passers. Proteins destined to the extracellular space, the plasma membrane or different compartments *en route* are transported along a set of specialized organelles, which constitute the secretory pathway. Transport between these organelles is mediated by coated vesicles, which selectively incorporate soluble and membrane-bound proteins as well as lipids and deliver them to the next destination, thereby playing a pivotal role in sorting and targeting. In recent years, the role of lipids in intracellular transport has evolved from a purely structural role towards a direct involvement in sorting events. Lipids are no longer viewed as a continuous “sea” but have been shown to segregate into distinct domains, which are enriched in specific lipids and contain a particular set of proteins. For this reason, lipid microdomains play a role not only in sorting and transport but also in signalling events because they act as rallying points for sets of proteins, which in concert fulfil specific functions. While the key players in intracellular protein transport are well characterized in molecular and structural detail, the post-genomic era enables us to search for molecular components of specialized and less characterized transport events. One of these is the intracellular transport of microdomain-associated peripheral membrane proteins, which are targeted to the intracellular leaflet of the plasma membrane. Membrane association of these proteins by virtue of dual lipid modifications is required for their targeting, however, the molecular machinery involved has to date remained elusive.

1.1 PROTEIN SECRETION AND ENDOCYTOSIS

Proteins destined for secretion reach the extracellular space by one of two mechanistically distinct pathways: the majority of secretory cargo travels along compartments of the classical secretory machinery, while a smaller number of

extracellular proteins are externalized by a variety of ER/Golgi-independent mechanisms.

1.1.1 The Classical Secretory Pathway

Proteinaceous cargo targeted to the plasma membrane, endosomal compartments and organelles of the secretory pathway begin their journey by crossing the membrane of the endoplasmic reticulum (ER). the N-terminal amino acids of such proteins usually contain a stretch of hydrophobic amino acids, which are exposed during protein translation at free ribosomes in the cytosol. Binding of the signal recognition particle (SRP) to both signal sequence and ribosome occurs co-translationally and results in an arrest of protein translocation (Ng et al., 1996; Ogg and Walter, 1995). By binding to the heterodimeric SRP receptor, SRP targets the ribosome-nascent-chain complex to the ER membrane, transfers it to the translocon and recyles back to the cytosol (Corsi and Schekman, 1996; Rapoport, 1992; Walter et al., 1984). The protein-conducting channel consists of the heterotrimeric Sec61 membrane-protein complex, which is able to fully translocate soluble proteins and can also release hydrophobic transmembrane domains into the lipid bilayer (Mandon et al., 2009; Rapoport et al., 2004). After positioning at the membrane, the ribosome resumes translation, threads the nascent polypeptide chain into the ER lumen and dissociates after translation is terminated (Corsi and Schekman, 1996). The ER plays an essential role in protein folding, maturation and oligomeric assembly. Proper folding is achieved by a variety of chaperones and folding enzymes. The formation of disulfide bonds is favoured by the oxidizing potential of the ER lumen and the presence of redox proteins and many proteins are further modified by the attachment of N-linked glycans (Kleizen and Braakman, 2004). A complex quality control machinery ensures that only properly folded proteins are exported from the ER. A general quality control mechanism recognizes partially and misfolded proteins and mediates their retention in the ER. This mechanism is based on the recognition of hydrophobic regions, free cysteines and a high tendency for aggregation. In addition, individual proteins can undergo specific quality control mechanisms (Ellgaard and Helenius, 2001, 2003). If proteins ultimately fail to fold correctly, they are retrotranslocated to the cytosol and marked for degradation by the proteasome by polyubiquitination (Tsai et al., 2002). COPII-coated vesicles mediate export from the ER to the Golgi apparatus. While most cargo seems to be non-

selectively included into ER-to-Golgi transport carriers (Wieland et al., 1987), specific proteins appear to be selectively enriched at ER exit sites (Antonny and Schekman, 2001; Klumperman, 2000). After fission, COPII-vesicles are rapidly uncoated and fuse with the ER-Golgi intermediate compartment (ERGIC) consisting of vesicular-tubular clusters (VTCs) (Kaiser and Ferro-Novick, 1998; Klumperman, 2000). Transport from the ERGIC compartment to the Golgi is most probably not mediated by small coated vesicles but involves microtubule-dependent translocation of VTCs to the Golgi, where they fuse with the *cis*-Golgi network (Glick, 2000). While COPII-coated vesicles mediate anterograde transport, retrograde transport from the ERGIC/Golgi to the ER and inter-Golgi transport occurs via COPI-coated vesicles (Barlowe, 2000). Formation of COP-coated transport vesicles is controlled by small GTPases, which upon binding to membranes trigger the assembly of heteromeric coat complexes. Coat assembly locally shapes the donor membrane, resulting in formation of buds and fission of vesicles (Wieland and Harter, 1999). In the case of COPII, the small GTPase Sar1p binds to the ER membrane in its GTP-bound state, associates with cargo and recruits components of the COPII coat. Binding of Sec23/24p to Sar1p is followed by subsequent recruitment of Sec13/31p, which in turn drives membrane curvature and vesicle fission (Antonny and Schekman, 2001; Barlowe, 2002; Fromme and Schekman, 2005). The nucleation point for COPI-vesicle formation is provided by the small GTPase Arf1, which in its GDP-bound form is a cytosolic protein. Nucleotide exchange factors trigger the exchange of GDP for GTP and thus induce a conformational change, resulting in the exposure of a covalently attached myristate, which enables Arf1 to bind to membranes (Rothman, 1994). Members of the p24-family seem to assist recruitment of Arf1 to membranes and Arf1-GTP and p24 together recruit the heptameric coatomer complex (consisting of the subunits α-COP, β-COP, β'-COP, γ-COP, δ-COP, ε-COP and ζ-COP) *en bloc* (Bethune et al., 2006). Retrograde cargo sorting requires binding of coatomer subunits to dilysine motifs in cytoplasmic tails of transmembrane proteins (Barlowe, 2000). Once secretory proteins have reached the Golgi apparatus, they are further modified and sorted to their final destinations. Due to its dynamics and complexity, transport within the Golgi has long been elusive and has led to the co-existence of two contradicting models (Warren and Malhotra, 1998). The anterograde vesicular trafficking model postulates that anterograde transport of cargo between different Golgi compartments containing specific enzymes is mediated by COPI-

coated vesicles (Palade, 1975). However, some experimental findings, such as the occurrence of Golgi enzymes in COPI-vesicles (Malsam et al., 2005; Martinez-Menarguez et al., 2001) and evidence that large cargo molecules do not leave Golgi cisternae during anterograde transport (Bonfanti et al., 1998) lead to the development of the cisternal-maturation model. In this model, anterograde cargo transport is mediated by the maturation of Golgi compartments, while retrograde transport in COPI-coated vesicles ensures the retrieval of enzymes to their steady-state localization (Warren and Malhotra, 1998). Very recent evidence that cargo exit is not related to its residency time within the Golgi (Patterson et al., 2008) lead to the formulation of the rapid-partitioning model. This model is based on partitioning of cargo and enzymes into two different lipid phases within each compartment, which enables cargo exit from any level of the Golgi. COPI-vesicles transport cargo and enzymes in both anterograde and retrograde direction (Jackson, 2009). In the current view, however, the trans-Golgi-network (TGN) acts as the main sorting station of proteins and lipids within the cell. A sophisticated sorting machinery mediates cargo delivery to the apical and basolateral plasma membrane (PM), endosomal compartments and specialized compartments in specialized cell types. Sorting and transport machinery components and membrane material from the endosomal system and the PM are recycled back to the TGN (De Matteis and Luini, 2008). In addition to their function in endocytosis, clathrin-coated vesicles (CCVs) mediate transport between the TGN and endosomes (Le Borgne and Hoflack, 1998a, b). Transport to the plasma membrane, however, is mediated by the formation of large pleiomorphic carriers (LPCs), which travel along microtubulues to their final destination (De Matteis and Luini, 2008; Luini et al., 2005). While sorting into endosome-directed CCVs depends on tyrosine and di-leucine-based motifs, which are recognized by adaptor proteins (Le Borgne and Hoflack, 1998a, b), sorting to the plasma membrane seems to depend on a complex interplay of a variety of sorting determinants. Apart from post-translational modifications, such as glycosylation, the affinity of proteins to membrane microdomains enriched in sphingolipids and cholesterol has an impact on sorting to the apical plasma membrane (Paladino et al., 2004; Schuck and Simons, 2004; Simons and van Meer, 1988).

1.1.2 Unconventional Protein Secretion

While the bulk of protein externalization involves transport along organelles of the secretory pathway, some proteins reach the extracellular space independently of the classical secretory machinery. These proteins do not contain hydrophobic signal peptides for ER import, never appear in compartments of the secretory pathway, do not carry ER/Golgi-dependent post-translational modifications and their export is not compromised by drugs blocking ER/Golgi-dependent protein secretion (Nickel, 2003, 2005; Nickel and Rabouille, 2009; Nickel and Seedorf, 2008). The group of unconventionally secreted proteins comprises viral proteins (HIV Tat, Herpes simplex VP22), chromatin-binding proteins (HMGB1), cytokines (interleukin-1, thioredoxin, macrophage migration inhibitory factor), parasitic surface proteins (*Leishmania* hydrophilic acylated surface protein B, HASPB), lectins of the extracellular matrix (Galectin-1 (Gal1) and Galectin-3 (Gal3)) as well as pro-angiogenic growth factors (fibroblast growth factors 1 and 2 (FGF-1 and FGF-2)). The *Leishmania* coat protein HASPB is targeted to the inner leaflet of the plasma membrane by a dual lipid modification and externalized directly across the plasma membrane in both parasites and Chinese hamster ovary (CHO) cells (Denny et al., 2000; Stegmayer et al., 2005). Experimental evidence in recent years has shown that FGF-2 and Gal1 directly translocate directly across the plasma membrane (Schafer et al., 2004; Seelenmeyer et al., 2005; Zehe et al., 2006), and that this translocation requires molecular trapping by extracellular receptors, such as beta-galactoside-containing molecules or cell surface heparan sulfate proteoglycans (HSPGs) that trap Gal1 and FGF-2, respectively (Seelenmeyer et al., 2005; Zehe et al., 2006). FGF-2 export requires recruitment to the inner leaflet of the plasma membrane by phosphatidylinositol-4,5-bisphosphate ($PI(4,5)P_2$) and can occur in a folded state (Backhaus et al., 2004; Torrado et al., 2009). Interestingly, export of an FGF-2 fusion protein containing a signal sequence by the classical secretory machinery results in post-translational modifications of FGF-2, which block the functionally important interaction with HSPGs, indicating that unconventional export is required for secretion of the functional protein (Wegehingel et al., 2008).

1.1.3 Endocytosis

In contrast to the delivery of proteins to the extracellular space, cells also interact with their environment by the uptake of nutrients and signalling molecules. Moreover, the protein and lipid composition of the plasma membrane is influenced and controlled by endocytic events. Clathrin-mediated endocytosis (CME) is initiated by binding of adaptor and accessory proteins to endocytic cargo. Recruitment of clathrin lattices results in the formation of clathrin-coated-pits (CCPs), which are separated from the donor membrane by dynamin-mediated vesicle fission (Doherty and McMahon, 2009; Ungewickell and Hinrichsen, 2007). Apart from CME, a wide variety of clathrin-independent mechanisms are known, such as caveolae-mediated and flotillin-dependent endocytosis, phagocytosis, macropinocytosis and other less characterized processes (Doherty and McMahon, 2009; Hansen and Nichols, 2009). Endocytic events mediated by the scission of flask-shaped PM invaginations called caveolae are especially interesting, because they are not only enriched in specific proteins, such as caveolin-1, but also selectively include lipids, such as cholesterol (Parton and Simons, 2007; Pelkmans and Helenius, 2002). Similar to processes in the secretory pathway, internalized molecules are sorted within the endosomal system and are transported to various target destinations, such as lysosomes and the Golgi apparatus or can be recycled back to the plasma membrane (Grant and Donaldson, 2009; Maxfield and McGraw, 2004).

1.2 Post-translational Lipid Modifications and their Effect on Protein Targeting

Hydrophobic post-translational modifications of different molecular composition play important roles in protein targeting by providing sorting signals to intracellular membranes and membrane microdomains.

1.2.1 Glycosylphosphatidylinositol-anchored Proteins

Glycosylphosphatidylinositol (GPI) anchors are complex glycolipid structures attached to the C-terminus of proteins, which mediate attachment to the outer leaflet of the plasma membrane (Nosjean et al., 1997). GPI anchors are synthesized by a series of sequential reactions in the ER membrane and have a complex structure, consisting of a phosphoethanolamine linker, a glycan core

and a phospholipid tail. Structural diversity is mainly achieved by various sugar modifications of the glycan core, however, this diversity could not be directly related to differential functions and the only confirmed biological role of all GPI anchors is to provide a stable membrane anchor (Paulick and Bertozzi, 2008). GPI-anchored proteins fulfil diverse cellular functions, such as signal transduction, hydrolytic activity and cell adhesion and are enriched in cholesterol-dependent microdomains at the cell surface. By promoting partitioning into lipid microdomains at the level of the TGN, GPI-anchors were proposed to provide a sorting signal for apical transport (Simons and Wandinger-Ness, 1990). This hypothesis was recently challenged by the detection of GPI-anchored proteins at the basolateral membrane that precedes their appearance at the apical membrane, suggesting a transcytotic mechanism (Polishchuk et al., 2004). Most data, however, support the hypothesis of direct microdomain-dependent transport from the TGN to the apical membrane in polarized cells, a process which could be reinforced by other sorting determinants, such as glycosylation and oligomerization (Schuck and Simons, 2006).

1.2.2 Protein Prenylation

Protein prenylation involves the covalent attachment of the 15-carbon farneysl or the 20-carbon geranylgeranyl isoprenoid to one or more cysteine residues at the C-termini of proteins containing a 'CAAX box' motif. This process occurs in the cytoplasm and is catalyzed by a farnesyl transferase and two distinct geranylgeranyltransferases, respectively (McTaggart, 2006). The largest family of prenylated proteins are small GTPases, which fulfil diverse functions in intracellular signaling events. Like other hydrophobic post-translational modifications, prenylation confers membrane binding. In most cases, a second signal, provided by a basic cluster or fatty-acylation, is required for stable attachment to the membrane (Resh, 2006b). For example, spatio-temporal regulation of farnesylated Ras is achieved by a dynamic acylation and de-acylation cycle (Rocks et al., 2006; Rocks et al., 2005).

1.2.3 N-Myristoylation of Proteins

N-myristoylation refers to the irreversible covalent attachment of the 14-carbon saturated fatty acid myristate to the N-terminal glycine residue of nascent

polypeptides, resulting in an increased membrane binding affinity (Johnson et al., 1994). Target proteins begin with Met-Gly and often contain a threonine or serine at position 5. After removal of the initiating methionine by methionine aminopeptidase, N-myristoyltransferase (NMT) co-translationally catalyzes the transfer of myristate from myristoyl-CoA to the glycine residue, creating an amide linkage (Farazi et al., 2001). In most cases, myristoylation alone is in most cases not sufficient to mediate stable membrane anchorage. According to the two-signal model for membrane binding of myristoylated proteins, a second signal for membrane binding is required to stabilize membrane association directed by the relatively weak binding energy of myristate (Resh, 1994, 1999).

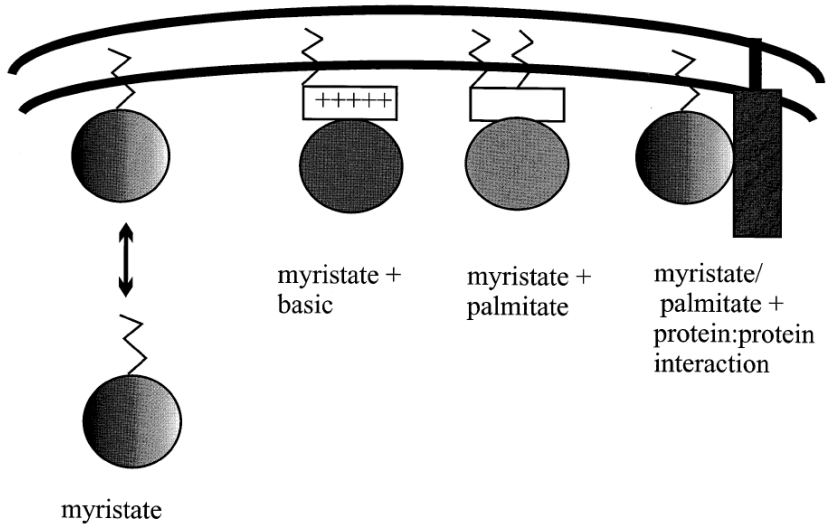

Figure 1 The two-signal model for membrane binding of myristoylated proteins. Stable membrane association of myristoylated proteins often requires a second membrane binding signal, such as electrostatic interactions (myristate + basic), a second acylation (myristate + palmitate) or interaction with membrane-bound proteins (myristate/palmitate + protein:protein interaction (from Resh, 1999).

Figure 1 illustrates how membrane anchoring of myristoylated proteins is stabilized by a second membrane-binding signal. The second motif for membrane binding can be provided by electrostatic interaction of basic amino acid clusters with negatively charged headgroups of phospholipids in the membrane bilayer ("myristate + basic"). This mechanism mediates stable membrane association of the non-receptor tyrosine kinase Src (Silverman and Resh, 1992). With the exception of Src itself, a "myristate + palmitate" signal in the SH4-domain directs stable membrane association of Src family kinases (SFKs) (Resh, 1994).

Myristoylation also provides the opportunity of reversible membrane attachment. In the case of myristoylated alanine-rich C kinase substrate

(MARCKS), the “myristate +basic” motif is used to achieve reversible membrane binding: MARCKS interacts with acidic membrane phospholipids via a basic amino acid cluster (Taniguchi and Manenti, 1993). Phosphorylation of MARCKS increases its negative charge, resulting in decreased phospholipid affinity and subsequent loss of membrane association (Thelen et al., 1991). A myristoyl switch, however, regulates reversible membrane association of Arf1: in the GDP-bound form, the myristoyl moiety of Arf1 is accommodated in a hydrophobic pocket and not accessible for interaction with membranes. Upon GTP-binding, myristate is exposed and directs membrane binding of Arf1 (Haun et al., 1993).

1.2.4 Protein Palmitoylation

Palmitoylation refers to the covalent attachment of the 16-carbon saturated fatty acid palmitate to cysteine residues by formation of a thioesther linkage. Like myristoylation, palmitoylation significantly increases membrane-binding affinity of peripheral membrane proteins. In contrast to myristoylation, however, this process is reversible and only recently has some of the enzymology of palmitate attachment been unravelled (Linder and Deschenes, 2007; Magee and Seabra, 2005; Mitchell et al., 2006). The group of palmitoylated proteins is diverse and includes membrane as well as cytosolic and viral proteins (Bijlmakers and Marsh, 2003). While palmitoylation of transmembrane proteins seems to affect stability of the target protein, it plays a distinct role in intracellular targeting and transport of cytosolic and peripheral membrane proteins and often coincides with the occurrence of myristoylation (1.2.3) or prenylation (1.2.2) (Linder and Deschenes, 2007). Stable membrane association of SFKs and Gα-subunits depends on co-translational myristoylation and subsequent palmitoylation (Alland et al., 1994; Kwong and Lublin, 1995; Shenoy-Scaria et al., 1993). Moreover, palmitoylation is not only required for general association with membranes, but directs partitioning of palmitoylated proteins into membrane microdomains (Zacharias et al., 2002), which is often a requirement for downstream signaling events to occur (Resh, 1999, 2006a, b).

1.2.4.1 Intracellular Sites of Protein Palmitoylation

Protein palmitoylation plays an important role in association of cytosolic or peripheral membrane proteins with intracellular membranes. Often,

palmitoylated proteins are additionally modified by myristoylation (1.2.3) or prenylation (1.2.2). Even though palmitoylation of a variety of proteins is required for their localization at the plasma membrane, the subcellular sites of palmitoylation as well as transport mechanisms to the plasma membrane are diverse. Even before individual protein acyl transferases (PATs) were identified, it was shown that PAT activity is localized to intracellular and plasma membranes and exhibited the biochemical characteristics of transmembrane proteins, such as resistance to extraction with high salt and sensitivity towards treatment with proteases and detergents (Das et al., 1997).

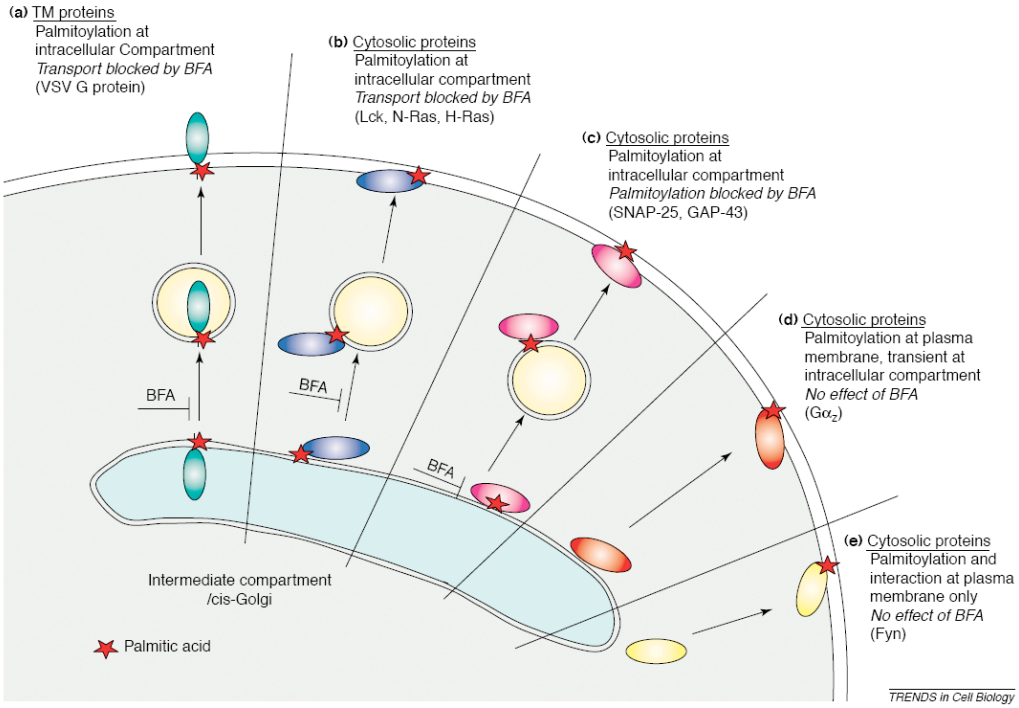

Figure 2 Transport pathways and intracellular sites of palmitoylation of different palmitoylated proteins. Palmitoylated proteins travel to the plasma membrane on diverse routes and are palmitoylated at distinct subcellular sites, as indicated by differential sensitivity to treatment with brefeldin A (BFA) (from Bijlmakers and Marsh, 2003).

Figure 3 shows intracellular transport pathways of different palmitoylated proteins. Pamitoylation of the transmembrane vesicular stomatitis virus glycoprotein (VSV-G) protein occurs early in the exocytic pathway, before aspartic-acid-linked oligosaccharides are trimmed, which implicates ERGIC or *cis*-Golgi as the site of palmitoylation (Bonatti et al., 1989). Consequently, VSV-G transport, but not its palmitoylation is blocked by treatment with brefeldin A (BFA) (Figure 2a). Dynamic regulation of Ras localization and activity is achieved by a palmitoylation cycle (Rocks et al., 2006; Rocks et al., 2005), with

palmitoylation occurring at an intracellular compartment, which is not the Golgi. The interaction of the SFK Lck with cytoplasmic domains of T-cell receptors depends on palmitoylation and occurs in the ERGIC or *cis*-Golgi compartment, indicating that the Lck PAT must be associated with compartments of the early secretory pathway (Bijlmakers and Marsh, 1999). Transport of both Ras and Lck but not their palmitoylation is sensitive to treatment with BFA (Figure 2b). Transport of peripheral membrane proteins in neurons is also dependent upon their palmitoylation state: SNAP-25, a t-SNARE acting in fusion of exocytic vesicles with the presynpatic membrane, and GAP-43, a crucial factor in axonal regeneration, are palmitoylated at their central domain and at the N-terminus, respectively. Both proteins accumulate at the TGN and are transported to the axon plasma membrane on exocytic vesicles. In contast to Ras and Lck, however, their palmitoylation is blocked by treatment with BFA, indicating that a functional Golgi system is required for efficient palmitoylation (Gonzalo and Linder, 1998) (Figure 2c). The heterotrimeric G-protein subunit $G\alpha_z$ interacts with intracellular membranes before it reaches the plasma membrane, however, neither transport nor palmitoylation are affected by treatment with BFA (Fishburn et al., 1999), indicating that palmitoylation occurs at the PM (Figure 2d). The SFK Fyn is targeted directly to the PM and does not require intracellular compartments for transport or palmitoylation (van't Hof and Resh, 1997; Wolven et al., 1997), rendering both processes insensitive to treatment with BFA (Figure 2e).

1.2.4.2 Enzymology of Protein Palmitoylation

The enzymology of protein palmitoylation is still poorly understood and only recently have protein acyl transferases (PATs) for specific proteins been identified. Even though non-enzymatic palmitoylation has been reported for several different proteins *in vitro* (Bano et al., 1998; Duncan and Gilman, 1996), this process is counteracted by the presence of acyl-CoA binding protein (ACBP) (Dunphy et al., 2000; Leventis et al., 1997), which is highly abundant in the cytosol. Therefore, non-enzymatic palmitoylation most likely and therefore most probably does not occur to a significant extent *in vivo*.

Figure 3 Enzymology of protein palmitoylation. Transfer of palmitate to target protein cysteines is catalyzed by a protein acyl transferase (PAT). Removal of palmitate is catalyzed by an acyl protein thioesterase (APT) (from Linder and Deschenes, 2007).

Enzymatic attachment of palmitate from palmitate-CoA to cysteine side chains of target proteins is catalyzed by a protein acyl transferase (PAT). Since PATs also have a high affinity for other long-chain fatty acids except palmitate, this process is often more correctly referred to as S-acylation. An acyl protein thioesterase (APT) catalyzes the removal of palmitate from palmitoylated proteins. Even before individual PATs were identified, it became clear that PAT activity occurs in different intracellular membranes, such as the PM, ER, Golgi, ERGIC and the vacuole in yeast (Bijlmakers and Marsh, 2003). Moreover, PAT activity of isolated membranes was resistant to high salt extraction and sensitive to treatment with detergent and proteases, indicating that the isolated PATs are transmembrane proteins (Das et al., 1997). The first two palmitoyltransferases were identified in 2002 in *S. cerevisiae*: Erf2p-Erf4p is an ER-resident PAT for yeast Ras (Lobo et al., 2002) and yeast casein kinase Yck2 is palmitoylated by Akr1, which localizes to the Golgi apparatus (Roth et al., 2002). Both proteins belong to a family of membrane proteins with a cysteine-rich DHHC (Asp-His-His-Cys) domain, which comprises more than 20 members in the human genome. Exactly how DHHC-containing proteins participate in the transfer of palmitate to the target enzyme is not yet clear, yet. First evidence, however, indicates that the DHHC-domain-containing proteins may serve as acyl-enzyme intermediate, because DHHC-proteins are autoacylated. Rapid autoacylation requires the cysteine residue in the DHHC domain, is labile and precedes the transfer of palmitate to target proteins (Mitchell et al., 2006).

1.3 Intracellular Transport of Acylated SH4-domain-containing Proteins

The SH4-domain consists of the N-terminal 16-20 amino acids of acylated peripheral membrane proteins and directs stable membrane anchoring by virtue of dual lipid modifications. These consist of an N-myristoylation, which is usually followed by a palmitoylation as a second membrane-binding motif (Resh, 1993, 1994, 1999).

1.3.1 Src Family Kinases

Src family kinases (SFKs) are non-receptor tyrosine kinases, which participate in multiple biological functions, such as cell adhesion, motility, proliferation and survival, and play an important role in several disease contexts (Ingley, 2008; Martin, 2001; Sandilands and Frame, 2008). The family consists of eight members, which are subgrouped into Lyn-related SFKs, i.e. Lyn, Hck, Lck, Blk, and Src-related SFKs, i.e. Src, Yes, Fyn, Fgr. While Src, Yes, Lyn and Fyn are widely expressed, Blk, Fgr, Hck and Lck expression is mainly restricted to hematopoietic cells (Thomas and Brugge, 1997).

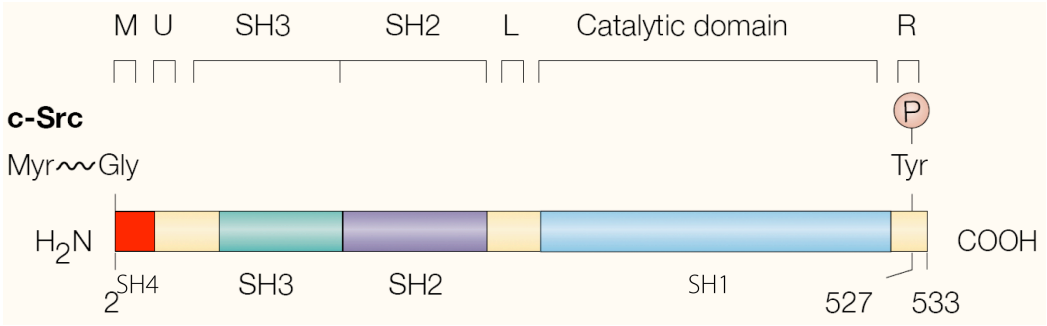

Figure 4 Src modular domain structure. Src contains an N-terminal SH4-domain for membrane anchorage followed by a unique domain, an SH3 and SH2-domain mediating protein-protein interactions and a catalytic domain (SH1) followed by a C-terminal regulatory region (adapted from Martin, 2001).

All SFKs share a similar modular structure consisting of six functional domains (Figure 4): The N-terminal acylated domain, termed src homology 4 (SH4) domain, confers membrane binding via one or several lipid modifications and, sometimes, basic clusters of amino acids (Resh, 1994) and is followed by a unique domain. The SH3-domain directs protein-protein interactions and has a high affinity for proline-rich motifs, while the SH2-domain mediates an interaction with phosphotyrosine residues (Koch et al., 1991). The SH1-domain, which contains an autophosphorylation site at Y416, is the kinase domain and mediates enzymatic activity. A C-terminal regulatory region contains a phosphorylation site and contributes to transition from the inactive "closed" to the active "open" confirmation, by docking of phosphorylated Y527 to the SH2-domain, rendering the molecule inactive. Dephosphorylation releases Y527 and results in a structural rearrangement to the active conformation, in which the catalytic domain autophosphorylates (Ingley, 2008).

Already in the early '90s, SFKs were known to interact with cellular membranes via lipid modifications in the N-terminal domain (Blenis and Resh, 1993; Resh, 1993; Silverman et al., 1993). Nowadays, it is clear that SH4-domains not only

confer unspecific binding to intracellular membranes, but that different SH4-domains direct targeting to distinct intracellular localizations and transport routes and mediate partitioning into membrane microdomains (Bijlmakers and Marsh, 2003; Resh, 1994, 1999, 2006a, b).

1.3.1.1 Src

Src is one of two Src family kinases (SFK) that are only myristoylated, but not palmitoylated. Instead, the second signal for membrane binding is provided by a basic cluster of amino acids (Koegl et al., 1994; Resh, 1994; Shenoy-Scaria et al., 1993). Src localizes to the cytosol and endosomes and is recruited to the plasma membrane upon external stimuli (Kaplan et al., 1992). Src-containing endosomes are RhoB-positive endosomes at the perinuclear recycling compartment and participate in actin-dependent delivery of Src to the plasma membrane (Sandilands et al., 2004). Introduction of two palmitoylation sites changes its intracellular trafficking behaviour reminiscent of those of Fyn (Sandilands et al., 2007) Moreover, Src was shown to rapidly cycle between the plasma membrane and late endosomes/lysosomes in living cells, most likely through cytosolic release (Kasahara et al., 2007).

1.3.1.2 Yes

The plasma-membrane-localized SFK Yes is myristoylated at glycine 2 and cysteine 3 (Koegl et al., 1994; Resh, 1994) and it is present in detergent-resistant membranes and the apical plasma membrane of polarized cells (Arreaza et al., 1994). Unlike Fyn and Src, Yes does not colocalize with RhoB/RhoD-positive endosomes (Sandilands et al., 2007). Similar to Lyn, however, Yes transport to the plasma membrane is blocked by low temperatures (19°C) and dominant-negative Rab11, indicating export via the TGN (Sato et al., 2009).

1.3.1.3 Fyn

The SFK Fyn is myristoylated at the N-terminal glycine and dually palmitoylated at cysteines 3 and 6 (Koegl et al., 1994; Resh, 1994, 1999). Fyn was shown to rapidly interact with the plasma membrane 5 minutes after biosynthesis and to partition into membrane microdomains 10-20 minutes after membrane binding

(van't Hof and Resh, 1997). Moreover, the reversibly palmitoylated SH4-domain of Fyn is sufficient to target chimeric proteins to the plasma membrane and to detergent-resistant membranes (Wolven et al., 1997) and is responsible for interaction with membrane microdomains containing GPI-anchored proteins (Shenoy-Scaria et al., 1993) and binding to the T-cell receptor (Timson Gauen et al., 1996). Interestingly, the aforementioned effects are dependent upon cysteine 3 but not cysteine 6. Deletion of the second palmitoylation site (cysteine 6) did not compromise membrane targeting, but instead resulted in export via the conventional secretory pathway, similar to the monopalmitoylated SFKs Yes and Lyn (Sato et al., 2009). Unlike Src, Fyn localizes to caveolae (Shenoy-Scaria et al., 1994). Abrogation of palmitoylation by mutagenesis of both palmitoylation sites rendered intracellular trafficking of Fyn similar to Src (Sandilands et al., 2007). Incorporation of unsaturated fatty acids into the SH4-domain of Fyn, which seems to occur *in vivo* (Liang et al., 2004), reduced its ability to partition into detergent-resistant microdomains (DRMs) and resulted in decreased signalling efficiency (Liang et al., 2001). Methylation of its N-terminal domain at lysine 7 and/or 9 is required for Fyn functionality and depends on previous acylation (Liang et al., 2004). Even though plasma membrane targeting of Fyn was shown to occur directly and not via the Golgi system, it requires intact actin filaments and, moreover, Fyn is present in RhoD-positive endosomes (Sandilands et al., 2007).

1.3.1.4 Lck

Lck is myristoylated at glycine 2 and dually palmitoylated at cysteines 3 and 5 (Koegl et al., 1994; Resh, 1994). Lck membrane association and incorporation into lipid microdomains is dependent upon palmitoylation. Whereas palmitoylation of Lck is essential for its specific association with DRMs, general membrane binding properties in these palmitoylation deficient mutants can be preserved by introduction of a basic amino acid cluster (Koegl et al., 1994). Mutation of serine 6 abolished acylation, rendering the mutant Lck incapable of interaction with membranes and unable to activate downstream effectors (Yasuda et al., 2000). Acylated Lck associates with membrane microdomains enriched in GPI-anchored proteins (Shenoy-Scaria et al., 1993).

1.3.2 *Leishmania Hydrophilic Acylated Surface Protein B*

Hydrophilic Acylated Surface Proteins (HASPs) are coat components of *Leishmania* parasites, which are exclusively expressed in infective stages of the parasite (Flinn et al., 1994). Deletion of the LcmDNA16 locus encoding the highly related genes HASPA1, HASPA2 and HASPB, does not interfere with virulence, however, these mutants show an increased sensitivity to complement-mediated lysis (McKean et al., 2001). Interestingly, HASPB localization at the surface of Leishmania parasites depends on its N-terminal SH4-domain, which is myristoylated at glycine 2 and palmitoylated at cysteine 5 (Denny et al., 2000). The SH4-domain of HASPB (SH4-HASPB) is able to direct acylation-dependent plasma membrane targeting of fluorescent reporter proteins in the mammalian system, too, indicating a conserved export mechanism (Stegmayer et al., 2005; Tournaviti et al., 2007; Tournaviti et al., 2009). Export to the extracellular leaflet of the mammalian plasma membrane occurs via direct translocation across the plasma membrane by an unknown mechanism (Stegmayer et al., 2005). Overexpression of SH4-HASPB-containing reporter proteins in Chinese hamster ovary (CHO) cells results in a dynamization of the plasma membrane, as indicated by the formation of highly dynamic non-apoptotic plasma membrane blebs (Tournaviti et al., 2007).

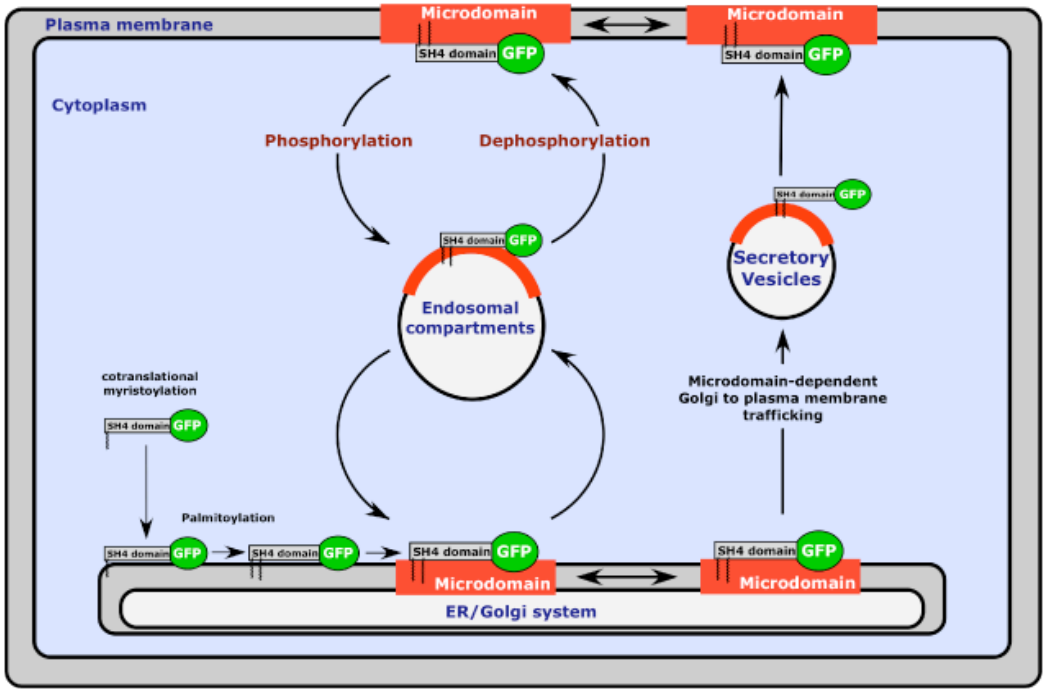

Figure 5 Intracellular trafficking of acylated Src kinases. Upon co-translational myristoylation, SH4-domains interact with the ER/Golgi system, where they are palmitoylated. Palmitoylated SH4-domains partition into intracellular microdomains and are exported to the plasma

membrane. A phosphorylation cycle regulates endocytosis and recycling without affecting microdomain association (modified from Nickel, 2005).

In an attempt to elucidate the intracellular transport route of HASPB, several mutants of its SH4-domain were analyzed with regard to their intracellular localization, acylation state and ability to partition into membrane microdomains, resulting in the model outlined in Figure 5. Mutation of the myristoylation site (SH4-HASPB-Δmyr) inhibits subsequent palmitoylation of SH4-HASPB and thus renders the protein incapable of interacting with cellular membranes, resulting in a cytosolic localization. Mutation of the palmitoylation site (SH4-HASPB-Δpal) results in association with perinuclear membranes and, unlike the wild-type SH4-domain, a failure to partition into membrane microdomains (Tournaviti et al., 2009). Most probably, myristoylation mediates association with intracellular membranes, for example the ER/Golgi system, harbouring the SH4-specific PAT. Palmitoylation directs partitioning into export-competent membrane microdomains, thereby enabling transport to the plasma membrane. In addition to the importance of the dual lipid modification, transport of SH4-HASPB is regulated by a phosphorylation at threonine 6 (Tournaviti et al., 2009). Phosphomimetic mutation of threonine 6 to glutamate (T6E) resulted in a perinuclear accumulation of the fully acylated protein, however, mutation of threonine 6 to alanine (T6A) did not disturb plasma-membrane localization. Association with detergent-resistant membranes (DRMs) was not affected by either mutation. This indicates an additional level of regulation by an endocytic step, which is independent of the association with membrane microdomains. Phosphorylation of the SH4-domain triggers internalization and sequestration in intracellular compartments, such as endosomes and the ER/Golgi system. Dephosphorylation of threonine 6 is required for recycling to the plasma membrane (Tournaviti et al., 2009).

1.4 MEMBRANE MICRODOMAINS

In 1972, Singer and Nicolson proposed the fluid-mosaic model of cellular membranes. According to this model, transmembrane proteins are randomly dispersed in a fluid lipid bilayer, enabling them to diffuse freely in a lateral direction (Singer and Nicolson, 1972). In the past 20 years, however, this model has been revised, resulting in a model, which postulates the existence of small and highly dynamic membrane domains, which are enriched in cholesterol and

sphingolipids, as well as in specific proteins (Simons and Ikonen, 1997; Simons and van Meer, 1988). These domains were henceforth called lipid or membrane "rafts" or, in recent years, microdomains (Pike, 2006).

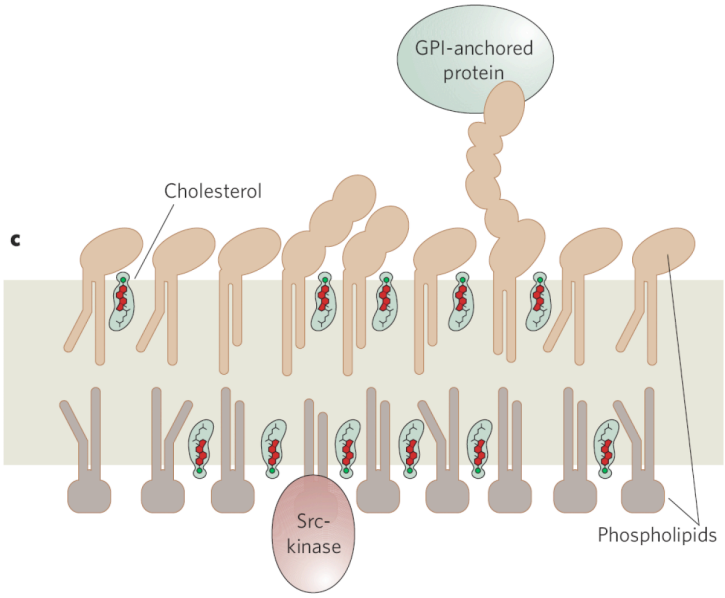

Figure 6 Model of an exemplary membrane microdomain or "raft". Membrane microdomains are enriched in cholesterol and sphingolipids in the outer leaflet of the plasma membrane and contain specific proteins, such as GPI-anchored proteins and acylated Src kinases (from Maxfield and Tabas, 2005).

Membrane microdomains are enriched in specific lipids, like cholesterol and sphingolipids, and contain specific proteins, such as GPI-anchored proteins attached to the outer leaflet and acylated Src kinases associated with the inner leaflet of the plasma membrane (Figure 6) (Maxfield and Tabas, 2005; Simons and Ikonen, 1997; Simons and van Meer, 1988; Simons and Wandinger-Ness, 1990).

The plasma membrane is characterized by a specific lipid composition: like all cellular membranes, it contains glycophospholipids and; in addition, is enriched in two additional lipid classes: sphingolipids consist of a ceramide backbone attached to a phosphocholine headgroup (sphingomyelin) or different carbohydrate structures (glycosphingolipids) and both acyl chains are often saturated. While sphingomyelin is present at 10-20 mol % in the plasma membrane, glycosphingolipids usually exist at low levels, with the exception of some specialized cell types (Munro, 2003). Cholesterol, the main sterol in vertebrates, is a rigid, four-ring structure and is present at 30-40 mol % at the plasma membrane (Munro, 2003). Plasma membrane microdomains are specifically enriched in sphingolipids and cholesterol. These lipids favour a more closely packed, "liquid-ordered" lateral organization, which significantly reduces

bilayer permeability without affecting lateral diffusion of lipids and proteins. In contrast, lateral diffusion between the liquid-ordered “raft”-phase and the surrounding liquid-disordered “non-raft”-phase is reduced, thereby creating different compartments within the plane of the same membrane (Brown and London, 1998; Simons and Ikonen, 1997). More recently, it has been proposed that small-scale rafts can coalesce upon protein-protein interactions (Schuck and Simons, 2004). Clustering of membrane microdomains allows a highly dynamic spatio-temporal regulation of the formation of larger platforms, which are implicated in diverse cellular functions, such as signal transduction, pathogen entry, virus formation and membrane trafficking (Brown and London, 1998; Schuck and Simons, 2004; Simons and Ikonen, 1997; Simons and Toomre, 2000).

Another interesting feature of the plasma membrane bilayer is the differential lipid composition of its two leaflets. Most sphingolipids and at least half of the cholesterol is present in the outer leaflet, while some glycerophospholipids, such as phosphatidylinositol (PI), phosphatidylethanolamine (PE) and phosphatidylserine (PS) are restricted to the inner leaflet of the plasma membrane (Bretscher, 1973). Membrane microdomains are mainly characterized by a specific enrichment of lipids and proteins of the outer leaflet but nevertheless incorporate proteins, such as acylated SFKs, which associate with the inner leaflet. It has been suggested that lateral organization of sphingolipids and cholesterol in the outer leaflet results in a similar organization of lipids and proteins in the inner leaflet, and that this is achieved by interdigitation of sphingolipid acyl chains with lipids of the inner leaflet (Simons and Ikonen, 1997; Simons and Toomre, 2000). Whether and how this inter-leaflet coupling is assisted by proteinaceous accessory factors, however, has to date remained elusive.

1.4.1 The Role of Membrane Microdomains in Signal Transduction

Membrane Microdomains have been implicated in a variety of signalling cascades, such as T-cell receptor and Ras signalling (Simons and Toomre, 2000). Due to the coupling of components in the outer and the inner leaflet of the plasma membrane, rafts have the intrinsic ability to transduce signals from the outside of the cell to the inside, by activating intracellular signalling cascades. Moreover, rafts selectively include or exclude proteins. Proteins, which are separated by their residence in different phases of the plasma membrane or in

different small-scale rafts, can interact only upon changes in their raft affinity or upon clustering of smaller rafts to form large signalling platforms, respectively.

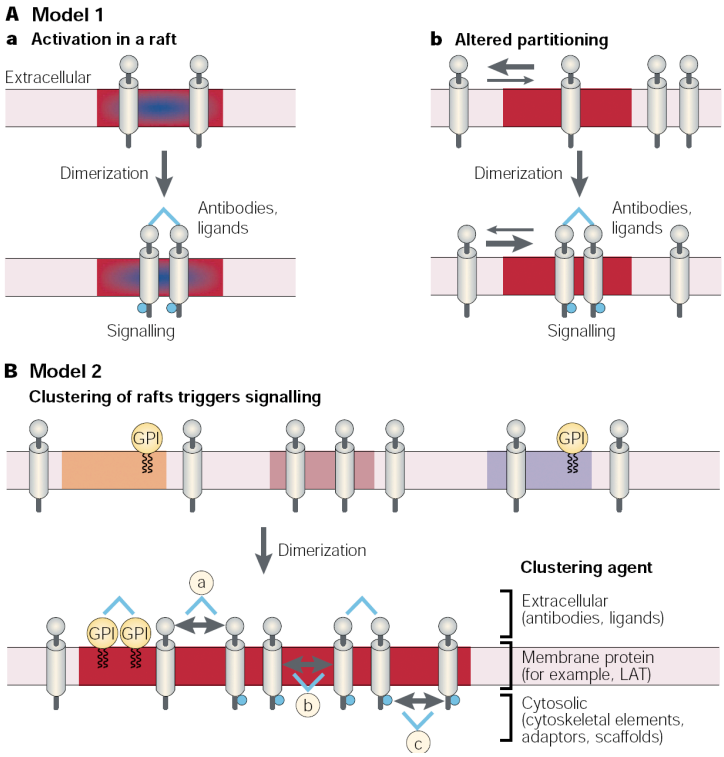

Figure 7 Models for the involvement of rafts in the initiation of signalling events. Signalling through rafts can occur in single rafts (model 1) and or through clustered rafts (model 2) and results in phosphorylation of receptors (from Simons and Toomre, 2000).

Figure 7 depicts different models of how rafts could be involved in the initiation of signalling events. Receptors within the same raft could be activated through ligand binding, resulting in receptor phosphorylation and downstream signalling events (panel A, Model 1a). Oligomerization of receptors upon ligand binding could lead to changes in raft affinity and result in extended residency times, thereby allowing downstream signalling events to occur (panel A, model 1b). Alternatively, clustering of small-scale rafts into large signalling platforms is induced by extracellular ligands (a), protein-protein interactions (b) or crosslinking through cytosolic elements (c) (panel B, model 2).

1.4.2 The Role of Membrane Microdomains in Intracellular Transport

Membrane microdomains play important roles in intracellular transport events, by sorting specific proteins into transport-competent compartments and,

moreover, could play a direct role in the formation of transport vesicles (Brown and London, 1998; Rajendran and Simons, 2005; Schuck and Simons, 2004; Simons and Ikonen, 1997; Simons and van Meer, 1988).

Caveolae were one of the first compartments to be identified as membrane microdomains with a role in intracellular transport (Brown and London, 1998). Caveolae are flask-shaped invaginations at the plasma membrane, which mediate clathrin-independent endocytosis (Parton and Richards, 2003; Parton and Simons, 2007; Pelkmans and Helenius, 2002). Caveolae contain the cholesterol-binding protein caveolin-1, which is also palmitoylated and, moreover, are enriched in cholesterol and sphingomyelin. They are involved in internalization of specific lipids and proteins, in particular GPI-anchored proteins and, are entry sites for viruses and toxins (Pelkmans and Helenius, 2002).

Membrane microdomains have long been implicated in apical sorting in polarized cells (Brown and London, 1998; Schuck and Simons, 2004; Simons and van Meer, 1988). While liquid-ordered microdomains are distributed within the liquid-disordered plasma membrane in non-polarized cells, they accumulate at the apical membrane in polarized epithelial cells and form a continuous raft membrane, in which liquid-disordered domains are embedded (Schuck and Simons, 2004). The basolateral membrane also contains rafts, but to a lesser extent (Simons and Ikonen, 1997). To create and maintain the two distinct membrane identities, a tightly regulated sorting machinery has to ensure targeted delivery and/or selective retention of apical and basolateral cargo. Basolateral sorting, however, is more accurate than apical sorting and mainly relies on sorting determinants in cytoplasmic or extracellular protein domains, which participate in receptor mediated cargo capture. Examples for basolateral sorting determinants are tyrosine-based and dileucine motifs, which can be recognized by different adaptor complexes. By contrast, GPI-anchors, some transmembrane domains, N- and O-glycosylation and protein oligomerization confer apical targeting (Schuck and Simons, 2004). Apical sorting motifs influence partitioning into membrane microdomains: GPI-anchored proteins have been shown to partition into membrane microdomains (Schuck and Simons, 2006; Simons and Wandinger-Ness, 1990). Liquid-ordered phases are thicker and more rigid than liquid-disordered phases and therefore preferentially accommodate a select group of membrane proteins with long transmembrane domains (Bretscher and Munro, 1993). Oligomerization of

proteins, which could, amongst other mechanisms, be mediated by the binding of multivalent lectins to glycosylated proteins, promotes affinity to and clustering of membrane rafts (Schuck and Simons, 2004). While the majority of cellular microdomains can be found at the plasma membrane, rafts also exist in biosynthetic and endocytic pathways. It has been postulated that rafts first form at the level of the Golgi, where sphingolipid biosynthesis is completed (Simons and Toomre, 2000).

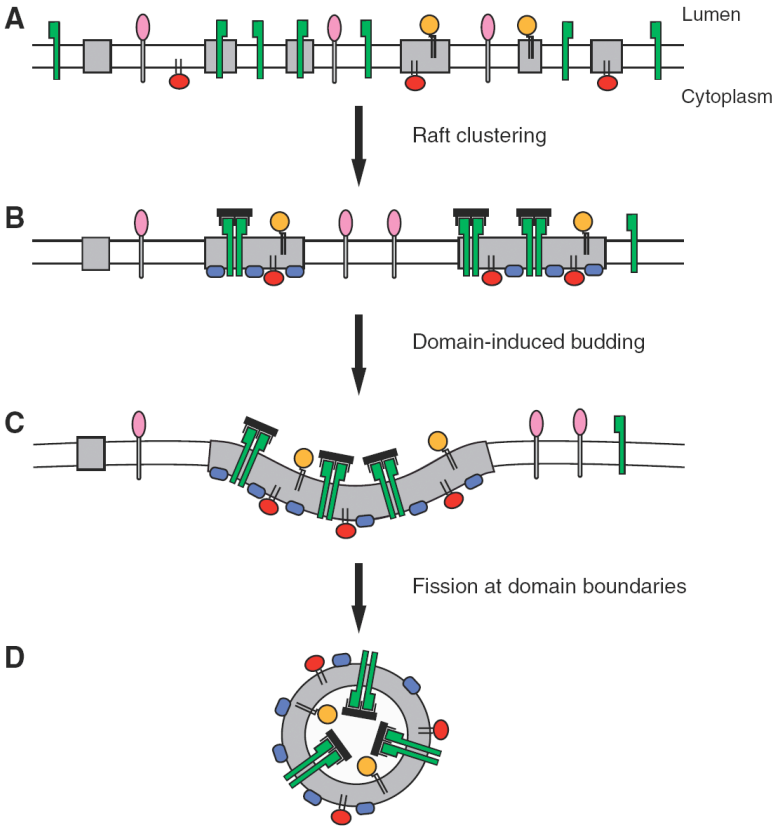

Figure 8 Raft clustering promotes domain-induced budding. Small-scale rafts (grey) include GPI-anchored proteins (yellow), the majority of acylated proteins (red) and some transmembrane proteins (green), while other transmembrane proteins (pink) are excluded. Raft clustering induced by binding of multimeric proteins (blue) increases affinity of partially associated proteins and promotes bud formation, resulting in vesicle fission at the domain boundary (from Schuck and Simons, 2004).

Apart from acting as sorting platforms, large microdomain assemblies could lead to the formation of transport vesicles (Figure 8): small-scale rafts containing apical cargo, such as GPI-anchored proteins (yellow), acylated proteins (red) and specific transmembrane proteins (green) are clustered, connecting the transport fates of incorporated proteins and lipids (panels A and B). This clustering is initiated by binding of multivalent proteins (blue) and results in an increased affinity of partially raft associated proteins to the liquid-

ordered phase. Microdomain affinity of transmembrane proteins is further increased by oligomerization upon crosslinking. Line tension at the interface between large liquid-ordered domains and the liquid-disordered phase is released by formation of buds and fission of vesicles with a raft-like vesicle composition (panels C and D). The existence of secretory vesicles with a raft-like lipid composition has recently been shown in *S. cerevisiae* (Klemm et al., 2009). In conclusion, apical sorting motifs mediate partitioning into newly formed raft clusters at the level of the Golgi, which promote formation of transport carriers by domain-induced budding (Schuck and Simons, 2004).

1.5 RNA INTERFERENCE

The discovery of RNA interference (RNAi) was a key finding to enable functional analyses of complete genomes in the post-genomic era. In 1998, Fire and colleagues discovered this type of post-transcriptional gene silencing (PTGS) induced by double-stranded RNAs (dsRNA) in *C. elegans* (Fire et al., 1998). RNA interference is involved in silencing of endogenous genes and acts in response to virus infection (Siomi and Siomi, 2009).

1.5.1 Molecular Mechanism of RNA Interference

In the past decade, the molecular mechanism mediating RNAi has been characterized in detail (Meister and Tuschl, 2004).

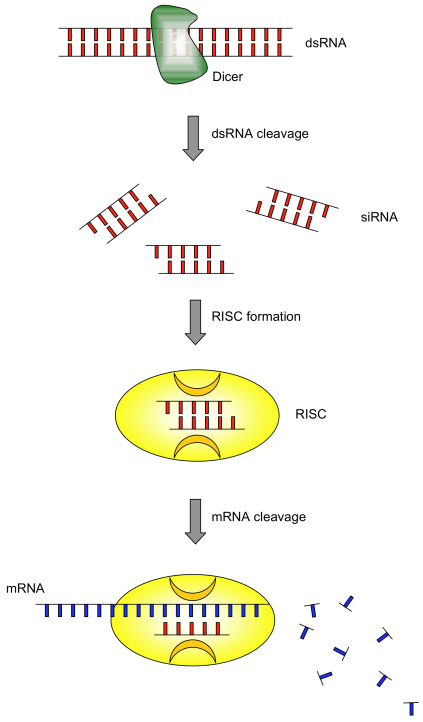

Figure 9 RNA interference pathway in mammalian cells. Dicer cleaves double-stranded RNA (dsRNA) precursors into small RNA duplexes. Short interfering RNAs (siRNAs) are incorporated into the RNA induced silencing complex (RISC) resulting in mRNA cleavage (from Mocellin and Provenzano, 2004).

Figure 9 illustrates the three-step process of RNAi-mediated gene silencing: A double-stranded RNA (dsRNA) precursor molecule is processed into small dsRNA molecules by the ribonuclease Dicer. The guide strand is incorporated into the RNA induced silencing complex (RISC) and serves as template for recognition of complementary mRNAs, resulting in gene silencing by mRNA cleavage (Meister and Tuschl, 2004; Mocellin and Provenzano, 2004; Siomi and Siomi, 2009).

RNAi is always initiated by the occurrence of dsRNA molecules. Naturally occurring dsRNA originates from RNA-templated RNA polymerization, e.g. from viruses, or hybridization of overlapping or complementary RNAs, e.g. genome-encoded stem-loop structures. These precursors are degraded by a member of the RNase III family of dsRNA-specific endonucleases called Dicer (Bernstein et al., 2001; Hutvagner et al., 2001; Lee et al., 2004). This results in the generation of short dsRNAs called short interfering RNAs (siRNAs) or, in the case of genome-encoded RNA stem-loops, microRNAs (miRNAs). They have a length of 21-25 nucleotides (nt) and contain 2-nt-overhangs at both 3'-ends. In systems biology, the application of synthetic siRNAs is used to trigger the RNAi pathway. After unwinding of siRNAs, only one strand, the guide strand, is incorporated

into a ribonucleoprotein complex called RISC (RNA induced silencing complex) (Hammond et al., 2000). The RISC effector complex then binds to complementary mRNAs in a sequence-specific manner and mediates mRNA cleavage in the middle of the complementary sequence region (Hutvagner and Zamore, 2002; Martinez et al., 2002).

1.5.2 RNAi-based Screening Approaches

Advances in RNA technologies in the last decade have enabled researchers to functionally analyze large numbers of gene products by targeted gene silencing using RNA interference. One of the first RNAi-based screening approaches was an analysis of all 2300 genes encoded on chromosome III in *C. elegans* regarding their role in cell division (Gonczy et al., 2000). 133 genes were identified to be necessary in cell division in *C. elegans* embryos. A genome-wide viability screen was performed in *Drosophila* cells (Boutros et al., 2004). In this screen, 438 essential genes were identified and functionally classified. A library consisting of more than 5000 enzymatically prepared siRNAs was used to identify genes required for cell division in HeLa cells (Kittler et al., 2004). 37 cell division genes were identified, which included inhibitors as well as accelerators. All kinases of the human genome were analyzed regarding their role in endocytosis, utilizing virus infectivity as an assay for functional endocytosis (Pelkmans et al., 2005). 128 kinases had an effect on vesicular stomatitis virus (VSV) infection and out of these, 92 played a role in clathrin-mediated endocytosis. 34 out of 39 kinases involved in simian virus (SV40) entry also affected caveolae/raft trafficking. In the past two years, several RNAi-based screens were conducted to unravel host factors for infectivity of human immunodeficiency virus (HIV) (Brass et al., 2008; Konig et al., 2008), West Nile virus (Krishnan et al., 2008), dengue virus (DV) (Rothwell et al., 2009; Sessions et al., 2009) and hepatitis C virus (HCV) (Vaillancourt et al., 2009).

1.6 AIM OF THIS THESIS

It is known that an N-terminal SH4-domain directs association with lipid microdomains at plasma membranes by virtue of a dual lipid modification, a myristoylation and one or several palmitoylations. Inhibition of palmitoylation has a severe effect on intracellular localization and results in accumulation at intracellular membranes, while loss of myristoylation inhibits subsequent

palmitoylation rendering the protein soluble. Even though the importance of fatty acylation of the SH4-domain is well characterized, little is known about the cellular machinery participating in intracellular transport of SH4-domain-containing proteins.

It was the aim of this thesis to identify components of the intracellular transport machinery of SH4-domain-containing proteins by an RNAi-based screening approach.

For this purpose, we established a cellular system expressing two distinct fluorescently tagged SH4-domains, which is compatible to a high-content microscopy-based RNAi screening platform.

Moreover, we established a suitable automated image analysis tool, which is able to identify single cells and determine the intracellular distribution of the fluorescent reporter proteins independently.

These tools were used to screen the ~ 22.000 genes of the human genome for gene products involved in intracellular targeting of the two SH4-domain-containing reporter proteins. Candidate gene products identified in a primary screen were validated by the application of independent siRNAs in a validation screen.

2 MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 Chemicals and Consumables

1 kb DNA standard	New England Biolabs, Frankfurt
2-bromopalmitate (2-BP)	Sigma-Aldrich Chemie GmbH, Steinheim
Agar	Becton Dickinson, Le Pont de Claix, France
Agarose electrophoresis grade	Invitrogen Ltd., Paisley, UK
Ampicillin sodium grade	Gerbu Biotechnik GmbH, Gaiberg
Aqua ad iniectabilia	Braun AG, Melsungen
Bovine Serum Albumine, fraction V (BSA)	Carl Roth GmbH, Karlsruhe
Cell Dissociation Buffer (CDB)	Invitrogen Ltd, Paisley, UK
Chloroquine	Sigma-Aldrich Chemie GmbH, Steinheim
CO_2-independent medium	Invitrogen Ltd, Paisley, UK
Collagen R	Serva Electrophoresis, Heidelberg
Dimethyl sulfoxide (DMSO)	J.T. Baker, Deventer, USA
D-MEM	Biochrom KG, Berlin
dNTP-Mix	Peqlab, Erlangen
Doxycyline	Clontech, Palo Alto, USA
Ethanol	Riedel-de Haen, Seelze
Ethidium Bromide	Carl Roth GmbH, Karlsruhe
Ethylene diamine tetraacetic acid (EDTA)	Merck KGaA, Darmstadt
Fetal calf serum (FCS)	PAA Laboratories GmbH, Cölbe
Glycerol	Carl Roth GmbH, Karlsruhe
Hoechst 33342	Sigma-Aldrich Chemie GmbH, Steinheim
L-Glutamine	Biochrom KG, Berlin
Lipofectamine 2000	Invitrogen Ltd, Paisley, UK
Lipofectin	Invitrogen Ltd, Paisley, UK
Magnesium Chloride ($MgCl_2$)	Applichem, Darmstadt
Oligonucleotide	Thermo Scientific, Ulm
Opti-MEM	Invitrogen Ltd, Paisley, UK
Penicillin/Streptomycin	Biochrom KG, Berlin
Potassium chloride (KCl)	Merck KGaA, Darmstadt
Potassium dihydrogen phosphate (KH_2PO_4)	Applichem, Darmstadt

siRNAs	Ambion, Applied Biosystems, Darmstadt
Sodium Chloride (NaCl)	J.T. Baker, Deventer, USA
Sodium hydrogen phosphate (Na_2HPO_4)	Applichem, Darmstadt
Tris	Carl Roth GmbH, Karlsruhe
Trypsin/EDTA	Biochrom KG, Berlin
Tryptone	Becton Dickinson, Le Pont du Claix, France
Xylencyanol FF	Serva Electrophoresis GmbH, Heidelberg
Yeast Extract	Becton Dickinson, Le Pont du Claix, France
Zeocin	Invitrogen Ltd, Paisley, UK

2.1.2 Enzymes

AmpliTaq polymerase	
Age I	New England Biolabs, Frankfurt
BamHI	New England Biolabs, Frankfurt
Nco I	New England Biolabs, Frankfurt
Not I	New England Biolabs, Frankfurt
Sal1	New England Biolabs, Frankfurt
Xba1	New England Biolabs, Frankfurt

2.1.3 Molecular Biological and Biochemical Kits

MBS Mammalian Transfection Kit	Stratagene, La Jolla, USA
Mycoplasma PCR Elisa Kit	Roche, Grenzach-Wyhlen
NucleoSpin Plasmid	Macherey-Nagel, Düren
NucleoBond PC 100	Macherey-Nagel, Düren
NucleoBond PC 500	Macherey-Nagel, Düren
TaKaRa DNA Ligation Kit	Takara Biochemical, Berkeley, USA
QIAquick Gel Extraction Kit	Qiagen GmbH, Hilden
QIAquick PCR Purification Kit	Qiagen GmbH, Hilden

2.1.4 Technical Devices

Bacterial incubator	Infors AG, Einsbach
Bacterial shaker Centromat R	Braun, Melsungen
Centrifuge 5417 R	Eppendorf, Hamburg
Centrifuge Megafuge 1.0 R	Kendro, Langenselbold

Confocal Microscope LSM 510	Zeiss, Göttingen
FACSAria	Becton Dickinson, Heidelberg
FlexDrop PLUS precision reagent dispenser	PerkinElmer, Massachusetts, USA
Fluorescence microscope Axiovert 200M	Zeiss, Göttingen
Gel Doc 2000	Bio-Rad Laboratories GmbH, München
Incubator Heraeus CO_2-Auto-Zero	Kendro, Langenselbold
Kern 474 balance	Kern, Albstadt
MICROLAB STAR liquid handling robot	Hamilton Robotics,
MiVac concentrator speed vac	Genevac, Ipswich, UK
NanoDrop ND-1000 spectrophotometer	Bio-Rad Laboratories GmbH, München
PCR Primus Advanced 26 and 96	Peqlab, Erlangen
Pipetboy acu	Integra Biosciences, Fernwald
Pipets	Gilson, Villiers-le-Bel, France
PowerPack 200 and 300	Bio-Rad Laboratories GmbH, München
Sartorius CP64 balance	Sartorius, Göttingen
Scan^R screening microscope	Olympus, Hamburg
Thermomixer compact and comfort	Eppendorf, Hamburg

2.2 MOLECULAR BIOLOGICAL METHODS

2.2.1 Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) was used to specifically amplify DNA inserts and introduce restriction sites required for cloning of various constructs.

PCR program:

Start	94°C	4 min
Amplification (30 cycles):		
Denaturation	94°C	45 sec
Hybridization	$T < T_m$	45 sec
Elongation	72°C	2 min
End	72°C	10 min
Store	4°C	∞

Standard PCR reaction:

50ng DNA template
5 µl 10x buffer
5 µl $MgCl_2$ (25mM)
1 µl dNTP mix (10 mM each)
1 µl forward primer (10 µM)
1 µl reverse primer (10 µM)
1 µl Taq polymerase (5 U/µl)
ad 50 µl H_2O

PCR products were purified using the QIAquick PCR purification kit (Qiagen) according to manufacturer's instructions. DNA was eluted with 30 µl H_2O.

Name	*Sequence*
GFP-Age for	TATATATATATAaccggtCTGGTGAGCAAGGGCG
GFP-Not rev	TATATATATATAgcggccgcTTACTTGTACAGCTCGTC
GFP mut for	GCTTCAGCCGCTACCCCGACCACATGAAGCAG
mCherry-Age for	ATATATATATATaccggtATGGTGAGCAAGGGCG
mCherry-Not rev	ATATATATATATgcggccgcTTACTTGTACAGCT
dRFP-Age for	ATATATATaccggtGATAAGGATCtGATGGTGGCCt
dRFP-Not rev	ATATATATgcggccgcTTCGAATTCCTACAG
pRevTRE2 for primer	ACCTCCATAGAAGACACCGGGACCG
pRevTRE2 rev primer	GCTAGCTTGCCAAACCTACAGGTGG

Table 1 Primers used for cloning of DNA constructs

2.2.2 Restriction Digests

Restriction DNA endonucleases (New England Biolabs) were used according to manufacturer's instructions. BSA was added to the reaction according to manufacturer's instructions. Digests were performed in 50 µl reaction volume at 37°C for 2 hours. 2 U enzymes were used per µg DNA.

2.2.3 Hybridization of single-stranded Oligonucleotides

Complementary single-stranded oligonucleotides (10 µM) were diluted 1:1000, 10 µl of each oligonucleotide added to the reaction mix, incubated at 95°C for 10 min and slowly cooled down to a temperature below 16°C.

Name	*Sequence*
BamHI-HASPBN18-Age for	gatccCGCCACCATGGGAAGTTCTTGTACAAAGGACTCCGCAAAGGAGCCCCAGAAGAGTGCTGATa
BamHI-HASPBN18-Age rev	ccggtATCAGCACTCTTCTGGGGCTCCTTTGCGGAGTCCTTTGTACAAGAACTTCCCATGGTGGCGg
BamHI-HASPBΔmyr-Age for	gatccCGCCACCATGgcaAGTTCTTGTACAAAGGACTCCGCAAAGGAGCCCCAGAAGAGTGCTGATa
BamHI-HASPBΔmyr-Age rev	ccggtATCAGCACTCTTCTGGGGCTCCTTTGCGGAGTCCTTTGTACAAGAACTtgcCATGGTGGCGg
BamHI-HASPBΔpal-Age for	gatccCGCCACCATGGGAAGTTCTgctACAAAGGACTCCGCAAAGGAGCCCCAGAAGAGTGCTGATa
BamHI-HASPBΔpal-Age rev	ccggtATCAGCACTCTTCTGGGGCTCCTTTGCGGAGTCCTTTGTagcAGAACTTCCCATGGTGGCGg
BamHI-YesN18-Age for	gatccCGCCACCATGGGCTGCATTAAAAGTAAAGAAAACAAAAGTCCAGCCATTAAATACAGACCTa
BamHI-YesN18-Age rev	ccggtAGGTCTGTATTTAATGGCTGGACTTTTGTTTTCTTTACTTTTAATGCAGCCCATGGTGGCGg
BamHI-YesΔmyr-Age for	gatccCGCCACCATGgccTGCATTAAAAGTAAAGAAAACAAAAGTCCAGCCATTAAATACAGACCTa
BamHI-YesΔmyr-Age rev	ccggtAGGTCTGTATTTAATGGCTGGACTTTTGTTTTCTTTACTTTTAATGCAggcCATGGTGGCGg
BamHI-YesΔpal-Age for	gatccCGCCACCATGGGCgccATTAAAAGTAAAGAAAACAAAAGTCCAGCCATTAAATACAGACCTa
BamHI-YesΔpal-Age rev	ccggtAGGTCTGTATTTAATGGCTGGACTTTTGTTTTCTTTACTTTTAATggcGCCCATGGTGGCGg

Table 2 Oligonucleotides used for cloning of DNA constructs

2.2.4 Ligation of DNA Fragments

Ligations were performed in a 1:1, 1:3 and 1:6 molar ratio of vector and insert using the TaKaRa ligation kit 2.1 according to manufacturer's instructions and were incubated at 16°C for 1.5 hours.

A 1:1 molar ratio can be calculated as follows:

50ng vector x bp of DNA insert / bp complete construct = ng insert

2.2.5 Transformation of Competent Bacteria

20 µl competent DH5α *E. coli* cells (Invitrogen) were mixed with 10 µl ligation reaction and on ice for 30 min. After heat shock treatment (20 sec at 37°C), cells were incubated for 2 min on ice, 970 µl LB medium added and incubated for 1 hour at 37°C (200 rpm) in a Thermomixer (Eppendorf). Cells were sedimented (5 min, 300 rpm), resuspended in 100 µl LB medium and spread on agar plates containing appropriate antibiotics for selection. Growth of colonies was allowed by overnight cultivation of agar plates supplemented with appropriate antibiotics at 37°C.

LB medium:

5 g	yeast extract
10 g	tryptone
5-10 g	NaCl
16 g	agar (only for agar plates)
ad 1 l	H_2O

2.2.6 Preparation of Plasmid DNA from Bacteria

LB medium supplemented with appropriate antibiotics (100 µg/ml ampicillin) for selection were inoculated with bacteria and incubated over night at 37°C (200-300 rpm). Plasmid DNA was isolated using different DNA isolation kits from Macherey Nagel according to manufacturer's instructions. DNA was reconstituted in H_2O and DNA concentrations were determined using a NanoDrop ND-1000 spectrophotometer.

2.2.7 Agarose Gel Electrophoresis

For the preparation of 1% agarose gels, corresponding amounts (0.5 g / 50 ml) were solved in 1x TAE buffer by heating in a microwave. After cooling down, ethidium bromide was added to a final concentration of 0.5 µg/ml. The gel was poured in casting chambers and combs were inserted for formation of sample slots.
Samples were mixed with sample buffer in a 1:5 dilution and loaded in sample slots on the agarose gel. 1kb DNA ladder (New England Biolabs) was used for size determination of DNA fragments. For this purpose, the stock solution was diluted 1:20 in water and 5x loading buffer (10 µl DNA ladder, 40 µl 5x sample buffer, 150 µl H_2O) and 10 µl loaded. Electrophoretic sample separation was performed in 1x TAE buffer at 120 V for 20 minutes and documented using a Gel Doc 2000 imaging system (Bio-Rad).

TAE buffer (50x)

242 g	Tris
57.1 ml	Glacial acidic acid
100 ml	EDTA (0.5 M, pH 8)
ad 1 l	H_2O

Sample buffer (5x)

0.25% (w/v)	Bromphenol blue
0.25% (w/v)	Xylencyanol FF
30% (w/v)	Glycerol

2.2.8 Cloning of DNA constructs

Molecular cloning refers to the insertion of DNA fragments in plasmid vectors and the clonal selection of the construct after transformation into *E. coli*. After

restriction digest of insert and plasmid vector with suitable restriction enzymes, fragments were separated by agarose gel electrophoresis and purified using a QiaQuick gel extraction kit (Qiagen) according to manufacturer's instructions and eluted into 30 µl H_2O. After ligation of vector and insert, plasmids were transformed into *E. coli*, which were plated on antibiotic-containing agar plates for selection of clones. Single colonies were picked and used to inoculate LB medium for preparation of plasmid DNA.

2.2.9 DNA sequencing

For validation of correct DNA sequences, DNA samples and appropriate primers were sent to a commercial sequencing company (GATC, Konstanz). Obtained sequences were compared to theoretical sequences using the freeware 4 Peaks (Mek&Tosj).

2.3 CELL CULTURE TECHNIQUES

2.3.1 Maintenance of Stable Cell Lines

Adherent HeLa and Kyoto cell lines were cultivated on culture dishes of different sizes D-MEM at 37°C with 5% CO_2. HEK cells were cultivated in D-MEM on collagen-coated plates at 37°C with 5% CO_2. Cell culture dishes were coated with collagen by rinsing with a 1mg/ml collagen solution (Serva) and drying over night. For splitting of cells, the cultivation medium was discarded, cells were washed with PBS and detached by incubation in Trypsin/EDTA (Biochrom KG) for 5 min at 37°C. To quench trypsin activity, fresh medium was added and cells were seeded at desired concentrations in new culture dishes containing fresh medium.

PBS (1x)

140 mM	NaCl
2.7 mM	KCl
10mM	Na_2HPO_4
1.8 mM	KH_2PO_4

HeLa / Kyoto cultivation medium:

10% (v/v)	fetal calf serum (FCS)
2mM	L-Glutamine
100 µg/ml	Penicillin/Streptomycin
in D-MEM	

2.3.2 Freezing and Thawing of Cells

For freezing, cells were cultivated on 10 cm dishes until 90% confluency. The cell culture medium was discarded, cells washed with PBS and detached by incubation in the presence of low amounts of trypsin/EDTA for 2 min at 37°C. Cells were resuspended in 3 ml freezing medium per 10 cm plate and transferred to two cryo tubes (Eppendorf) (1.5 ml each). Tubes were incubated for 24-48 hours at -80°C in cryo boxes to ensure a slow decrease in temperature. Samples were transferred to liquid nitrogen for long-term storage. For thawing, cryo tubes were incubated in a 37°C waterbath to ensure fast thawing and the thawed cell solution was transferred to 13.5 ml prewarmed cultivation medium. Cells were sedimented at 200 g for 3 min, resuspended in fresh prewarmed medium and plated on 10 cm dishes.

Freezing medium:

20% (v/v)	fetal calf serum (FCS)
10% (v/v)	DMSO
2mM	L-Glutamine
100 µg/ml	Penicillin/Streptomycin
in D-MEM	

2.3.3 Lipofectin-based Transfection of Mammalian Cells

To generate stable Kyoto cell lines expressing the viral receptor MCAT-1, cells were transfected with the pcDNA3.1/Zeo-MCAT-1 construct using Lipofectin reagent (Invitrogen) according to manufacturer's instructions. 48 hours post-transfection, cells, which incorporated the construct, were selected by cultivation in the presence of 600 µg/ml Zeocin (Invitrogen) for two weeks.

2.3.4 Retroviral Transduction of Mammalian Cells

The MBS mammalian transfection kit (Stratagene) was used according to manufacturer's instructions for retroviral transduction of mammalian cells. The

procedure involves the production of plasmid-containing retroviral particles after transfection of HEK 293T cells with two vectors required for virus production (pVPack-GP and pVPack-Eco) and the "cargo" plasmid. Supernatant containing the viral particles was harvested and used for transduction of MCAT-1-expressing HeLa or Kyoto cells. 48 hours post-transduction, cells were either selected by cultivation in the presence of 1000 µg/ml G418 for two weeks for stable insertion of pFB(Neo)-rtTA-M2 or sorted for inducible reporter gene expression from the pRev-TRE2 vector by several rounds of sorting using a FACSAria flow cytometer (Becton Dickinson).

2.3.5 Fluorescence-activated Cell Sorting

Fluorescence-activated cell sorting (FACS) was performed to select for cell lines, which stably incorporated various reporter molecules and expresses these in a doxycyline-regulated manner. Cells were cultivated for 48 hours in the presence of or without 1 µg/ml doxycycline (for "bright" and "dark" sort, respectively) to a final confluency of 90%, washed with PBS and detached using Cell Dissociation Buffer (CDB, Invitrogen) and reconstituted in cold sorting medium containing 2 µg/ml propidium iodide (PI) to stain dead cells. Subsequently, cells were passed through a cell strainer cap (Becton Dickinson) and sorted according to fluorescence intensities employing a FACSAria flow cytometer (Becton Dickinson). To obtain cell lines which stably express reporter proteins in a tightly doxycyline-regulated manner, cells were subjected to a triple sorting procedure: fluorescent cells were collected in the first "bright" sort, which was followed by sorting of non-fluorescent cells in a "dark" sort. A final "bright" sort was performed and resulted in the selection of stable, doxycyline-inducible cell pools. Single clones were selected in the second bright sort.

Sorting medium:

5%	CDB
0.2%	FCS
2mM	L-Glutamine
100µM	Pen/Strep
in D-MEM	

2.4 RNA TECHNIQUES

2.4.1 Liquid-phase Transfection of Mammalian Cells with siRNAs

Adherent or detached HeLa cells were transfected with siRNAs (Ambion) in 24-well plates using oligofectamine (Invitrogen) with slight modifications of the manufacturer's instructions. For each well, 50 pmol siRNA (30-50µM) was diluted in 50 µl Opti-MEM (Invitrogen) and mixed with 12 µl oligofectamine, which was diluted 1:4 in Opti-MEM (3 µl oligofectamine, 9 µl Opti-MEM). The siRNA-oligofectamine mix was incubated for at least 20 minutes at room temperature and then slowly added to detached or adherent HeLa cells in 450 µl Opti-MEM (24-well plate, 60% confluency). Cells were incubated for 4-6 hours ate 37°C, before the medium was replaced by normal growth medium. 36 hours post-transfection, cells were splitted onto 8-well LabTek imaging chambers (Nunc) at 40-60% confluency and incubated for additional 12-24 hours (48 hours or 60 hours total knockdown time, respectively) in the presence of 1 µg/ml doxycyline. A comprehensive list of siRNAs used is attached in the appendix (5.4)

2.4.2 Reverse Transfection of Mammalian Cells with siRNAs

For reverse transfection with siRNAs, HeLa cells were seeded on siRNA arrays or coated multi-well plates containing an immobilized siRNA/transfection-reagent mix. For 48 hours total knockdown time on 384-spot siRNA arrays printed in 1-well LabTek imaging chambers (Nunc), 120.000 cells were seeded in 3.5 ml cultivation medium on each LabTek. For 48 hours knockdown time on siRNA-coated 96-well plates (Greiner), 4.000 cells were seeded in 100 µl D-MEM using a FlexDrop system (Perkin Elmer). For 48-60 hours total knockdown time on siRNA-coated 8-well-well LabTek imaging chambers (Nunc), 4.000 cells were seeded in 3.5 ml cultivation medium on each LabTek. Medium was replaced 36 hours post-transfection to prewarmed D-MEM containing 1 µg/ml doxycyline. A comprehensive list of siRNAs used is attached in the appendix (5.4)

2.4.3 Preparation of siRNA-coated multi-well Plates

The following protocol can be used for parallel coating of 12 96-well plate replicas (Erfle et al., 2008). 554 µl sucrose/Opti-MEM solution were

supplemented with 646 µl Lipofectamine 2000 (Invitrogen) solution, which was previously diluted 1:2 in H_2O if silencer select siRNAs (Ambion) were used, to prepare a transfection reagent mix. For each siRNA sample, 5 µl (3 µM) silencer select siRNA (Ambion) or 5 µl (30 µM) silencer siRNA (Ambion), was added to 6.5 µl of the transfection mix and incubated for at least 30 minutes at room temperature. 7.25 µl fibronectin/gelatine solution were added to yield a final volume of 18.75 µl. The usage of fibronectin is this step is optional and can be omitted by using a 0.2% gelatine solution without fibronectin. 15 µl of this siRNA-containing transfection mixture were diluted in 750 ml H_2O, and 50 µl of the diluted mixture were transferred to one well of each of the 12 96-well plates (Greiner). Liquid handling was performed on a MICROLAB STAR robot (Hamilton). The solution was dried in a MiVac concentrator (Genevac) between 37°C and 52°C for 45-90 minutes. Coated multi-well plates can be stored ready-to-use at room temperature in the presence of drying pearls (Fluka) for several months. A comprehensive list of siRNAs used is attached in the appendix (5.4)

Sucrose/Opti-MEM solution:

1.37 g sucrose
in 10 ml Opti-MEM (solve without shaking)

0.2% gelatine solution:

0.2 g gelatine
in 100 ml H_2O (heat to 56°C for 20 min, filter through 0.45 µm filter)

Fibronectin/gelatine solution:

10 µl fibronectin (human plasma, 0.1%, Sigma)
10 ml 0.2% gelatine solution

2.4.4 Preparation of siRNA-coated LabTeks

The following protocol can be used for parallel coating of 8 8-well LabTek replicas. For each siRNA sample, 3 µl sucrose/Opti-MEM solution were supplemented with 3.5 µl Lipofectamine 2000 (Invitrogen). When silencer select siRNAs (Ambion) were used instead of silencer siRNAs (Ambion), Lipofectamine 2000 was diluted 1:2 diluted in H_2O. After addition of 5 µl (3 µM) silencer select or 5 µl (30 µM) silencer siRNA (Ambion), the siRNA/Lipofectamine-mix was incubated for at least 20 minutes at room temperature. 7.25 µl 0.2 % gelatine solution were added to yield a final volume of 18.75 µl. 16 µl of this siRNA-

containing mixture were diluted in 800 ml H_2O, and 100 µl of the diluted mixture were transferred to one well of an 8-well LabTek. The solution was dried in a MiVac concentrator (Genevac) at 45°C for 30 minutes. Coated LabTeks can be stored ready-to-use at room temperature in the presence of drying pearls (Fluka) for several months. A comprehensive list of siRNAs used is attached in the appendix (5.4)

2.5 LIVE-CELL FLUORESCENCE MICROSCOPY

2.5.1 Widefield and Confocal Fluorescence Microscopy

For imaging of live cells, medium was replaced by imaging medium containing 0.2 µg/ml Hoechst 33342 (Sigma) 30 minutes before imaging, for staining of cell nuclei and cells were subsequently subjected to live-cell imaging on a Axiovert 200M fluorescence microscope (Zeiss) or a LSM 510 confocal microscope (Zeiss).

Imaging medium:

10% (v/v)	FCS
2mM	L-Glutamine
100 µg/ml	Pen/Strep

in CO_2-independent medium with or without phenol-red (Invitrogen)

2.5.2 Automated Widefield Fluorescence Microscopy

For automated live-cell microscopy, the first spot on siRNA arrays was marked from the bottom with a permanent marker to determine the starting position. For imaging of live cells, medium was replaced by imaging medium containing 0.2 µg/ml Hoechst 33342 (Sigma) 30 minutes before imaging, for staining of cell nuclei and cells were subsequently subjected to live-cell imaging on a Scan^R screening station (Olympus).

Microscope settings:

Gradient autofocus
2-5 min exposure time, 2x2 binning
DAPI filter cube, UPLFLN or UPLSAPO 10x objective
Coarse AF: 20 layers, 4 µm step width, ± 40 µm scan range
Fine AF: 20 layers, 1.5 µm step width, ± 15 µm scan range
Fluorescence microscopy
1 layer, 1 cycle, 1x1 binning
DAPI/GFP/Cy3 filter cube, UPLFLN or UPLSAPO 10x objective
Well manager:
384 siRNA arrays: 1125 µm row and column spacing
96-well plates: 2 images per well

2.6 MICROSCOPY-BASED RNAI SCREENING

2.6.1 Primary Screening on 384-spot siRNA Arrays

The genome-wide primary screen was conducted by reverse transfection of a human silencer siRNA library (Ambion) consisting of ~ 52.000 siRNAs targeting ~ 22.000 genes (ensembl version 27) on previously spotted siRNA arrays (Erfle et al., 2007; Erfle and Pepperkok, 2007; Erfle et al., 2004; Neumann et al., 2006; Simpson et al., 2007). HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells were thawn, treated for one week with mycoplasma removal agent (Serotec) according to manufacturer's instructions and tested one week later for mycoplasma contamination using the Mycoplasma PCR Elisa Kit (Roche) according to manufacturer's instructions. Cells were then splitted in different concentrations on 5 cm dishes and used for reverse transfection when a confluency of ~ 80% was reached. For each experiment, duplicates of 2 different siRNA arrays were used in parallel. 120.000 cells were seeded in 3.5 ml cultivation medium on each 384-spot siRNA array. Medium was replaced 36 hours post-transfection to prewarmed D-MEM containing 1 µg/ml doxycyline. 47.5 hours post-transfection, medium was replaced by prewarmed imaging medium (without phenol-red, Invitrogen) containing 0.2µg/ml Hoechst 33342 (Sigma). 30 minutes later, LabTeks were loaded onto a Scan^R screening microscope (Olympus) in an alternating order (position 1: array 1 – replica 1, position 2: array 2 – replica 1, position 3: array 1 – replica 2, position 4: array 2 – replica 2) and live-cell imaging in the Hoechst, GFP and Cherry channel was performed (duration: 4-6 hours). LabTeks 1-49 contained siRNA targeting ~ 5.600 genes from the "druggable" genome, LabTeks 60-159 contained siRNAs

targeting genes from the “non-druggable” genome. Acquired images were analyzed using the single-cell image analysis tool.

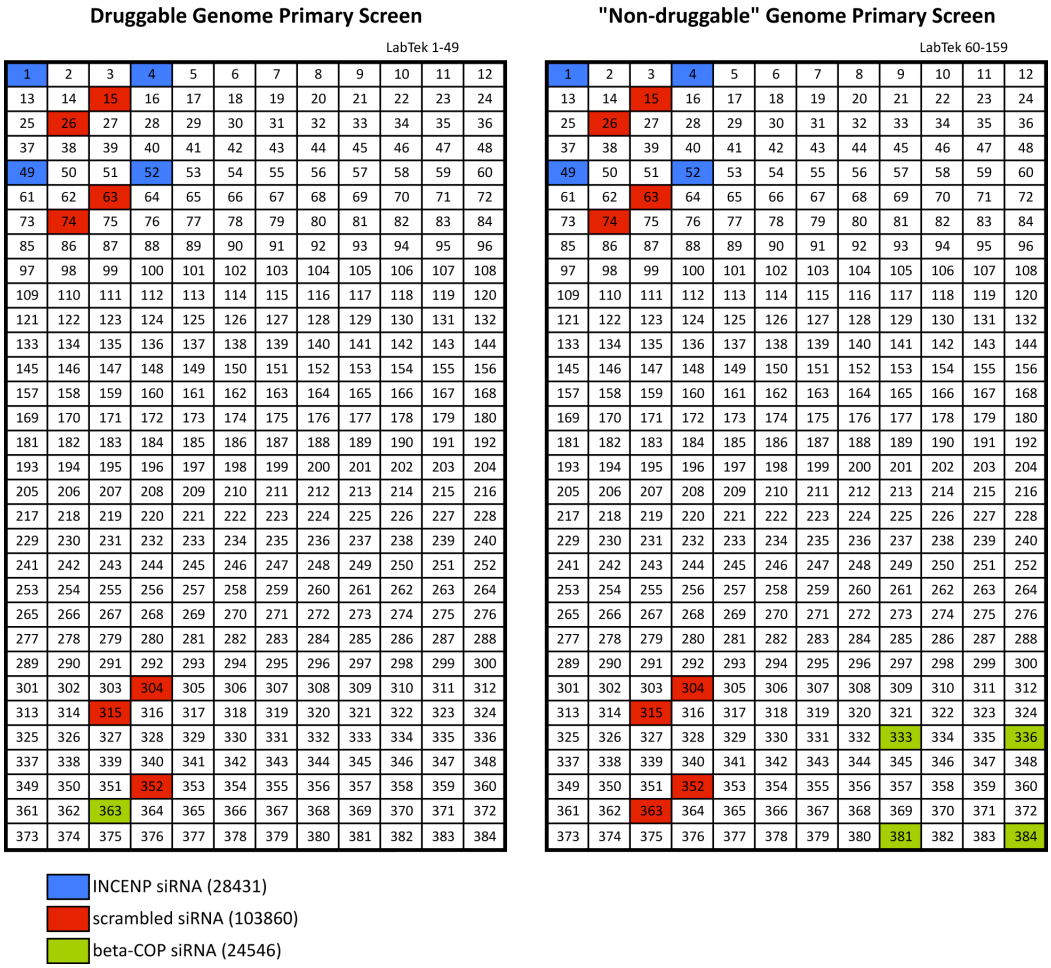

Figure 10 Layout of siRNA arrays used in the primary screen. Positions of control-siRNAs are highlighted.

2.6.2 *Validation Screening on siRNA-coated multi-well Plates*

The genome-wide secondary screen was conducted by reverse transfection of a validation library consisting of 776 silencer select siRNAs (Ambion) targeting 311 genes (ensembl version 53) on siRNA-coated 96-well plates (Erfle et al., 2008). HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells were thawn, treated for one week with mycoplasma removal agent (Serotec) according to manufacturer’s instructions and tested one week later for mycoplasma contamination using the Mycoplasma PCR Elisa Kit (Roche) according to manufacturer’s instructions. Cells were then splitted in different concentrations on 10 cm dishes and used for reverse transfection when a confluency of ~ 80% was reached. For each experiment, four different 96-well plates (Greiner) were

used in parallel. In total, four replicas were performed from each multi-well plate. 4.000 cells were seeded in 100 µl D-MEM using a FlexDrop system (Perkin Elmer). Medium was replaced 36 hours post-transfection to prewarmed D-MEM containing 1 µg/ml doxycyline. 47.5 hours post-transfection, medium was replaced by prewarmed imaging medium (without phenol-red, Invitrogen) containing 0.2µg/ml Hoechst 33342 (Sigma). 30 minutes later, 96-well plates were subsequently loaded onto a Scan^R screening microscope (Olympus) and live-cell imaging in the Hoechst, GFP and Cherry channel was performed (duration per plate: 45 min). Acquired images were analyzed using the single-cell image analysis tool. Plates I-V included siRNAs targeting hits identified in the "druggable" genome, and plates VI-VIII and TW included siRNAs targeting hits identified in the "non-druggable" genome. Acquired images were analyzed using the single-cell image analysis tool.

Druggable Genome Validation Screen

Plate I-V

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12

"Non-druggable" Genome Validation Screen

Plate VI-VIII, TW

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12

INCENP siRNA (s7424) | beta-COP siRNA (s3371) | alpha-COP siRNA (s3369)
scrambled siRNA (s813) | beta-COP siRNA (s3372)

Figure 11 Layout of siRNA-coated 96-well plates used in the validation screen. Positions of control-siRNAs are highlighted.

2.7 Analysis of High-content Screening Data

2.7.1 Image Analysis with the Automated Single-cell Analysis Tool

Images were analyzed using the automated image analysis tool established in this study (Remmele et al., 2008, see appendix). The analysis was performed on Macintosh computer (Apple) in the Windows partition (Bootcamp). The

software tool is programmed as GUI in Matlab R2007b (the MathWorks). For analysis of data from the primary screen, GUI V8 was used. Data were copied using the “copyCellFiles” command and analysis started using the “CellAnalyse” command. For analysis of data from the validation screen, MATLABtool_validation was used. Data were copied using the “copyVaidationCellFiles” command and analysis started using the “CellAnalyse” command. In the validation tool, the standard deviation of the whole cell area was neglected for determination of the “strong intracellular retention” phenotype. Results of the cell classification were stored as txt-files.

2.7.2 Spreadsheet Analysis of Cell Classification Results

Results of the cell classification were subjected to further analysis in excel (Microsoft). In short, the percentage of each “phenotype” was determined for each image. In the validation screen, cells from both images of the same well were counted together in further analysis. From percentages of one complete array, we calculated the median and the standard deviation and used these to transform each percentage into a score, which for each well is the distance in standard deviations to the median of the complete LabTek. In the validation screen, median and standard deviation from all negative controls on the same plate were used. The hit threshold for SH4-HASPB-GFP was adjusted to a score of 2.5 in one hit class on both replicas in the primary screen and 2 in one hits class in 3 our of four or three out of two replicas in the secondary screen. In the druggable genome, SH4-HASPB-GFP hits were included in the validation screen when the hit threshold was above 5 on one replica only. The hit threshold for SH4-Yes-Cherry was adjusted to a score of 3.5 in the primary screen and 2 in the secondary screen.

3 RESULTS

Proteins containing an N-terminal SH4-domain localize to the cytoplasmic leaflet of lipid microdomains at the plasma membrane. So far, the underlying molecular machinery that mediates targeting of these proteins to the plasma membrane has been elusive. In this work we aimed to identify gene products involved in intracellular trafficking of SH4-domain-containing reporter proteins to the plasma membrane. For this purpose, we established an experimental system to screen the complete human genome for components involved in this process. This required the generation of a human cell line for RNAi-mediated knockdown of genes, which is compatible with a high-content microscopy-based screening platform. Furthermore, a crucial prerequisite was the establishment of a suitable evaluation method, which enabled us to analyze acquired data in a reliable, fast, unbiased and quantitative fashion. Using these tools, we applied a genome-wide siRNA-library to screen for gene products involved in intracellular targeting and transport of SH4-domain-containing proteins and were able to validate 60 hits with siRNAs of independent sequence.

3.1 Generation and Characterization of Human Cell Lines Expressing SH4-domain-containing Reporter Proteins

In previous studies performed in our laboratory, Chinese hamster ovary (CHO) cells inducibly expressing the SH4-domain of *Leishmania* hydrophilic acylated surface protein B (HASPB) were used to study the export mechanism of HASPB (Stegmayer et al., 2005; Tournaviti et al., 2009). CHO cells expressing fusion proteins of SH4-HASPB or the SH4-domains of endogenous Src kinase family members were used to study SH4-mediated membrane dynamization (Tournaviti et al., 2007). For our purpose, however, we required a human model cell line, which inducibly expresses SH4-domain-containing proteins, since the genome-wide siRNA library is designed to target all genes of the human genome.

3.1.1 Generation of SH4-domain-containing Reporter Constructs

We designed reporter proteins consisting of the N-terminal SH4-domains of *Leishmania* HASPB (SH4-HASPB) and the human Src kinase Yes (SH4-Yes) fused

to fluorescent proteins of different excitation and emission spectra. All SH4-domains used in this study are comprised of the N-terminal 18 amino acids of the respective protein. All constructs contain a single KOZAK sequence upstream of the SH4 domain and were cloned into the pRevTRE2-vector to enable retroviral transduction, subsequent stable insertion into the genome and inducible expression upon doxycyline induction.

SH4-HASPB	GFP
ATGGGAAGTTCTTGTACAAAGGACTCCGCAAAGGAGCCCCAGAAGAGTGCTGAT (M)GSSCTKDSAKEPQKSAD	Insert length 780bp Protein size 28.9 kD

SH4-HASPB-Δmyr	GFP
ATGgcaAGTTCTTGTACAAAGGACTCCGCAAAGGAGCCCCAGAAGAGTGCTGAT (M)ASSCTKDSAKEPQKSAD	Insert length 780bp Protein size 28.9 kD

SH4-HASPB-Δpal	GFP
ATGGGAAGTTCTgctACAAAGGACTCCGCAAAGGAGCCCCAGAAGAGTGCTGAT (M)GSSATKDSAKEPQKSAD	Insert length 780bp Protein size 28.9 kD

SH4-HASPB	Cherry
ATGGGAAGTTCTTGTACAAAGGACTCCGCAAAGGAGCCCCAGAAGAGTGCTGAT (M)GSSCTKDSAKEPQKSAD	Insert length 771bp Protein size 28.7 kD

SH4-HASPB	dRFP
ATGGGAAGTTCTTGTACAAAGGACTCCGCAAAGGAGCCCCAGAAGAGTGCTGAT (M)GSSCTKDSAKEPQKSAD	Insert length 1464bp Protein size 55.1 kD

SH4-Yes	GFP
ATGGGCTGCATTAAAAGTAAAGAAAACAAAAGTCCAGCCATTAAATACAGACCT (M)GCIKSKENKSPAIKYRP	Insert length 780bp Protein size 29.1 kD

SH4-Yes-Δmyr	GFP
ATGgccTGCATTAAAAGTAAAGAAAACAAAAGTCCAGCCATTAAATACAGACCT (M)ACIKSKENKSPAIKYRP	Insert length 780bp Protein size 29.1 kD

SH4-Yes-Δpal	GFP
ATGGGCgccATTAAAAGTAAAGAAAACAAAAGTCCAGCCATTAAATACAGACCT (M)GAIKSKENKSPAIKYRP	Insert length 780bp Protein size 29.1 kD

SH4-Yes	Cherry
ATGGGCTGCATTAAAAGTAAAGAAAACAAAAGTCCAGCCATTAAATACAGACCT (M)GCIKSKENKSPAIKYRP	Insert length 771bp Protein size 28.9 kD

ATGGGCTGCATTAAAAGTAAAGAAAACAAAAGTCCAGCCATTAAATACAGACCT	Insert length	1464bp
(M)GCIKSKENKSPAIKYRP	Protein size	55.3 kD

Table 3 Overview of the DNA and protein sequences and sizes of the different SH4-domain-containing reporter proteins generated in this study. Underlined amino acids are acylated: Glycine 2 is myristoylated in SH4-HASPB and SH4-Yes, Cysteine 5 is palmitoylated in SH4-HASPB and Cysteine 3 is palmitoylated in SH4-Yes. Amino acids and base triplets highlighted in red are mutations from the corresponding wild-type SH4-domain and result in loss of acylation.

The N-terminal 18 amino acids of HASPB and Yes constitute SH4-domains (SH4-HASPB and SH4-Yes, respectively) and were used to generate SH4-domain-containing reporter proteins. After removal of the N-terminal methionine, a myristoyl group is attached to the alpha-amino-group of glycine 2 via an amide bond in both SH4-HASPB and SH4-Yes. Subsequent attachment of a palmitoyl group via a thioester-linkage occurs at cysteine 5 in SH4-HASPB and cysteine 3 in SH4-Yes. Based on these constructs, mutants were designed that were deficient in both, myristoylation and palmitoylation (mutation of glycine 2 to alanine in SH4-HASPB-Δmyr and SH4-Yes-Δmyr), or palmitoylation alone without affecting myristoylation (mutation of cysteine 3 and cysteine 5 in SH4-HASPB-Δpal and SH4-Yes-Δpal). All constructs contain a single KOZAK sequence upstream of the SH4-domain, whereas the N-terminal methionine was removed from GFP. All constructs contain a *BamHI* cleavage site upstream of the SH4-domain, an *AgeI* cleavage site between the SH4-domain and the fluorescent reporter protein and a *NotI* cleavage site behind the fluorescent reporter and were cloned into the multiple cloning site (MCS) of the retroviral vector pRevTRE2.

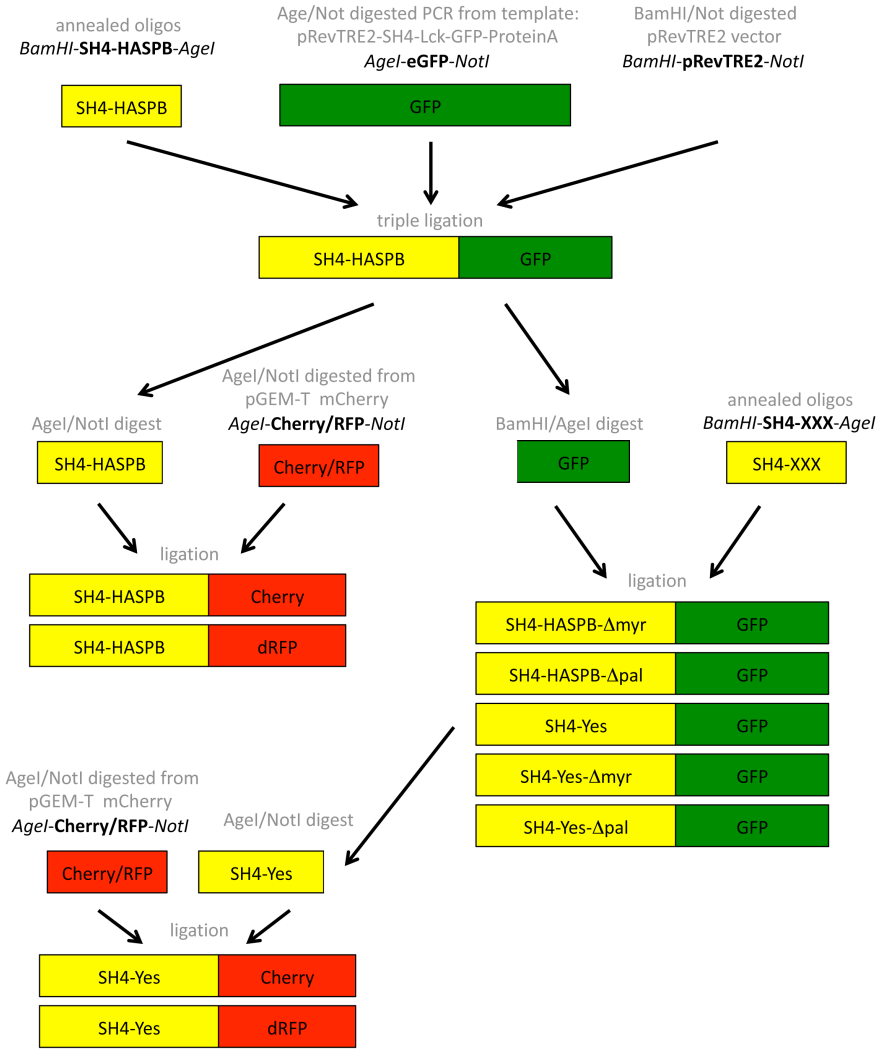

Figure 12 Schematic overview of the cloning strategy for the generation of reporter proteins containing different (wild-type and mutant) SH4-domains fused to fluorescent proteins.

As depicted in Figure 12, SH4-domain-containing reporter constructs (see Table 3) were cloned into the MCS of pRevTRE2. In order to conveniently swap fluorophores and SH4-domains, the initial construct was generated by a triple ligation. Subsequently, removal of the SH4-domain was accomplished by sequential digestion with the restriction enzymes *BamHI* and *AgeI* to allow substitution by other SH4-domains by ligation of annealed oligonucleotides. Removal of the fluorophore was achieved by sequential digestion with the restriction enzymes *AgeI* and *NotI*. For enzymatic digestion, PCR products of fluorophores were subcloned into the pGEM-T vector, removed by sequential digestion with the restriction enzymes *AgeI* and *NotI* and ligated into the *AgeI/NotI*-digested SH4-domain-containing-pRevTRE2 vector.

3.1.2 *Generation of Kyoto Cell Lines Expressing SH4-domain-containing Reporter Proteins*

Due to the reduced potential for cross-contaminations by cell migration between neighbouring siRNA-containing spots, a strongly adherent HeLa cell line termed Kyoto was used in previous screens on siRNA arrays (Neumann et al., 2006, Simpson et al., 2006). To take advantage of the low migration behaviour of these cells in our genome-wide RNAi screen, we generated doxycyline-inducible Kyoto cell lines expressing fluorescent SH4-domain-containing reporter proteins.

In a first step, in order to express the retroviral surface receptor MCAT-1 (murine cationic amino acid transporter 1), the cells were transfected with pcDNA3.1/Zeo-MCAT-1 using Lipofectin reagent (Invitrogen). Subsequently, cells were cultivated in the presence of 600 µg/ml Zeocin (Invitrogen) for two weeks to select for cells, which display the viral surface receptor at the cell surface ($Kyoto_{MCAT}$). In the o allow the induction of protein expression with doxycyline, $Kyoto_{MCAT}$ cells were then transduced with pFB(Neo)-rtTA-M2 using the MBS mammalian transfection kit (Stratagene) according to manufacturer's instructions and were subsequently selected for the incorporation of the doxycyline-sensitive transactivator rtTA-M2 ($Kyoto_{MCAT-TAM2}$) by cultivation in the presence of 1000 µg/ml G418. In the final step, $Kyoto_{MCAT-TAM2}$ cells were transduced with different SH4-domain containing fluorescent reporter proteins (listed in Table 3) in the pRevTRE2 vector using the MBS mammalian transfection kit (Stratagene). The pRevTRE2 vector is a retroviral vector containing a tet-response element, which is a target for rtTA-M2 transactivator binding. Following retroviral transduction, cells were induced by addition of 1 µg/ml doxycyline for 48 hours and subjected to fluorescent activated cell sorting (FACS) using a FACSAria flow cytometer (Becton Dickinson) to isolate cell pools and single clones expressing the fluorescent proteins of interest. Seven days later, isolated cells were subjected to a "dark sort" without induction of protein expression by doxycycline to select cells that express the reporter molecules in an inducible, rather than a constitutive manner. An additional "bright sort" after cultivation in the presence of 1 µg/ml doxycyline for 48 hours was performed to exclude cells, which had not stably incorporate the reporter constructs. This sorting procedure usually yields homogeneous cell pools, which express the protein of interest in a tightly regulated, doxycyline-dependent manner. Single clones were isolated in the first "bright sort". Cell lines, which

simultaneously express two fluorescent reporter proteins were first transduced with and sorted for the Cherry-containing reporter protein and subsequently transduced with and sorted for the GFP-containing reporter construct.

3.1.3 *Characterization of Kyoto Cell Lines Expressing SH4-domain-containing Reporter Proteins by Flow Cytometry*

To characterize reporter protein expression levels and homogeneity of isolated cell lines, cells were cultivated for 48 hours without or in the presence of 1 µg/ml doxycyline, detached using PBS/EDTA and analyzed by flow cytometry (FACSCalibur, Becton Dickinson).

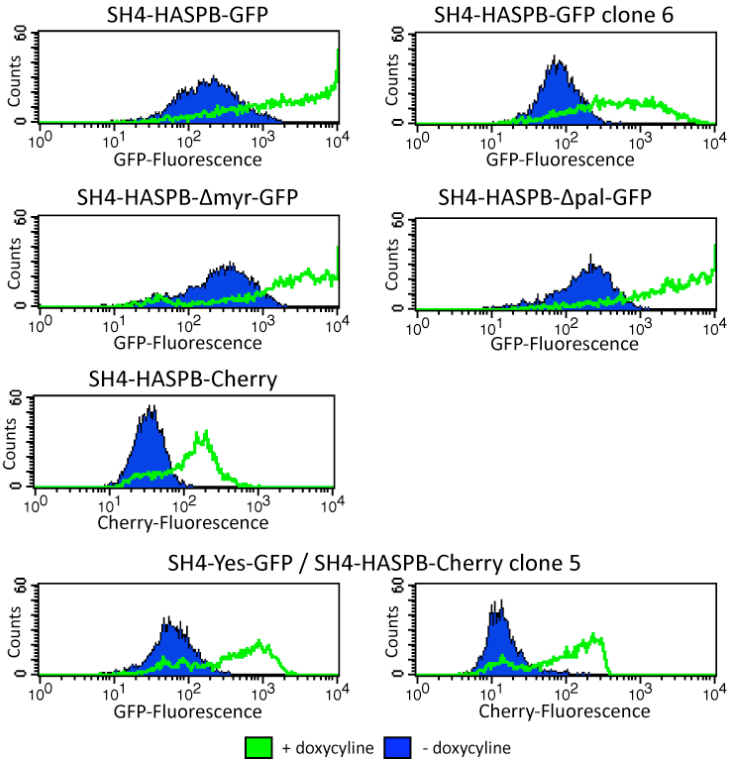

Figure 13 Flow cytometric analysis of Kyoto cell lines expressing different SH4-HASPB-domain-containing reporter proteins in a doxycyline-inducible manner. Stable Kyoto cell lines expressing wild-type and mutant SH4-HASPB-fusion proteins were cultivated in the absence (blue filled profile) and presence (green empty profile) of 1 µg/ml doxycyline for 48 hours, detached, and analyzed by flow cytometry using a FACSCalibur flow cytometer (Becton Dickinson). Autofluorescence of $Kyoto_{MCAT\text{-}TAM2}$ cells was adjusted to 10^1.

Figure 13 depicts the expression of wild-type and mutant SH4-HASPB-domain-containing proteins in Kyoto cells before and after induction of protein expression by cultivation in the presence of 1 µg/ml doxycyline for 48 hours. Autofluorescence of $Kyoto_{MCAT\text{-}TAM2}$ cells was adjusted to 10^1. Most cell lines show low levels of constitutive reporter protein expression as indicated by

fluorescence levels above 10^1 in the absence of doxycyline. In all cases reporter protein expression was increased by at least one order of magnitude after cultivation in the presence of 1 µg/ml doxycyline. Induction of reporter protein expression, however, does not result in homogeneous reporter expression but yields a widely distributed expression pattern.

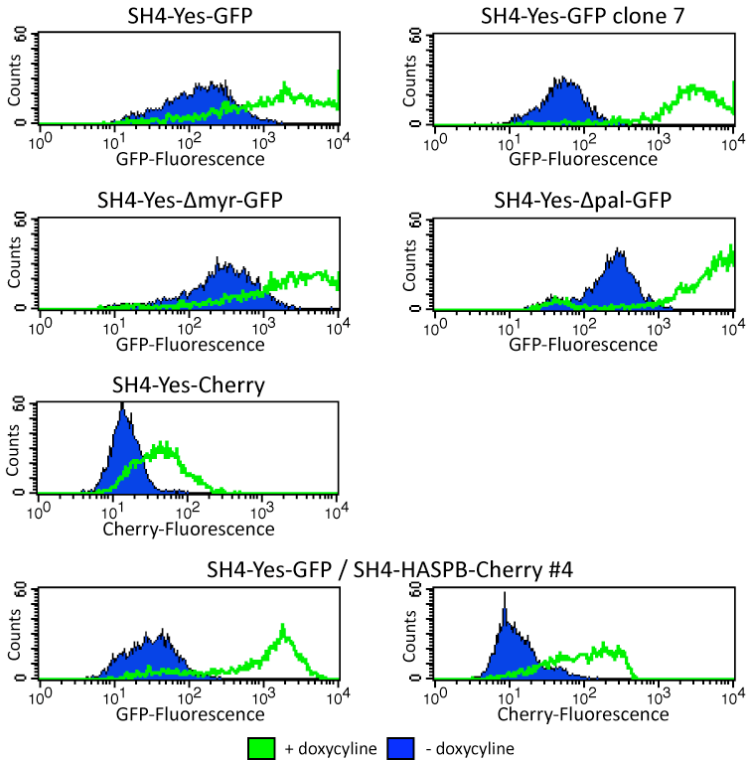

Figure 14 Flow cytometric analysis of Kyoto cell lines expressing different SH4-Yes-domain-containing reporter proteins in a doxycyline-inducible manner. Stable Kyoto cell lines expressing wild-type and mutant SH4-Yes-fusion proteins were cultivated in the absence (blue filled profile) and presence (green empty profile) of 1µg/ml doxycyline for 48 hours, detached, and analyzed by flow cytometry using a FACSCalibur flow cytometer (Becton Dickinson). Autofluorescence of $Kyoto_{MCAT\text{-}TAM2}$ cells was adjusted to 10^1.

Figure 14 depicts the expression of wild-type and mutant SH4-Yes-domain-containing reporter proteins in Kyoto cells before and after induction of reporter protein expression by cultivation in the presence of 1µg/ml doxycyline. Autofluorescence of $Kyoto_{MCAT\text{-}TAM2}$ cells was adjusted to 10^1. Most cell lines show low levels of constitutive reporter protein expression as indicated by fluorescence levels above 10^1 in the absence of doxycyline. In all cases reporter protein expression was increased by at least one order of magnitude after cultivation in the presence of 1µg/ml doxycyline for 48 hours. Induction of reporter protein expression, however, does not result in homogeneous reporter expression but yields a widely distributed expression pattern.

Taken together, the wide expression profile of SH4-domain-containing reporter proteins in stable Kyoto cell lines after doxycyline induction appeared not suitable for application in the genome-wide RNAi screen due to large intensity differences between individual cells.

Therefore, we decided to generate HeLa cell lines expressing SH4-domain-containing reporter proteins expressing reporter proteins in a more homogeneous manner.

3.1.4 Generation of HeLa Cell Lines Expressing SH4-domain-containing Reporter Proteins

SH4-domain-containing reporter proteins listed in Table 3 were transduced into HeLa cells containing the retroviral receptor MCAT-1 and the doxycyline-sensitive transactivator rtTA-M2 ($HeLa_{MCAT\text{-}TAM2}$). Cells were submitted to a triple fluorescent activated cell sorting procedure as described in section 3.1.2 to isolate homogeneous, tightly regulated, doxycyline-inducible cell pools. Single clones were isolated in the first “bright sort”. Cell lines, which simultaneously express two fluorescent reporter proteins were first transduced with and sorted for the Cherry-containing reporter protein and subsequently transduced with and sorted for the GFP-containing reporter construct.

3.1.5 Characterization of HeLa Cell Lines Expressing SH4-domain-containing Reporter Proteins by Flow Cytometry

To further characterize the homogeneity and expression levels of the isolated cell lines, cells were cultivated for 48 hours without or in the presence of 1 µg/ml doxycyline, detached using PBS/EDTA and analyzed by flow cytometry (FACSCalibur, Becton Dickinson).

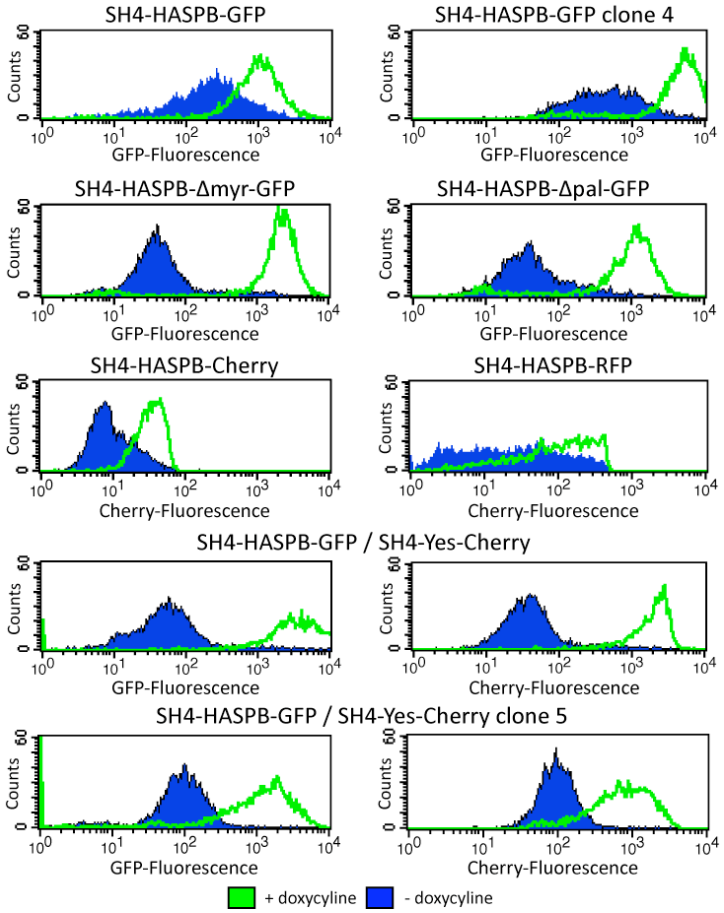

Figure 15 Flow cytometric analysis of HeLa cell lines expressing different SH4-HASPB-domain-containing reporter proteins in a doxycyline-inducible manner. Stable HeLa cell lines expressing wild-type and mutant SH4-HASPB-fusion proteins were cultivated in the absence (blue filled profile) and presence (green empty profile) of 1 µg/ml doxycyline for 48 hours, detached, and analyzed by flow cytometry using a FACSCalibur flow cytometer (Becton Dickinson). Autofluorescence of $HeLa_{MCAT\text{-}TAM2}$ cells was adjusted to 10^1.

Figure 15 depicts the expression of wild-type and mutant SH4-HASPB-domain-containing reporter proteins in HeLa cells before and after induction of protein expression by cultivation in the presence of 1 µg/ml doxycyline for 48 hours. Autofluorescence of $HeLa_{MCAT\text{-}TAM2}$ cells was adjusted to 10^1. All cell lines (with the exception of SH4-HASPB-Cherry) show low levels of constitutive reporter protein expression as indicated by fluorescence levels above 10^1 in the absence of doxycyline. However, in contrast to the Kyoto cells, reporter protein expression was now homogeneously increased by one to two orders of magnitude (with the exception of HeLa SH4-HASPB-RFP) after induction with doxycyline. We decided to use clone #5 expressing SH4-HASPB-GFP/SH4-Yes-Cherry to perform the genome-wide RNAi screen since this cell line showed and homogeneous reporter protein expression upon doxycyline induction.

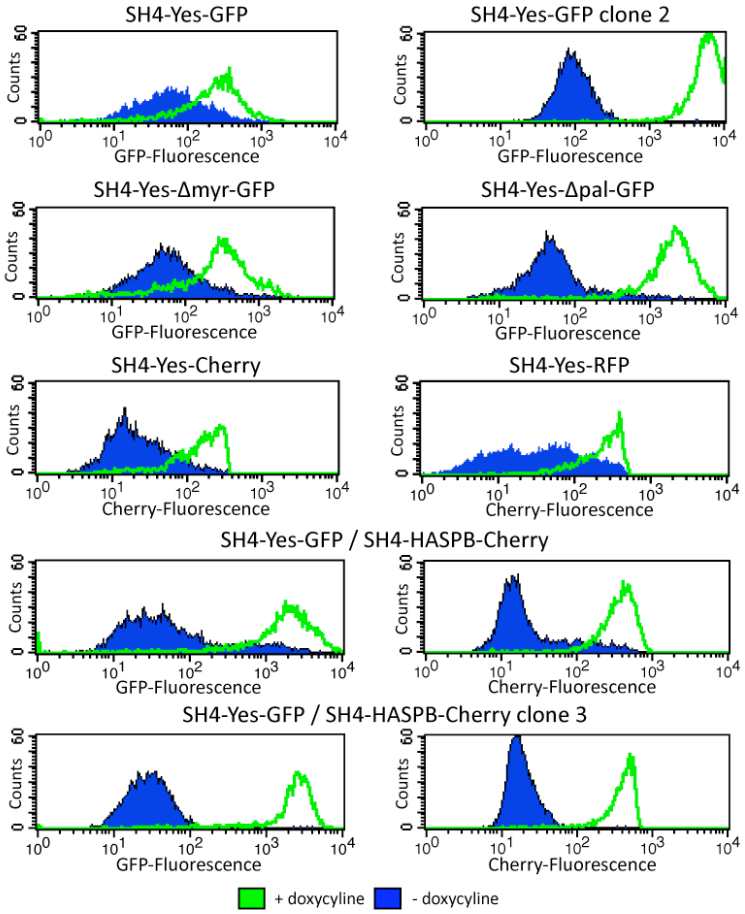

Figure 16 Flow cytometric analysis of HeLa cell lines expressing different SH4-Yes-domain-containing reporter proteins in a doxycyline-inducible manner. Stable HeLa cell lines expressing wild-type and mutant SH4-Yes-fusion proteins were cultivated in the absence (blue filled profile) and presence (green empty profile) of 1 µg/ml doxycyline for 48 hours, detached, and analyzed by flow cytometry using a FACSCalibur flow cytometer (Becton Dickinson). Autofluorescence of $HeLa_{MCAT\text{-}TAM2}$ cells was adjusted to 10^1.

Figure 16 depicts the expression of wild-type and mutant SH4-Yes-domain-containing reporter proteins in HeLa cells before and after induction of reporter protein expression by cultivation in the presence of 1 µg/ml of doxycyline for 48 hours. Autofluorescence of $HeLa_{MCAT\text{-}TAM2}$ cells was adjusted to 10^1. Most cell lines show low levels of constitutive reporter protein expression as indicated by fluorescence levels above 10^1 in the absence of doxycyline. However, in all cases reporter protein expression was homogeneously increased by one to two orders of magnitude treatment with doxycyline.

To validate the intracellular distribution of wild-type and mutant SH4-domain-containing reporter proteins we observed in CHO cells, we examined the respective HeLa cell lines by widefield and confocal microscopy.

3.1.6 *Characterization of HeLa Cell Lines Expressing SH4-domain-containing Reporter Proteins by Confocal and Widefield Microscopy*

Previous studies by our laboratory have shown that the N-terminal SH4-domain of HASPB is sufficient to mediate the localization of an SH4-HASPB-GFP fusion protein to the plasma membrane of CHO and HeLa cells (Stegmayer et al., 2005; Tournaviti et al., 2007; Tournaviti et al., 2009) and the partitioning into detergent-resistant membranes (Tournaviti et al., 2009). The Exchange of glycine 2 for alanine (SH4-HASPB-Δmyr) prevents myristoylation by deletion of the attachment site and the subsequent palmitoylation of the SH4-domain (Tournaviti et al., 2009), which requires previous myristoylation, most probably to confer localization of the SH4-domain to the compartment containing the respective palmitoyl transferase. This renders the fusion protein incapable of association with membranes, resulting in a cytoplasmic localization in CHO cells (Stegmayer et al., 2005). Mutation of cysteine 5 to alanine (SH4-HASPB-Δpal) abolishes the target site for palmitate attachment and consequently results in a loss of palmitoylation without an impact on myristoylation efficiency (Tournaviti et al., 2009). Reporter proteins containing the C5A mutation do not reach the plasma membrane and accumulate at intracellular membranes in close proximity to the nucleus, which most probably represent the Golgi apparatus and endosomal structures (Stegmayer et al., 2005; Tournaviti et al., 2007; Tournaviti et al., 2009). Moreover, non-palmitoylated SH4-HASPB-Δpal cannot be recovered in detergent-resistant membranes, indicating a failure to segregate into lipid microdomains (Tournaviti et al., 2009). To determine the subcellular localization of the SH4-domain-containing reporter proteins in HeLa cell lines generated in this study, cells were analyzed by widefield and confocal microscopy after induction with 1 µg/ml doxycyline for 24 hours.

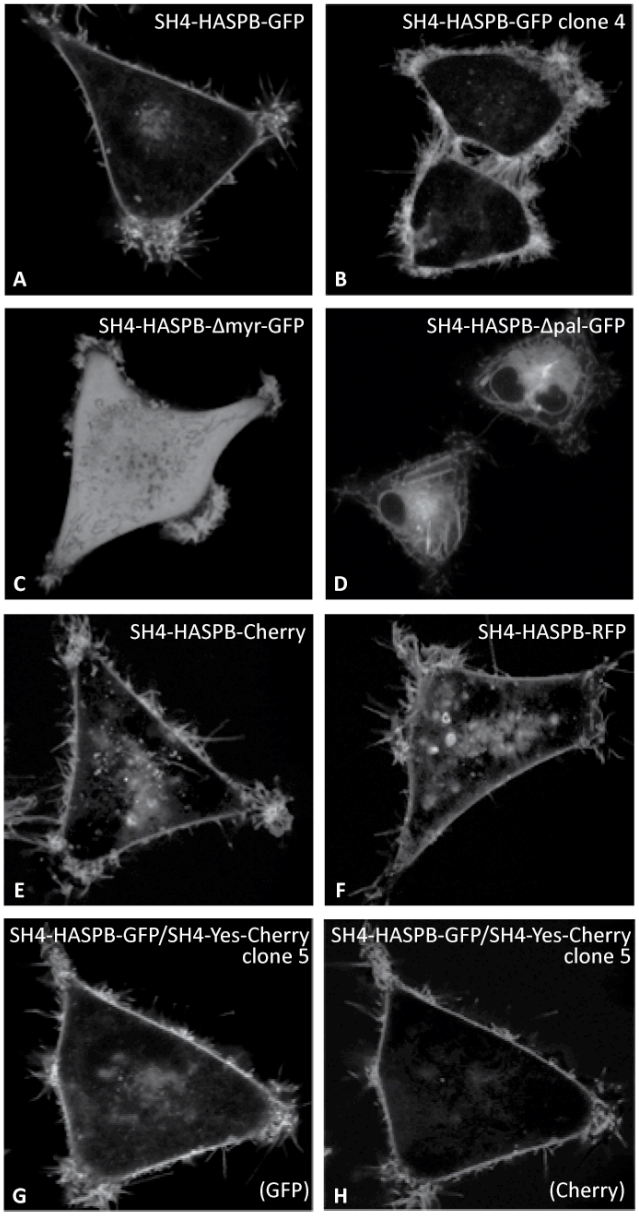

Figure 17 Confocal analysis of the subcellular localization of wild-type and mutant SH4-HASPB-containing reporter proteins in stable HeLa cell lines. Stable HeLa cells expressing SH4-HASPB-containing fluorescent reporter proteins were cultivated on 8-well LabTek imaging chambers for 24 hours in the presence of 1 µg/ml doxycyline, transferred to CO_2-independent medium and analyzed by live-cell confocal microscopy (63x magnification).

HeLa cell lines expressing SH4-HASPB-containing reporter proteins were cultivated on 8-well LabTek imaging chambers (Nunc) in the presence of 1 µg/ml doxycyline for 24 hours and analyzed by confocal microscopy in CO_2-independent medium (Invitrogen). As expected, the dually acylated N-terminal SH4-domain of HASPB (SH4-HASPB, Figure 17, panels A, B, E, F, G) is sufficient to mediate localization of different fluorescent reporter proteins to the plasma membrane in stable HeLa cell lines. A minor fraction localizes to a perinuclear

region, probably representing biosynthetic material interacting with perinuclear membranes *en route* to the plasma membrane or recycled material. Removal of the myristoylation site (SH4-HASPB-Δmyr-GFP, Figure 17, panel C) results in a cytosolic localization due to the inability of the protein to interact with membranes, while deletion of the palmitoylation site (SH4-HASPB-Δpal-GFP, Figure 17, panel D) results in a perinuclear accumulation. By contrast, both SH4-HASPB-GFP and SH4-Yes-Cherry (Figure 17, panels G and H) co-localize at the plasma membrane as expected.

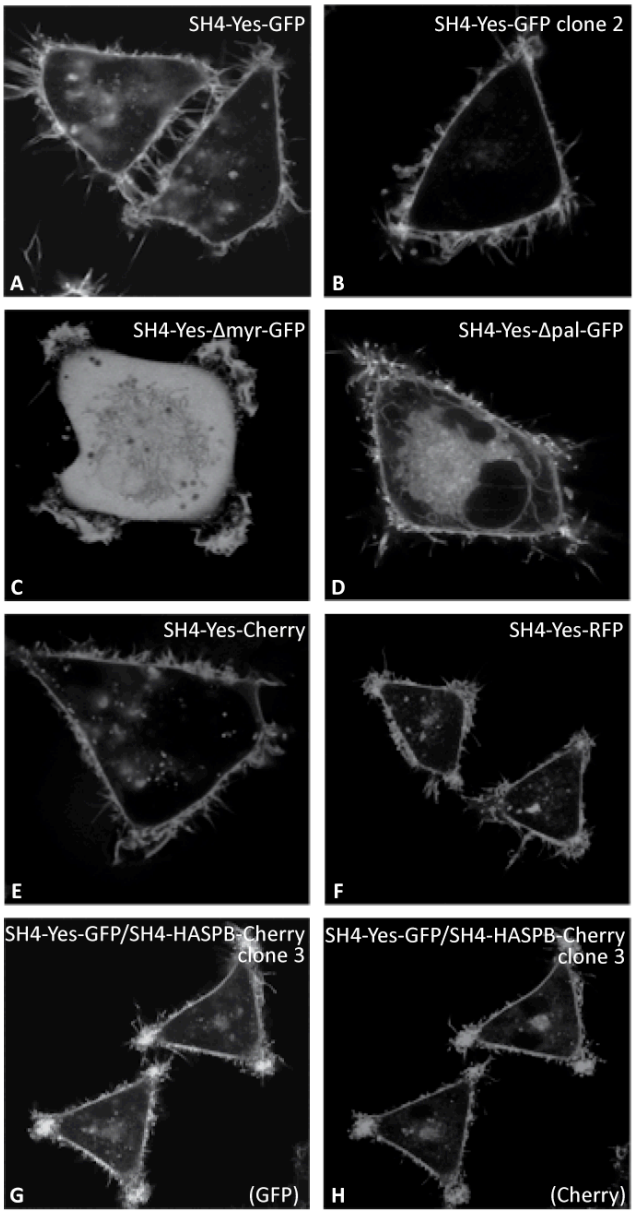

Figure 18 Confocal analysis of the subcellular localization of wild-type and mutant SH4-Yes-containing reporter proteins in stable HeLa cell lines. Stable HeLa cells expressing SH4-Yes-containing fluorescent reporter proteins were cultivated on 8-well LabTek imaging chambers for

24 hours in the presence of 1 µg/ml doxycyline, transferred to CO_2-independent medium and analyzed by live-cell confocal microscopy (63x magnification).

HeLa cell lines expressing SH4-Yes-containing reporter proteins were cultivated on 8-well LabTek imaging chambers (Nunc) in the presence of 1 µg/ml doxycyline for 24 hours and analyzed by confocal microscopy in CO_2-independent medium (Invitrogen). As expected, the dually acylated N-terminal SH4-domain of Yes (SH4-Yes, Figure 18, panels A, B, E, F, H) is sufficient to mediate localization of different fluorescent reporter proteins to the plasma membrane in stable HeLa cell lines. Removal of the myristoylation site (SH4-Yes-Δmyr-GFP, Figure 18, panel C) results in a cytosolic localization, while deletion of the palmitoylation site (SH4-Yes-Δpal-GFP, Figure 18, panel D) results in a perinuclear accumulation. In contrast to the localization of SH4-HASPB-Δpal-GFP (Figure 17, panel D), however, a significant fraction of the reporter protein reaches the plasma membrane, indicating that transport to the plasma membrane is not completely abolished. These findings indicate that the SH4-domain of Yes is less sensitive to loss of palmitoylation with regard to the effect on subcellular localization. Moreover, it may point towards the presence of one or several other sorting determinants, which may not be present or less dominant in SH4-HASPB. Moreover, these findings may indicate that SH4-HASPB and SH4-Yes are employing at least partially distinct cellular machineries. Wild-type SH4-Yes-GFP and SH4-HASPB-Cherry (Figure 18, panels G and H) co-localize at the plasma membrane as expected.

Having established human cell lines showing the expected phenotypes in confocal microscopy, we now analyzed whether the observed phenotypes can be detected at low magnification since the high-content screening platform is based on automated high-throughput widefield microscopy. Hence, we examined the localization of SH4-domain-containing reporter proteins using low magnification microscopy.

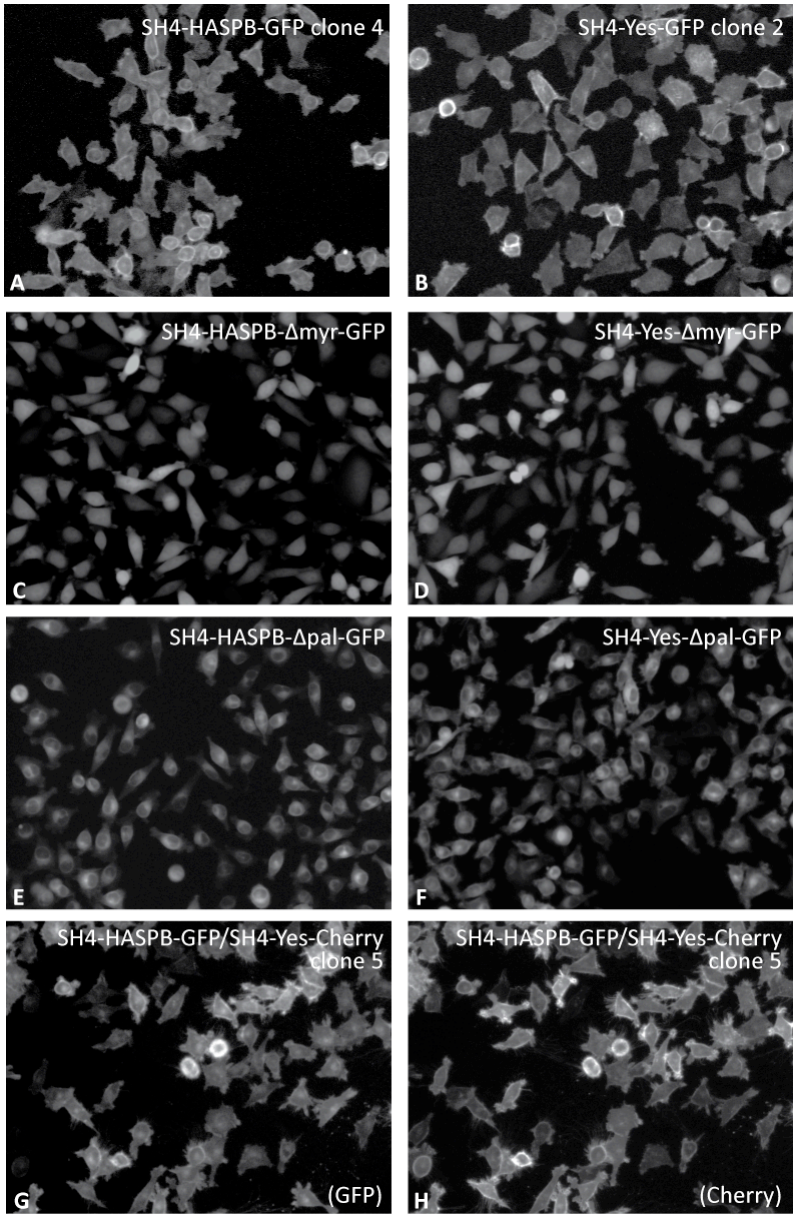

Figure 19 Widefield analysis of the subcellular localization of wild-type and mutant SH4-domain-containing reporter proteins in stable HeLa cell lines. Stable HeLa cells expressing different SH4-domain-containing fluorescent reporter proteins were cultivated on 8-well LabTek imaging chambers for 24 hours in the presence of 1 µg/ml doxycyline, transferred to CO_2-independent medium and analyzed by live-cell widefield microscopy (cutout of 10x magnification image).

Figure 19 shows HeLa cell lines expressing different SH4-domain-containing reporter proteins, which were cultivated on 8-well LabTek imaging chambers (Nunc) in the presence of 1 µg/ml doxycyline for 24 hours and analyzed by widefield microscopy in CO_2-independent medium (Invitrogen). The subcellular localization, as determined by confocal microscopy in Figure 17 and Figure 18 is clearly distinguishable. SH4-HASPB and SH4-Yes (Figure 19, panels A, B, G, H) confer plasma membrane localization of fluorescent reporter proteins. However, projection of upper and lower plasma membranes in the widefield

microscope results in a more homogeneous cell staining of wild-type SH4-domain-containing proteins (Figure 19, panels A, B, G, H). Due to the pronounced plasma membrane staining at the cell boundaries, however, this localization is clearly distinguishable from the cytosolic localization of SH4-Δmyr-containing reporter proteins (Figure 19, panels C and D).

Mutation of the myristoylation site (SH4-HASPB-Δmyr-GFP, Figure 19, panel C and SH4-Yes-Δmyr-GFP, panel D) results in a cytosolic localization, while removal of the palmitoylation site (SH4-HASPB-Δpal-GFP, Figure 19, panel E and SH4-Yes-Δpal-GFP, panel F) results in a perinuclear accumulation.

3.1.7 Inhibition of Palmitoylation in HeLa Cell Lines Expressing SH4-domain-containing Reporter Proteins

To exclude that the mislocalization of the palmitoylation-deficient mutants is a direct consequence of the lack of the palmitoyl moiety, we wanted to confirm our previous results by abolishing the palmitoylation of the wild-type protein which should result in a retention of SH4-domain-containing reporter proteins at intracellular membranes. A widely established method to inhibit palmitoylation *in vivo* and *in vitro* is treatment with the non-metabolizable palmitate analogue 2-bromopalmitate (2-BP). Cells treated with 2-BP are blocked in their capability to incorporate palmitate into target proteins, which has been shown for a variety of proteins (Resh 2006). We treated stable HeLa cell lines expressing SH4-HASPB- and SH4-Yes-containing reporter proteins with increasing amounts of 2-BP and investigated the effect on localization of the fluorescent reporter proteins by confocal and widefield microscopy.

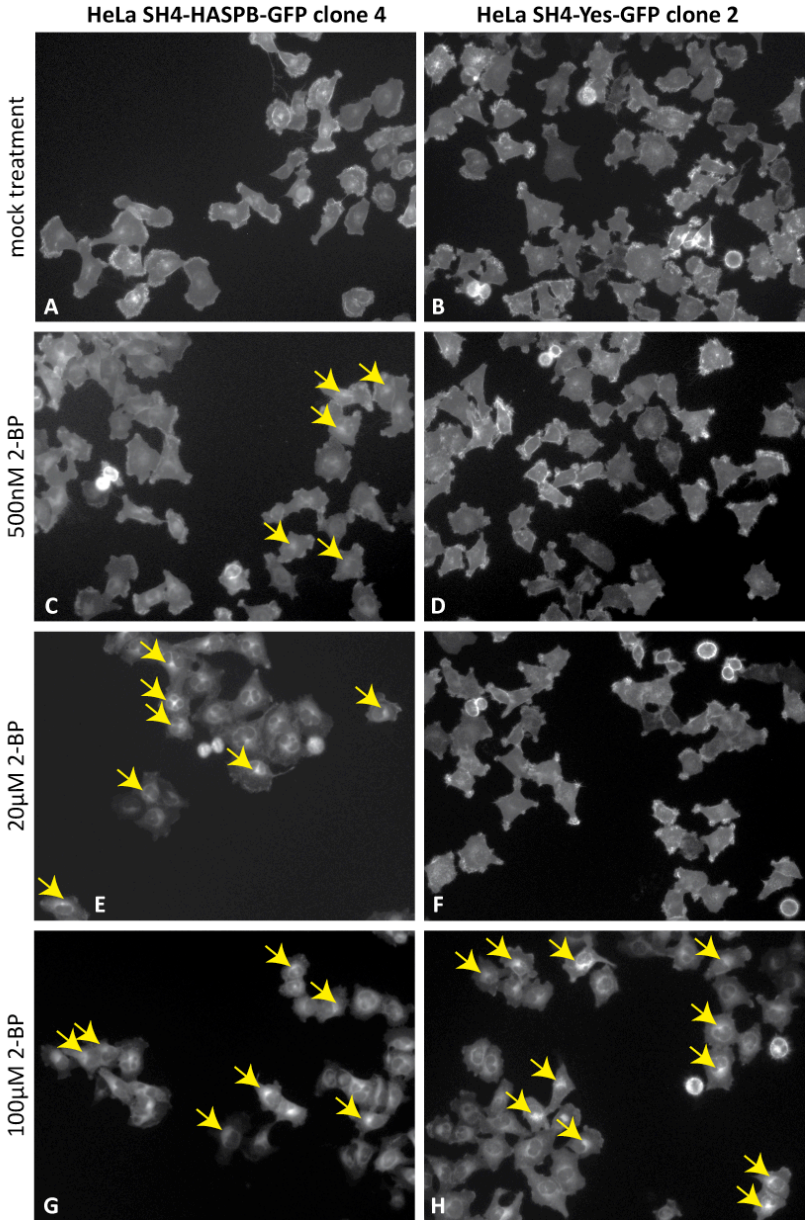

Figure 20 Widefield analysis of the subcellular localization of SH4-HASPB-GFP and SH4-Yes-GFP in stable HeLa cell lines after treatment with the palmitoylation inhibitor 2-bromopalmitate. Stable HeLa cells expressing SH4-domain-containing fluorescent reporter proteins were cultivated on 8-well LabTek imaging chambers for 12 hours in the presence of 1 µg/ml doxycyline and treated with DMSO (mock) or increasing concentrations of 2-bromopalmitate (2-BP), transferred to CO_2-independent medium and analyzed by live-cell widefield microscopy (cutout of 10x magnification image).

Stable HeLa cell lines expressing SH4-HASPB-GFP or SH4-Yes-GFP were cultivated for 12 hours in the presence of 1 µg/ml doxycyline and increasing amounts of the palmitoylation inhibitor 2-bromopalmitate (2-BP, Sigma). As depicted in Figure 20 (panels A and B), mock treatment with DMSO does not affect localization of either reporter protein. Cultivation in the presence of 500 nM 2-BP results in a weak intracellular retention of SH4-HASPB-GFP (Figure 20, panel C) but does not affect localization of SH4-Yes-Cherry (Figure 20, panel D).

Cultivation in the presence of increasing amounts of the inhibitor (20 µM and 100 µM) resulted in complete intracellular retention of SH4-HASPB-GFP (Figure 20, panels E and G). The localization of SH4-Yes-Cherry was only affected by cultivation in the presence of high concentrations of 2-BP (100 µM, Figure 20, panel H) and resulted in a partial intracellular retention at perinuclear membranes.

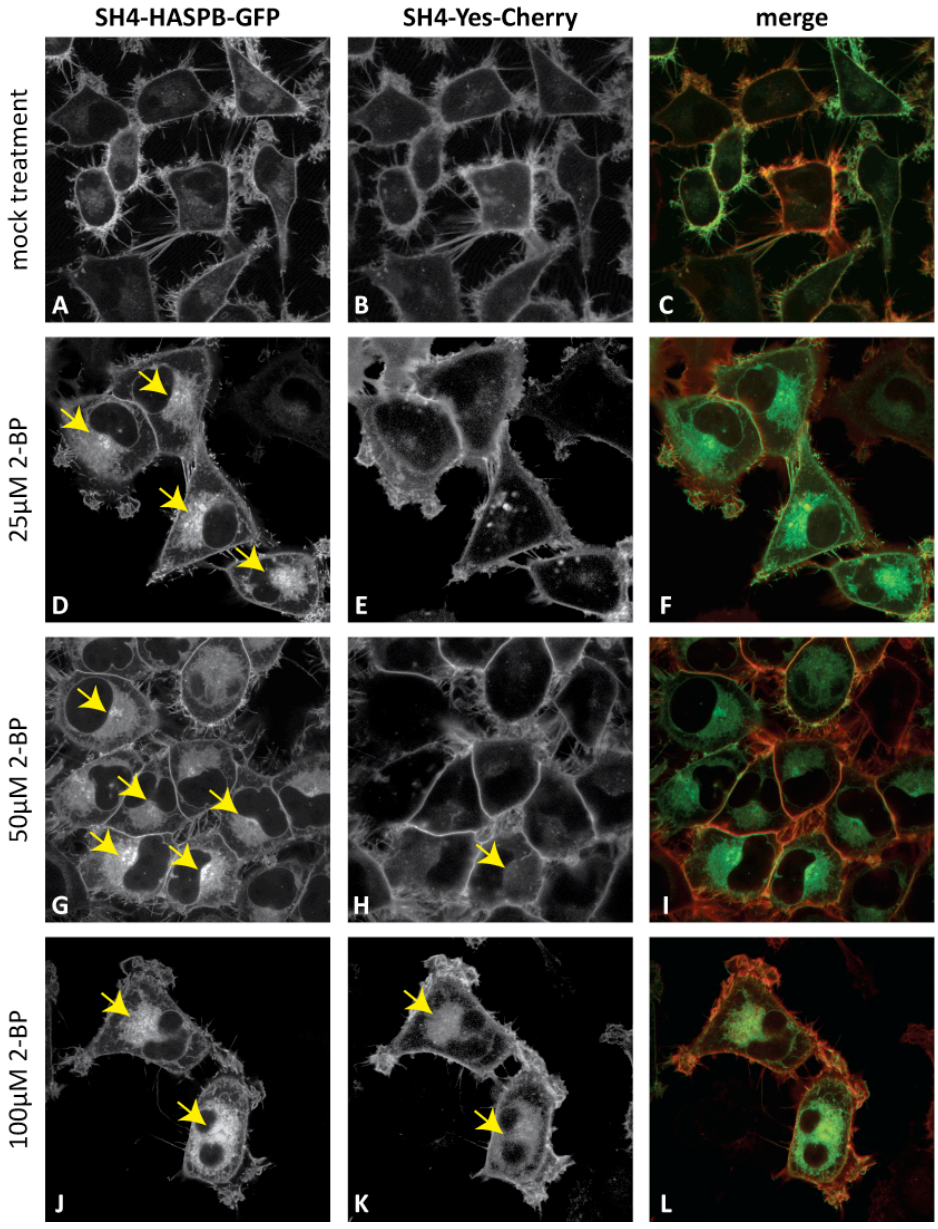

Figure 21 Confocal analysis of the subcellular localization of SH4-HASPB-GFP and SH4-Yes-Cherry in a stable HeLa cell line after treatment with the palmitoylation inhibitor 2-bromopalmitate. A stable HeLa cell line simultaneously expressing SH4-HASPB-GFP and SH4-Yes-Cherry was cultivated on 8-well LabTek imaging chambers for 12 hours in the presence of 1 µg/ml doxycyline and treated with DMSO (mock) or increasing concentrations of 2-bromopalmitate, transferred to CO_2-independent medium and analyzed by live-cell confocal microscopy (63x magnification).

A stable HeLa cell line simultaneously expressing SH4-HASPB-GFP and SH4-Yes-Cherry (HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5) was cultivated for 12 hours in the presence of 1 µg/ml doxycyline and increasing amounts of the palmitoylation inhibitor 2-bromopalmitate (2-BP, Sigma). Consistent with the results depicted in Figure 20, confocal microscopy reveals the differential potential of 2-BP in exerting intracellular retention of SH4-domain-containing proteins as shown in Figure 21. While SH4-HASPB-GFP is already fully retained at intracellular membranes after cultivation in the presence of 25 µM 2-BP (Figure 21, panels D and F) and at higher concentrations (Figure 21, panels G, I, J, L), localization of SH4-Yes-Cherry is only mildly affected after treatment with 50 µM 2-BP (Figure 21, panels H and I). A pronounced intracellular retention of SH4-Yes-Cherry can only be observed after incubation in the presence of 100 µM 2-BP (Figure 21, panels K and L), but does not result in a complete loss of plasma-membrane-associated material.

Consistent with the mutational analysis of the two SH4-domains used in this study (compare 3.1.6), SH4-HASPB is much more sensitive to interference with palmitoylation efficiency than SH4-Yes, as indicated by the differential effects of 2-BP on the two distinct SH4-domains. These results support the hypothesis that targeting and transport of SH4-Yes depends to a lesser extent on palmitoylation than SH4-HASPB and may, therefore, employ at least partially different cellular machineries and sorting determinants.

In the next step, we wanted to examine the compatibility of the selected clonal cell line HeLa SH4-HASPB-GFP/Yes-Cherry #5 with the high-content screening platform. This platform requires an efficient reverse transfection of cells cultivated on siRNA arrays, successful automated focussing and acquisition of sufficient quality images showing the expected localization of the reporter proteins. For this purpose, cells were reverse transfected by cultivation on live-cell imaging chambers containing spots of immobilized control siRNAs and subjected to automated live-cell microscopy (for a more detailed description see 3.3) to examine effects of the siRNA treatment.

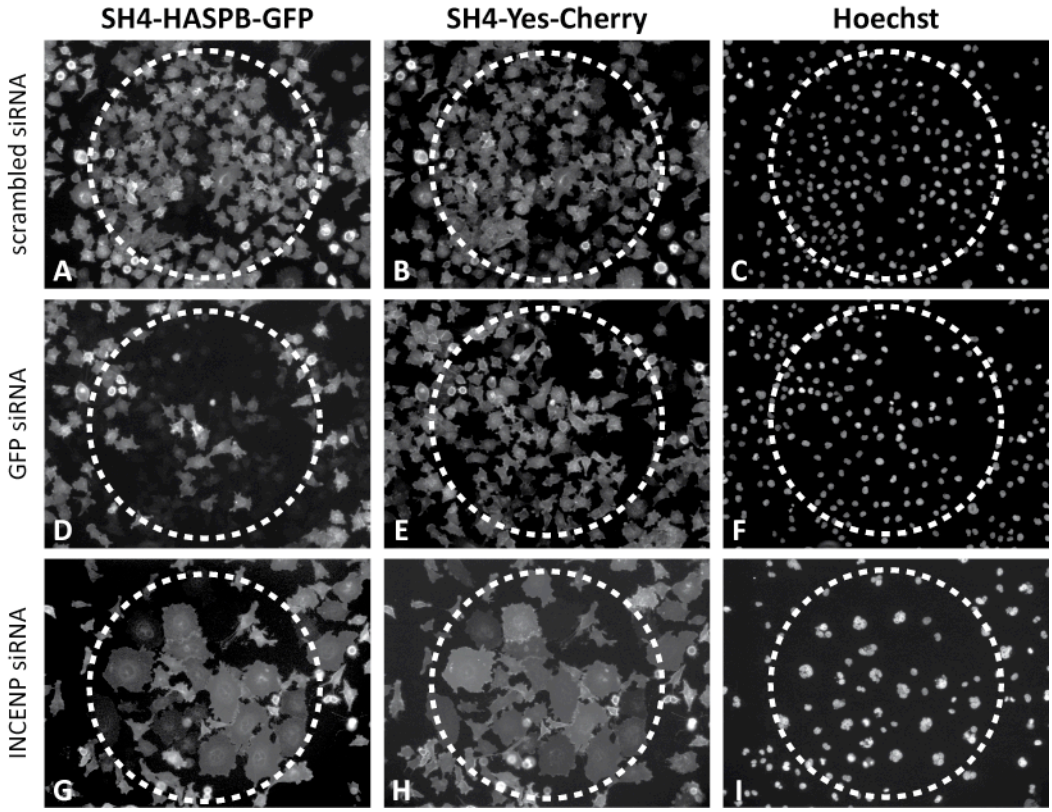

Figure 22 Automated widefield fluorescence microscopy of HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells cultivated on siRNA arrays for 48 hours. HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells were seeded on LabTek imaging chambers containing spots with immobilized siRNA and induced for induction of reporter protein expression by cultivation in the presence of 1 µg/ml doxycyline after 36 hours for additional 12 hours. 48h post-transfection cells were subjected to automated fluorescence microscopy (10x magnification).

Figure 22 shows HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells, which were cultivated on siRNA arrays for 48 hours. Reporter protein expression was induced 36 hours post-transfection by cultivation in the presence of 1 µg/ml doxycyline for additional 12 hours. Cells were transferred to phenol-red-free CO_2-independent medium containing 0.2 µg/ml Hoechst 33342 for staining of ell nuclei and subjected to automated live-cell microscopy. Reverse transfection of a scrambled siRNA (Figure 22, panels A-C) does not result in phenotypic changes. Reverse transfection of an siRNA directed against GFP (Figure 22, panels D-F) results in an efficient and specific downregulation of SH4-HASPB-GFP (Figure 22, panel D) but does not affect expression of SH4-Yes-Cherry (Figure 22, panel E). Downregulation of the inner centromere protein antigens 135/155kDa (INCENP, Figure 22, panels G-I) results in a mitotic arrest leading to the accumulation of large, multinucleated cells (Figure 22, panel I).

Taken together, our findings show that stable HeLa cell lines expressing wild-type SH4-domain-containing fluorescent fusion proteins exhibit the expected

plasma membrane localization of the reporter proteins. Failure of acylation either due to mutation of the target amino acids or by pharmacological inhibition results in intracellular retention of reporter proteins and this redistribution can be clearly visualized in confocal as well as widefield microscopy. In conclusion, cell lines generated during this study are compatible with a microscopy-based readout system for the genome-wide RNAi screen. To address transport of SH4-domain-containing proteins from a generalized point of view, we decided to perform the screen with a HeLa cell line simultaneously expressing two distinct SH4-domain-containing reporter proteins. Based on our characterization by flow cytometry (Figure 15), confocal (Figure 17, panels G and H) and widefield microscopy (Figure 19, panels G and H), we decided to use the clonal HeLa cell line expressing SH4-HASPB-GFP and SH4-Yes-Cherry #5. This cell line showed the expected co-localization of both reporter proteins at the plasma membrane and was sensitive to treatment with the palmitoylation inhibitor 2-BP (Fluka) (Figure 21), exhibiting concentration-dependent and distinct effects of 2-BP on the subcellular localization of SH4-HASPB-GFP and SH4-Yes-Cherry. Moreover, the clonal cell line is efficiently reverse transfected by cultivation on siRNA arrays and displays the expected phenotypic changes, which makes it compatible with the high-content screening platform.

3.2 ESTABLISHMENT OF AN AUTOMATED IMAGE ANALYSIS TOOL

Performance of high-content microscopy-based screens results in a huge accumulation of image-based data, posing the problem of efficient, unbiased and quantitative analysis. Therefore, we established an automated image analysis tool in collaboration with Steffen Remmele (AG Hesser, BioQuant, Heidelberg University, described in Remmele et al., 2008, see appendix).

3.2.1 Requirements for an Automated Image Analysis Tool

The goal of our genome-wide RNAi screen is to identify gene products, which are involved in intracellular targeting and transport of SH4-domain-containing proteins. Based on mutational studies (Figure 17, Figure 18, Figure 19) and experiments with a palmitoylation inhibitor (Figure 20, Figure 21) we expect the occurrence of two different phenotypes upon interference with the intracellular transport machinery of SH4-domain-containing proteins: Knockdown of gene products involved in myristoylation of the SH4-domain and potential accessory

factors supporting the myristoylation-dependent transport to and association with perinuclear membranes would result in a cytosolic accumulation of the respective SH4-domain-containing reporter protein, resembling the phenotypes of SH4-HASPB-Δmyr-GFP and SH4-Yes-Δmyr-GFP expressing cell lines (see Figure 17, panel C, Figure 18, panel C, Figure 19, panels C and D). In contrast, knockdown of gene products involved in efficient palmitoylation of the respective SH4-domain, incorporation of the dually acylated SH4-domain into lipid microdomains, factors involved in clustering of small SH4-domain-containing lipid microdomains and components of the post-Golgi transport machinery would result in the retention of SH4-domain-containing proteins at intracellular membranes resembling the phenotypes of SH4-HASPB-Δpal-GFP and SH4-Yes-Δpal-GFP expressing cell lines (see Figure 17, panel D, Figure 18, panel D, Figure 19, panels E and F) and cell lines expressing wild-type SH4-domains after treatment with palmitoylation inhibitors (see Figure 20, Figure 21). The hallmark of both such phenotypic changes is the redistribution of fluorescence intensity from the plasma membrane to intracellular regions in individual cells. A suitable automated image analysis tool therefore has to fulfil several tasks during analysis of an image: it has to identify all individual cells, determine their shape and subsequently has to segment each cell into different subcellular compartments (such as plasma membrane, cytoplasm and nucleus). Afterwards, it has to quantify the intensity distribution of both reporter proteins within distinct intracellular compartments. Each cell then has to be classified independently according to its intracellular intensity distribution of each reporter protein. In addition, the software tool has to cope with the limitations of high-throughput data, i.e. images of bad quality (e.g. out-of-focus images), noise, varying brightness of images and individual cells and local contaminations.

Tasks of an automated image analysis tool

- Batch analysis of large numbers of images
- Determination of image quality
- Identification of all individual cells on one image
- Determination of cell shape independent of reporter intensity distribution
- Segmentation of cells into three different compartments:
    i. Cell Membrane
    ii. Cytoplasm
    iii. Nucleus
- Intensity feature extraction
    i. Independently for each cell
    ii. Independently for the GFP and Cherry channel
- Classification according to the distribution of reporter intensities
    i. Independently for each cell
    ii. Independently for the GFP and Cherry channel
- Results output for further analysis in spreadsheet application
    i. Complete result of cell classification
    ii. Result of image quality control
    iii. Positional information on the siRNA-array

Challenges of an automated image analysis tool

- Large amounts of data require fast processing time
- Low quality of screening data
    i. Out-of-focus images
    ii. Noise
    iii. Local contaminations
    iv. Varying image brightness
- Variance in cell features
    i. Varying cell brightness on the same image
    ii. Touching cells in areas of high cell density
    iii. Cell segmentation has to work for different reporter intensity distributions (“phenotypes”)
    iv. Occurrence of abnormally sized and shaped cells

In conclusion, we require a very robust system comprised of a combination of quantitative and qualitative analyses, which contains a quality control mechanism, a cell segmentation tool, a quantitative measurement device and independent cell classification tools for each reporter.

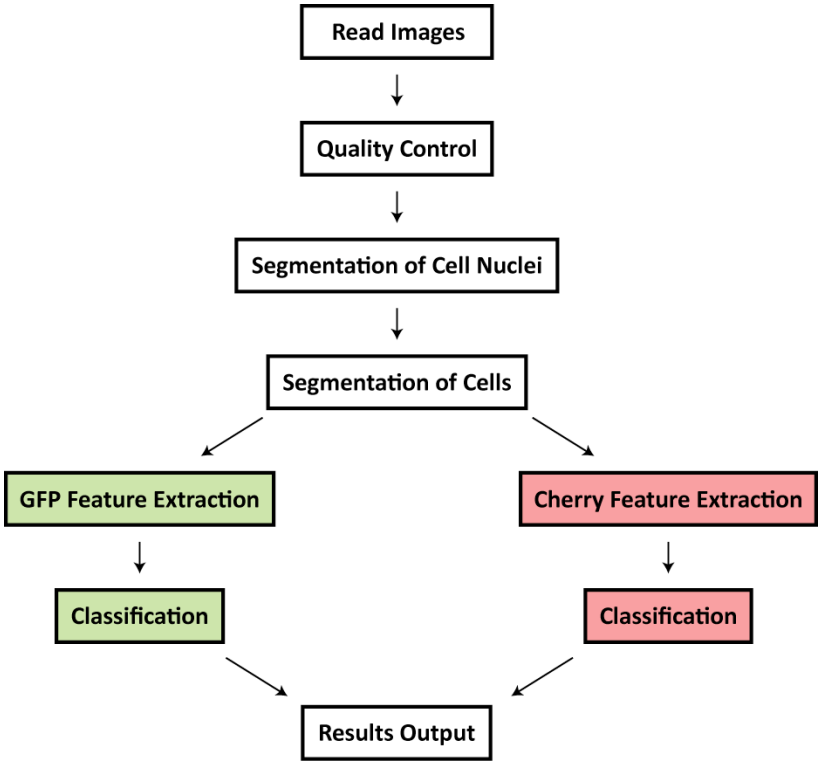

Figure 23 Schematic overview of the image analysis strategy of the automated image analysis tool. The first step of the image analysis is a quality control mechanism, followed by segmentation of cell nuclei in the Hoechst channel image. Using nuclei as a seed point, cells are segmented in the GFP and Cherry channel images and intensity features are extracted independently for GFP and Cherry fluorescence. Each cell is classified according to intensity features and results are issued in separate text files for each channel (adapted from Remmele et al., 2008).

Figure 23 outlines the image analysis strategy of the automated image analysis tool established in collaboration with Steffen Remmele (AG Hesser, BioQuant, Heidelberg University, described in Remmele at al., 2008, see appendix). This Matlab-based software tool recognizes the three corresponding images (GFP channel, Cherry channel, Hoechst channel) acquired at the same position of an siRNA array by identical file names with channel extensions and conducts a parallel analysis. The Quality control mechanism is designed to recognize out-of-focus images. According to the appearance of high frequency bands after Fourier transformation, images are assigned a quality grade from 3 (excellent) to 0 (insufficient). Based on the Hoechst channel, cell nuclei are identified and used as seed points to identify and segment individual cells in the GFP and Cherry channel images, which are subsequently subdivided into three different compartments: membrane, cytoplasm and nucleus (for a more detailed description, see 3.2.2. and Remmele et al., 2008, see appendix). GFP and Cherry intensity features are extracted from the corresponding GFP and Cherry channel

in areas of individual cells. Classification of individual cells is performed according to the distribution of intensities between the three intracellular compartments. The classification and quality control results for an analyzed batch of images is issued in two separate text files for the GFP and Cherry channel images and can be used for further analysis in spreadsheet applications.

3.2.2 Software-based Cell Identification, Segmentation and Classification

An absolutely crucial step during automated image analysis is the correct identification and segmentation of individual cells, because this step provides the basis for intensity feature extractions from cellular areas and therefore has a direct impact on the quality and reliability of the analysis results. We established a very robust cell segmentation tool, which reliably identifies and segments cells even under difficult circumstances, such as high cell densities, variable brightness of individual cells on the same image, occurrence of different intensity distribution phenotypes (i.e. membrane-associated, cytoplasmic and perinuclear localization of the fluorescent reporter protein) and local contaminations.

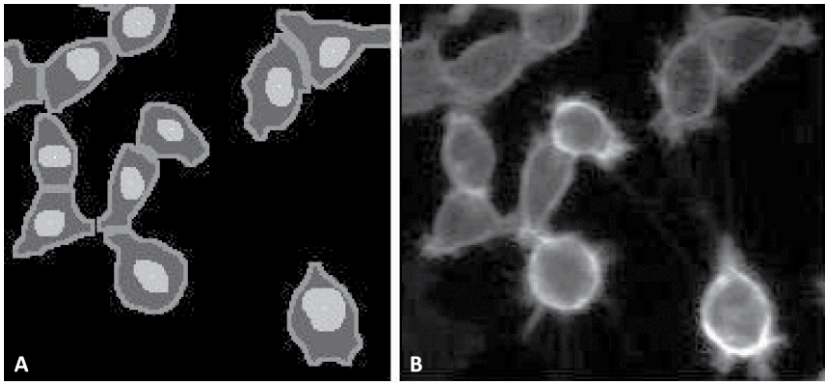

Figure 24 Cell segmentation by the automated image analysis tool. Cell segmentation result (panel A) of a multi-cell GFP channel image with plasma membrane phenotype (panel B). Cell nuclei are highlighted in light grey, the plasma membrane is labelled in medium grey and the cytoplasm is coloured in dark grey (panel A). Touching cells are reliably identified and segmented by the software tool (adapted from Remmele et al., 2008).

Figure 24 shows the segmentation results of a GFP channel image of cells with a plasma membrane phenotype. Cell segmentation works reproducible and reliable for the three expected reporter intensity distributions (termed as plasma membrane, cytoplasm and perinuclear phenotype, hereafter) even in

areas of higher cell density with touching cells. Based on the Hoechst channel images, cell nuclei are identified by an adaptive thresholding approach. Starting from high intensities, the threshold is decreased in variable step sizes. The segmentation tool generates connected regions and reduces the threshold parameter until the area of the connected regions converges the expected size of cell nuclei. Areas of local contamination and abnormally sized cell nuclei are excluded from further analysis. Of particular importance was the development of a robust and reliable cell segmentation tool, which is invariant to variable cell brightness within the same image and can cope with the three expected reporter intensity distributions, even though these have to be used to determine the boundaries of each cell. In our image analysis tool, cell nuclei are used as a seed point for a classifier-enhanced region-growing approach in the GFP and Cherry channel images. The region growing is stopped, when two criteria are fulfilled: (1) a large brightness difference between inner and outer pixels occurs and (2) the brightness of the outer pixels is below a local brightness threshold for this specific cell. This local threshold value is determined by subtracting a suitable fixed threshold from the lowest reporter intensity within the nucleus area and is therefore different for each cell. As a consequence, this method is more reliable with regard to differential brightness of individual cells within the same image and therefore poses a dramatic improvement compared to well-established object-identification tools, which rely on a fixed threshold. This approach works for all three expected phenotypes: if the fluorescent reporter protein localizes to the cytoplasm, there are no large brightness differences and the region growing is continued until fluorescence intensity drops dramatically at the transition between cell and extracellular space and both criteria are fulfilled: a large brightness difference occurs and the extracellular pixels are darker than the local threshold value. If the fluorescent reporter protein localizes to the plasma membrane, no large brightness differences occur inside the cell and the region growing stops when intensity drops below the local region threshold at the transition between intracellular and extracellular space. The biggest challenge is the correct determination of the cell shape in the case of a perinuclear accumulation of the fluorescent reporter protein. In this case, high gradient magnitudes occur within the cell due to the local accumulation of the reporter protein, which fulfils the first criterion. However, within the nuclear area, GFP and Cherry fluorescence are low and therefore the sensitive local brightness criterion constitutes a

suitable method to discriminate between the low intensities in cytoplasmic and membrane areas and the extracellular space. As a consequence, the region growing is only stopped if the extracellular space is reached and the pixel brightness lies below the local threshold value. This method works very reliably and is further supported by an additional evaluation step examining cell shape and size. If one of these does not match the expected values, the object is regarded as a local contamination or abnormal cell and is excluded from further analysis. Cells, whose nuclei are positioned too close together or lie close to the edges of the image, are also excluded from further analysis to avoid artefacts due to overlapping or only partially depicted cells. The three subcellular compartments are determined as follows: the nucleus is defined by the segmentation results of the Hoechst channel image. The cell membrane is defined by a specific pixel width touching the cell boundary defined in the region-growing approach. The cytoplasm is comprised of the residual area of the cell, which is not attributed to the nucleus or the cell membrane. Intensity features are extracted for each of these three intracellular compartments as well as of the complete cell area.

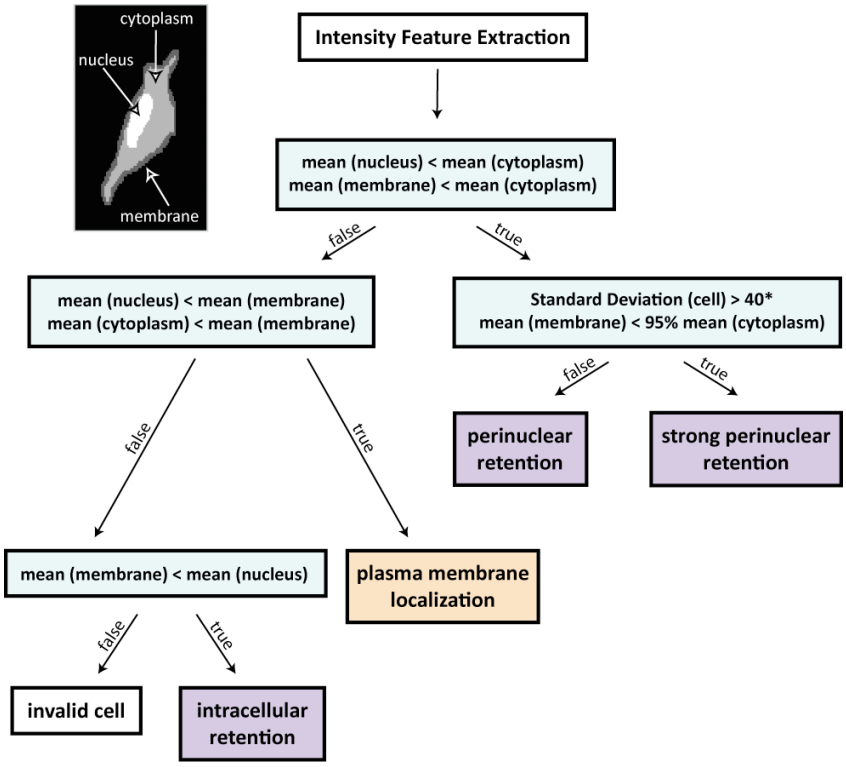

Figure 25 Decision tree for phenotype classification by the automated software tool. Intensity features are extracted for three subcellular compartments (nucleus, cytoplasm, membrane) and used for classification of individual cells. Highlighted in purple are phenotypes regarded as hits. The orange highlighted plasma membrane phenotype represents the normal phenotype of the

HeLa cell line used in the screen. Cells classified as invalid are neglected for further analysis. The criterion marked with an asterisk (*) was only used for analysis of the primary, but not the validation screen (adapted from Remmele et al., 2008).

Figure 25 shows a schematic overview of the decision function used for classification of individual cells. Intensity features of the complete cell area and the three different subcellular compartments (nucleus, cytoplasm and membrane) are extracted independently from the GFP and Cherry images. Theoretical considerations as well as empirical analyses showed that the standard deviation of the complete cell and the mean intensity values of each subcellular compartment were relevant to identify three different hit classes: (1) strong perinuclear retention (2) perinuclear retention and (3) intracellular retention. Perinuclear retention is characterized by a high cytoplasmic mean intensity compared to the nucleus and the membrane. The perinuclear retention is regarded as strong when the standard deviation within the complete cell area lies above a fixed threshold value (due to large differences in the intracellular brightness) and the mean intensity of the membrane is smaller than 95% of the intensity in the cytoplasm. A more generally distributed intracellular retention (e.g. cytoplasmic localization) is characterized by a higher intensity within the nuclear area compared to the membrane area. A hallmark of plasma membrane localization of the fluorescent reporter proteins is a high intensity at the membrane compared to the nuclear and cytoplasmic area.

3.2.3 Validation of the Automated Image Analysis Tool

The automated image analysis tool described in section 3.2.2 was examined with regard to its ability to identify phenotypic changes evoked by the expression of mutant SH4-domain-containing reporter proteins (compare Figure 17, Figure 18, Figure 19) and treatment of a stable HeLa cell line expressing wild-type SH4-HASPB-GFP after treatment with the palmitoylation inhibitor 2-BP (compare Figure 20 and Figure 21).

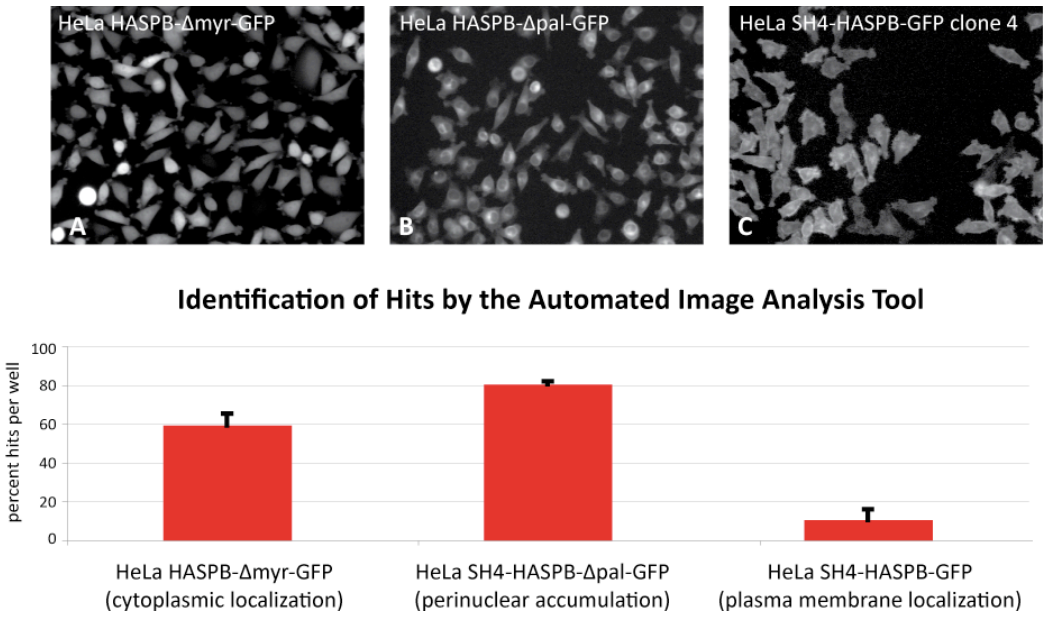

Figure 26 Recognition of different phenotypes by the automated image analysis tool. Widefield images (10x magnification) of HeLa cell lines expressing SH4-HASPB-Δmyr-GFP (panel A), SH4-HASPB-Δpal-GFP (panel B) and SH4-HASPB-GFP (panel C) were analyzed using the automated image analysis tool. For cytoplasmic localization, 60% of the cells were identified as hits, for the perinuclear accumulation 80% of the cells were identified as hits and below 10% of the cells exhibiting a plasma membrane phenotype were classified as hits.

Figure 26 shows an automated image analysis result of HeLa cells expressing HASPB-Δmyr-GFP (cytoplasmic localization, panel A), HASPB-Δpal-GFP (perinuclear accumulation, panel B) and SH4-HASPB-GFP (plasma membrane localization, panel C). The software was able to identify 60% of the HASPB-Δmyr-GFP-expressing cells as hits. 80% of the HASPB-Δpal-GFP expressing cells were classified as hits, while less than 10% of the SH4-HASPB-GFP-expressing cells were regarded as hits.

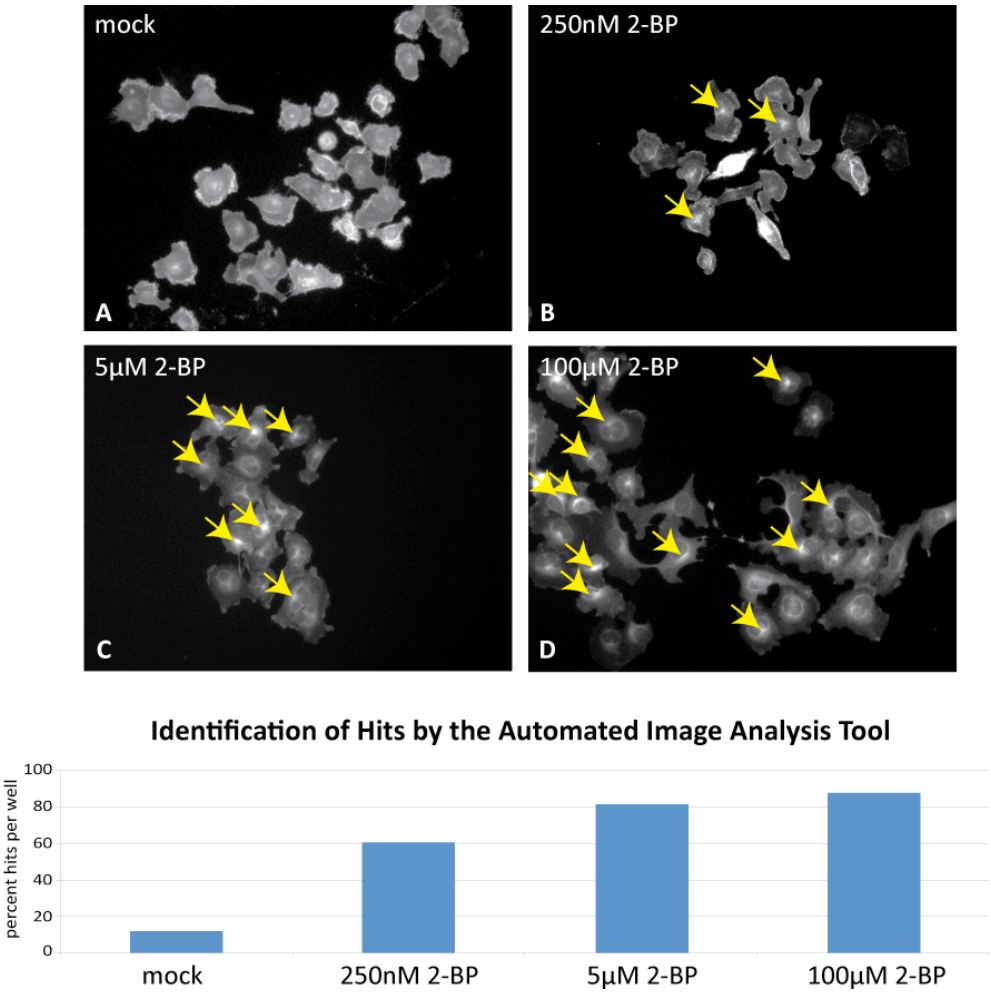

Figure 27 Recognition of different phenotypes evoked by treatment with the palmitoylation inhibitor 2-BP by the automated image analysis tool. HeLa cells expressing SH4-HASPB-GFP were cultivated for 24 hours in the presence of 1 µg/ml doxycyline and increasing amounts of 2-BP. Widefield images (cutout of 10x magnification image) of these cells were analyzed using the automated image analysis tool. Less than 15% of the mock treated cells were regarded as hits. Up to 85% of cells treated with 2-BP were recognized as hits in a concentration-dependent manner.

HeLa cells expressing SH4-HASPB-GFP were cultivated in the presence of 1 µg/ml doxycyline and different concentrations of the palmitoylation inhibitor 2-BP (Sigma) for 24 hours and subjected to widefield microscopy (Figure 27, panels A-D, cutout of 10x magnification image). Acquired images were analyzed using the automated image analysis tool described above (3.2.2). The software recognized less than 15% of the mock treated cells as hits. While low concentrations of the inhibitor resulted in weak perinuclear retention of SH4-HASPB-GFP (Figure 27, panel B) increasing amounts of the inhibitor resulted in a pronounced intracellular accumulation of SH4-HASPB-GFP (Figure 27, panels C and D). This effect is clearly reflected in the analysis of the corresponding

images, as an increasing fraction of cells is identified as hit when exposed to higher inhibitor concentrations.

Taken together, these data show that the automated image analysis tool established during this study (in collaboration with Steffen Remmele, AG Hesser, BioQuant, Heidelberg University, described in Remmele et al., 2008) meets the requirements for our experimental system. The tool represents a fast, reliable, robust and unbiased system for analysis of fluorescent widefield images that are the output of the high-content screening platform. The software is able to reliably identify and segment cells despite several limitations in the acquired screening data (such as low resolution, bad quality, local contaminations and areas of high cell density) and regardless of the subcellular localization of the respective SH4-domain-containing reporter protein. Moreover, it is a suitable tool for the identification of phenotypic changes with regard to the localization of SH4-domain-containing reporter proteins.

3.3 GENOME-WIDE MICROSCOPY-BASED RNAI-SCREENING

The aim of this study was to identify molecular components involved in intracellular targeting and transport of SH4-domain-containing proteins to the plasma membrane. RNA interference (RNAi) is a nowadays well-established tool for the specific and rapid silencing of gene expression at the post-transcriptional level by application of synthetic small interfering RNAs (siRNAs) in cultured mammalian cells. Recent advances in siRNA methodologies provide the possibility to specifically silence large numbers of genes in a short period of time and assess phenotypic changes by fluorescence microscopy (Erfle et al., 2004; Neumann et al., 2006; Simpson et al., 2007). To make use of this advanced technology, we generated and characterized a stable HeLa cell line, which simultaneously expresses fluorescent fusion proteins of the SH4-domain of *Leishmania* hydrophilic acylated surface protein B (SH4-HASPB-GFP) and the Src kinase Yes (SH4-Yes-Cherry). This clonal HeLa cell line (termed HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 hereafter) has been thoroughly characterized with regard to expression (Figure 15) and subcellular localization of both reporter proteins, which co-localize at the plasma membrane (Figure 17, panels G and H, Figure 19, panels G and H). The usage of this cell line has the additional benefit of addressing two similar but distinct SH4-domains under the same experimental conditions, thereby enabling us to discriminate between effects on general SH4-domain-dependent intracellular transport (if localization of both reporter proteins is affected) and factors which are involved in transport of only one of the two SH4-domain-containing fluorescent fusion proteins used. Furthermore, we established an automated image analysis tool, which is compatible with data acquired during high-content screening and is able to distinguish between different intracellular localizations of both reporter proteins independently (see 3.2).

3.3.1 Experimental Procedure of the Genome-wide RNAi Screen

Our screen was performed in collaboration with Rainer Pepperkok and the Mitocheck project group at the European Molecular Biology Laboratory (EMBL, Heidelberg). A pre-existing high-content screening platform based on a combination of siRNA arrays and automated live-cell fluorescence microscopy of cultured mammalian cells was used to successfully screen large numbers of

genes in different microscopy-based assays (Neumann et al., 2006; Simpson et al., 2007). In this system, the application of siRNA arrays enables simultaneous reverse transfection of cell lines with 384 individual siRNAs, which are immobilized on live-cell imaging chambers (Erfle et al., 2007; Erfle and Pepperkok, 2007; Erfle et al., 2004). These arrays are compatible with an automated fluorescence microscopy approach and therefore enable the screening of the complete human genome in a limited time.

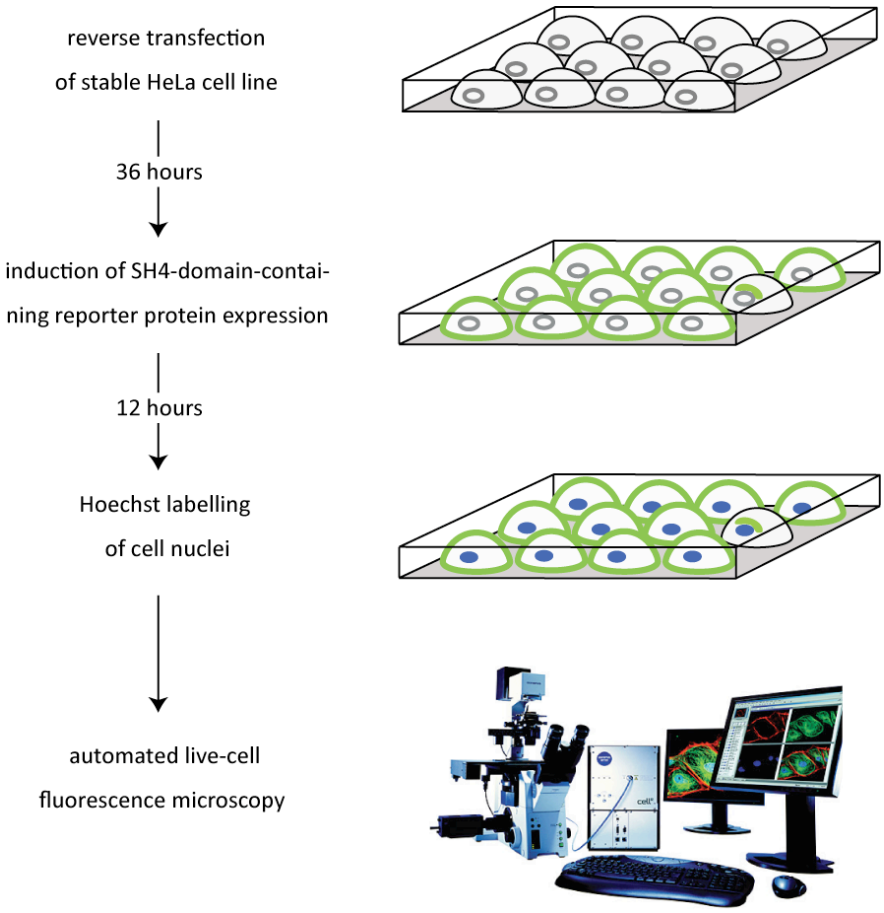

Figure 28 Experimental procedure of siRNA-mediated gene silencing and automated live-cell microscopy applied in the primary genome-wide RNAi screen. HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells are cultivated on 384-spot siRNA arrays for 36 hours before induction of protein expression by cultivation in the presence of 1µg/ml doxycyline for further 12 hours. 48 hours post-transfection, cell nuclei are stained with Hoechst 33342 and cells are subjected to automated live-cell fluorescence microscopy.

Figure 28 shows a schematic diagram of the experimental procedure of siRNA-mediated gene silencing and automated live cell microscopy as applied in the primary genome-wide RNAi screen. HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells are cultured on siRNA arrays to reverse transfect 384 different siRNAs

simultaneously. After 36 hours knockdown-time, reporter protein expression is induced by cultivation in the presence of 1 µg/ml doxycyline for additional 12 hours. 48 hours post-transfection, cells are transferred to phenol-red-free CO_2-independent medium (Invitrogen) containing 0.2 µg/ml Hoechst 33342 (Invitrogen) to stain cell nuclei. Three images per siRNA spot (GFP, Cherry and Hoechst channel) are acquired by automated live-cell fluorescence microscopy on fully automated Scan^R screening microscopes (Olympus Biosystems); the Hoechst nuclear stain is used for automatic focusing. Acquired images are analyzed by the automated image analysis tool described in 3.2.

Results of the cell classification generated by the software were analyzed as follows: considering the low probability of hits on each LabTek and the variance between individual LabTeks during the screening procedure, each well was analyzed with regard to the occurrence of hit phenotypes (i.e. strong perinuclear retention, perinuclear retention and intracellular retention) on the same LabTek. This approach is invariant to experimental variations between different LabTeks and is robust to outliers in single wells. We determined the percentage of cells displaying one of the three hit phenotypes on each position (i.e. on every GFP and Cherry channel image) and then calculated the standard deviation (StDev) and the median of the incidence rate of each hit phenotype of all 384 spots on the same LabTek. These two values were used to calculate the distance of the respective spot's percentage to the median in standard deviations of the same LabTek. SiRNA spots were regarded as hits when this score was larger than 2.5 in the case of SH4-HASPB-GFP and larger than 3.5 in the case of SH4-Yes-Cherry in any of the three hit classes on both replicas of the same LabTek. Gene products targeted in the "druggable genome" library, which had a score above 5 on only one replica for SH4-HASPB-GFP, were also included in the validation screen. Images containing less than 25 cells were disregarded in the analysis.

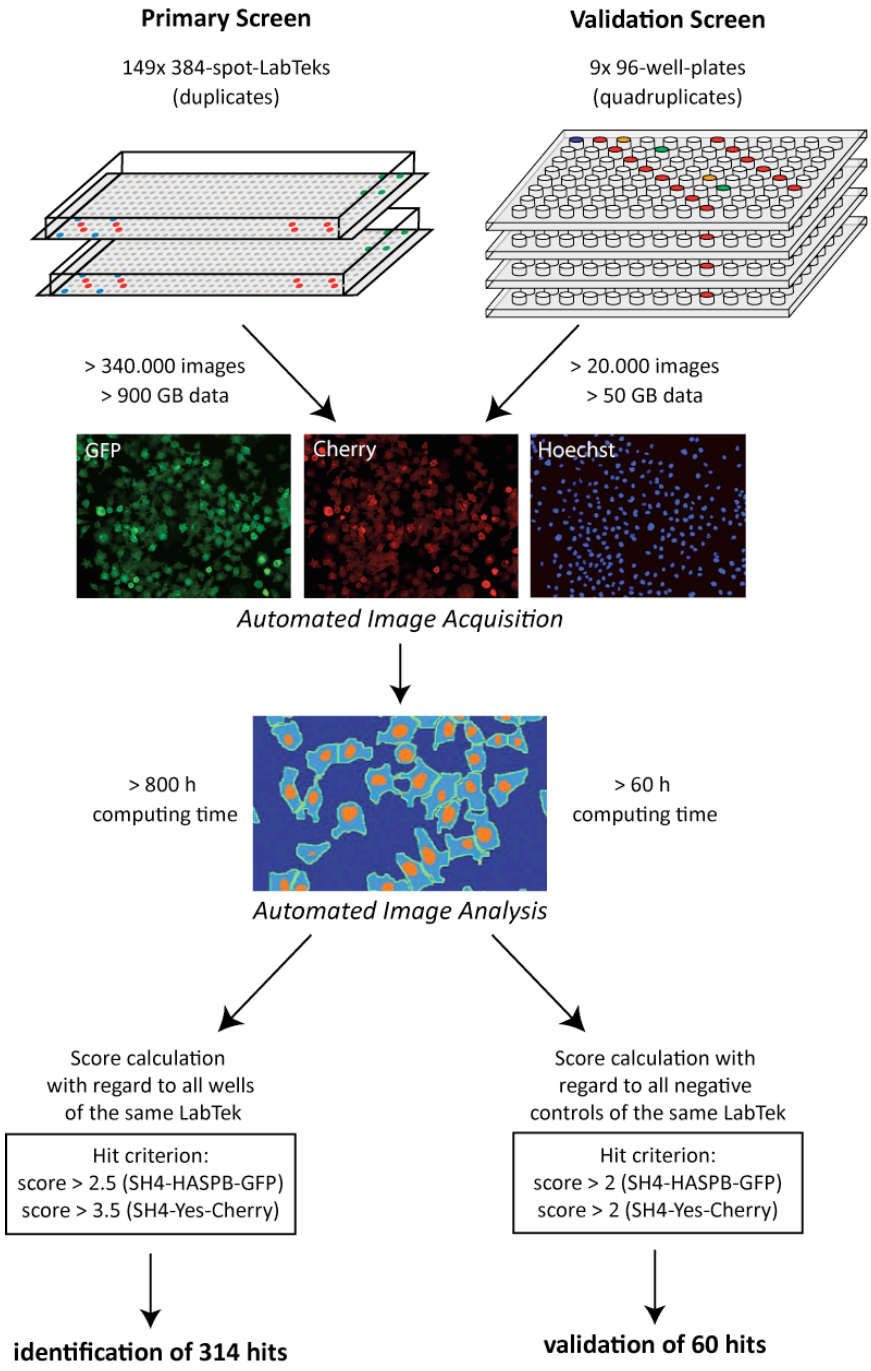

Figure 29 Workflow of the genome-wide RNAi screen. Reverse transfection of siRNA was achieved by culturing cells on siRNA arrays (primary screen) or siRNA-coated multi-well plates (validation screen). Colours represent positions of control siRNAs (red: scrambled siRNA, green: COPB siRNA, orange: COPA siRNA, blue: INCENP siRNA). 48 hours post-transfection, cells were submitted to automated live-cell fluorescence microscopy and images analyzed with an automated image analysis tool. This procedure resulted in the identification of 314 genes in the primary screen and the validation of 60 hits in the validation screen.

Figure 29 (left side) shows a diagram of the workflow of the genome-wide primary screen on duplicates of 149 different 384-spot siRNA arrays. More than 340.000 images comprising over 900 GB data are acquired by automated

fluorescence microscopy on Scan^R screening microscopes (Olympus Biosystems). Data analysis by the automated image analysis tool requires more than 800 hours computing time. Results of the cell classification of single spots are considered with regard to cell classification results of all spots on the same LabTek using distances to the median in standard deviations as hit score. Hit thresholds of 2.5 for SH4-HASPB-GFP and 3.5 for SH4-Yes-Cherry resulted in the identification of 314 hits in the primary screen. For the validation screen (Figure 29, right side), independent siRNAs for these genes are immobilized on 9 different 96-well plates. Firstly, this has the advantage that two images per spot can be acquired, doubling the number of cells to be analyzed. Secondly, potential cross-contamination between neighbouring siRNA spots by cell migration or diffusion of siRNAs can be excluded. During the validation screen, more than 20.000 images are acquired, comprising over 50 GB of data. Images are analyzed by the automated image analysis tool, which requires more than 60 hours computing time. Differences in the primary data, due to acquisition on different microscopes, required slight changes in the cell classification tool, where the standard deviation of the cell intensity was disregarded. In contrast to the primary screen, classification results from the validation screen were considered with regard to classification results of negative controls, due to the higher probability of hit occurrences. By regarding scores above 2 as hits (for both SH4-HASPB-GFP and SH4-Yes-Cherry), we were able to validate 60 gene products, whose downregulation had an impact on intracellular transport of SH4-domain-containing proteins.

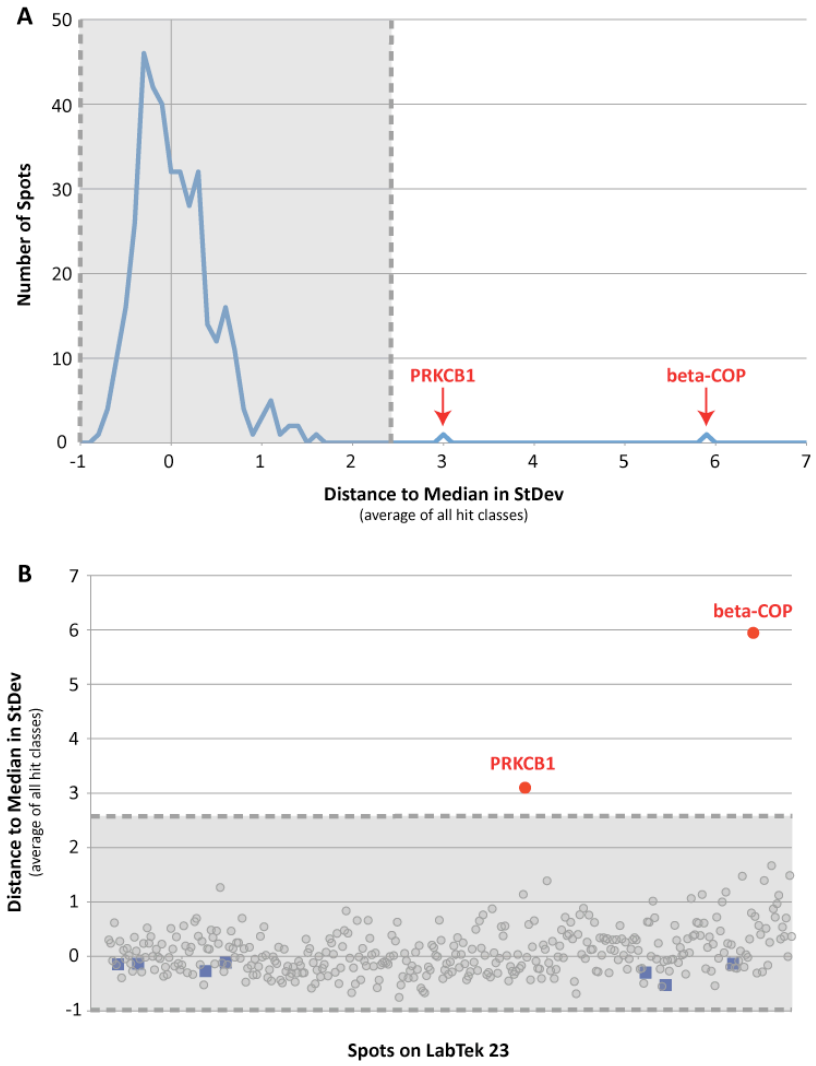

Figure 30 Exemplary SH4-HASPB-GFP analysis results for LabTek 23. Cell classification results of SH4-HASPB-GFP were transformed to distances to the median in standard deviations of the complete LabTek (average values of all hit classes on both replicas are shown). Spots were considered as hits when lying above a threshold of 2.5 standard deviations above median. The negative area is highlighted in grey, spots containing scrambled siRNA are marked as blue squares (panel B) and hits are highlighted in red (panels A and B).

Figure 30 shows an exemplary analysis result for the intracellular distribution of SH4-HASPB-GFP on LabTek 23. In this case, averages of all hit classes of both replicas are shown. Panel A shows the distribution of distances to the median in standard deviations across LabTek 23. Score values above 2.5 (outside the grey shaded area) are considered as hits. In this example, the coatomer subunit beta-COP (COPB) and protein kinase C subunit beta 1 (PRKCB1) were identified and could subsequently also be validated in the validation screen. In panel B, scores for all individual spots on LabTek 23 are displayed. Spots containing non-silencing siRNAs (highlighted as blue squares) lie well within the negative area (shaded in grey), while the scores of the two validated hits beta-COP and PRKCB1 lie well above the threshold of 2.5.

3.3.2 *Microscopy-based RNAi Screen of the Human Genome*

The genome-wide library (based on ensembl version 27) applied in the primary screen consists of two complementary parts: LabTek 1-49 comprises silencer siRNAs (Ambion) targeting ~ 5.000 genes of the so-called druggable genome. This subset of the human genome consists of gene products, which are able to bind drug-like molecules (Hopkins and Groom, 2002). LabTek 60-159 comprises silencer siRNAs (Ambion) targeting the remaining ~ 17.000 genes of the complete human genome. The complete library consists of ~ 52.000 silencer siRNAs (Ambion). This includes around 2.2% non-silencing siRNAs as negative controls and around 0.9% validated beta-COP siRNAs, which we identified to have an impact on SH4-domain-dependent trafficking and could therefore be used as a positive control in our assay.

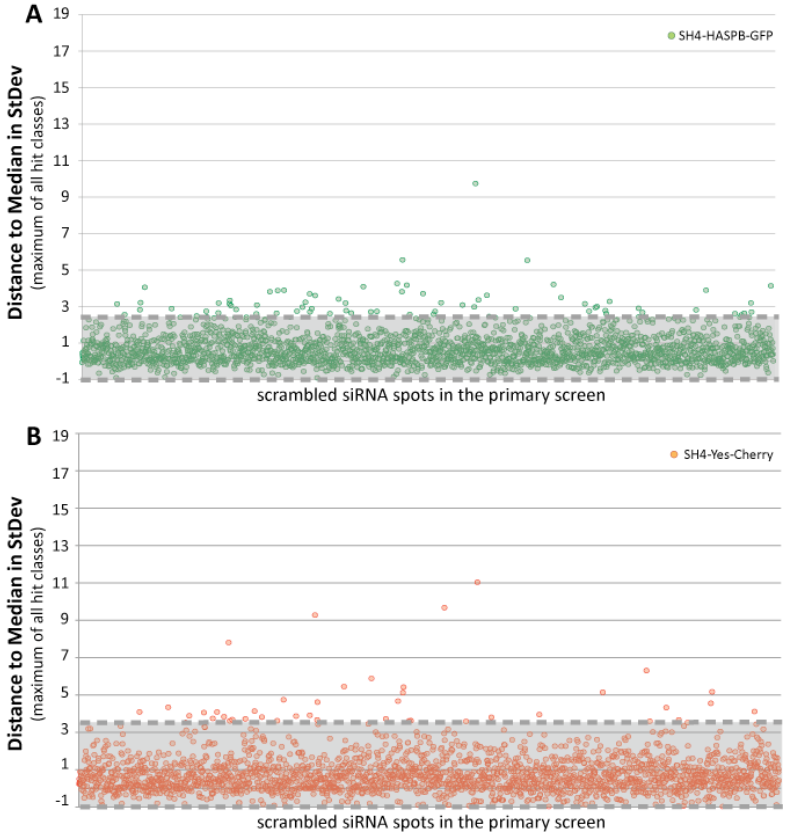

Figure 31 Hit scores of all negative controls in the primary screen. The maximum hit scores for the three hit classes of all siRNA spots containing non-silencing siRNAs are shown for SH4-HASPB-GFP (panel A) and SH4-Yes-Cherry (panel B). Score values below the hit thresholds 2.5 (SH4-HASPB-GFP) and 3.5 (SH4-Yes-Cherry) lie within the grey shaded area.

Figure 31 shows diagrams of all spots containing negative controls in the primary screen. HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells were cultivated on 384-spot siRNA arrays for 36 hours before protein expression was induced by 1 µg/ml doxycyline for additional 12 hours. Cells were subjected to

automated fluorescence microscopy 48 hours post-transfection and acquired data were analyzed using the automated image analysis tool. Panel A displays the highest hit score (distance to median in standard deviations) for SH4-HASPB-GFP of the three hit classes for every individual spot containing non-silencing (scrambled) siRNA in the primary screen. Scores below the hit threshold of 2.5 lie within the grey shaded area. Panel B displays the highest hit score (distance to median in standard deviations) for SH4-Yes-Cherry of the three hit classes for every individual spot containing non-silencing (scrambled) siRNA in the primary screen. Scores below the hit threshold of 2.5 lie within the grey shaded area. In both panels, data points outside the grey shaded area indicate the occurrence of false positive values. However, only a very low percentage of negative control spots were false positive on both replicas, indicating the robustness of the experimental system.

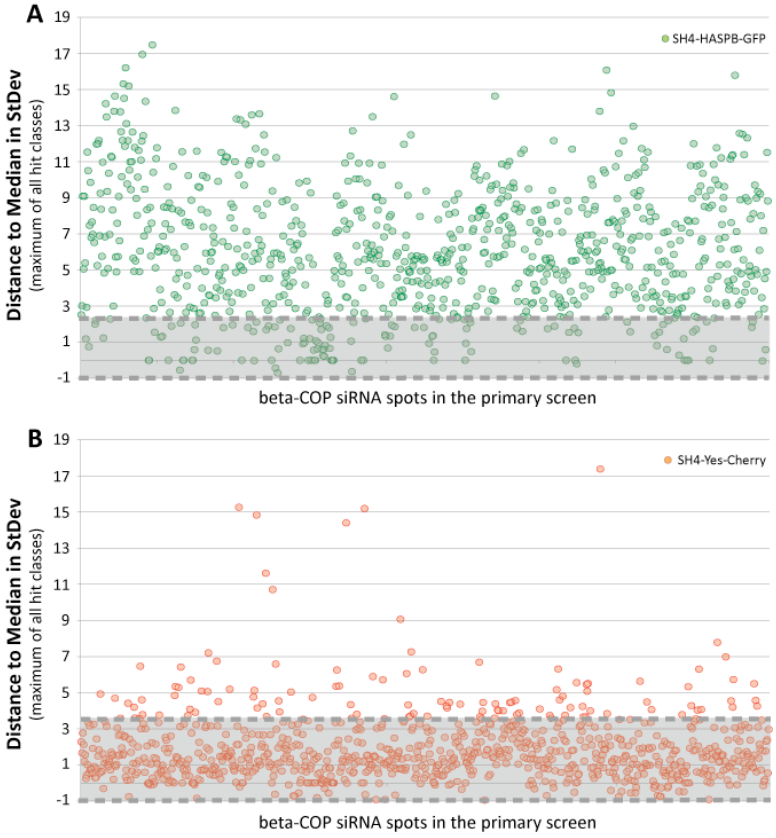

Figure 32 Hit scores of all beta-COP siRNA containing spots in the primary screen. The maximum hit scores for the three hit classes of all siRNA spots containing beta-COP siRNAs is shown for SH4-HASPB-GFP (panel A) and SH4-Yes-Cherry (panel B). The negative score values below 2.5 (SH4-HASPB-GFP) and 3.5 (SH4-Yes-Cherry) are shaded in grey.

Figure 32 shows diagrams of all spots containing beta-COP siRNAs in the primary screen. HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells were cultivated on 384-spot siRNA arrays for 36 hours before protein expression was induced

by 1 µg/ml doxycyline for additional 12 hours. Cells were subjected to automated fluorescence microscopy 48 hours post-transfection and acquired data were analyzed using the automated image analysis tool. Panel A displays the highest hit score (distance to median in standard deviations) for SH4-HASPB-GFP of the three hit classes for every individual spot containing beta-COP siRNA in the primary screen. Scores above the hit threshold of 2.5 lie outside the grey shaded area. These spots were used as positive controls, because the vast majority of beta-COP knockdowns resulted in a pronounced effect on intracellular localization of SH4-HASPB-GFP (see Figure 40 for further analysis), as indicated by the frequent occurrence of hit scores above 2.5. Panel B displays the highest hit score (distance to median in standard deviations) for SH4-Yes-Cherry of the three hit classes for every individual spot containing beta-COP siRNA in the primary screen. Scores below the hit threshold of 2.5 lie within the grey shaded area. Comparison of panels A and B indicates that knockdown of beta-COP has distinct effects on intracellular transport of the two distinct SH4-domains used in this screen, with a pronounced effect on the intracellular localization SH4-HASPB-GFP and a very mild effect on SH4-Yes-Cherry localization (see Figure 40 for further analysis). These data once more illustrate the advantage of using two similar but distinct reporter proteins in parallel to determine general and specific factors in SH4-domain-dependent trafficking.

Genome-wide Library				
total library	**targeted genes: ~ 22.000**		**total number of siRNAs: ~ 52.000**	
scrambled siRNAs	**total number**	**2286**	**2.2%**	**of complete screen**
	total valid number	2270	99.3%	
	false positives (HASPB-GFP)	70	3.1%	of negative controls
	false positives (Yes-Cherry)	44	1.9%	
beta-COP siRNAs	**total number**	**898**	**0.9%**	**of complete screen**
	total valid number	873	97.2%	
	false negatives (HASPB-GFP)	122	14.0%	of COPB siRNAs
	negatives (Yes-Cherry)	754	86.4%	

Identification of Hits				
complete genome	**total number of identified hits**	**314**	**1.4%**	**of complete library**
	HASPB-GFP-specific hits	191	60.8%	
	Yes-Cherry-specific hits	87	27.7%	of all hits
	general SH4-domain hits	36	11.5%	
druggable genome	**total number of identified hits**	**184**	**0.8%**	**of druggable library**
	HASPB-GFP-specific hits	111	60.3%	
	Yes-Cherry-specific hits	55	29.9%	of druggable-genome hits
	general SH4-domain hits	18	9.8%	
"non-druggable" genome	**total number of identified hits**	**130**	**0.6%**	**of non-druggable library**
	HASPB-GFP-specific hits	80	61.5%	
	Yes-Cherry-specific hits	32	24.6%	of non-druggable-genome hits
	general SH4-domain hits	18	13.8%	
	identified with 2 siRNAs	12	9.2%	

Figure 33 Overview of the results from the genome-wide primary RNAi screen.

An overview of the results of the genome-wide primary screen is given in Figure 33. 2.2% of the ~ 52.000 siRNA targeting the ~ 22.000 genes of the human genome were negative controls. 99.3% of these were valid, i.e. image quality was sufficient and more than 25 cells were present on each spot. Only 3.1% of valid individual negative controls reached a score value above 2.5 for SH4-HASPB-GFP and only 1.9% had a score value above 3.5 for SH4-Yes-Cherry and were consequently false positive. 86% of the valid beta-COP siRNA-containing spots showed a hit score above 2.5 for SH4-HASPB-GFP, but only 13.6% were above the hit threshold of 3.5 for SH4-Yes-Cherry. For the identification of hits, only spots showing hit scores above the hit threshold on both replicas were considered as hits. In the druggable genome, we included spots, which displayed a high score above 5 for SH4-HASPB-GFP in only one replica. In total, 314 genes in the primary screen matched the hit criteria and were selected for further analysis. Of these, the majority (60.8%) had a specific impact on the intracellular localization of SH4-HASPB-GFP, while 27.7% were exclusively

affecting SH4-Yes-Cherry transport. 11.5% disturbed plasma-membrane localization of both reporters.

3.3.3 *Microscopy-based RNAi Validation Screen*

The library (based on ensembl version 53) used for validation of the hits identified in the primary screen was comprised of 776 silencer select siRNAs (Ambion), which were targeted against 311 genes identified in the primary screen. 3 genes were excluded from the validation screen due to lack of suitable siRNAs. 14.8 % of these were negative controls, 3.3% of siRNAs directed against beta-COP or alpha-COP were included as positive controls.

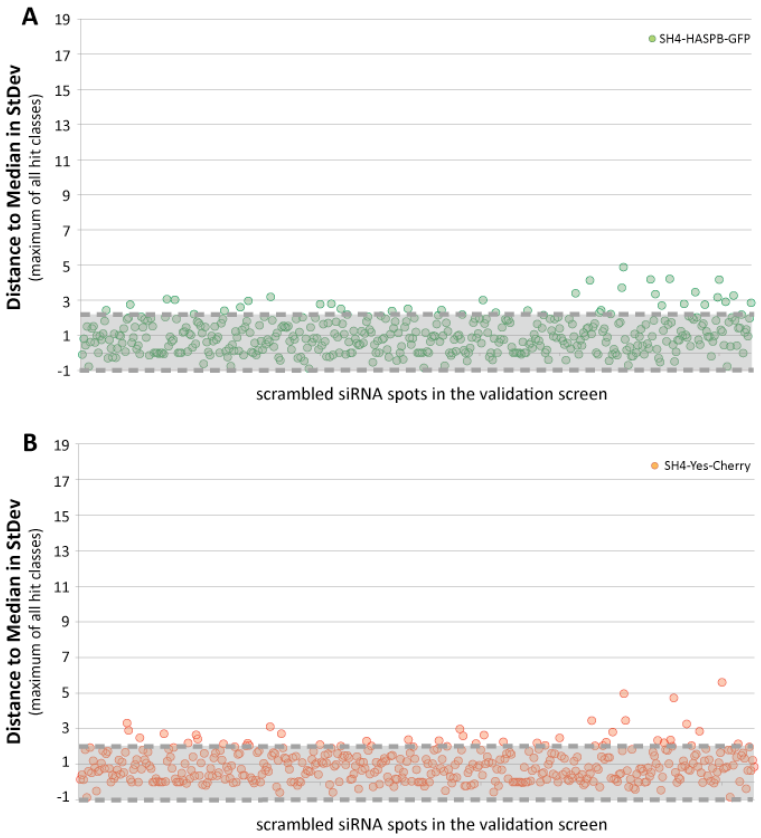

Figure 34 Hit scores of all negative controls in the validation screen. The maximum hit scores for the three hit classes of all siRNA spots containing non-silencing siRNAs are shown for SH4-HASPB-GFP (panel A) and SH4-Yes-Cherry (panel B). Score values below the hit threshold 2 (SH4-HASPB-GFP and SH4-Yes-Cherry) lie within the grey shaded area.

Figure 34 shows diagrams of all spots containing negative controls in the validation screen. HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells were cultivated on siRNA-coated 96-well plates for 36 hours before protein expression was induced by 1 µg/ml doxycyline for additional 12 hours. Cells were subjected to automated fluorescence microscopy 48 hours post-transfection and acquired

data were analyzed using the automated image analysis tool. Panel A displays the highest hit score (distance to median in standard deviations) for SH4-HASPB-GFP of the three hit classes for every individual spot containing non-silencing (scrambled) siRNA in the validation screen. Scores below the hit threshold of 2 lie within the grey shaded area. Panel B displays the highest hit score (distance to median in standard deviations) for SH4-Yes-Cherry of the three hit classes for every individual spot containing non-silencing (scrambled) siRNA in the primary screen. Scores below the hit threshold of 2 lie within the grey shaded area. In both panels, data points outside the grey shaded area indicate the occurrence of false positive values. However, only a very low percentage of negative control spots were false positive on both replicas, indicating the robustness of the experimental system.

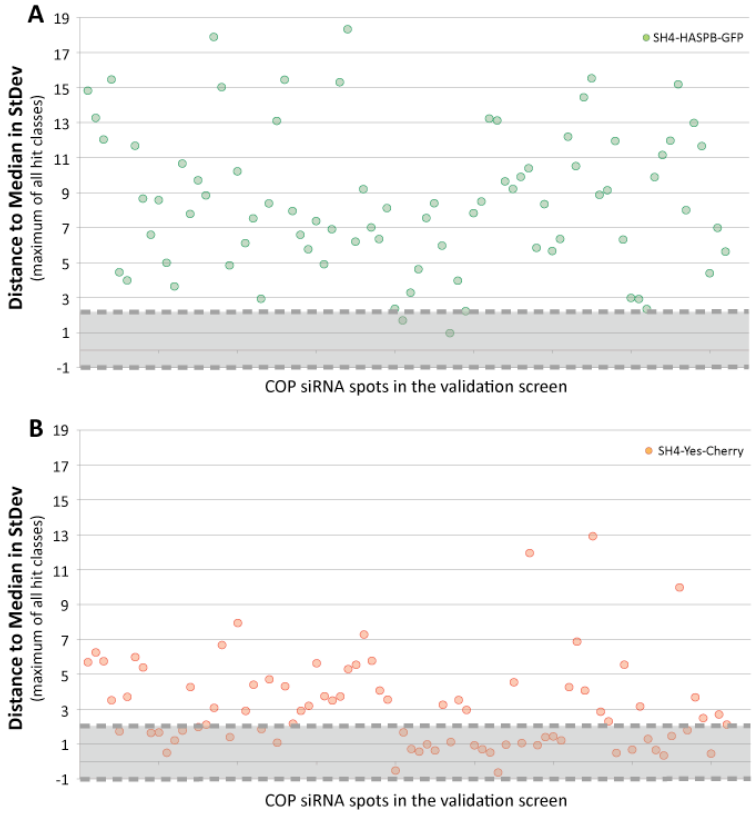

Figure 35 Hit scores of all alpha- and beta-COP siRNA-containing spots in the validation screen. The maximum hit scores for the three hit classes of all siRNA spots containing alpha- or beta-COP siRNAs are shown for SH4-HASPB-GFP (panel A) and SH4-Yes-Cherry (panel B). Score values above the hit threshold 2 (SH4-HASPB-GFP and SH4-Yes-Cherry) lie outside the grey-shaded area.

Figure 35 shows diagrams of all spots containing alpha-CP and beta-COP siRNAs in the validation screen. HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells were cultivated on siRNA-coated 96-well plates for 36 hours before protein expression was induced by 1 µg/ml doxycyline for additional 12 hours. Cells

were subjected to automated fluorescence microscopy 48 hours post-transfection and acquired data were analyzed using the automated image analysis tool. Panel A displays the highest hit score (distance to median in standard deviations) for SH4-HASPB-GFP of the three hit classes for every individual spot containing alpha-COP and beta-COP siRNAs in the validation screen. Scores above the hit threshold of 2 lie outside the grey shaded area. Like in the primary screen, the vast majority of alpha-COP and beta-COP knockdowns resulted in a pronounced effect on intracellular localization of SH4-HASPB-GFP (see Figure 40 for further analysis). Panel B displays the highest hit score (distance to median in standard deviations) for SH4-Yes-Cherry of the three hit classes for every individual spot containing alpha-COP and beta-COP siRNAs in the validation screen. Scores below the hit threshold of 2 lie within the grey shaded area. In accordance with data from the primary screen, comparison of panels A and B indicates that knockdown of coatomer subunits has distinct effects on intracellular transport of the two distinct SH4-domains used in this screen, with a pronounced effect on the intracellular localization SH4-HASPB-GFP and a milder effect on SH4-Yes-Cherry localization (see Figure 40 for further analysis). In comparison to data from the primary screen (see Figure 32), however, more individual knockdowns of coatomer subunits resulted in scores above the hit threshold for SH4-Yes-Cherry.

Genome-wide Validation Library				
total library	**targeted genes: 311**			**total number of siRNAs: 776**
scrambled siRNAs	**total number**	**420**	**13.5%**	**of complete screen**
	total valid number	401	95.5%	
	false positives (HASPB-GFP)	53	12.6%	of negative controls
	false positives (Yes-Cherry)	45	10.7%	
COP siRNAs	**total number**	**88**	**2.8%**	**of complete screen**
	total valid number	82	93.2%	
	false negatives (HASPB-GFP)	2	2.3%	of COP siRNAs
	negatives (Yes-Cherry)	35	39.8%	

Validation of Hits				
complete genome	**total number of validated hits**	**60**	**19.3%**	**of complete validation library**
	HASPB-GFP-specific hits	19	31.7%	
	Yes-Cherry-specific hits	18	30.0%	of all validated hits
	general SH4-domain hits	23	38.3%	
druggable genome	**total number of identified hits**	**21**	**35.0%**	**of druggable validation library**
	HASPB-GFP-specific hits	7	33.3%	
	Yes-Cherry-specific hits	5	23.8%	of validated druggable-genome hits
	general SH4-domain hits	9	42.9%	
"non-druggable" genome	**total number of identified hits**	**39**	**12.5%**	**of "non-druggable" library**
	HASPB-GFP-specific hits	12	30.8%	
	Yes-Cherry-specific hits	13	33.3%	of validated "non-druggable"-genome hits
	general SH4-domain-specific hits	14	35.9%	

Figure 36 Overview of the results from the RNAi validation screen.

An overview of the results of the genome-wide validation screen is given in Figure 36. 13.5% of the 776 siRNAs targeting 311 of the genes identified in the primary screen were negative controls. 95.5% of these were valid, i.e. image quality was sufficient and more than 25 cells were present on each spot. 12.6% of valid individual negative controls reached a score value above 2 for SH4-HASPB-GFP and only 10.7% had a score value above 2 for SH4-Yes-Cherry. 97.7% of the valid alpha- and beta-COP siRNA-containing spots showed a hit score above 2 for SH4-HASPB-GFP, 60.2% were above the hit threshold for SH4-Yes-Cherry. Because of the comparably high occurrence of false positive negative controls, replicas were excluded from the analysis if less than 70% of all negative controls of this replica showed scores below 2. Three to four replicas of each plate were valid and wells were regarded as hits when 2 out of

3 or 3 out of 4 replicas of the respective well showed a score above the hit threshold. In total, 60 genes in the validation screen matched the hit criteria. Of these, around one third had a specific impact on the intracellular localization of SH4-HASPB-GFP, one third was exclusively affecting SH4-Yes-Cherry transport and one third disturbed plasma-membrane localization of both reporters. These data indicate that the validation screen was more sensitive towards changes in the intracellular localization of SH4-Yes-Cherry, maybe due to the application of the effective silencer select siRNAs (Ambion).

Gene symbol	Gene Name	Gene ID	RefSeq ID	SH4-HASPB	SH4-Yes
A2M	alpha-2-macroglobulin	2	NM_000014	yes	yes
ACADVL	acyl-Coenzyme A dehydrogenase, very long chain	37	NM_000018	yes	yes
ACY3	aspartoacylase (aminocyclase) 3	91703	NM_080658	yes	yes
ATP6AP2	ATPase, H+ transporting, lysosomal accessory protein 2	10159	NM_005765	no	yes
BAI3	brain-specific angiogenesis inhibitor 3	577	NM_001704	yes	no
C11orf38	chromosome 11 open reading frame 38	399967	NM_212555	yes	yes
C12orf36	chromosome 12 open reading frame 36	283422	NM_182558	yes	no
C16orf71	chromosome 16 open reading frame 71	146562	NM_139170	yes	no
CACNA1S	calcium channel, voltage-dependent, L type, alpha 1S subunit	779	NM_000069	no	yes
COPA	coatomer protein complex, subunit alpha	1314	NM_004371	yes	no
COPB1	coatomer protein complex, subunit beta 1	1315	NM_016451	yes	no
CDH26	cadherin-like 26	60437	NM_021810	yes	yes
CHERP	calcium homeostasis endoplasmic reticulum protein	10523	NM_006387	yes	yes
DBH	dopamine beta-hydroxylase (dopamine beta-monooxygenase)	1621	NM_000787	no	yes
EIF4A3	eukaryotic translation initiation factor 4A, isoform 3	9775	NM_014740	yes	yes
ENSG00000176584	NULL	0	ENST00000388832	yes	yes
FAM108A1	family with sequence similarity 108, member A1	81926	NM_001130111	yes	no
FLJ21767	hypothetical protein FLJ21767	401331	XM_001132988	yes	no
GALNTL4	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-like 4	374378	NM_198516	yes	no
GBL	G protein beta subunit-like	64223	NM_022372	yes	yes
hCG_1640785	hCG1640785	401206	XM_937380	yes	yes
HNRNPA1	heterogeneous nuclear ribonucleoprotein A1	3178	NM_002136	no	yes
IL7R	interleukin 7 receptor	3575	XM_001127146	yes	no
ITGAV	integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51)	3685	NM_002210	yes	no
ITGB1	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	3688	NM_002211	yes	yes
KIF20B	kinesin family member 20B	9585	NM_016195	yes	no
LILRB5	leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 5	10990	NM_006840	yes	no
LOC387787	similar to CG9804-PA	387787	XM_370636	no	yes
LOC390688	CDC37-like	390688	NM_001080829	no	yes
LOC441377	similar to ribosomal protein S26	441377	XM_496991	no	yes
LOC442160	similar to ribosomal protein L21	442160	XM_001128707	no	yes
LOC641768	similar to 40S ribosomal protein S26	641768	XM_935907	yes	yes
LOC644191	similar to hCG15685	644191	XM_930072	yes	yes
MAGEB1	melanoma antigen family B, 1	4112	NM_002363	yes	no
MAK10	MAK10 homolog, amino-acid N-acetyltransferase subunit (S. cerevisiae)	60560	NM_024635	yes	yes
MBD3L2	methyl-CpG binding domain protein 3-like 2	125997	NM_144614	yes	no
MGC70863	similar to RPL23AP7 protein	284942	NM_203302	yes	yes
MVD	mevalonate (diphospho) decarboxylase	4597	NM_002461	no	yes
NAE1	NEDD8 activating enzyme E1 subunit 1	8883	NM_001018159	yes	yes
OR10V1	olfactory receptor, family 10, subfamily V, member 1	390201	NM_001005324	no	yes
PGCP	plasma glutamate carboxypeptidase	10404	NM_016134	yes	no
PHF5A	PHD finger protein 5A	84844	NM_032758	yes	yes
POLR2F	polymerase (RNA) II (DNA directed) polypeptide F	5435	NM_021974	(yes)	yes
PRKCA	protein kinase C, alpha	5578	NM_002737	yes	yes
PRKCB1	protein kinase C, beta 1	5579	NM_212535	yes	yes
PRPF31	PRP31 pre-mRNA processing factor 31 homolog (S. cerevisiae)	26121	NM_015629	no	yes
PRPF8	PRP8 pre-mRNA processing factor 8 homolog (S. cerevisiae)	10594	NM_006445	yes	yes
PTPRD	protein tyrosine phosphatase, receptor type, D	5789	NM_001040712	yes	no
RAN	RAN, member RAS oncogene family	5901	NM_006325	no	yes
RPL18	ribosomal protein L18	6141	NM_000979	no	yes
RPL21	ribosomal protein L21	6144	NM_000982	no	yes
SF3A1	splicing factor 3a, subunit 1, 120kDa	10291	NM_001005409	no	yes
SHC1	SHC (Src homology 2 domain containing) transforming protein 1	6464	NM_003029	yes	yes
SMPDL3B	sphingomyelin phosphodiesterase, acid-like 3B	27293	NM_014474	yes	no
SOBP	sine oculis binding protein homolog (Drosophila)	55084	NM_018013	yes	yes
SREBF2	sterol regulatory element binding transcription factor 2	6721	NM_004599	yes	no
TGFBR1	transforming growth factor, beta receptor I (activin A receptor type II-like kinase, 53kDa)	7046	NM_004612	no	yes
TMEM126A	transmembrane protein 126A	84233	NM_032273	no	yes
YAP1	Yes-associated protein 1, 65kDa	10413	NM_006106	yes	no
ZNF326	zinc finger protein 326	284695	NM_182975	no	yes

Table 4 Comprehensive list of validated hits from the genome-wide RNAi screen.

Table 4 shows the comprehensive list of 60 gene products, which we identified and validated to be involved in intracellular transport of SH4-domain-containing reporter proteins. Additional information about siRNAs used for the silencing of these genes can be found in the appendix (5.2).

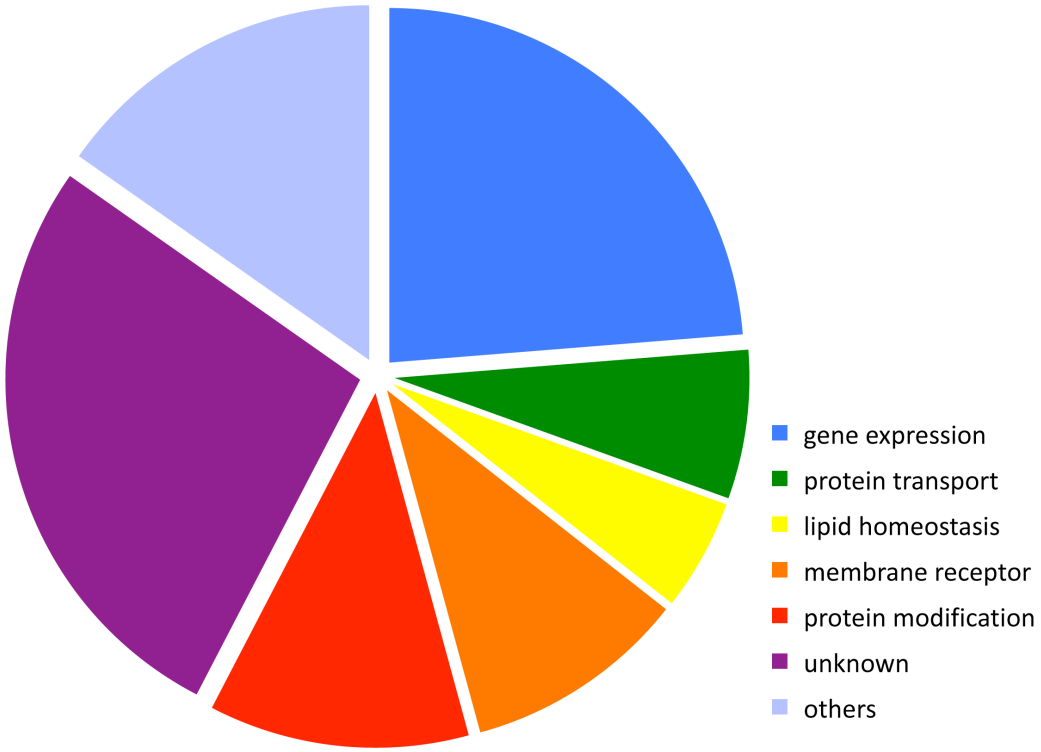

Figure 37 Functional classification of the 60 validated hits from the genome-wide RNAi screen

Figure 37 displays a functional classification of the 60 gene products identified and validated in our genome-wide RNAi screen. A major fraction of proteins, such as polymerases, splicing factors and ribosomal proteins, are involved in gene expression and therefore most probably exert secondary effects on the subcellular localization of SH4-domain-containing proteins. Interestingly, we were able to identify factors involved in intracellular protein transport (e.g. subunits of the heptameric coatomer complex) and lipid homeostasis. These findings might support our hypothesis that membrane microdomains of a specific lipid composition could play a role in targeting and transport of the SH4-domain to the plasma membrane. Quite surprisingly, we found a large number of surface receptors. A significant number of proteins we identified is involved in protein modifications (such as kinases and phosphatases) and might therefore play a regulatory role in intracellular trafficking of acylated proteins. The largest fraction of proteins has to date not been characterized in molecular detail. The remaining proteins are involved in other functions, such as ion homeostasis and signalling.

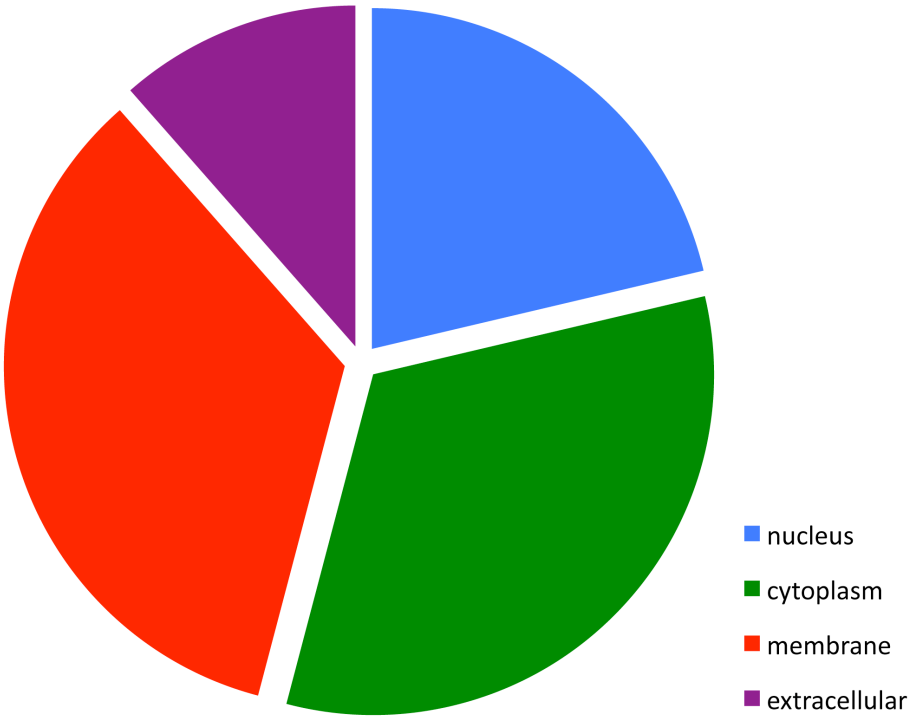

Figure 38 Subcellular localization of the 60 validated hits from the genome-wide RNAi screen

Promisingly, the vast majority of proteins identified and validated in our genome-wide RNAi screen localize to the cytoplasm or to intracellular or plasma membranes. This renders these molecules capable to directly interact with SH4-domain-containing proteins or regulatory and structural components of the unidentified transport machinery involved in trafficking of acylated proteins. Nuclear proteins most probably exert secondary effects and are involved in gene expression. Extracellular proteins cannot directly interact with the SH4-domain or transport entities containing SH4-domain-containing proteins, and could, therefore, influence localization of our reporter proteins by indirect mechanisms.

3.4 Examples of Screening Data for Validated Hits

Original data from the primary and validation screen are often difficult to interpret manually, since it is difficult to detect minor changes in the intracellular distribution of the fluorescent reporter proteins by eye. Some of the 60 validated hits, however, show a clear intracellular retention of SH4-domain-containing proteins after treatment with the respective siRNAs.

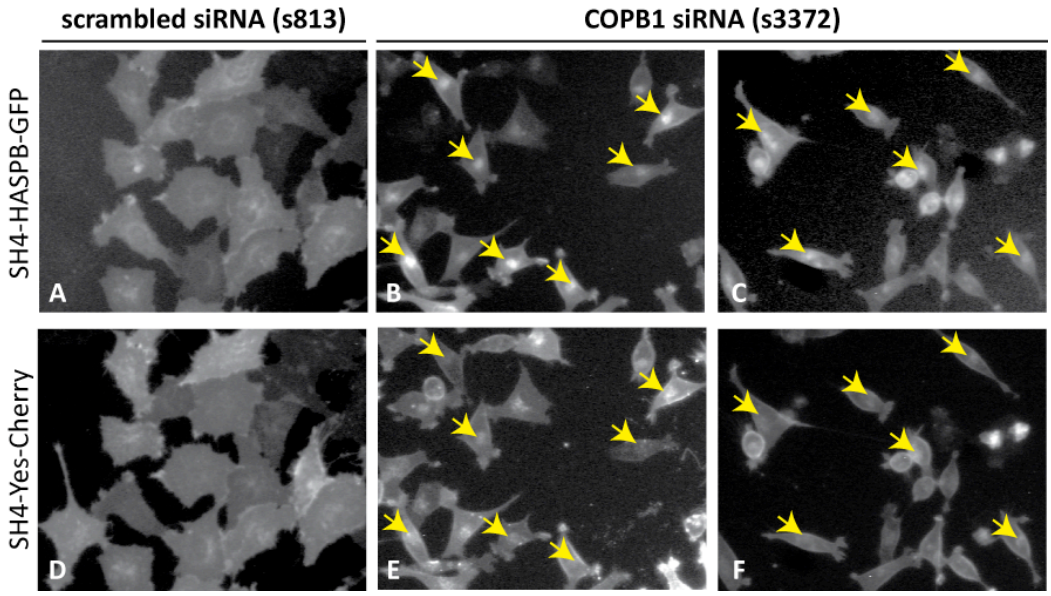

Figure 39 Original images of scrambled-siRNA-treated and COPB1-siRNA-treated cells acquired during the validation screen. Treatment of HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells with scrambled siRNA does not affect subcellular localization of SH4-HASPB-GFP (panel A) and SH4-Yes-Cherry (panel D). Knockdown of beta-COP results in strong perinuclear retention of SH4-HASPB-GFP (panels B and C) but only has a mild effect on SH4-Yes-Cherry localization (panels E, F; cutouts of 10x magnification images).

Figure 39 shows original data from the validation screen. HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells were cultivated on siRNA-coated 96-well plates for 36 hours before protein expression was induced by 1µg/ml doxycyline for additional 12 hours. Cells were then transferred to phenol-red-free CO_2-independent medium containing 0.2µg/ml Hoechst 33342 for nuclear staining and were subjected to automated fluorescence microscopy 48 hours post-transfection. Reverse transfection with scrambled siRNA does not affect the plasma-membrane-localization of both reporter proteins (panels A and D). Knockdown of beta-COP results in a pronounced perinuclear retention of SH4-HASPB-GFP (panels B and C), while localization of SH4-Yes-Cherry is only mildly affected (panels E and F). Due to the essential function of the COPI-coat in intracellular protein transport, cells undergo apoptosis if knockdown times of coatomer subunits are extended to more than 52 hours.

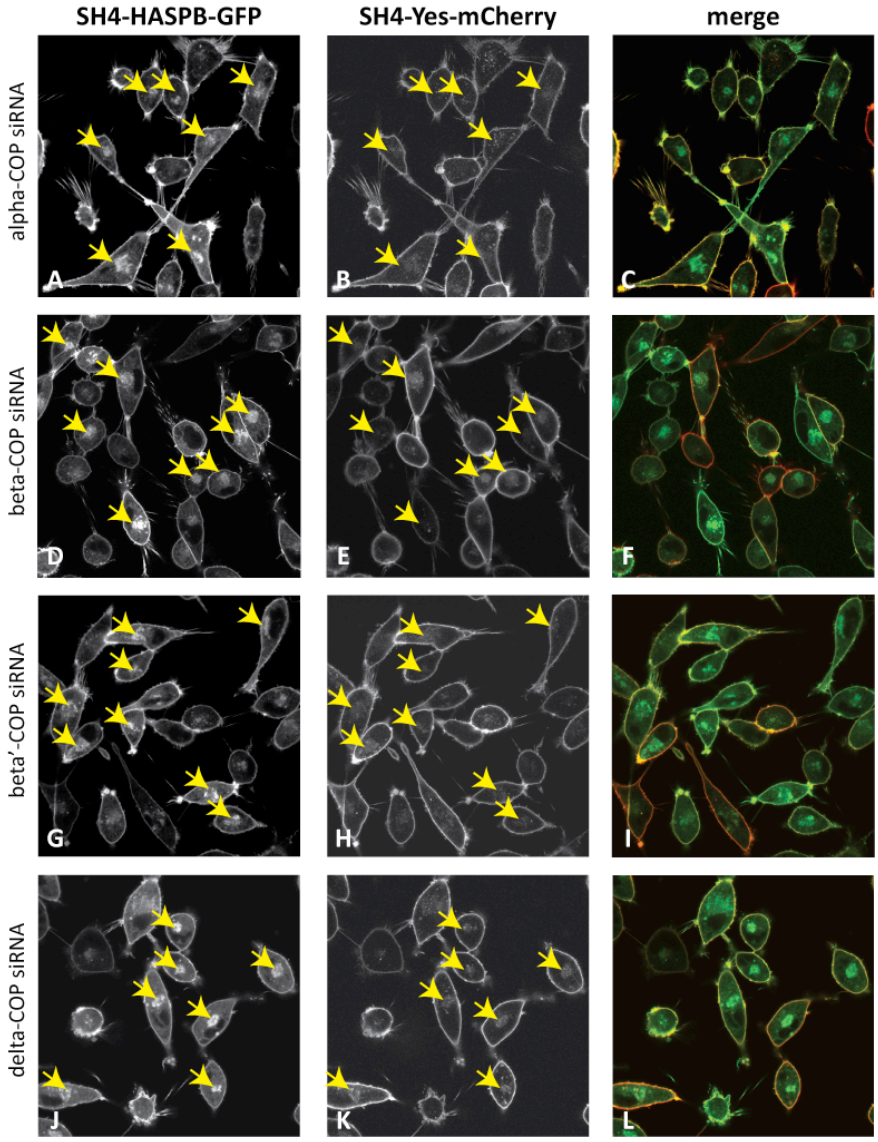

Figure 40 Knockdown of distinct coatomer subunits affects intracellular retention of SH4-domain-containing reporter proteins to different extents. Knockdown of alpha-COP (panels A, B, C), beta-COP (panels D, E, F), beta'-COP (panels G, H, I) and delta-COP (panels J, K, L) in HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 results in a strong intracellular retention of SH4-HASPB-GFP (panels A, D, G, J) but only mildly affects SH4-Yes-Cherry localization (panels B, E, H, K, confocal microscopy, 63x magnification).

Figure 40 depicts HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells after knockdown of different subunits of the heptameric coatomer complex. SiRNAs were liquid-phase transfected using oligofectamine (Invitrogen) and cultivated for 36 hours before protein expression was induced by cultivation in the presence of 1µg/ml doxycyline for additional 12 hours. Cells were analyzed by confocal microscopy 48 hours post-transfection. Panels A, D, G and J show unambiguously that knockdown of one of the coatomer subunits alpha-COP, beta-COP, beta'-COP and delta-COP results in a very pronounced perinuclear retention of SH4-HASPB-GFP. The intracellular localization of SH4-Yes-Cherry

(panels B, E, H and K) is not or only mildly affected. Due to downregulation of other COPI-coat components upon knockdown of individual coatomer subunits (Tao Wang, unpublished data), our data indicate that the coatomer protein complex is involved in intracellular transport of SH4-domain-containing proteins to the plasma membrane. Whether this effect is due to a direct (e.g. transport of SH4-domain-containing proteins in COPI-coated transport vesicles) or indirect role (e.g. mislocalization of an accessory factor for SH4-domain-dependent protein transport after destabilization of coatomer) remains to be elucidated. However, these data pose a proof-of-principle for the identification of proteins involved in intracellular protein transport by RNAi-based screening. At the same time, siRNAs directed against beta-COP present on the 384-spot siRNA arrays could be used as positive controls. Silencer select siRNAs (Ambion) directed against beta-COP and alpha-COP were used as positive controls in the validation screen.

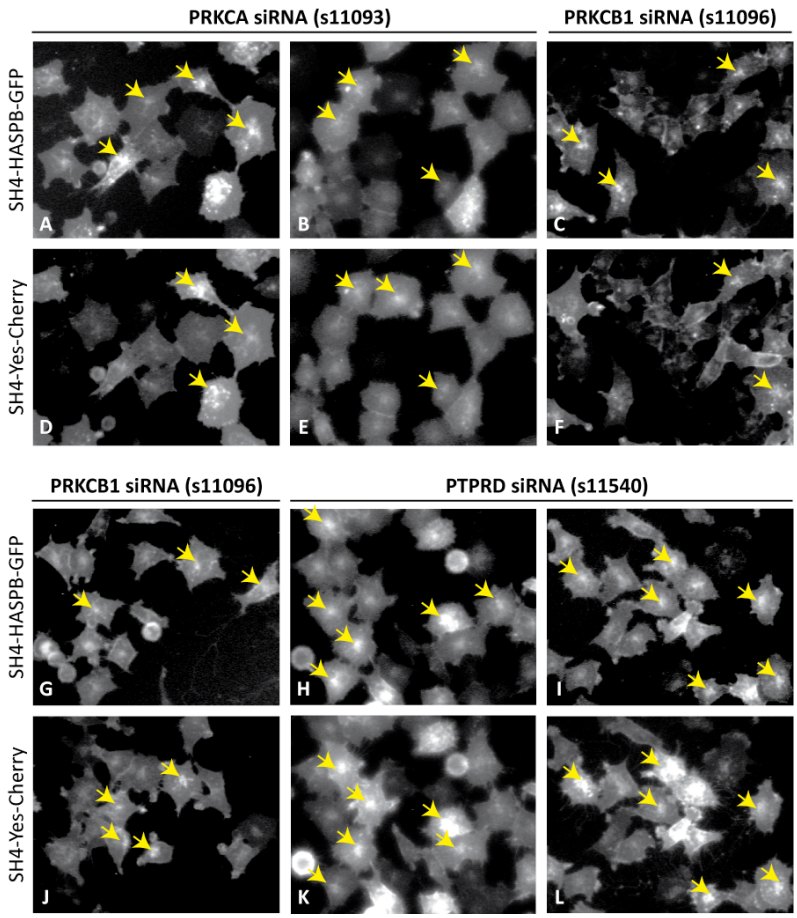

Figure 41 Original images of HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells treated with siRNAs against kinases and phosphatases validated in the genome-wide RNAi screen. Treatment of HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells with siRNAs directed against protein kinase C (PKC) family members (PRKCA: protein kinase C, alpha, panels A, B, D, E; PRKCB1: protein kinase C, beta, panels C, G, F, J) and protein tyrosine phosphatase, receptor

type, D (PTPRD, panels H, I, K, L) affects subcellular localization of both SH4-HASPB-GFP and SH4-Yes-Cherry (cutouts of 10x magnification images).

Figure 41 shows original data from the validation screen. HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells were cultivated on siRNA-coated 96-well plates for 36 hours before protein expression was induced by 1µg/ml doxycyline for additional 12 hours. Cells were then transferred to phenol-red-free CO_2-independent medium containing 0.2µg/ml Hoechst 33342 for nuclear staining and were subjected to automated fluorescence microscopy 48 hours post-transfection. Knockdown of two distinct members of the protein kinase C (PKC) family (PRKCA: protein kinase C, alpha and PRCB1: protein kinase C, beta) results in a perinuclear accumulation of both SH4-HASPB-GFP and SH4-Yes-Cherry (panels A-G, J). Members of the PKC family of proteins are serine-/threonine-specific kinases. PRKCA participates in a variety of cellular processes, such as cell adhesion and cell cycle checkpoint. PRKCB1 is, amongst others, involved in B cell activation, induction of apoptosis and endothelial cell proliferation. Treatment with siRNAs directed against PTPRD (protein tyrosine phosphatase, receptor type, D) affects plasma membrane localization of both SH4-domain-containing reporter proteins (panels H, I, K, L). Members of the protein tyrosine phosphatase (PTP) family of proteins are involved in various processes, such as cell growth, cell differentiation and mitotic cycle. Previous studies from our laboratory have shown that direct phosphorylation of threonines in the SH4-domain is a regulatory mechanism for recycling between the plasma membrane and endosomal compartments (Tournaviti et al., 2009).

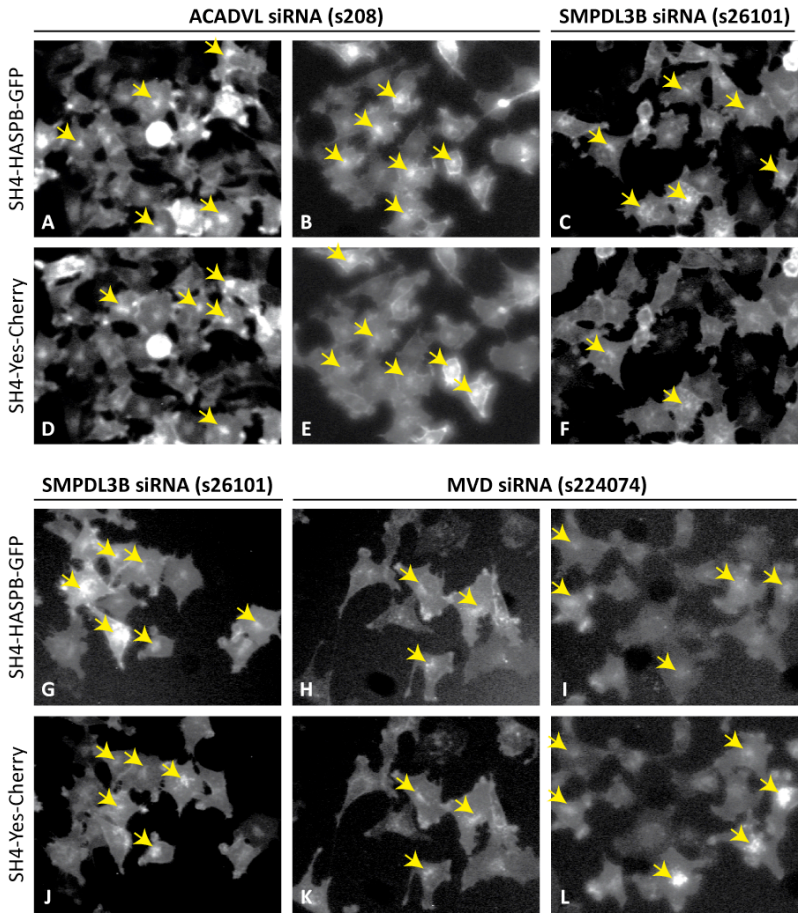

Figure 42 Original images of HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells treated with siRNAs against lipid homeostatic enzymes validated in the genome-wide RNAi screen. Treatment of HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells with siRNAs directed against acyl-Coenzyme A dehydrogenase, very long chain (ACADVL, panels A, B, D, E), sphingomyelin phopshodiesterase, acid-like 3B (SMPDL3B, panels C, G, F, J) and mevalonate (diphospho) decarboxylase (MVD, panels H, I, K, L) affects subcellular localization of both SH4-HASPB-GFP and SH4-Yes-Cherry (cutouts of 10x magnification images).

Figure 42 shows original data from the validation screen. HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells were cultivated on siRNA-coated 96-well plates for 36 hours before protein expression was induced by 1µg/ml doxycyline for additional 12 hours. Cells were then transferred to phenol-red-free CO_2-independent medium containing 0.2µg/ml Hoechst 33342 for nuclear staining and were subjected to automated fluorescence microscopy 48 hours post-transfection. Knockdown of enzymes involved in lipid homeostasis results in a perinuclear accumulation of both SH4-HASPB-GFP and SH4-Yes-Cherry. Acyl-coenzyme-A dehydrogenase, very long chain (ACADVL, knockdowns shown in panels A, B, D and E) is an enzyme of the inner mitochondrial membrane, which is specific to long-chain and very-long-chain fatty acids. It catalyzes the first step of the mitochondrial fatty-acid beta-oxidation pathway. Sphingomyelin phosphodiesterase, acid-like 3B (SMPDL3B, knockdowns shown in panels C, G,

F, J) catalyzes the hydrolysis of membrane sphingomyelin to phosphatidylcholine and ceramide. Mevalonate (diphospho) dehydrogenase (MVD, knockdowns shown in panels H, I, K, L) decarboxylates mevalonate pyrophosphate and produces isopentenyl pyrophosphate during early steps of cholesterol biosynthesis.

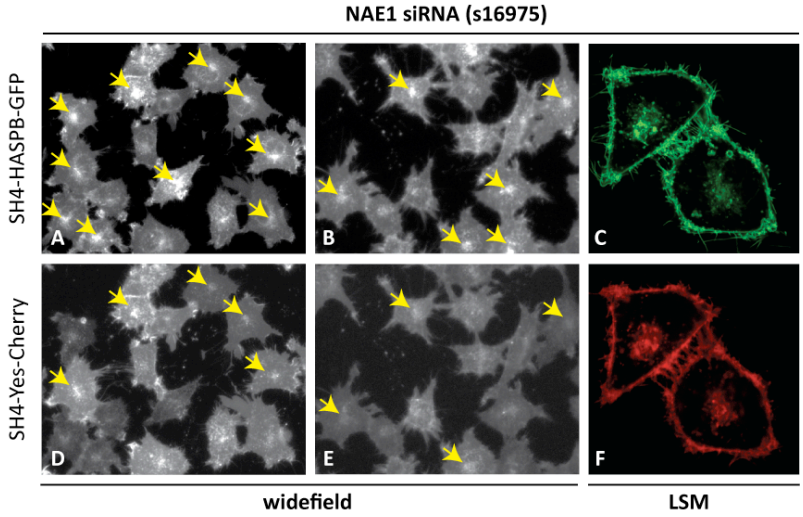

Figure 43 Widefield and confocal analysis of HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells treated with siRNAs against nedd-activating enzyme 1 (NAE1) for 60h. Treatment of HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells with siRNAs directed against nedd-activating enzyme 1 (NAE1), affects subcellular localization of SH4-HASPB-GFP (panels A-C) and, to a lesser extent, SH4-Yes-Cherry (panels D-F) as shown by widefield microscopy (panels A, B, D, E; cutouts of 10x magnification images) and confocal microscopy (63x magnification).

Figure 43 shows widefield and confocal microscopy images of HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5, which were cultivated on siRNA-coated LabTeks for 36 hours before protein expression was induced by 1µg/ml doxycyline for additional 24 hours. Cells were then transferred to phenol-red-free CO_2-independent medium containing 0.2µg/ml Hoechst 33342 for nuclear staining and were subjected to widefield and confocal microscopy 60 hours post-transfection. NAE1 could not be validated during the validation screen. This is most probably du to the fact that after 48 hours of knockdown, effects on the localization of the acylated reporter proteins was only mild, while extended knockdown times resulted in a significant intracellular retention of SH4-HASPB-GFP and, to a lesser extent, SH4-Yes-Cherry. NAE1 is involved in activation of the ubiquitin-like protein NEDD8, thus participating in the neddylation pathway, and interacts with amyloid precursor protein (APP) in membrane microdomains.

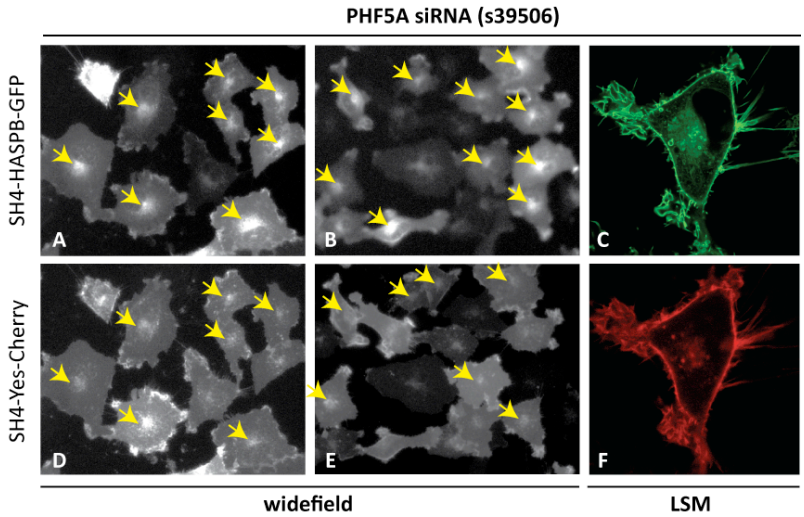

Figure 44 Widefield and confocal analysis of HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells treated with siRNAs against PHD finger protein 5A (PHF5A) for 60h. Treatment of HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells with siRNAs directed against PHF5A has a pronounced effect on subcellular localization of SH4-HASPB-GFP (panels A-C) and, to a lesser extent, SH4-Yes-Cherry (panels D-F) as shown by widefield microscopy (panels A, B, D, E; cutouts of 10x magnification images) and confocal microscopy (63x magnification).

Figure 44 shows widefield and confocal microscopy images of HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5, which were cultivated on siRNA-coated LabTeks for 36 hours before protein expression was induced by 1µg/ml doxycyline for additional 24 hours. Cells were then transferred to phenol-red-free CO_2-independent medium containing 0.2µg/ml Hoechst 33342 for nuclear staining and subjected to widefield and confocal microscopy 60 hours post-transfection. Knockdown of PHD finger protein 5A (PHF5A) results in a strong perinuclear accumulation of SH4-HASPB-GFP and, to a lesser extent, SH4-Yes-Cherry. PHF5A is a nuclear protein, which is a subunit of the splicing factor 3b complex.

4 DISCUSSION

Since the early '90s, it is known that the presence of an N-terminal SH4-domain directs targeting to intracellular membranes and that this targeting mechanism depends on lipid modifications within the SH4-domain (Blenis and Resh, 1993). While a co-translational myristoylation at the N-terminal glycine is in many cases sufficient to mediate association with intracellular membranes, subsequent palmitoylation is required for correct targeting to membrane microdomains at the plasma membrane and intracellular compartments (Tournaviti et al., 2009; Zacharias et al., 2002). While many SH4-domain-containing proteins associate with and are palmitoylated at compartments of the secretory partway (Bijlmakers and Marsh, 2003), little is known about the intracellular machinery required for transport to the plasma membrane. The fact that association of proteins with membrane microdomains is an apical sorting determinant in polarized cells (Schuck and Simons, 2004) led to the hypothesis that the partitioning of SH4-domain-containing proteins into export-competent lipid microdomains at intracellular compartments is a prerequisite for export to the plasma membrane.

It was the aim of this study to identify molecular components of the intracellular transport machinery of acylated proteins. For this purpose, we established a human model cell line, which inducibly expresses two distinct fluorescently tagged SH4-domains. In addition, we established an automated image-analysis tool, which is able to independently determine the intracellular localization of the two reporter proteins. These tools enabled us to perform a genome-wide, microscopy-based RNAi screen for the identification of gene products, whose downregulation results in a failure of plasma membrane targeting of one or both of the SH4-domain-containing reporter proteins. To validate candidate genes identified in the primary screen, we performed a small-scale validation screen using siRNAs of independent sequence.

4.1 Generation and Characterization of a Human Model Cell Line for Microscopy-based RNAi Screening

The application of a human siRNA library on a high-content microscopy-based screening platform required a human model cell line expressing fluorescent SH4-domain-containing reporter proteins. Therefore, we generated a clonal

HeLa cell line, which stably expresses two distinct SH4-domain-containing reporter proteins upon induction with doxycyline (HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5). SH4-HASPB-GFP and SH4-Yes-Cherry are myristoylated at the N-terminal glycine and palmitoylated at cysteine 5 and cysteine 3, respectively (Table 3). Reporter protein expression is homogeneously induced in the presence of doxycyline (Figure 15), and results in co-localization of both reporter proteins at the plasma membrane, which is clearly visible in confocal (Figure 17) as well as widefield microscopy (Figure 19). This approach enabled us to not only search for factors involved in intracellular transport of one specific SH4-domain-containing protein, but rather to extend our investigation to a more comprehensive view on SH4-domain-dependent intracellular protein transport.

As shown before, plasma membrane localization of SH4-HASPB and SH4-Yes in HeLa cells critically depends on the acylation state of the SH4-domain: mutagenesis of the myristoylation site also blocks palmitoylation and hence results in a cytosolic accumulation of SH4-domain-containing proteins (Figure 17, Figure 18 and Figure 19). When palmitoylation is prevented, either by mutagenesis of the target cysteine, or by treatment with a palmitate analogue, both SH4-domains accumulate at intracellular membranes. However, the two reporters show distinct sensitivities towards loss of palmitoylation: mutagenesis of the target cysteine in SH4-HASPB results in a complete loss of plasma-membrane-associated material, while a reduced but substantial amount of reporter protein still reaches the plasma membrane after deletion of the palmitoylation site of SH4-Yes (Figure 17, Figure 18). The finding that SH4-HASPB accumulates at intracellular membranes after treatment with low concentrations of a palmitoylation inhibitor, while SH4-Yes transport to the plasma membrane is blocked only at high concentrations of the same inhibitor corroborates these data (Figure 20). Interestingly, this effect is also clearly visible in HeLa SH4-HASPB-GFP/SH4-Yes-Cherry #5 cells (Figure 21), indicating that this model cell line is indeed suitable for detection of differential effects on the two expressed reporter proteins.

Our model cell line was efficiently transfected by cultivation on siRNA arrays and showed the expected phenotypes upon knockdown of control genes (Figure 22).

4.2 ESTABLISHMENT OF AN AUTOMATED IMAGE ANALYSIS TOOL

The acquisition of large amounts of high-content data during microscopy-based screening procedures poses the problem of how to analyze data efficiently, quantitatively and in an unbiased manner. A suitable analysis tool has to identify cells, which, upon knockdown of individual genes, are defective in transport of SH4-domain-containing proteins to the plasma membrane and hence accumulate reporter proteins at intracellular membranes or in the cytosol. The hallmark of both such phenotypic changes is the redistribution of fluorescence intensity from the plasma membrane to intracellular compartments. Therefore, an appropriate analysis tool has to fulfil several tasks: it has to identify single cells, determine the intracellular distribution of reporter protein fluorescence between subcellular compartments and classify cells according to intracellular intensity distributions. For the identification of individual cells, this poses several problems: reporter fluorescence intensities of single cells vary, so that fixed threshold approaches cannot be applied. Moreover, even though identification of the cell shape depends on reporter protein fluorescence, it has to work independently from the distribution of the fluorescence intensities between different cellular compartments, while resolution of images is low and the quality often compromised. To meet these requirements, we established a single-cell analysis tool, which uses cell nuclei as seed point for a classifier-enhanced region growing approach. This approach is based on local fluorescence intensity thresholds and hence adapts to varying cell brightness. Moreover, the identification of cell boundaries works robustly regardless of the intracellular intensity distributions of individual cells. After identification of the cell area, each cell is segmented into three different compartments: the nucleus (which corresponds to Hoechst-stained objects), the plasma membrane (which correlates to a fixed width of pixels inside the cell boundary) and the cytoplasm (which corresponds to the remaining cell area) (Figure 24). GFP and Cherry intensity features for each area are extracted independently and used for classification of individual cells (Figure 25). The automated image analysis tool was able to robustly identify different reporter protein localizations, as evoked by expression of mutants of the SH4-domain (Figure 26) or treatment with a palmitoylation inhibitor (Figure 27). In conclusion, the automated single-cell image analysis tool established during this study provides a powerful tool for reliable analysis of high-content imaging data acquired during the genome-wide RNAi screen.

4.3 Performance of a Genome-wide Microscopy-based RNAi Screen

To identify cellular factors involved in intracellular transport of SH4-domain-containing proteins, we conducted a genome-wide microscopy-based RNAi screen. To silence all ~ 22.000 genes of the human genome, HeLa cells were reverse transfected with a genome-wide siRNA library consisting of more than 50.000 siRNAs by cultivation of cells on 384-spot siRNA arrays. After 36 hours, simultaneous expression of SH4-HASPB-GFP and SH4-Yes-Cherry was induced by cultivation in the presence of doxycyline for another 12 hours. 48 hours post-transfection, cell nuclei were stained and cells subjected to automated live-cell microscopy. In the primary screen, 149 siRNA arrays were analyzed in duplicates. Results from the automated image analysis were transformed into scores with regard to analysis results from the same siRNA array. Application of suitable score thresholds resulted in the identification of 314 gene products, whose downregulation resulted in a redistribution of either one of the two reporter proteins. In a validation screen, these candidate genes were analyzed by reverse transfection of independent siRNAs in multi-well plates. Classification results of 9 plates in quadruplicates were transformed to scores with regard to negative controls. The application of an adapted hit threshold resulted in the validation of 60 gene products, which had an effect on localization of one or both reporter proteins (Figure 29). Of these, a significant amount is involved in functions, which at first sight seem to be unrelated to intracellular transport events, such as gene expression and thus probably exert secondary effects. However, we were also able to identify components of the intracellular transport machinery and enzymes involved in lipid homeostasis and protein modifications, which could play direct roles in the intracellular transport of SH4-domain-containing proteins. More than three quarters of all hits localize to the cytoplasm and the nucleus, which theoretically renders these proteins capable to directly interact with SH4-domain-containing proteins or membrane microdomains and transport carriers that include these as cargo.

4.3.1 *Potential Role of Coatomer in Intracellular Transport of Acylated Proteins*

The first gene product identified in our genome-wide RNAi screen was beta-COP, which is one of seven subunits of the heptameric coatomer protein complex. COPI-coated vesicles mediate retrograde transport from the Golgi to the ER and traffic between Golgi cisternae (Rothman and Wieland, 1996). While beta-COP and alpha-COP were identified in the genome-wide RNAi screen, further analysis showed that other coatomer subunits also have a pronounced effect on intracellular localization of SH4-HASPB-GFP but do not or only mildly affect SH4-Yes-Cherry localization (Figure 40). Analysis of knockdown efficiencies showed that downregulation of one coatomer subunit resulted in decreased protein levels of other coatomer subunits (Tao Wang, personal communication) and thus most probably led to destabilization of the coatomer complex. Indeed, a prolonged knockdown of different coatomer subunits induces apoptosis (data not shown). The effect of coatomer knockdown on the intracellular localization of SH4-HASPB-GFP poses several possibilities: the involvement of coatomer subunits in intracellular transport of SH4-HASPB-GFP could indicate that the acylated protein associates with membranes of the ERGIC/Golgi and for this reason its localization is directly affected by COPI-dependent sorting events. Alternatively, since interference with the COPI-coat affects intracellular sorting of a wide variety of cargo proteins, coatomer destabilization might indirectly affects SH4-HASPB-GFP transport, for example by causing a mislocalization of trafficking factors essential for SH4-HASPB-GFP surface delivery. Even though the molecular role of coatomer in intracellular transport of SH4-domain-containing proteins is unknown, its identification poses a strong proof-of-principle for our screening procedure. By analysis of microscopy-based data acquired with a high-content screening platform with a newly established image analysis tool, we were able to identify several components of a multiprotein complex that plays distinct roles in intracellular transport of two different SH4-domain-containing reporter proteins.

4.3.2 *Potential Role of Kinases and Phosphatases in Intracellular Transport of Acylated Proteins*

The downregulation of two protein kinase C (PKC) isoforms, PKC alpha (PRKCA/PRKCα) and beta 1 (PRKCB1/PRKCβI), and one phosphatase, protein

tyrosine phosphatase (PTPRD), causes a significant accumulation of SH4-domain-containing reporter proteins at perinuclear membranes (Figure 41).

Members of the protein kinase C (PKC) family are serine/threonine kinases, whose biological functionality is regulated at the level of their catalytic activity and spatial localization. Activation requires a “maturation” step, which involves phosphorylation of internal serine and threonine residues and the presence of several activators, including diacylglycerol (DAG) and Ca^{2+}. Upon activation, PKC translocates to specific intracellular compartments, for example the plasma membrane and the Golgi apparatus (Shirai and Saito, 2002).

PKCα, a classical PKC, is ubiquitously expressed and activated by various stimuli, for example by tyrosine kinase receptors or hydrolysis of phospholipids (Nakashima, 2002; Nishizuka, 1992). Kinase activity of PKCα requires phosphorylation of three conserved residues within the kinase domain (Hansra et al., 1999). Upon stimulation, PKCα translocates from the cytosol to the nucleus and to membranes, in particular the plasma membrane, where DAG is produced (Tanimura et al., 2002). Since PKCα is involved in cell adhesion, motility, proliferation and survival, it is not surprising that elevated PKCα activity has been implicated in tumorigenesis (Michie and Nakagawa, 2005).

PKCβI is also a classical Ca^{2+} and DAG-dependent isoform of PKC and originates from the PKCβ gene, which gives rise to PKCβI and PKCβII by alternative splicing of C-terminal exons (Ono et al., 1987). Similar to PKCα, PKCβI is widely expressed and translocates to membranes upon activation (Kawakami et al., 2002). PKCβI plays well-established roles in immune-receptor and insulin receptor signaling and is involved in the activation of JNK and NFκB, which regulate cytokine gene expression and cell survival (Kawakami et al., 2002).

The role of PKC isoforms on intracellular transport of SH4-domain-containing proteins is not known. The presence of serine and (in the case of SH4-HASPB) threonine residues poses the possibility that SH4-domains are a substrate of PKC. Indeed, it was recently shown that phosphorylation and dephosphorylation of the SH4-domains of HASPB and Yes provides an additional level of regulation by an internalization and recycling switch (Tournaviti et al., 2009). In the current model, phosphorylation of the SH4-domain is required for endocytosis, while dephosphorylation triggers recycling to the plasma membrane (Figure 5). This hypothesis therefore contradicts the involvement of PKC in the phosphorylation step, since a loss of phosphorylation by knockdown of PKC isoforms should result in a reduced internalization rate and hence

stabilize the association of acylated proteins with the plasma membrane. However, it is possible that phosphorylation by PKC isoforms has a distinct role in intracellular transport of acylated proteins. One possibility is a PKC-mediated regulatory mechanism based on phosphorylation of factors, which are required for plasma membrane localization of SH4-domain-containing proteins. Phosphorylation of such proteins could promote transport to the plasma membrane or inhibit re-internalization of acylated proteins by endocytosis. Due to the wide variety of processes in which PKC seems to be involved, however, secondary effects cannot be excluded.

Protein tyrosine phosphatase, receptor type, D (PTPRD) is a member of the protein tyrosine phosphatase (PTP) family consisting of a large extracellular domain, a single transmembrane segment and two tandem intracytoplasmic domains and forms a homodimer at the plasma membrane (Pulido et al., 1995). Even though little is known about downstream functions of PTPRD, transmembrane PTPs have been implicated in mediating integrin-cytoskeleton association via an interaction with catenins and thus may play a role in neurite outgrowth (Kypta et al., 1996). Since the phosphorylation switch regulating endocytosis and recycling of SH4-domains depends on phosphorylation and dephosphorylation of serine and/or threonine residues, a potential direct involvement of PTPRD in this process is unlikely. Moreover, SH4-HASPB does not contain any tyrosine residues, thus making a direct dephosphorylation impossible. Therefore, the effect of downregulation of PTPRD seems to be based on an indirect mechanism and most probably does not involve dephosphorylation of the SH4-domain.

4.3.3 Potential Role of lipid-homeostatic Enzymes in Intracellular Transport of Acylated Proteins

Given our hypothesis that intracellular transport of SH4-domain-containing proteins requires partitioning into export-competent membrane microdomains at the level of the Golgi, the identification of three enzymes involved in lipid homeostasis was especially intriguing (Figure 42).

One of the enzymes involved in lipid metabolism identified in our genome-wide screen is mevalonate (diphospho) decarboxylase (MVD). MVD is a cytosolic homodimer and involved in early steps of cholesterol biosynthesis, where it catalyzes the conversion of mevalonate pyrophosphate into isopentenyl

pyrophosphate and dehydrates its substrate under ATP consumption (Hogenboom et al., 2002; Hogenboom et al., 2003; Toth and Huwyler, 1996). The structure of MVD showed that it is a member of the GHMP (galactokinase, homoserine kinase, mevalonate kinase, phosphomevalonate kinase) kinase family (Bonanno et al., 2001). Asp 302 and Lys 18 constitute the active site (Krepkiy and Miziorko, 2004). In an RNAi-based approach, MVD was identified as a key component in dengue virus replication (Rothwell et al., 2009). Cholesterol is one of the main constituents of lipid rafts and has been proposed to fill voids between lipids and enhance tight packing of the liquid-ordered microdomain by a tight interaction with sphingomyelin (Brown and London, 1998; Simons and Ikonen, 1997). Inhibition of the cholesterol pathway by downregulation of MVD may result in cholesterol depletion and could thus interfere with the biogenesis of cholesterol-enriched lipid microdomains. As a result, formation and clustering of intracellular microdomains, which could direct sorting and transport of acylated proteins to the plasma membrane, could be inhibited by downregulation of MVD and thus cause an intracellular accumulation of SH4-domain-containing reporter proteins.

Acyl-Coenzyme A dehydrogenase, very long chain (ACADVL) is one of five members of the family of acyl-CoA dehydrogenases (ACADs), which catalyze the first step of the mitochondrial fatty-acid beta-oxidation pathway. In contrast to the other ACADs, which are soluble homotetramers, ACADVL is a homodimer of 70 kDa subunits, which is associated with the inner mitochondrial membrane (Souri et al., 1998). A C-terminal loop, which is not present in the other ACADs, is involved in binding to the mitochondrial membrane. The ACADVL structure allows for binding of acyl chains up to 24 carbons in length (McAndrew et al., 2008), however, ACADVL shows the highest specificity to palmitoyl-CoA (Izai et al., 1992). While the beta-oxidation of long-chain fatty acids is an essential pathway for production of energy equivalents in the organism and loss of ACADVL activity results in cardiomyopathy in patients (Aoyama et al., 1993), the effects in a single-cell system are most probably governed by a deregulation of intracellular lipid homeostasis. Long-chain saturated fatty acidy are essential components of lipid species enriched in membrane microdomains (Simons and Ikonen, 1997; Simons and van Meer, 1988). Moreover, palmitoyl-CoA is the substrate for protein acyl transferases (PATs) and therefore essential for target protein palmitoylation (Bijlmakers and Marsh, 2003). Even though a decreased level of long-chain fatty acid degradation might provide more substrate for

biogenesis of raft lipids and palmitoylated proteins, it most probably severely affects the tightly controlled homeostasis of different lipid species within the cell. Potential effects on clustering of microdomains at the level of the Golgi, which may lead to intracellular accumulation of SH4-domain-containing reporter proteins, therefore do not seem unlikely. Moreover, an overflow of palmitoyl-CoA in the cytosol is most probably counteracted by sequestration by acyl CoA binding protein (ACBP) and thus does not necessarily provide an efficient substrate for target palmitoylation. Potential feedback loops changing long-chain fatty-acid synthesis rates can also not be excluded. However, it is difficult to easily infer a direct relation between the downregulation of ACADVL and its potential effect on targeting and transport of acylated proteins.

Sphingomyelin phosphodiesterase, acid-like 3B (SMPDL3B) catalyzes the hydrolysis of sphingomyelin (SM) into phosphocholine and ceramide. The family of sphingomyelinases (SMases) contains one lysosomal and one secreted acid form, the latter being dependent on Zn^{2+} for enzymatic activity (Marchesini and Hannun, 2004). Both forms are encoded by the same gene and can be differentially targeted to lysosomes or through the Golgi secretory pathway (Schissel et al., 1998b; Tabas, 1999). Even though its enzymatic activity is highest at acidic pH, cell surface SMase is able to hydrolyze raft-associated sphingomyelin to ceramide at physiological pH (Schissel et al., 1998a). SMase activity results in formation and clustering of ceramide-enriched microdomains and causes shedding of vesicles *in vitro* (Nurminen et al., 2002). Formation of ceramide-enriched microdomains at the plasma membrane is triggered by different extracellular cues, for example bacterial or viral infection, which trigger SMase appearance at the cell surface, or by extracellular treatment with sphingomyelinase or ceramide (Grassme et al., 1997; Gulbins and Kolesnick, 2003). Moreover, it has been shown that CD40 and CD95 initiate release of acid sphingomyelinase, which results in generation of ceramide and subsequent clustering of the receptors in ceramide-enriched membrane platforms (Grassme et al., 2001a; Grassme et al., 2002; Grassme et al., 2001b). These data clearly show that ceramide-enriched domains at the plasma membrane are generated by extracellular sphingomyelinase activity and that these microdomains participate in a variety of physiological signalling events, most probably by promoting selective clustering of signalling molecules. The identification of a sphingomyelinase in our screen implies the importance of ceramide-enriched plasma membrane domains for plasma membrane localization of SH4-domain-

containing proteins, since downregulation of sphingomyelinase by treatment with specific siRNAs may inhibit formation of such clustered domains at the cell surface. However, this effect is exerted directly at the extracellular leaflet of the plasma membrane and has no direct implication in the formation of functional SH4-domain-containing membrane microdomains at the level of the Golgi. One possible explanation for this effect might be that plasma membrane targeting is prevented by an unknown feedback mechanism from plasma membrane platforms to intracellular microdomains, preventing the formation of export-competent domains. Another possibility would be an accelerated rate of endocytosis of SH4-domain-containing reporter proteins. Even though no direct mechanism for the intracellular accumulation of our reporter can be conferred from our finding, it is an interesting observation that interference with the outer leaflet of membrane microdomains has a direct influence on proteins associated with the cytosolic leaflet of membrane microdomains.

Even though it is at first sight challenging to infer direct roles of the three enzymes, it is clear that interference with the cellular lipid homeostasis has a pronounced effect on the intracellular distribution of acylated proteins. These findings may corroborate our hypothesis, that surface delivery of SH4-domain-containing proteins is dependent on partitioning into membrane microdomains.

4.3.4 Potential Role of NAE1 in Intracellular Transport of Acylated Proteins

The involvement of Nedd-activating enzyme 1 (NAE1) in intracellular transport of acylated proteins was identified in additional experiments with extended knockdown times, where NAE1 knockdown resulted in a significant intracellular retention of the acylated reporter proteins (Figure 43). NAE1 was first characterized as a novel binding partner of the cytoplasmic domain of amyloid precursor protein (APP) and was hence called APP binding protein 1 (APP-BP1) (Chow et al., 1996). Soon, it was shown that NAE1 participates in the neddylation pathway (Gong and Yeh, 1999). Protein neddylation refers to the covalent attachment of the ubiquitin-like protein NEDD8 to target proteins. Though NEDD8 is essential in most organisms, the functional consequences of neddylation for target proteins are not entirely clear and neddylation has been implicated to trigger structural changes of target proteins, to prevent association with binding partners, or to recruit NEDD8-interacting enzymes

(Rabut and Peter, 2008). NAE1 forms a heterodimer with hUba3 and activates NEDD8 by a conserved E1-like mechanism (Bohnsack and Haas, 2003; Osaka et al., 1998). NAE1 co-localizes with APP in membrane microdomains, an effect that is enhanced in Alzheimer's disease. The ability of NAE1 to induce neuronal apoptosis is dependent on interaction with the APP cytosolic domain and requires a functional neddylation pathway, indicating that APP-BP1 recruitment to membrane microdomains by APP initiates the neddylation pathway and thus mediates neuronal apoptosis (Chen et al., 2003; Chen et al., 2000). Interestingly, transcriptional activity of the intracellular domain of APP is regulated by neddylation (Lee et al., 2008). The interaction of APP and NAE1 has also been shown to activate Rab5-dependent endocytosis in a neddylation-independent manner (Laifenfeld et al., 2007). NAE1 may influence intracellular transport of acylated proteins in several ways: NAE1 could co-localize with SH4-domain-containing proteins in lipid rafts and thus increase their affinity for liquid-ordered domains, a process which should in turn be dependent on the presence of APP in the same microdomain. Another possibility is a potential involvement of protein neddylation. Even though it seems unlikely that the SH4-domain is neddylated itself, other factors required for the localization of acylated proteins at the plasma membrane could be regulated by covalent attachment of NEDD8. Neddylation of such proteins could be required for efficient surface delivery of SH4-domain-containing proteins, or could prevent their internalization. A more indirect effect could be caused by elevated levels of Rab5-mediated endocytosis, which could result in accelerated internalization of plasma-membrane-associated SH4-proteins.

4.3.5 Potential Role of PHF5A in Intracellular Transport of Acylated Proteins

PHD finger protein 5A (PHF5A) was validated in the genome-wide RNAi screen and analyzed in further experiments with extended knockdown times, where its knockdown resulted in a strikingly pronounced intracellular retention of SH4-HASPB-GFP and, to a lesser extend, SH4-Yes-Cherry (Figure 44).

PHF5A is a small nuclear protein containing a PHD domain that acts as a chromatin-associated protein (Trappe et al., 2002a), and is essential for morphogenetic development in *C. elegans* (Trappe et al., 2002b). Later, PHF5A was shown to interact with ATP-dependent helicases at its N-terminus and

subunits of the arginine-serine-rich domains of splicing factors at its C-terminus (Rzymski et al., 2008).

At first sight, these findings seem to make a direct role of PHF5A in intracellular transport of SH4-domain-containing proteins very unlikely and thus may point towards a pleiotropic effect.

However, in recent years, dual functions have been observed for several proteins in different cellular compartments: Sec13, a component of the COPII-coat, which mediates cargo traffic between the ER and the Golgi complex, was shown to be an essential part of a nuclear pore subcomplex (Siniossoglou et al., 1996). High mobility group box 1 (HMGB1) is a chromatin-associated protein promoting protein association with specific DNA targets (Yang and Tracey, 2005) and is additionally released by an unconventional mechanism, acting as a late mediator in systemic inflammations (Yang et al., 2005). Thioredoxin is an antioxidant participating in the reduction of intracellular proteins (Holmgren, 1989) and in addition is secreted to the extracellular space (Rubartelli et al., 1992). These findings indicate that some proteins fulfil differential functions at different localizations within or outside the cell. Whether PHF5A also possesses such dual functionality, or whether the observed effect on the steady-state localization of acylated reporter proteins is due to secondary effects, however, remains to be elucidated.

4.4 CONCLUSION AND FUTURE PERSPECTIVES

In this study, we performed a genome-wide microscopy-based RNAi screen. The establishment of an automated single-cell analysis tool enabled us to identify 60 gene products involved in intracellular targeting and transport of SH4-domain-containing proteins. We were able to identify components of the intracellular trafficking machinery and showed that potential effects on homeostasis of specific raft lipids have an influence on the steady-state plasma membrane localization of acylated proteins, which may corroborate our hypothesis that intracellular transport is indeed dependent on association with membrane microdomains. It will be an interesting challenge to unravel the molecular mechanisms underlying the observed intracellular accumulations of reporter proteins and thus to shed light on the intracellular transport machinery required for surface delivery of acylated proteins.

For validation of our microscopy-based data, the establishment of a biochemical assay is in progress. This assay is based on subcellular membrane fractionation and can be used to quantify retention of SH4-domain-containing reporter proteins on intracellular membranes under knockdown conditions.

In addition, secondary assays will be applied to characterize the role of identified gene products in the intracellular transport of acylated proteins. For example, isolation of detergent resistant membranes (DRMs) will be used to determine and distinguish specific effects on the DRM-association of SH4-domain-containing reporter proteins and a more general interference with microdomain composition and integrity. Pleiotropic effects could be excluded by analyzing export efficiencies of conventional and unconventional secretory cargo. In the case of NAE1, the role of APP and a potential involvement of the neddylation pathway will be further investigated.

Recent findings from out laboratory have shown that the steady-state localization of acylated reporter proteins is not only dependent on surface delivery, but is additionally regulated by an internalization and recycling step (Tournaviti et al., 2009). The discovery of a mutant version of the SH4-domain of HASPB, SH4-HASPB-T6A, which cannot be phosphorylated and therefore does not undergo endocytosis, provides a powerful tool to distinguish between RNAi-mediated effects on surface delivery of acylated reporter proteins and intracellular retention following an internalization step. Moreover, extended knockdown times and an improved siRNA library could yield a more comprehensive view on the molecular machineries involved in intracellular transport of SH4-domain-containing reporter proteins.

5 APPENDIX

5.1 ABBREVIATIONS

2-BP	2-bromopalmitate
APT	acyl protein thioesterase
ACBP	acyl-CoA binding protein
ACAD	acyl-CoA dehydrogenase
ACADVL	acyl-CoA dehydrogenase, very long chain
ATP	adenosine triphosphate
ARF	ADP-ribosylation factor
COPA	alpha-COP
N-terminal	amino-terminal
APP	amyloid precursor protein
APP-BP1	amyloid precursor protein binding protein 1
COPB	beta-COP
BFA	brefeldin A
C. elegans	*Caenorhabditis elegans*
C-terminal	carboxy-terminal
CHO	Chinese hamster ovary
CCP	clathrin-coated pit
CCV	clathrin-coated vesicle
CME	clathrin-mediated endocytosis
COP	coat protein
DV	dengue virus
DNA	desoxyribonucleic acid
dNTP	desoxyribonucleotide triphosphate
DRM	detergent-resistant membrane
DAG	diacylglycerol
dsRNA	double-stranded RNA
D-MEM	Dulbecco's modified eagle medium
ER	endoplasmic reticulum
ERGIC	ER/Golgi intermediate compartment
E. coli	*Escherichia coli*
Erf	essential recombination function
FCS	fetal calf serum
FGF	fibroblast growth factor
FACS	fluorescence activated cell sorting
Gal	galectin
GPI	glycosylphosphatidylinositol
GFP	green fluorescent protein
GAP	growth associated protein
GTP	guanosine triphosphate
HSPG	heparan sulfate proteoglycan
HCV	hepatitis C virus
HMG	high mobility group

HEK	human embryonic kidney
HIV	human immunodeficiency virus
HASP	hydrophilic acylated surface protein
kb	kilobase
LPC	large pleiomorphic carrier
mRNA	messenger RNA
MVD	mevalonate (diphodpho) dehydrogenase
miRNA	micro RNA
MCAT	murine cationic transporter
MCS	mutliple cloning site
MARCKS	myristoylated alanine-rich C kinase substrate
NMT	N-myristoyltransferase
NAE1	nedd-activating enzyme 1
NEDD8	neural precursor cell expressed, developmentally downregulated 8
nt	nucleotide
PBS	phosphate buffered saline
PE	phosphatidylethanolamine
PI	phosphatidylinositol
$PI(4,5)P_2$	phosphatidylinositol-4,5-bisphosphate
PS	phosphatidylserine
PM	plasma membrane
PCR	polymerase chain reaction
PAT	protein acyl transferase
PKC	protein kinase C
PRKCA	protein kinase C alpha
PRKCB1	protein kinase C beta 1
PTP	protein tyrosine phosphatase
RNA	ribonucleic acid
RNAi	RNA interference
RISC	RNA-induced silencing complex
rpm	rounds per minute
S. cerevisiae	*Saccharomyces cerevisiae*
siRNA	short interfering RNA
SRP	signal recognition particle
SV40	simian virus 40
SNARE	SNAP receptor
SNAP	soluble NSF attachment protein
SMPDL3B	sphingomyelin phosphodiesterase, acid-like 3B
SMase	sphingomyelinase
SFK	Src family kinase
SH	Src homology
TGN	*trans*-Golgi network
U	unit
VSV-G	vesicular stomatitis virus glycoprotein
VTC	vesicular-tubular cluster
VP22	virus protein 22
Yck	yeast casein kinase

5.2 Comprehensive List of Cell Lines Generated in this Study

Cell Line	*Insert*	*Size*
HeLa SH4-HASPB-GFP	SH4-HASPB-GFP	28.9 kD
HeLa SH4-HASPB-GFP #6	SH4-HASPB-GFP	28.9 kD
HeLa SH4-HASPB-Δmyr-GFP	SH4-HASPB-Δmyr-GFP	28.9 kD
HeLa SH4-HASPB-Δpal-GFP	SH4-HASPB-Δpal-GFP	28.9 kD
HeLa SH4-HASPB-Cherry	SH4-HASPB-Cherry	28.7 kD
HeLa SH4-HASPB-dRFP	SH4-HASPB-dRFP	55.1 kD
HeLa SH4Yes-GFP	SH4-Yes-GFP	29.1 kD
HeLa SH4-Yes-GFP #7	SH4-Yes-GFP	29.1 kD
HeLa SH4-Yes-Δmyr-GFP	SH4-Yes-Δmyr-GFP	29.1 kD
HeLa SH4-Yes-Δpal-GFP	SH4-Yes-Δpal-GFP	29.1 kD
HeLa SH4-Yes-Cherry	SH4-Yes-Cherry	28.9 kD
HeLa SH4-Yes-dRFP	SH4-Yes-dRFP	55.3 kD
HeLa SH4-HASPB-GFP / SH4-Yes-Cherry	SH4-HASPB-GFP SH4-Yes-Cherry	28.9 kD 28.9 kD
HeLa SH4-HASPB-GFP / SH4-Yes-Cherry #5	SH4-HASPB-GFP SH4-Yes-Cherry	28.9 kD 28.9 kD
HeLa SH4-Yes-GFP / SH4-HASPB-Cherry	SH4-Yes-GFP SH4-HASPB-Cherry	29.1 kD 28.7 kD
HeLa SH4-Yes-GFP / SH4-HASPB-Cherry #3	SH4-Yes-GFP SH4-HASPB-Cherry	29.1 kD 28.7 kD
Kyoto SH4-HASPB-GFP	SH4-HASPB-GFP	28.9 kD
Kyoto SH4-HASPB-GFP #6	SH4-HASPB-GFP	28.9 kD
Kyoto SH4-HASPB-Δmyr-GFP	SH4-HASPB-Δmyr-GFP	28.9 kD
Kyoto SH4-HASPB-Δpal-GFP	SH4-HASPB-Δpal-GFP	28.9 kD
Kyoto SH4-HASPB-Cherry	SH4-HASPB-Cherry	28.7 kD
Kyoto SH4-HASPB-dRFP	SH4-HASPB-dRFP	55.1 kD
Kyoto SH4Yes-GFP	SH4-Yes-GFP	29.1 kD
Kyoto SH4-Yes-GFP #7	SH4-Yes-GFP	29.1 kD
Kyoto SH4-Yes-Δmyr-GFP	SH4-Yes-Δmyr-GFP	29.1 kD
Kyoto SH4-Yes-Δpal-GFP	SH4-Yes-Δpal-GFP	29.1 kD
Kyoto SH4-Yes-Cherry	SH4-Yes-Cherry	28.9 kD
Kyoto SH4-Yes-dRFP	SH4-Yes-dRFP	55.3 kD
Kyoto SH4-Yes-GFP / SH4-HASPB-Cherry #3	SH4-Yes-GFP SH4-HASPB-Cherry	29.1 kD 28.7 kD
Kyoto SH4-Yes-GFP / SH4-HASPB-Cherry #5	SH4-Yes-GFP SH4-HASPB-Cherry	29.1 kD 28.7 kD

5.3 COMPREHENSIVE LIST OF GENES AND SIRNAS VALIDATED IN THE GENOME-WIDE RNAI SCREEN

Gene symbol	Gene Name	Gene ID	RefSeq ID	siRNA ID	sense siRNA sequence	SH4-HASPB-GFP	SH4-Yes-mCherry
A2M	alpha-2-macroglobulin	2	NM_000014	s819	CAGUGAUGGUUAAGAACGAtt	yes	yes
ACADVL	acyl-Coenzyme A dehydrogenase, very long chain	37	NM_000018	s208	GAUUGUCAAUGAACAGUUUtt	yes	yes
ACY3	aspartoacylase (aminocyclase) 3	91703	NM_080658	s40752	GCACCUGCUUAAUCGCGAAtt	yes	yes
ATP6AP2	ATPase, H+ transporting, lysosomal accessory protein 2	10159	NM_005765	s19790	GGUCUGUUGUUUUCCGAAAtt	no	yes
BAI3	brain-specific angiogenesis inhibitor 3	577	NM_001704	s1878	GAAUGAGCUUAGCAAUCCAtt	yes	no
C11orf38	chromosome 11 open reading frame 38	399967	NM_212555	s53217	GGCAGAGCAUGUAUCAUUAtt	yes	yes
C12orf36	chromosome 12 open reading frame 36	283422	NM_182558	s49233	GACACUUCAUGCUCCACUAtt	yes	no
C16orf71	chromosome 16 open reading frame 71	146562	NM_139170	s44876	AAACCUCCCUGAUUCCAGAtt	yes	no
CACNA1S	calcium channel, voltage-dependent, L type, alpha 1S subunit	779	NM_000069	s2297	GGACCCAAAUAAACCGGAAtt	no	yes
COPA	coatomer protein complex, subunit alpha	1314	NM_004371	s3369	GGAUGAGAAAACUCGCUUUtt	yes	no
COPB1	coatomer protein complex, subunit beta 1	1315	NM_016451	s3372	GGAGAUGUAAAGUCAAAGAtt	yes	no
CDH26	cadherin-like 26	60437	NM_021810	s34067	CAUACACUCUACUCCUGCUtt	yes	yes
CHERP	calcium homeostasis endoplasmic reticulum protein	10523	NM_006387	s20627	GCUCGGAGAAGAGAACAAAtt	yes	yes
DBH	dopamine beta-hydroxylase (dopamine beta-monooxygenase)	1621	NM_000787	s3947	GGAGAAAGGUGGUCACAGUtt	no	yes
EIF4A3	eukaryotic translation initiation factor 4A, isoform 3	9775	NM_014740	s18876 s18877	GGAUAUUCAGGUUCGUGAAtt CGAGCAAUCAAGCAGAUCAtt	yes yes	yes yes
ENSG00000176584	NULL	0	ENST00000388832	s225386	CCAUCAUGACUGUACCACAtt	yes	yes
FAM108A1	family with sequence similarity 108, member A1	81926	NM_001130111	s230886	ACGUGAAAACGGAAAUUAAtt	yes	no
FLJ21767	hypothetical protein FLJ21767	401331	XM_001132988	s196448	AAGCGCCCUCAUCAAGUUAtt	yes	no
GALNTL4	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-like 4	374378	NM_198516	s51536	CAUCAAAUAUGACAACUUUtt	yes	no
GBL	G protein beta subunit-like	64223	NM_022372	s34579	GCACAUCCGCAUGUAUGAUtt	yes	yes
hCG_1640785	hCG1640785	401206	XM_937380	s53508	GGACAAGCUCAAUAACUUAtt	yes	yes
HNRNPA1	heterogeneous nuclear ribonucleoprotein A1	3178	NM_002136	s6710 s6710	GAAUAAUGGUACCAGAUAAtt GAAUAAUGGUACCAGAUAAtt	no no	yes yes
IL7R	interleukin 7 receptor	3575	XM_001127146	s7325	GAUCCUAUCUUACUAACCAtt	yes	no
ITGAV	integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51)	3685	NM_002210	s7569	GGUCUUGAAGUGUACCCUAtt	yes	no
ITGB1	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	3688	NM_002211	s7575 s7576	GGAGGAAUGUUACACGGCUtt GCAGUUGGUUUUGCGAUUAtt	yes yes	yes yes
KIF20B	kinesin family member 20B	9585	NM_016195	s18420 s18421	GCAAGUUAAAGAAUAUCGAtt CAAACGUAUUAGUUCAGCAtt	yes yes	no no
LILRB5	leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 5	10990	NM_006840	s21637	AGAUAGACACUUUCUUUUUtt	yes	no
LOC387787	similar to CG9804-PA	387787	XM_370636	s226411	AGAUUGUCAUUUCUCAAUUtt	no	yes
LOC390688	CDC37-like	390688	NM_001080829	s52851	CACUGAGCGAUGAUGAAGAtt	no	yes
LOC441377	similar to ribosomal protein S26	441377	XM_496991	s54340	GGAGCUGAGUUCUUAAAGAtt	no	yes
LOC442160	similar to ribosomal protein L21	442160	XM_001128707	s54494	GUUGUAAACAAACAAGUUAtt	no	yes
LOC641768	similar to 40S ribosomal protein S26	641768	XM_935907	s54843	AGCUGUAUGUGAAGCUACAtt	yes	yes
LOC644191	similar to hCG15685	644191	XM_930072	s55596	CUUCCCAAACUGUAUGUGAtt	yes	yes

Gene symbol	Gene Name	Gene ID	RefSeq ID	siRNA ID	sense siRNA sequence	SH4-HASPB-GFP	SH4-Yes-mCherry
MAGEB1	melanoma antigen family B, 1	4112	NM_002363	s8455	CAUUCUACGUAAGUAUAAAtt	yes	no
MAK10	MAK10 homolog, amino-acid N-acetyltransferase subunit (S. cerevisiae)	60560	NM_024635	s34135	GGACGGCAAAGUACGUAAAtt	yes	yes
MBD3L2	methyl-CpG binding domain protein 3-like 2	125997	NM_144614	s42925	AGAAGAAACGAGAAAAUCCAtt	yes	no
				s42925	AGAAGAAACGAGAAAAUCCAtt	yes	no
MGC70863	similar to RPL23AP7 protein	284942	NM_203302	s49770	CCACCCUGAUUUGUCCUGAtt	yes	yes
MVD	mevalonate (diphospho) decarboxylase	4597	NM_002461	s224074	CACCAGGACCAGUUAAAAAtt	no	yes
NAE1	NEDD8 activating enzyme E1 subunit 1	8883	NM_001018159	s16975	CGAAUACCUAAAACGUAUAtt	yes	yes
OR10V1	olfactory receptor, family 10, subfamily V, member 1	390201	NM_001005324	s52738	CAUUCUCACCUGACCCUGAtt	no	yes
PGCP	plasma glutamate carboxypeptidase	10404	NM_016134	s20353	GCUUCUCAUGGGAUCAAAAtt	yes	no
PHF5A	PHD finger protein 5A	84844	NM_032758	s39506	UGUGAUUUGUGUGACUCCUAUtt	yes	yes
POLR2F	polymerase (RNA) II (DNA directed) polypeptide F	5435	NM_021974	s224297	CAUUCGCCGUUACCUGCCAtt	yes	yes
				s10809	CCAGAAGCGAAUCACCACAtt	no	yes
PRKCA	protein kinase C, alpha	5578	NM_002737	s11093	GCUCCACACUAAAUCCGCAtt	yes	yes
PRKCB1	protein kinase C, beta 1	5579	NM_212535	s11096	GGAUGAAACUGACCGAUUUtt	yes	yes
PRPF31	PRP31 pre-mRNA processing factor 31 homolog (S. cerevisiae)	26121	NM_015629	s25122	AGUUCAUCCGGGAUAAGUAtt	no	yes
PRPF8	PRP8 pre-mRNA processing factor 8 homolog (S. cerevisiae)	10594	NM_006445	s20796	CCCUACAUGUGAACAACGAtt	yes	yes
				s20797	GCAGAUACAUUGAUCGCAUtt	yes	yes
PTPRD	protein tyrosine phosphatase, receptor type, D	5789	NM_001040712	s11540	GAAUAGAUACGCGAAUGUAtt	yes	no
RAN	RAN, member RAS oncogene family	5901	NM_006325	s11768	CGAGAGUUACUUACAAGAAtt	no	yes
				s11769	GGAUAUUAAGGACAGGAAAtt	no	yes
RPL18	ribosomal protein L18	6141	NM_000979	s194744	GGCUGUUGGUCAAGUUAUAtt	no	yes
				s194745	AAACUAACCCUGGAUCCUAtt	no	yes
RPL21	ribosomal protein L21	6144	NM_000982	s12186	CACUCUAAGAGCCGAGAUAtt	no	yes
				s12187	GGAAGAGUCUACAAUGUUAtt	no	yes
SF3A1	splicing factor 3a, subunit 1, 120kDa	10291	NM_001005409	s20117	CACCAGGUCUGGAUAUUGAtt	no	yes
SHC1	SHC (Src homology 2 domain containing) transforming protein 1	6464	NM_003029	s12811	CAUGCAAUCUAUCUCAUUUtt	yes	yes
SMPDL3B	sphingomyelin phosphodiesterase, acid-like 3B	27293	NM_014474	s26101	GCAUCAUCGCGUCAUAGCAtt	yes	no
SOBP	sine oculis binding protein homolog (Drosophila)	55084	NM_018013	s30136	GAAUAACUGCGAACUACUUtt	yes	yes
SREBF2	sterol regulatory element binding transcription factor 2	6721	NM_004599	s28	GGAUGAUGCAAAGGUCAAAtt	yes	no
TGFBR1	transforming growth factor, beta receptor I (activin A receptor type II-like kinase, 53kDa)	7046	NM_004612	s14073	GGUACUACGUUGAAAGACUtt	no	yes
TMEM126A	transmembrane protein 126A	84233	NM_032273	s38695	CAUCUUAAGUUACUGGAUUtt	no	yes
YAP1	Yes-associated protein 1, 65kDa	10413	NM_006106	s20367	AGAGAUACUUCUUAAAUCAtt	yes	no
ZNF326	zinc finger protein 326	284695	NM_182975	s195972	UAUCAGGGCUUUAAUGGAAtt	no	yes

5.4 COMPREHENSIVE LIST OF SIRNAS USED FOR FURTHER ANALYSES

Gene Symbol	Full Gene Name	Gene ID	siRNA ID	sense siRNA sequence	antisense siRNA sequence
COPA	coatomer protein complex, subunit alpha	1314	s3369	GGAUGAGAAAACUCGCUUUtt	AAAGCGAGUUUUCUCAUCCtt
COPA	coatomer protein complex, subunit alpha	1314	14600	GGCAUUGACUUCCAUAAGCtt	GCUUAUGGAAGUCAAUGCCtc
COPA	coatomer protein complex, subunit alpha	1314	146404	CCUUGGAUCCUGACUAGUUtt	AACUAGUCAGGAUCCAAGGtc
COPB1	coatomer protein complex, subunit beta 1	1315	24546	GGAUCUUCAACAUCCUAAUtt	AUUAGGAUGUUGAAGAUCCtt
COPB1	coatomer protein complex, subunit beta 1	1315	s3371	GGUCUGUCAUGCUAAUCCAtt	UGGAUUAGCAUGACAGACCtt
COPB1	coatomer protein complex, subunit beta 1	1315	s3372	GGAGAUGUAAAGUCAAAGAtt	UCUUUGACUUUACAUCUCCtt
COPB2	coatomer protein complex, subunit beta 2 (beta prime)	9276	137699	GCGGAUGACAUGCAGAUUAtt	UAAUCUGCAUGUCAUCCGCtc
COPB2	coatomer protein complex, subunit beta 2 (beta prime)	9276	13770	GCCUUUCAUUCUAACUAGCtt	GCUAGUUAGAAUGAAAGGCtg
COPD/ARCN1	coatomer protein complex, subunit delta	372	10234	GGGCUUAUUAGCAGCUUUUtt	AAAAGCUGCUAAUAAGCCCtc
COPD/ARCN1	coatomer protein complex, subunit delta	372	10330	GGCUCUUCUCAAGAGUGAUtt	AUCACUCUUGAGAAGAGCCtt
GFP	green fluorescent protein		4626	CAAGCUGACCCUGAAGUUCtt	GAACUUCAGGGUCAGCUUGtt
INCENP	inner centromere protein antigens 135/155kDa	3619	28431	GGAGAAGAAGAAGCAGAUUtt	AAUCUGCUUCUUCUUCUCCtc
INCENP	inner centromere protein antigens 135/155kDa	3619	s7424	AGUCCUUUAUUUAAGCGCAAtt	UUGCGCUUAAUAAAGGACUtc
(scrambled)	Silencer Negative Control #1		103860	AGUACUGCUUACGAUACGGtt	CCGUAUCGUAAGCAGUACUtt
(scrambled)	Silencer Select Negative Control #1 siRNA (SiSel_NC1)		s813	UAACGACGCGACGACGUAAtt	UUACGUCGUCGCGUCGUUAtt

5.5 "AUTOMATED CELL ANALYSIS TOOL FOR A GENOME-WIDE RNAI SCREEN" (REMMELE ET AL., 2008)

Automated Cell Analysis Tool for a genome-wide RNAi screen

Steffen Remmele[1,3], Julia Ritzerfeld[2], Walter Nickel[2], Jürgen Hesser[1,3,4]

[1]Institute for Computational Medicine, University of Heidelberg, Germany
[2]Heidelberg University Biochemistry Centre, Germany
[3]Bioquant, University of Heidelberg, Germany
[4]Experimental Radiooncology, Medical Centre Mannheim, University of Heidelberg, Germany

Abstract— The amount of data in high-throughput microscopy can often exceed the range of Terabytes to even Petabytes and thus manual analysis is impossible. In this paper, a new robust and accurate image analysis tool is presented that automatically processes data of an RNAi-based screen. The screen investigates intracellular trafficking and targeting of acylated Src kinases. The main issue of the image processing task is the automatic detection and classification of cells from different phenotypes; which is non-trivial. Furthermore, the tool is brightness invariant allowing for processing cells of different staining quality, and it provides a quality control that copes with typical defects during preparation and acquisition. Technically, we further develop the standard region growing-based segmentation to one that uses a rule-based classification as stopping criterion and thus allows for a flexible use on different staining of cells. The tool has been applied to an RNAi-screen with approximately 1 Terabyte image data. The positive control of the screen is validated to 97,5% manually and by 91,25% with our software tool.

Index Terms — high-throughput fluorescense microscopy, automated analysis, phenotype recognition

I. Introduction

An important task in biomedical science is the decryption of the biological functions of human genes; which are mostly unknown. The identification of the genes and their corresponding functions are important for fundamental research as well as for the development of pharmaceutic agents. It is possible to reduce the expression of target genes by RNA interference (RNAi). RNAi was discoverd by Fire and Mello [1] and with this screening method all known genes can be silenced separately and existing functions of the genes can be observed. For a complete and systematic analysis for a specific biological function a genome-wide high-throughput RNAi screen has to be done. Such a screen requires a robust and automated methodology like that described in [2]. Besides, it provides a huge amount of data to be analysed; which renders a manual analysis impossible. Therefore, suitable tools are required. On one hand side, these tools have similar structure and functionalities like e.g. cell segmentation. On the other hand side, however, since each RNAi screen investigates a specific biological function, the visual representation of the cells is different; which leads to problem-specific segmentation techniques. In addition, the subsequent classification is problem specific, too.

The analysis tool we present in this paper was developed for a genome-wide RNAi screen investigating intracellular trafficking and targeting of acylated Src kinases. For this specific investigation it is necessary to differentiate between three phenotypes: plasma membrane, cytoplasm and Golgi. Figure 1 shows a confocal image of these three different phenotypes. In fact, there are different intensities within the Golgi class and for the purpose of a more accurate analysis an additonal class named Golgi strong was added. The analysed screen includes 149 LabTeks in duplicates with 384 spots per LabTek. For each spot an image dataset containing three channels is acquired. The Hoechst channel contains the cell nuclei while the acutal cells in appearance of the three phenotypes can be seen in both the GFP and the mCherry channel. These two channels are the result of the simultaneous expression of two different Src kinases labeled with GFP and mCherry and have to be analysed separately.

Corresponding author: Steffen Remmele, email: steffen.remmele@ziti.uni-heidelberg.de

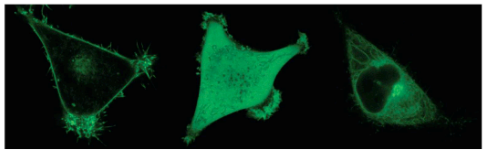

Fig. 1. confocal image of phenotypes (left to right: plasma membrane, cytoplasm, Golgi)

This work is organized as follows: In Sec. II, a short summary of the state of the art of cell segmentation and automated analysis tools is given. In Sec. III our analysis tool is described in detail while Sec. IV contains an overview about the experiments used to evaluate the quality of the tool and the results of this evaluation. A summary and discussion concludes the paper in Sec. V.

II. State of the art

Various approaches for cell nuclei or whole cell segmentation have been published in recent years. Fernandez et al. [3] focussed on cell nuclei segmentaion and tried to detect directly neighbouring cell nuclei in order to separate them. In this approach, dominant or concave points on the binary contour of the region are detected and considered to be points where the nuclei can be split. In [4] Lezoray and Cardot described an active contour technique for the segmentation of cell nuclei. A whole cell segmentation using region growing and adaptive segmentation methods was presented by Anoraganingrum [5]. Wu et al. [6] used a two step method

for coarse and fine segmentation which applied hierarchical thresholding for whole cell segmentation. In [7] the segmentation is done by a thresholding method and adaptive fuzzy c-means clustering. Alternatively, there are several approches using a watershed algorithm for cell nuclei or whole cell segmentation (see [8] [9] and references therein). In addition Wählby [8] is trying to separate cell aggregates by merging or deleting small regions in a second step following the watershed segmentation. The approach of Metzler et al. [10] is able to separate touching cells, too. In this multi-scale approach, mathematical morphology operators are used. The cell segmentation in [11] is realized by a neural network approach but for this approach the cell contours have to be emphasied. Elter et al. [12] present a three-step segmentation called maximum-intensity-linking closely related to classic watershed methods. Here, the neighborhood of each pixel is evaluated; which requires a smoothing preprocessing step in order to be robust against noise. In addition, a postprocessing step merges oversegmented regions.

Harder et al. [13] presented an automated image analysis method focussing on cell nuclei. The method segments, tracks and classifies cell nuclei automatically and distinguishes the nuclei into different mitotic phases. In this approach, multi-cell 3D images are analysed while a 2D projection is used for segmentation and tracking. The segmentation of the cell nuclei is based on a local adaptive thresholding approach and the classification is using a Support Vector Machine. In [14] an analysis tool for high-throughput microscopy is described. The underlying RNAi screen investigates genes involved in Hepatitis C and Dengue virus replication. In a first step the cell nuclei are segmented by a gradient thresholding and in the following step the infection level is determined by investigating the neighbourhood of the cell nuclei. In order to determine the neighbourhood of the nuclei, different approaches from dilation to region growing were tested.

III. Methods

The input data in our analysis tool consists of three-channel microscopy images. These images are multi-cell 2D fluorescence images where one channel is dedicated for the cell nuclei and the other two channels are assigned for the cell signals. One of these three-channel images is shown in figure 2. In order to perform an analysis for the genome-wide RNAi screen, it is necessary to count the number of cells in each image and to differentiate the phenotypes shown in the GFP and the mCherry channel. Our tool processes many LabTeks at once by a batch process; which is controlled by a Matlab GUI. The results of the analysis are stored in two text files, one for the GFP and one for the mCherry results. This way, the results can be easily imported into a spreadsheet application for further evaluation.

A. Analysis algorithm

For the analysis of each three-channel image, the strategy sketched in figure 3 is taken.

I. First, the analysis tool reads the three-channel image which is stored in three different files.

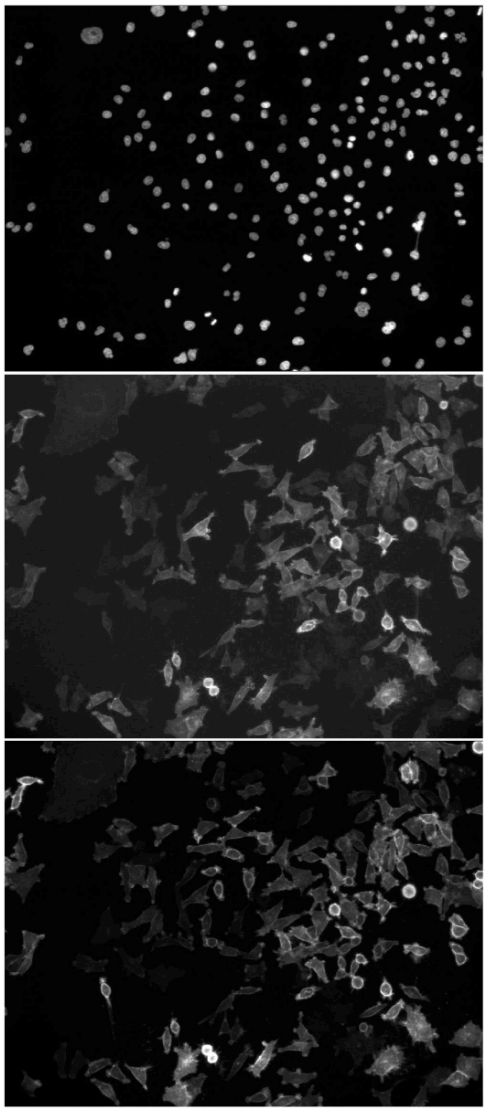

Fig. 2. Cell image with Hoechst(top), GFP(middle) and mCherry(bottom) channel

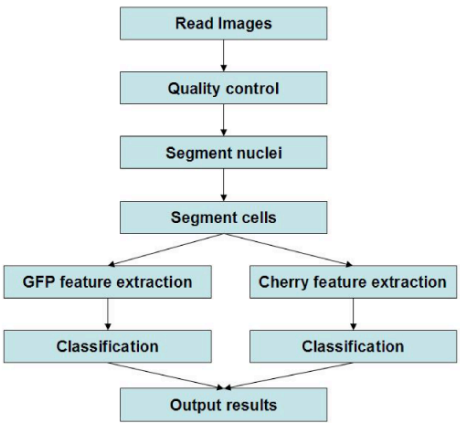

Fig. 3. Image analysis strategy

II. Next, a quality control of the three channels is performed by analyzing the quality of each channel separately. We consider three reasons for image degradation, out-of-focus, noise and local contamination. Hereby, image degradation by noise in the image is coped by our robust technique. Furthermore, we consider areas of the images corrupted by contaminations. These degradations are only local and do not influence the image quality in general. As our analysis shows, it is no problem to detect and skip such areas later. Finally, the focus of the microscope is determined by an auto focus functionality that is not always reliable. To measure this degradation, the image channels are transformed in the Fourier space and Out-of-focus images are, hereby, detected by their lacking in high frequency bands. In general, we characterize the images by four quality levels. The quality rating of each image channel is added to the analysis result and in this way a weighting of the results is possible.

III. The segmentation of the cells is the third step and this functionality of the algorithm is essential for an accurate classification. The task is to segment all cells accurately irrespective of the cell's phenotype. Since the brightness of the cells can vary in the GFP and mCherry channel, the segmentation has to be invariant to fluorescence brightness as well. Thus, the algorithm is realized in two steps: first the cell nuclei are determined in the Hoechst channel being the base for the cell segmentation in the second step. A description of the segmentation is given in detail in the next section.

IV: In the next step the classification is performed on the basis of the segmentation result. After segmenting the cells, the number of cells in the image is known and it is now necessary to determine each cell's phenotype. There are two separate classifications operating in the same way, one for the GFP channel and one for the mCherry channel. Initially, the segmentation result for each cell is verified and afterwards a feature extraction for each cell is performed. On basis of these features the cell is classified and assigned to the appropriate phenotype class. If the cell does not fit in one of three phenotype classes or the segmentation for this cell cannot be used for a classification, the cell will be marked as invalid. In this way, it can be assured that only those cells will be considered in the result of the analysis where it is certain that the classification is valid. The section classification contains the details for the feature extraction and the decision functionality.

V: Finally, it is necessary to output the results in a way useful for further evaluation of the RNAi screen. The output has to contain the complete result of the classification together with the quality report. Furthermore, it is indispensable to know from which spot of which LabTek the results are coming from. In order to assure an easy identification the file name of the image given by the microscope is used with the number of the LabTek. Then, all analysis results are listed and the quality measurements for both the Hoechst channel and the image channel containing the cell signal (GFP or mCherry) are added. The results are stored in text files and an import into a spreadsheet application is possible.

B. Segmentation

As mentioned above, the challenge of segmentation is to achieve a highly accurate result despite several limitations in the image acquisition step; which are

- low resolution in widefield images,
- out-of-focus images due to failures of auto focus,
- noise,
- varying cell brightness (non-uniform fluorscense),
- local contaminations.

In addition, the cell segmentation should be able to cope with different phenotypes shown in figure 1.

Under these circumstances, approaches using only edge or gradient information would be useful for a cytoplasm cell or a well-defined plasma membrane cell but fail for a Golgi cell, especially due to many internal structures found. In addition, for plasma membrane cells, one cannot assume that the membrane is always displayed as a closed structure and thus internal and external areas are not distinguishable. For similar reasons, watershed transforms used in many other algorithms for cell segmentation are not successful here.

Therefore, we concentrate on a new informed region based technique. As a starting point we use the Hoechst channel unambiguously displaying the nuclei of cells, being independent of the phenotype:

Since the intensity values of the cell nuclei in the Hoechst channel are within a certain range, threshold techniques are applied. An evaluation of several Hoechst images shows that the brightness in this image channel can vary as well and that a fixed threshold will not lead to a satisfactory segmentation for all Hoechst images. For this reason, an adaptive thresholding is applied: Starting with a high value, the threshold is decreased until an "appropriate result" is achieved. For each threshold parameter the segmentation generates connected regions of

a certain size. An appropriate result is reached when the average size of these regions correspond to the assumed size of the cell nuclei. The threshold is decreased with a variable step-size becoming smaller according to the proximity of the current segmentation to the assumed result. In case of holes in the nucleus regions, they are filled in a postprocessing step. Hereby, non-segmented pixels within a connected region of the nuclei are assigned the segmented cell nuclei as well.

In addition, local contaminations have to be detected and excluded from the segmentation result. The main indicator for a contamination is the size of the segmented cell nuclei: Since a cell nucleus has a certain maximal size, segmented regions being larger than this size are thus considered as contaminations and excluded consequently. In this way, an artifact-free segmentation result for the cell nuclei can be obtained.

The following cell segmentation process can either operate on the GFP or the mCherry channel. Both image channels contain the same cells and thus the segmentation can be realized on both channels. To verify this behavior, segmentation tests are run and both channels give identical results.

As basic technique for the segmentation we use a classifier-enhanced region growing. It requires a seed point and a stopping criterion. The latter is realized by a classifier, in our case a rule-based classifier. For seed points we use the cell nuclei region segmented from the Hoechst channel and in addition this region serves as source for information used in the classifier. Let the minimal local brightness in the nucleus region be t_1. A stopping criterion is fulfilled, if both following rules are maintained. If one rule is violated the region growing is not stopped.

- Rule 1: The brightness of a new pixel must be below $t_1 - \varepsilon$, where $\varepsilon > 0$ is a suitable threshold.
- Rule 2: The gradient magnitude of a new pixel has to be above a threshold $\mathfrak{z}$.

This classifier-enhanced region growing approach is suitable for all three phenotypes and, in addition, it is fluorescence brightness invariant. The approach is successful for bright and dark cells since the local brightness is considered in the segmentation. The classifier allows the segmentation of all three phenotypes: The cytoplasm cell can be segmented because (a) inside the cell there are only low gradient magnitudes and therefore rule 2 is violated, and (b) at the border of the cell lower intensity values occur and there is a larger brightness difference between inner and outer pixels. The Golgi can be segmented since, despite of high gradient magnitudes inside the cells, the brightness of the pixels within the cell is beyond the threshold $t_1 - \varepsilon$ of rule 1. Finally, the plasma membrane cells possess low gradient values inside the cells too and the borders of the cells are recognized correctly even it the membrane is not displayed as a closed structure.

In fact, the segmented cells have to be inspected before they can be classified. Local contaminations or weak and noisy intensity cell signals can cause infeasible segmentation results. These results are detected by evaluating the size and the shape of the segmented cells and should thus be excluded. In addition, a valid classification result cannot be guaranteed if a cell is too close to the edge of the image. Besides, if two cell nuclei are too close to each other it is likely that the cells overlap and a segmentation is not possible in this case. Again, by excluding these cells from the segmentation output, we achieve results free of artifacts.

C. Classification

The classification for the GFP channel is independent from the mCherry channel and vice versa. These channels contain images of the same cells and the segmentation result is valid for both of them while the feature extraction as well as the decision functionality are completely separated. Both channels are simultaneous expressions of two different protein kinases and contain the same phenotypes. Thus, the same features can be used for the classification. In addition, a suitable decision functionality for one channel is applied to the other channel as well.

The basis of the classification is the segmentation result. This result is used to divide the cell into complementary parts: nucleus, boundary, and interior (cytoplasm). The nucleus of the cell is already characterized by the segmentation result of the Hoechst channel. In order to determine the cell boundary, a modified morphological gradient (difference between original and eroded image) is applied. Finally, the interior is the difference between cell region on one side, and boundary unified with nucleus on the other side. For each cell and its constituent parts the mean value and standard deviation of the corresponding pixels are feature candidates for the phenotype classification. During our tests, we found the following features being relevant:

- standard deviation of cell,
- mean value of nucleus,
- mean value of boundary,
- mean value ofi nterior.

It is important for the phenotype classification to be invariant to the fluorescence brightness of the different cells. Therefore, the decision functionality is defined using mostly proportions between the different extracted features. As already mentioned in Sec I, for the Golgi phenotype an subtype Golgi strong was added for the purpose of a more accurate evaluation of the results. This subtype contains cells with a highly developed Golgi phenotype and therefore the definition for this class is more restricted.

The classification of the phenotypes is performed by a decision functionality displayed as a decision tree in figure 4. In rare cases, there are cells changing their phenotype when the image is acquired. Such cells do not fit into one of the phenotype classes and they are thus marked as invalid. The rules for the decision functionality are the result of an intensive analysis of the extracted features for all phenotypes. This analysis contains the evaluation of more than three thousand cells of different phenotypes.

IV. Results

We have evaluated our cell analysis tool by comparing the results of the analysis tool with those manually acquired by

results for all phenotypes are shown in the figures 7, 8 and 9. In all three figures the original data is displayed in the top row. On the right side the Hoechst channel can be seen which is used as basis for the cell nuclei segmentation while the GFP channel used for the cell segmentation is shown on the left side. In the bottom row the segmentation result is presented on the left side and the regions for nucleus, boundary and interior (cytoplasm) which are the basis for the classification are marked with different gray values. On the right side the stopping criterion for the region growing is illustrated. This image contains information about the regions where the gradient magnitude of the GFP image is above t_2 (rule 2) as well as information about the region where the intensity values of the cells are too low according to the minimal local brightness s_1 of the cells (rule 1). The region growing is stopped when both rules are fulfilled and these areas are marked as red. The yellow regions present the areas where the gradient magnitude is above t_2 and therefore rule 2 is fulfilled while rule 1 is still violated. The areas where the intensity values are low enough that rule 1 is fulfilled and rule 2 is violated arecyan. Finally, in the blue areas both rules are violated.

The segmentation result for all phenotypes is as expected. The segmentation for the cytoplasm cells in figure 7 is dominated by theblue areas in the centre of the cells where both rules necessary for the stopping criterion are violated. The actual edges of the cells are located where the intensity values of the GFP channel fall below the threshold determined by the local brightness of the cell (rule 1). Besides, the gradient magnitude is above t_2 and therefore rule 2 is already fulfilled in the yellow regions next to the edges of the cell. In the Golgi cells displayed in figure 8 there are also some areas where both rules are violated(blue regions) but in many areas at least one rule is fulfilled. Especially in the lower cell it can be seen that both rules are necessary for the segmentation of this phenotype. It would not be possible to segment such cells by using only one of the two criteria, the intensity values of the cells or the gradient magnitude. Due to the internal structures of the Golgi cells the gradient magnitude is above t_2 in many areas inside the cells and therefore rule 2 is fulfilled here (yellow regions). In addition, in the upper right part of the cell, the intensity values are already too low according to the local brightness of the cell and rule 1 is fulfilled while the gradient magnitude is still too low and rule 2 is violated(cyan regions). The collaboration of both rules allows the accurate segmentation of this phenotype. In fact, both rules are also necessary to segment the plasma membrane cell (figure 9). There are cyan and yellow regions inside of the cell where one of both rules is fulfilled although thecyan regions are not as large as in the Golgi cell in figure 8. In summary, the classifier-enhanced region growing approach is able to segment all phenotypes properly and, in addition, it is invariant to the fluorescence brightness of the segmented cells.

The classification for the cells in the figures 7, 8 and 9 is correct. All phenotypes are detected as expected. In addition, an expert verified the classification by evaluating the result for about a thousand cells manually.

There is a positive control in the RNAi screen on 80

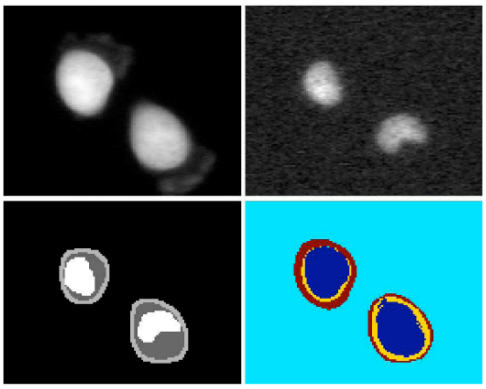

Fig. 7. Cytoplasm cell segmentation. Original images are shown in the top row (GFP on the left side and Hoechst on the right side). In the bottom row the different parts of the segmented cells are displayed on the left side while an image of the stopping criterion for the region growing can be seen on the right side.

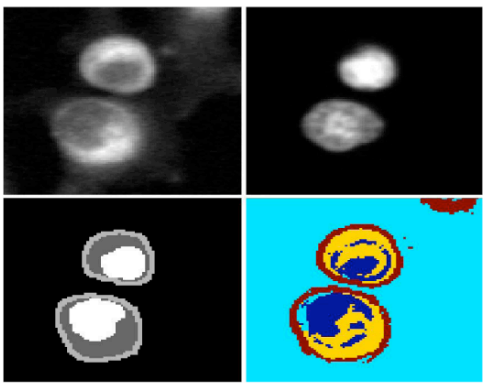

Fig. 8. Golgi cell segmentation. Top row: original images (GFP left and Hoechst right). Bottom row: Sections of the segmented cell(left) and stopping criterion for region growing(right)

LabTeks that is used to evaluate the cell analysis tool too. Both the results of a manual evaluation of the positive control and the results of the cell analysis tool are listed in table I.

The software tool recognizes 93,6% of the positive control being rated as valid by a manual analysis of an expert. In this case the interpretation of the results from the software tool is done by evaluating the distribution of the different phenotypes in the images. In addition the manual analysis of an expert is influenced by the location of the cells and therefore by the local distriution of the different phenotypes. In this way an image with a local conglomeration of a specific phenotype can be rated as valid in the manual analysis even if the total appearence of this phenotype is not dominating. This information is not considered in the software tool. Despite

results for all phenotypes are shown in the figures 7, 8 and 9. In all three figures the original data is displayed in the top row. On the right side the Hoechst channel can be seen which is used as basis for the cell nuclei segmentation while the GFP channel used for the cell segmentation is shown on the left side. In the bottom row the segmentation result is presented on the left side and the regions for nucleus, boundary and interior (cytoplasm) which are the basis for the classification are marked with different gray values. On the right side the stopping criterion for the region growing is illustrated. This image contains information about the regions where the gradient magnitude of the GFP image is above t_2 (rule 2) as well as information about the region where the intensity values of the cells are too low according to the minimal local brightness t_1 of the cells (rule 1). The region growing is stopped when both rules are fulfilled and these areas are marked as red. The yellow regions present the areas where the gradient magnitude is above t_2 and therefore rule 2 is fulfilled while rule 1 is still violated. The areas where the intensity values are low enough that rule 1 is fulfilled and rule 2 is violated arecyan. Finally, in the blue areas both rules are violated.

The segmentation result for all phenotypes is as expected. The segmentation for the cytoplasm cells in figure 7 is dominated by theblue areas in the centre of the cells where both rules necessary for the stopping criterion are violated. The actual edges of the cells are located where the intensity values of the GFP channel fall below the threshold determined by the local brightness of the cell (rule 1). Besides, the gradient magnitude is above t_2 and therefore rule 2 is already fulfilled in the yellow regions next to the edges of the cell. In the Golgi cells displayed in figure 8 there are also some areas where both rules are violated(blue regions) but in many areas at least one rule is fulfilled. Especially in the lower cell it can be seen that both rules are necessary for the segmentation of this phenotype. It would not be possible to segment such cells by using only one of the two criteria, the intensity values of the cells or the gradient magnitude. Due to the internal structures of the Golgi cells the gradient magnitude is above t_2 in many areas inside the cells and therefore rule 2 is fulfilled here (yellow regions). In addition, in the upper right part of the cell, the intensity values are already too low according to the local brightness of the cell and rule 1 is fulfilled while the gradient magnitude is still too low and rule 2 is violated(cyan regions). The collaboration of both rules allows the accurate segmentation of this phenotype. In fact, both rules are also necessary to segment the plasma membrane cell (figure 9). There are cyan and yellow regions inside of the cell where one of both rules is fulfilled although thecyan regions are not as large as in the Golgi cell in figure 8. In summary, the classifier-enhanced region growing approach is able to segment all phenotypes properly and, in addition, it is invariant to the fluorescence brightness of the segmented cells.

The classification for the cells in the figures 7, 8 and 9 is correct. All phenotypes are detected as expected. In addition, an expert verified the classification by evaluating the result for about a thousand cells manually.

There is a positive control in the RNAi screen on 80

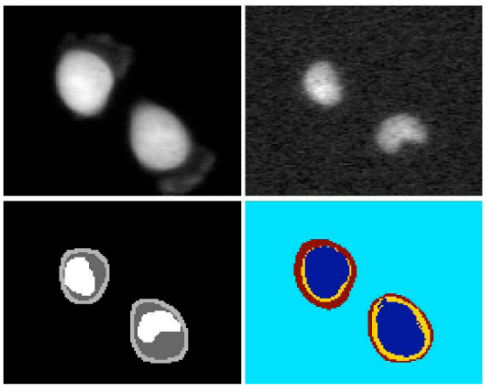

Fig. 7. Cytoplasm cell segmentation. Original images are shown in the top row (GFP on the left side and Hoechst on the right side). In the bottom row the different parts of the segmented cells are displayed on the left side while an image of the stopping criterion for the region growing can be seen on the right side.

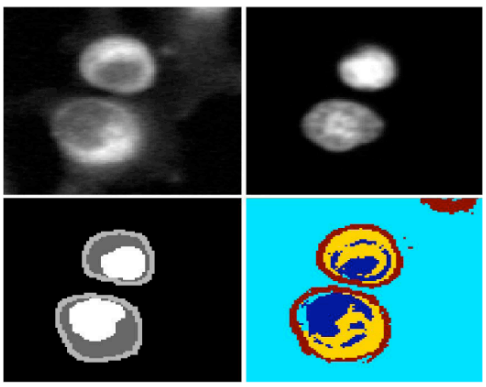

Fig. 8. Golgi cell segmentation. Top row: original images (GFP left and Hoechst right). Bottom row: Sections of the segmented cell(left) and stopping criterion for region growing(right)

LabTeks that is used to evaluate the cell analysis tool too. Both the results of a manual evaluation of the positive control and the results of the cell analysis tool are listed in table I.

The software tool recognizes 93,6% of the positive control being rated as valid by a manual analysis of an expert. In this case the interpretation of the results from the software tool is done by evaluating the distribution of the different phenotypes in the images. In addition the manual analysis of an expert is influenced by the location of the cells and therefore by the local distriution of the different phenotypes. In this way an image with a local conglomeration of a specific phenotype can be rated as valid in the manual analysis even if the total appearence of this phenotype is not dominating. This information is not considered in the software tool. Despite

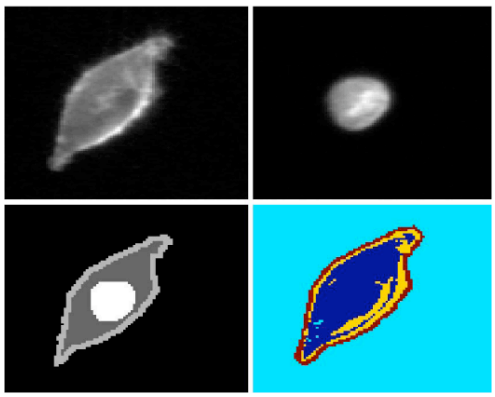

Fig. 9. Plasma membrane cell segmentation. Top row: original images (GFP left and Hoechst right). Bottom row: Sections of the segmented cell(left) and stopping criterion for region growing(right)

	amount off ound positive controls	propotion (in percentage)
manual recognition	78	97,5%
software tool	73	91,25%

TABLE I

EVALUATION OF POSITIVE CONTROL WITH 80 LAB TEK

this limitation the result of the cell analysis tool is satisfying compared with a manual analysis.

V. Conclusion

We have presented an automated cell analysis tool designed for an RNAi screen investigating intracellular trafficking and targeting of acylated Src kinases. Our analysis tool segments the images of the screen using an classifier-enhanced region growing. This approach yields a reliable segmentation of different phenotypes and is especially invariant to fluorescence brightness of the segmented cells. Using features from complementary parts of the cells a rule based classification is applied. Furthermore, the quality of the images is analyzed and added to the analysis result. Both the positive control included in the RNAi screen and the extensive manual analysis by an expert prove the reliability and the robustness of our analysis tool. Our software is well suited for the analysis of the genome-wide RNAi screen and the results allow a complete evaluation of the screen.

References

[1] Fire A., Xu S., Montgomery M., Kostas S., Driver S., and Mello C.C., "Potent and specific genetic interference by double-stranded rna in caenorhabditis elegans," in Nature, 1998, vol. 391, pp. 806–811.

[2] Erfle Holger, Simpson Jeremy, Bastiaens Philippe, and Pepperkok Rainer, "sirna cell arrays for high-content screening microscopy," in Biotechnology 2004, vol. 37, pp. 454–462.

[3] G. Fernandez, Murat Kunt, and J.-P. Zryd, "A new plant cell image segmentation algorithm," in ICIAP , 1995, pp. 229–234.

[4] V. Meas-Yedid, Florence Cloppet, A. Roumier, A. Alcover, J-C Olivo-Marin, and G. Stamon, "Quantitative microscopic image analysis by active contours," in IAPR-Vision Interface , 2001, pp. 277–284.

[5] Dwi Anoraganingrum, "Cell image segmentation: Global vs. local adaptive segmentation," in SSM Kassel , 1999.

[6] Kenong Wu, D. Gauthier, and M.D. Levine, "Live cell image segmentation," in Biomedical Engineering, IEEE Transactions on , 1995, vol. 42, pp. 1–12.

[7] Tuan D. Pham, Denis I. Crane, Tuan H. Tran, and Tam H. Nguyen, "Extraction of fluorescent cell puncta by adaptive fuzzy segmentation," in Bioinformatics,, 2004, vol. 20, pp. 2189–2196.

[8] Carolina Waehlby, Joakim Lindblad, Mikael Vondrus, Ewert Bengtsson, and Lennart Bjrksten, "Algorithm for cytoplasm segmentation of fluorescence labelled cells," in Analytical Cellular Pathology, 2002, vol. 24, pp. 101–111.

[9] Olivier Lezoray and Hubert Cardot, "Cooperation of color pixel classification schemes and color watershed: a study for microscopical images," in IEEE Transactions on Image Processing , 2002, vol. 11,, pp. 783–789.

[10] Volker Metzler, Thomas Lehmann, Hans Bienert, Khosrow Mottaghy, and Klaus Spitzer, "Scale-independent shape analysis for quantitative cytology using mathematical morphology," in Computers in Medicine and Biology, 2000, pp. 135–151.

[11] Tim Nattkemper, Heiko Wersing, Walter Schubert, and Helge Ritter, "A neural network architecture for automatic segmentation of fluorescence micrographs," in ESANN , 2000, pp. 177–182.

[12] Matthias Elter, Volker Daum, and Thomas Wittenberg, "Maximum-intensity-linking for segmentation of fluorescence-stained cells," in Proc MIAAB , 2006, pp. 46–50.

[13] Nathalie Harder et al, "Automated analysis of the mitotic phases of human cells in 3d fluorescence microscopy image sequences," in MICCAI , 2006, pp. 840–848.

[14] Petr Matula et al, "Automated analysis of rnai screens of virus infected cells based on immunofluorescence microscopy," in Proc BVM , 2008, pp. 453–457.

6 REFERENCES

Alland, L., Peseckis, S.M., Atherton, R.E., Berthiaume, L., and Resh, M.D. (1994). Dual myristylation and palmitylation of Src family member p59fyn affects subcellular localization. J Biol Chem *269*, 16701-16705.

Antonny, B., and Schekman, R. (2001). ER export: public transportation by the COPII coach. Curr Opin Cell Biol *13*, 438-443.

Aoyama, T., Uchida, Y., Kelley, R.I., Marble, M., Hofman, K., Tonsgard, J.H., Rhead, W.J., and Hashimoto, T. (1993). A novel disease with deficiency of mitochondrial very-long-chain acyl-CoA dehydrogenase. Biochem Biophys Res Commun *191*, 1369-1372.

Arreaza, G., Melkonian, K.A., LaFevre-Bernt, M., and Brown, D.A. (1994). Triton X-100-resistant membrane complexes from cultured kidney epithelial cells contain the Src family protein tyrosine kinase p62yes. J Biol Chem *269*, 19123-19127.

Backhaus, R., Zehe, C., Wegehingel, S., Kehlenbach, A., Schwappach, B., and Nickel, W. (2004). Unconventional protein secretion: membrane translocation of FGF-2 does not require protein unfolding. J Cell Sci *117*, 1727-1736.

Bano, M.C., Jackson, C.S., and Magee, A.I. (1998). Pseudo-enzymatic S-acylation of a myristoylated yes protein tyrosine kinase peptide in vitro may reflect non-enzymatic S-acylation in vivo. Biochem J *330 (Pt 2)*, 723-731.

Barlowe, C. (2000). Traffic COPs of the early secretory pathway. Traffic *1*, 371-377.

Barlowe, C. (2002). COPII-dependent transport from the endoplasmic reticulum. Curr Opin Cell Biol *14*, 417-422.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.

Bethune, J., Wieland, F., and Moelleken, J. (2006). COPI-mediated transport. J Membr Biol *211*, 65-79.

Bijlmakers, M.J., and Marsh, M. (1999). Trafficking of an acylated cytosolic protein: newly synthesized p56(lck) travels to the plasma membrane via the exocytic pathway. J Cell Biol *145*, 457-468.

Bijlmakers, M.J., and Marsh, M. (2003). The on-off story of protein palmitoylation. Trends Cell Biol *13*, 32-42.

Blenis, J., and Resh, M.D. (1993). Subcellular localization specified by protein acylation and phosphorylation. Curr Opin Cell Biol *5*, 984-989.

Bohnsack, R.N., and Haas, A.L. (2003). Conservation in the mechanism of Nedd8 activation by the human AppBp1-Uba3 heterodimer. J Biol Chem *278*, 26823-26830.

Bonanno, J.B., Edo, C., Eswar, N., Pieper, U., Romanowski, M.J., Ilyin, V., Gerchman, S.E., Kycia, H., Studier, F.W., Sali, A.*, et al.* (2001). Structural genomics of enzymes involved in sterol/isoprenoid biosynthesis. Proc Natl Acad Sci U S A *98*, 12896-12901.

Bonatti, S., Migliaccio, G., and Simons, K. (1989). Palmitylation of viral membrane glycoproteins takes place after exit from the endoplasmic reticulum. J Biol Chem *264*, 12590-12595.

Bonfanti, L., Mironov, A.A., Jr., Martinez-Menarguez, J.A., Martella, O., Fusella, A., Baldassarre, M., Buccione, R., Geuze, H.J., Mironov, A.A., and Luini, A. (1998). Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. Cell *95*, 993-1003.

Boutros, M., Kiger, A.A., Armknecht, S., Kerr, K., Hild, M., Koch, B., Haas, S.A., Paro, R., and Perrimon, N. (2004). Genome-wide RNAi analysis of growth and viability in Drosophila cells. Science *303*, 832-835.

Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Xavier, R.J., Lieberman, J., and Elledge, S.J. (2008). Identification of host proteins required for HIV infection through a functional genomic screen. Science *319*, 921-926.

Bretscher, M.S. (1973). Membrane structure: some general principles. Science *181*, 622-629.

Bretscher, M.S., and Munro, S. (1993). Cholesterol and the Golgi apparatus. Science *261*, 1280-1281.

Brown, D.A., and London, E. (1998). Functions of lipid rafts in biological membranes. Annu Rev Cell Dev Biol *14*, 111-136.

Chen, Y., Liu, W., McPhie, D.L., Hassinger, L., and Neve, R.L. (2003). APP-BP1 mediates APP-induced apoptosis and DNA synthesis and is increased in Alzheimer's disease brain. J Cell Biol *163*, 27-33.

Chen, Y., McPhie, D.L., Hirschberg, J., and Neve, R.L. (2000). The amyloid precursor protein-binding protein APP-BP1 drives the cell cycle through the S-M checkpoint and causes apoptosis in neurons. J Biol Chem *275*, 8929-8935.

Chow, N., Korenberg, J.R., Chen, X.N., and Neve, R.L. (1996). APP-BP1, a novel protein that binds to the carboxyl-terminal region of the amyloid precursor protein. J Biol Chem *271*, 11339-11346.

Corsi, A.K., and Schekman, R. (1996). Mechanism of polypeptide translocation into the endoplasmic reticulum. J Biol Chem *271*, 30299-30302.

Das, A.K., Dasgupta, B., Bhattacharya, R., and Basu, J. (1997). Purification and biochemical characterization of a protein-palmitoyl acyltransferase from human erythrocytes. J Biol Chem *272*, 11021-11025.

De Matteis, M.A., and Luini, A. (2008). Exiting the Golgi complex. Nat Rev Mol Cell Biol *9*, 273-284.

Denny, P.W., Gokool, S., Russell, D.G., Field, M.C., and Smith, D.F. (2000). Acylation-dependent protein export in Leishmania. J Biol Chem *275*, 11017-11025.

Doherty, G.J., and McMahon, H.T. (2009). Mechanisms of endocytosis. Annu Rev Biochem *78*, 857-902.

Duncan, J.A., and Gilman, A.G. (1996). Autoacylation of G protein alpha subunits. J Biol Chem *271*, 23594-23600.

Dunphy, J.T., Schroeder, H., Leventis, R., Greentree, W.K., Knudsen, J.K., Silvius, J.R., and Linder, M.E. (2000). Differential effects of acyl-CoA binding protein on enzymatic and non-enzymatic thioacylation of protein and peptide substrates. Biochim Biophys Acta *1485*, 185-198.

Ellgaard, L., and Helenius, A. (2001). ER quality control: towards an understanding at the molecular level. Curr Opin Cell Biol *13*, 431-437.

Ellgaard, L., and Helenius, A. (2003). Quality control in the endoplasmic reticulum. Nat Rev Mol Cell Biol *4*, 181-191.

Erfle, H., Neumann, B., Liebel, U., Rogers, P., Held, M., Walter, T., Ellenberg, J., and Pepperkok, R. (2007). Reverse transfection on cell arrays for high content screening microscopy. Nat Protoc *2*, 392-399.

Erfle, H., Neumann, B., Rogers, P., Bulkescher, J., Ellenberg, J., and Pepperkok, R. (2008). Work flow for multiplexing siRNA assays by solid-phase reverse transfection in multiwell plates. J Biomol Screen *13*, 575-580.

Erfle, H., and Pepperkok, R. (2007). Production of siRNA- and cDNA-transfected cell arrays on noncoated chambered coverglass for high-content screening microscopy in living cells. Methods Mol Biol *360*, 155-161.

Erfle, H., Simpson, J.C., Bastiaens, P.I., and Pepperkok, R. (2004). siRNA cell arrays for high-content screening microscopy. Biotechniques *37*, 454-458, 460, 462.

Farazi, T.A., Waksman, G., and Gordon, J.I. (2001). The biology and enzymology of protein N-myristoylation. J Biol Chem *276*, 39501-39504.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Fishburn, C.S., Herzmark, P., Morales, J., and Bourne, H.R. (1999). Gbetagamma and palmitate target newly synthesized Galphaz to the plasma membrane. J Biol Chem *274*, 18793-18800.

Flinn, H.M., Rangarajan, D., and Smith, D.F. (1994). Expression of a hydrophilic surface protein in infective stages of Leishmania major. Mol Biochem Parasitol *65*, 259-270.

Fromme, J.C., and Schekman, R. (2005). COPII-coated vesicles: flexible enough for large cargo? Curr Opin Cell Biol *17*, 345-352.

Glick, B.S. (2000). Organization of the Golgi apparatus. Curr Opin Cell Biol *12*, 450-456.

Gonczy, P., Echeverri, C., Oegema, K., Coulson, A., Jones, S.J., Copley, R.R., Duperon, J., Oegema, J., Brehm, M., Cassin, E*., et al.* (2000). Functional genomic analysis of cell division in C. elegans using RNAi of genes on chromosome III. Nature *408*, 331-336.

Gong, L., and Yeh, E.T. (1999). Identification of the activating and conjugating enzymes of the NEDD8 conjugation pathway. J Biol Chem *274*, 12036-12042.

Gonzalo, S., and Linder, M.E. (1998). SNAP-25 palmitoylation and plasma membrane targeting require a functional secretory pathway. Mol Biol Cell *9*, 585-597.

Grant, B.D., and Donaldson, J.G. (2009). Pathways and mechanisms of endocytic recycling. Nat Rev Mol Cell Biol *10*, 597-608.

Grassme, H., Gulbins, E., Brenner, B., Ferlinz, K., Sandhoff, K., Harzer, K., Lang, F., and Meyer, T.F. (1997). Acidic sphingomyelinase mediates entry of N. gonorrhoeae into nonphagocytic cells. Cell *91*, 605-615.

Grassme, H., Jekle, A., Riehle, A., Schwarz, H., Berger, J., Sandhoff, K., Kolesnick, R., and Gulbins, E. (2001a). CD95 signaling via ceramide-rich membrane rafts. J Biol Chem *276*, 20589-20596.

Grassme, H., Jendrossek, V., Bock, J., Riehle, A., and Gulbins, E. (2002). Ceramide-rich membrane rafts mediate CD40 clustering. J Immunol *168*, 298-307.

Grassme, H., Schwarz, H., and Gulbins, E. (2001b). Molecular mechanisms of ceramide-mediated CD95 clustering. Biochem Biophys Res Commun *284*, 1016-1030.

Gulbins, E., and Kolesnick, R. (2003). Raft ceramide in molecular medicine. Oncogene *22*, 7070-7077.

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature *404*, 293-296.

Hansen, C.G., and Nichols, B.J. (2009). Molecular mechanisms of clathrin-independent endocytosis. J Cell Sci *122*, 1713-1721.

Hansra, G., Garcia-Paramio, P., Prevostel, C., Whelan, R.D., Bornancin, F., and Parker, P.J. (1999). Multisite dephosphorylation and desensitization of conventional protein kinase C isotypes. Biochem J *342 (Pt 2)*, 337-344.

Haun, R.S., Tsai, S.C., Adamik, R., Moss, J., and Vaughan, M. (1993). Effect of myristoylation on GTP-dependent binding of ADP-ribosylation factor to Golgi. J Biol Chem *268*, 7064-7068.

Hogenboom, S., Romeijn, G.J., Houten, S.M., Baes, M., Wanders, R.J., and Waterham, H.R. (2002). Absence of functional peroxisomes does not lead to deficiency of enzymes involved in cholesterol biosynthesis. J Lipid Res *43*, 90-98.

Hogenboom, S., Wanders, R.J., and Waterham, H.R. (2003). Cholesterol biosynthesis is not defective in peroxisome biogenesis defective fibroblasts. Mol Genet Metab *80*, 290-295.

Holmgren, A. (1989). Thioredoxin and glutaredoxin systems. J Biol Chem *264*, 13963-13966.

Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834-838.

Hutvagner, G., and Zamore, P.D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. Science *297*, 2056-2060.

Ingley, E. (2008). Src family kinases: regulation of their activities, levels and identification of new pathways. Biochim Biophys Acta *1784*, 56-65.

Izai, K., Uchida, Y., Orii, T., Yamamoto, S., and Hashimoto, T. (1992). Novel fatty acid beta-oxidation enzymes in rat liver mitochondria. I. Purification and properties of very-long-chain acyl-coenzyme A dehydrogenase. J Biol Chem *267*, 1027-1033.

Jackson, C.L. (2009). Mechanisms of transport through the Golgi complex. J Cell Sci *122*, 443-452.

Johnson, D.R., Bhatnagar, R.S., Knoll, L.J., and Gordon, J.I. (1994). Genetic and biochemical studies of protein N-myristoylation. Annu Rev Biochem *63*, 869-914.

Kaiser, C., and Ferro-Novick, S. (1998). Transport from the endoplasmic reticulum to the Golgi. Curr Opin Cell Biol *10*, 477-482.

Kaplan, K.B., Swedlow, J.R., Varmus, H.E., and Morgan, D.O. (1992). Association of p60c-src with endosomal membranes in mammalian fibroblasts. J Cell Biol *118*, 321-333.

Kasahara, K., Nakayama, Y., Kihara, A., Matsuda, D., Ikeda, K., Kuga, T., Fukumoto, Y., Igarashi, Y., and Yamaguchi, N. (2007). Rapid trafficking of c-Src, a non-palmitoylated Src-family kinase, between the plasma membrane and late endosomes/lysosomes. Exp Cell Res *313*, 2651-2666.

Kawakami, T., Kawakami, Y., and Kitaura, J. (2002). Protein kinase C beta (PKC beta): normal functions and diseases. J Biochem *132*, 677-682.

Kittler, R., Putz, G., Pelletier, L., Poser, I., Heninger, A.K., Drechsel, D., Fischer, S., Konstantinova, I., Habermann, B., Grabner, H.*, et al.* (2004). An endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division. Nature *432*, 1036-1040.

Kleizen, B., and Braakman, I. (2004). Protein folding and quality control in the endoplasmic reticulum. Curr Opin Cell Biol *16*, 343-349.

Klemm, R.W., Ejsing, C.S., Surma, M.A., Kaiser, H.J., Gerl, M.J., Sampaio, J.L., de Robillard, Q., Ferguson, C., Proszynski, T.J., Shevchenko, A.*, et al.* (2009). Segregation of sphingolipids and sterols during formation of secretory vesicles at the trans-Golgi network. J Cell Biol *185*, 601-612.

Klumperman, J. (2000). Transport between ER and Golgi. Curr Opin Cell Biol *12*, 445-449.

Koch, C.A., Anderson, D., Moran, M.F., Ellis, C., and Pawson, T. (1991). SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. Science *252*, 668-674.

Koegl, M., Zlatkine, P., Ley, S.C., Courtneidge, S.A., and Magee, A.I. (1994). Palmitoylation of multiple Src-family kinases at a homologous N-terminal motif. Biochem J *303 (Pt 3)*, 749-753.

Konig, R., Zhou, Y., Elleder, D., Diamond, T.L., Bonamy, G.M., Irelan, J.T., Chiang, C.Y., Tu, B.P., De Jesus, P.D., Lilley, C.E.*, et al.* (2008). Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. Cell *135*, 49-60.

Krepkiy, D., and Miziorko, H.M. (2004). Identification of active site residues in mevalonate diphosphate decarboxylase: implications for a family of phosphotransferases. Protein Sci *13*, 1875-1881.

Krishnan, M.N., Ng, A., Sukumaran, B., Gilfoy, F.D., Uchil, P.D., Sultana, H., Brass, A.L., Adametz, R., Tsui, M., Qian, F.*, et al.* (2008). RNA interference screen for human genes associated with West Nile virus infection. Nature *455*, 242-245.

Kwong, J., and Lublin, D.M. (1995). Amino-terminal palmitate or polybasic domain can provide required second signal to myristate for membrane binding of p56lck. Biochem Biophys Res Commun *207*, 868-876.

Kypta, R.M., Su, H., and Reichardt, L.F. (1996). Association between a transmembrane protein tyrosine phosphatase and the cadherin-catenin complex. J Cell Biol *134*, 1519-1529.

Laifenfeld, D., Patzek, L.J., McPhie, D.L., Chen, Y., Levites, Y., Cataldo, A.M., and Neve, R.L. (2007). Rab5 mediates an amyloid precursor protein signaling pathway that leads to apoptosis. J Neurosci *27*, 7141-7153.

Le Borgne, R., and Hoflack, B. (1998a). Mechanisms of protein sorting and coat assembly: insights from the clathrin-coated vesicle pathway. Curr Opin Cell Biol *10*, 499-503.

Le Borgne, R., and Hoflack, B. (1998b). Protein transport from the secretory to the endocytic pathway in mammalian cells. Biochim Biophys Acta *1404*, 195-209.

Lee, M.R., Lee, D., Shin, S.K., Kim, Y.H., and Choi, C.Y. (2008). Inhibition of APP intracellular domain (AICD) transcriptional activity via covalent conjugation with Nedd8. Biochem Biophys Res Commun *366*, 976-981.

Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell *117*, 69-81.

Leventis, R., Juel, G., Knudsen, J.K., and Silvius, J.R. (1997). Acyl-CoA binding proteins inhibit the nonenzymic S-acylation of cysteinyl-containing peptide sequences by long-chain acyl-CoAs. Biochemistry *36*, 5546-5553.

Liang, X., Lu, Y., Wilkes, M., Neubert, T.A., and Resh, M.D. (2004). The N-terminal SH4 region of the Src family kinase Fyn is modified by methylation and heterogeneous fatty acylation: role in membrane targeting, cell adhesion, and spreading. J Biol Chem *279*, 8133-8139.

Liang, X., Nazarian, A., Erdjument-Bromage, H., Bornmann, W., Tempst, P., and Resh, M.D. (2001). Heterogeneous fatty acylation of Src family kinases with polyunsaturated fatty acids regulates raft localization and signal transduction. J Biol Chem *276*, 30987-30994.

Linder, M.E., and Deschenes, R.J. (2007). Palmitoylation: policing protein stability and traffic. Nat Rev Mol Cell Biol *8*, 74-84.

Lobo, S., Greentree, W.K., Linder, M.E., and Deschenes, R.J. (2002). Identification of a Ras palmitoyltransferase in Saccharomyces cerevisiae. J Biol Chem *277*, 41268-41273.

Luini, A., Ragnini-Wilson, A., Polishchuck, R.S., and De Matteis, M.A. (2005). Large pleiomorphic traffic intermediates in the secretory pathway. Curr Opin Cell Biol *17*, 353-361.

Magee, T., and Seabra, M.C. (2005). Fatty acylation and prenylation of proteins: what's hot in fat. Curr Opin Cell Biol *17*, 190-196.

Malsam, J.r., Satoh, A., Pelletier, L., and Warren, G. (2005). Golgin Tethers Define Subpopulations of COPI Vesicles. Science *307*, 1095-1098.

Mandon, E.C., Trueman, S.F., and Gilmore, R. (2009). Translocation of proteins through the Sec61 and SecYEG channels. Curr Opin Cell Biol *21*, 501-507.

Marchesini, N., and Hannun, Y.A. (2004). Acid and neutral sphingomyelinases: roles and mechanisms of regulation. Biochem Cell Biol *82*, 27-44.

Martin, G.S. (2001). The hunting of the Src. Nat Rev Mol Cell Biol *2*, 467-475.

Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell *110*, 563-574.

Martinez-Menarguez, J.A., Prekeris, R., Oorschot, V.M., Scheller, R., Slot, J.W., Geuze, H.J., and Klumperman, J. (2001). Peri-Golgi vesicles contain retrograde but not anterograde proteins consistent with the cisternal progression model of intra-Golgi transport. J Cell Biol *155*, 1213-1224.

Maxfield, F.R., and McGraw, T.E. (2004). Endocytic recycling. Nat Rev Mol Cell Biol *5*, 121-132.

Maxfield, F.R., and Tabas, I. (2005). Role of cholesterol and lipid organization in disease. Nature *438*, 612-621.

McAndrew, R.P., Wang, Y., Mohsen, A.W., He, M., Vockley, J., and Kim, J.J. (2008). Structural basis for substrate fatty acyl chain specificity: crystal structure of human very-long-chain acyl-CoA dehydrogenase. J Biol Chem *283*, 9435-9443.

McKean, P.G., Denny, P.W., Knuepfer, E., Keen, J.K., and Smith, D.F. (2001). Phenotypic changes associated with deletion and overexpression of a stage-regulated gene family in Leishmania. Cell Microbiol *3*, 511-523.

McTaggart, S.J. (2006). Isoprenylated proteins. Cell Mol Life Sci *63*, 255-267.

Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. Nature *431*, 343-349.

Michie, A.M., and Nakagawa, R. (2005). The link between PKCalpha regulation and cellular transformation. Immunol Lett *96*, 155-162.

Mitchell, D.A., Vasudevan, A., Linder, M.E., and Deschenes, R.J. (2006). Protein palmitoylation by a family of DHHC protein S-acyltransferases. J Lipid Res *47*, 1118-1127.

Mocellin, S., and Provenzano, M. (2004). RNA interference: learning gene knock-down from cell physiology. J Transl Med *2*, 39.

Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Genes Dev *16*, 720-728.

Munro, S. (2003). Lipid rafts: elusive or illusive? Cell *115*, 377-388.

Nakashima, S. (2002). Protein kinase C alpha (PKC alpha): regulation and biological function. J Biochem *132*, 669-675.

Neumann, B., Held, M., Liebel, U., Erfle, H., Rogers, P., Pepperkok, R., and Ellenberg, J. (2006). High-throughput RNAi screening by time-lapse imaging of live human cells. Nat Methods *3*, 385-390.

Ng, D.T., Brown, J.D., and Walter, P. (1996). Signal sequences specify the targeting route to the endoplasmic reticulum membrane. J Cell Biol *134*, 269-278.

Nickel, W. (2003). The mystery of nonclassical protein secretion. A current view on cargo proteins and potential export routes. Eur J Biochem *270*, 2109-2119.

Nickel, W. (2005). Unconventional secretory routes: direct protein export across the plasma membrane of mammalian cells. Traffic *6*, 607-614.

Nickel, W., and Rabouille, C. (2009). Mechanisms of regulated unconventional protein secretion. Nat Rev Mol Cell Biol *10*, 148-155.

Nickel, W., and Seedorf, M. (2008). Unconventional mechanisms of protein transport to the cell surface of eukaryotic cells. Annu Rev Cell Dev Biol *24*, 287-308.

Nishizuka, Y. (1992). Intracellular signaling by hydrolysis of phospholipids and activation of protein kinase C. Science *258*, 607-614.

Nosjean, O., Briolay, A., and Roux, B. (1997). Mammalian GPI proteins: sorting, membrane residence and functions. Biochim Biophys Acta *1331*, 153-186.

Nurminen, T.A., Holopainen, J.M., Zhao, H., and Kinnunen, P.K. (2002). Observation of topical catalysis by sphingomyelinase coupled to microspheres. J Am Chem Soc *124*, 12129-12134.

Ogg, S.C., and Walter, P. (1995). SRP samples nascent chains for the presence of signal sequences by interacting with ribosomes at a discrete step during translation elongation. Cell *81*, 1075-1084.

Ono, Y., Kikkawa, U., Ogita, K., Fujii, T., Kurokawa, T., Asaoka, Y., Sekiguchi, K., Ase, K., Igarashi, K., and Nishizuka, Y. (1987). Expression and properties of two types of protein kinase C: alternative splicing from a single gene. Science *236*, 1116-1120.

Osaka, F., Kawasaki, H., Aida, N., Saeki, M., Chiba, T., Kawashima, S., Tanaka, K., and Kato, S. (1998). A new NEDD8-ligating system for cullin-4A. Genes Dev *12*, 2263-2268.

Palade, G. (1975). Intracellular Aspects of the Process of Protein Synthesis. Science *189*, 867.

Paladino, S., Sarnataro, D., Pillich, R., Tivodar, S., Nitsch, L., and Zurzolo, C. (2004). Protein oligomerization modulates raft partitioning and apical sorting of GPI-anchored proteins. J Cell Biol *167*, 699-709.

Parton, R.G., and Richards, A.A. (2003). Lipid rafts and caveolae as portals for endocytosis: new insights and common mechanisms. Traffic *4*, 724-738.

Parton, R.G., and Simons, K. (2007). The multiple faces of caveolae. Nat Rev Mol Cell Biol *8*, 185-194.

Patterson, G.H., Hirschberg, K., Polishchuk, R.S., Gerlich, D., Phair, R.D., and Lippincott-Schwartz, J. (2008). Transport through the Golgi apparatus by rapid partitioning within a two-phase membrane system. Cell *133*, 1055-1067.

Paulick, M.G., and Bertozzi, C.R. (2008). The glycosylphosphatidylinositol anchor: a complex membrane-anchoring structure for proteins. Biochemistry *47*, 6991-7000.

Pelkmans, L., Fava, E., Grabner, H., Hannus, M., Habermann, B., Krausz, E., and Zerial, M. (2005). Genome-wide analysis of human kinases in clathrin- and caveolae/raft-mediated endocytosis. Nature *436*, 78-86.

Pelkmans, L., and Helenius, A. (2002). Endocytosis via caveolae. Traffic *3*, 311-320.

Pike, L.J. (2006). Rafts defined: a report on the Keystone Symposium on Lipid Rafts and Cell Function. J Lipid Res *47*, 1597-1598.

Polishchuk, R., Di Pentima, A., and Lippincott-Schwartz, J. (2004). Delivery of raft-associated, GPI-anchored proteins to the apical surface of polarized MDCK cells by a transcytotic pathway. Nat Cell Biol *6*, 297-307.

Pulido, R., Krueger, N.X., Serra-Pages, C., Saito, H., and Streuli, M. (1995). Molecular characterization of the human transmembrane protein-tyrosine phosphatase delta. Evidence for tissue-specific expression of alternative human transmembrane protein-tyrosine phosphatase delta isoforms. J Biol Chem *270*, 6722-6728.

Rabut, G., and Peter, M. (2008). Function and regulation of protein neddylation. 'Protein modifications: beyond the usual suspects' review series. EMBO Rep *9*, 969-976.

Rajendran, L., and Simons, K. (2005). Lipid rafts and membrane dynamics. J Cell Sci *118*, 1099-1102.

Rapoport, T.A. (1992). Transport of proteins across the endoplasmic reticulum membrane. Science *258*, 931-936.

Rapoport, T.A., Goder, V., Heinrich, S.U., and Matlack, K.E. (2004). Membrane-protein integration and the role of the translocation channel. Trends Cell Biol *14*, 568-575.

Resh, M.D. (1993). Interaction of tyrosine kinase oncoproteins with cellular membranes. Biochim Biophys Acta *1155*, 307-322.

Resh, M.D. (1994). Myristylation and palmitylation of Src family members: the fats of the matter. Cell *76*, 411-413.

Resh, M.D. (1999). Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. Biochim Biophys Acta *1451*, 1-16.

Resh, M.D. (2006a). Palmitoylation of ligands, receptors, and intracellular signaling molecules. Sci STKE *2006*, re14.

Resh, M.D. (2006b). Trafficking and signaling by fatty-acylated and prenylated proteins. Nat Chem Biol *2*, 584-590.

Rocks, O., Peyker, A., and Bastiaens, P.I. (2006). Spatio-temporal segregation of Ras signals: one ship, three anchors, many harbors. Curr Opin Cell Biol *18*, 351-357.

Rocks, O., Peyker, A., Kahms, M., Verveer, P.J., Koerner, C., Lumbierres, M., Kuhlmann, J., Waldmann, H., Wittinghofer, A., and Bastiaens, P.I. (2005). An acylation cycle regulates localization and activity of palmitoylated Ras isoforms. Science *307*, 1746-1752.

Roth, A.F., Feng, Y., Chen, L., and Davis, N.G. (2002). The yeast DHHC cysteine-rich domain protein Akr1p is a palmitoyl transferase. J Cell Biol *159*, 23-28.

Rothman, J.E. (1994). Mechanisms of intracellular protein transport. Nature *372*, 55-63.

Rothman, J.E., and Wieland, F.T. (1996). Protein sorting by transport vesicles. Science *272*, 227-234.

Rothwell, C., Lebreton, A., Young Ng, C., Lim, J.Y., Liu, W., Vasudevan, S., Labow, M., Gu, F., and Gaither, L.A. (2009). Cholesterol biosynthesis modulation regulates dengue viral replication. Virology *389*, 8-19.

Rubartelli, A., Bajetto, A., Allavena, G., Wollman, E., and Sitia, R. (1992). Secretion of thioredoxin by normal and neoplastic cells through a leaderless secretory pathway. J Biol Chem *267*, 24161-24164.

Rzymski, T., Grzmil, P., Meinhardt, A., Wolf, S., and Burfeind, P. (2008). PHF5A represents a bridge protein between splicing proteins and ATP-dependent helicases and is differentially expressed during mouse spermatogenesis. Cytogenet Genome Res *121*, 232-244.

Sandilands, E., Brunton, V.G., and Frame, M.C. (2007). The membrane targeting and spatial activation of Src, Yes and Fyn is influenced by palmitoylation and distinct RhoB/RhoD endosome requirements. J Cell Sci *120*, 2555-2564.

Sandilands, E., Cans, C., Fincham, V.J., Brunton, V.G., Mellor, H., Prendergast, G.C., Norman, J.C., Superti-Furga, G., and Frame, M.C. (2004). RhoB and actin polymerization coordinate Src activation with endosome-mediated delivery to the membrane. Dev Cell *7*, 855-869.

Sandilands, E., and Frame, M.C. (2008). Endosomal trafficking of Src tyrosine kinase. Trends Cell Biol *18*, 322-329.

Sato, I., Obata, Y., Kasahara, K., Nakayama, Y., Fukumoto, Y., Yamasaki, T., Yokoyama, K.K., Saito, T., and Yamaguchi, N. (2009). Differential trafficking of

Src, Lyn, Yes and Fyn is specified by the state of palmitoylation in the SH4 domain. J Cell Sci *122*, 965-975.

Schafer, T., Zentgraf, H., Zehe, C., Brugger, B., Bernhagen, J., and Nickel, W. (2004). Unconventional secretion of fibroblast growth factor 2 is mediated by direct translocation across the plasma membrane of mammalian cells. J Biol Chem *279*, 6244-6251.

Schissel, S.L., Jiang, X., Tweedie-Hardman, J., Jeong, T., Camejo, E.H., Najib, J., Rapp, J.H., Williams, K.J., and Tabas, I. (1998a). Secretory sphingomyelinase, a product of the acid sphingomyelinase gene, can hydrolyze atherogenic lipoproteins at neutral pH. Implications for atherosclerotic lesion development. J Biol Chem *273*, 2738-2746.

Schissel, S.L., Keesler, G.A., Schuchman, E.H., Williams, K.J., and Tabas, I. (1998b). The cellular trafficking and zinc dependence of secretory and lysosomal sphingomyelinase, two products of the acid sphingomyelinase gene. J Biol Chem *273*, 18250-18259.

Schuck, S., and Simons, K. (2004). Polarized sorting in epithelial cells: raft clustering and the biogenesis of the apical membrane. J Cell Sci *117*, 5955-5964.

Schuck, S., and Simons, K. (2006). Controversy fuels trafficking of GPI-anchored proteins. J Cell Biol *172*, 963-965.

Seelenmeyer, C., Wegehingel, S., Tews, I., Kunzler, M., Aebi, M., and Nickel, W. (2005). Cell surface counter receptors are essential components of the unconventional export machinery of galectin-1. J Cell Biol *171*, 373-381.

Sessions, O.M., Barrows, N.J., Souza-Neto, J.A., Robinson, T.J., Hershey, C.L., Rodgers, M.A., Ramirez, J.L., Dimopoulos, G., Yang, P.L., Pearson, J.L.*, et al.* (2009). Discovery of insect and human dengue virus host factors. Nature *458*, 1047-1050.

Shenoy-Scaria, A.M., Dietzen, D.J., Kwong, J., Link, D.C., and Lublin, D.M. (1994). Cysteine3 of Src family protein tyrosine kinase determines palmitoylation and localization in caveolae. J Cell Biol *126*, 353-363.

Shenoy-Scaria, A.M., Gauen, L.K., Kwong, J., Shaw, A.S., and Lublin, D.M. (1993). Palmitylation of an amino-terminal cysteine motif of protein tyrosine kinases p56lck and p59fyn mediates interaction with glycosyl-phosphatidylinositol-anchored proteins. Mol Cell Biol *13*, 6385-6392.

Shirai, Y., and Saito, N. (2002). Activation mechanisms of protein kinase C: maturation, catalytic activation, and targeting. J Biochem *132*, 663-668.

Silverman, L., and Resh, M.D. (1992). Lysine residues form an integral component of a novel NH2-terminal membrane targeting motif for myristylated pp60v-src. J Cell Biol *119*, 415-425.

Silverman, L., Sudol, M., and Resh, M.D. (1993). Members of the src family of nonreceptor tyrosine kinases share a common mechanism for membrane binding. Cell Growth Differ *4*, 475-482.

Simons, K., and Ikonen, E. (1997). Functional rafts in cell membranes. Nature *387*, 569-572.

Simons, K., and Toomre, D. (2000). Lipid rafts and signal transduction. Nat Rev Mol Cell Biol *1*, 31-39.

Simons, K., and van Meer, G. (1988). Lipid sorting in epithelial cells. Biochemistry *27*, 6197-6202.

Simons, K., and Wandinger-Ness, A. (1990). Polarized sorting in epithelia. Cell *62*, 207-210.

Simpson, J.C., Cetin, C., Erfle, H., Joggerst, B., Liebel, U., Ellenberg, J., and Pepperkok, R. (2007). An RNAi screening platform to identify secretion machinery in mammalian cells. J Biotechnol *129*, 352-365.

Singer, S.J., and Nicolson, G.L. (1972). The fluid mosaic model of the structure of cell membranes. Science *175*, 720-731.

Siniossoglou, S., Wimmer, C., Rieger, M., Doye, V., Tekotte, H., Weise, C., Emig, S., Segref, A., and Hurt, E.C. (1996). A novel complex of nucleoporins, which includes Sec13p and a Sec13p homolog, is essential for normal nuclear pores. Cell *84*, 265-275.

Siomi, H., and Siomi, M.C. (2009). On the road to reading the RNA-interference code. Nature *457*, 396-404.

Souri, M., Aoyama, T., Hoganson, G., and Hashimoto, T. (1998). Very-long-chain acyl-CoA dehydrogenase subunit assembles to the dimer form on mitochondrial inner membrane. FEBS Lett *426*, 187-190.

Stegmayer, C., Kehlenbach, A., Tournaviti, S., Wegehingel, S., Zehe, C., Denny, P., Smith, D.F., Schwappach, B., and Nickel, W. (2005). Direct transport across the plasma membrane of mammalian cells of Leishmania HASPB as revealed by a CHO export mutant. J Cell Sci *118*, 517-527.

Tabas, I. (1999). Secretory sphingomyelinase. Chem Phys Lipids *102*, 123-130.

Taniguchi, H., and Manenti, S. (1993). Interaction of myristoylated alanine-rich protein kinase C substrate (MARCKS) with membrane phospholipids. J Biol Chem *268*, 9960-9963.

Tanimura, A., Nezu, A., Morita, T., Hashimoto, N., and Tojyo, Y. (2002). Interplay between calcium, diacylglycerol, and phosphorylation in the spatial and temporal regulation of PKCalpha-GFP. J Biol Chem *277*, 29054-29062.

Thelen, M., Rosen, A., Nairn, A.C., and Aderem, A. (1991). Regulation by phosphorylation of reversible association of a myristoylated protein kinase C substrate with the plasma membrane. Nature *351*, 320-322.

Thomas, S.M., and Brugge, J.S. (1997). Cellular functions regulated by Src family kinases. Annu Rev Cell Dev Biol *13*, 513-609.

Timson Gauen, L.K., Linder, M.E., and Shaw, A.S. (1996). Multiple features of the p59fyn src homology 4 domain define a motif for immune-receptor tyrosine-based activation motif (ITAM) binding and for plasma membrane localization. J Cell Biol *133*, 1007-1015.

Torrado, L.C., Temmerman, K., Muller, H.M., Mayer, M.P., Seelenmeyer, C., Backhaus, R., and Nickel, W. (2009). An intrinsic quality-control mechanism ensures unconventional secretion of fibroblast growth factor 2 in a folded conformation. J Cell Sci *122*, 3322-3329.

Toth, M.J., and Huwyler, L. (1996). Molecular cloning and expression of the cDNAs encoding human and yeast mevalonate pyrophosphate decarboxylase. J Biol Chem *271*, 7895-7898.

Tournaviti, S., Hannemann, S., Terjung, S., Kitzing, T.M., Stegmayer, C., Ritzerfeld, J., Walther, P., Grosse, R., Nickel, W., and Fackler, O.T. (2007). SH4-domain-induced plasma membrane dynamization promotes bleb-associated cell motility. J Cell Sci *120*, 3820-3829.

Tournaviti, S., Pietro, E.S., Terjung, S., Schafmeier, T., Wegehingel, S., Ritzerfeld, J., Schulz, J., Smith, D.F., Pepperkok, R., and Nickel, W. (2009). Reversible phosphorylation as a molecular switch to regulate plasma membrane targeting of acylated SH4 domain proteins. Traffic *10*, 1047-1060.

Trappe, R., Ahmed, M., Glaser, B., Vogel, C., Tascou, S., Burfeind, P., and Engel, W. (2002a). Identification and characterization of a novel murine multigene family containing a PHD-finger-like motif. Biochem Biophys Res Commun *293*, 816-826.

Trappe, R., Schulze, E., Rzymski, T., Frode, S., and Engel, W. (2002b). The Caenorhabditis elegans ortholog of human PHF5a shows a muscle-specific expression domain and is essential for C. elegans morphogenetic development. Biochem Biophys Res Commun *297*, 1049-1057.

Tsai, B., Ye, Y., and Rapoport, T.A. (2002). Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. Nat Rev Mol Cell Biol *3*, 246-255.

Ungewickell, E.J., and Hinrichsen, L. (2007). Endocytosis: clathrin-mediated membrane budding. Curr Opin Cell Biol *19*, 417-425.

Vaillancourt, F.H., Pilote, L., Cartier, M., Lippens, J., Liuzzi, M., Bethell, R.C., Cordingley, M.G., and Kukolj, G. (2009). Identification of a lipid kinase as a host factor involved in hepatitis C virus RNA replication. Virology *387*, 5-10.

van't Hof, W., and Resh, M.D. (1997). Rapid plasma membrane anchoring of newly synthesized p59fyn: selective requirement for NH2-terminal myristoylation and palmitoylation at cysteine-3. J Cell Biol *136*, 1023-1035.

Walter, P., Gilmore, R., and Blobel, G. (1984). Protein translocation across the endoplasmic reticulum. Cell *38*, 5-8.

Warren, G., and Malhotra, V. (1998). The organisation of the Golgi apparatus. Curr Opin Cell Biol *10*, 493-498.

Wegehingel, S., Zehe, C., and Nickel, W. (2008). Rerouting of fibroblast growth factor 2 to the classical secretory pathway results in post-translational modifications that block binding to heparan sulfate proteoglycans. FEBS Lett *582*, 2387-2392.

Wieland, F., and Harter, C. (1999). Mechanisms of vesicle formation: insights from the COP system. Curr Opin Cell Biol *11*, 440-446.

Wieland, F.T., Gleason, M.L., Serafini, T.A., and Rothman, J.E. (1987). The rate of bulk flow from the endoplasmic reticulum to the cell surface. Cell *50*, 289-300.

Wolven, A., Okamura, H., Rosenblatt, Y., and Resh, M.D. (1997). Palmitoylation of p59fyn is reversible and sufficient for plasma membrane association. Mol Biol Cell *8*, 1159-1173.

Yang, H., and Tracey, K.J. (2005). High mobility group box 1 (HMGB1). Crit Care Med *33*, S472-474.

Yang, H., Wang, H., Czura, C.J., and Tracey, K.J. (2005). The cytokine activity of HMGB1. J Leukoc Biol *78*, 1-8.

Yasuda, K., Kosugi, A., Hayashi, F., Saitoh, S., Nagafuku, M., Mori, Y., Ogata, M., and Hamaoka, T. (2000). Serine 6 of Lck tyrosine kinase: a critical site for Lck myristoylation, membrane localization, and function in T lymphocytes. J Immunol *165*, 3226-3231.

Zacharias, D.A., Violin, J.D., Newton, A.C., and Tsien, R.Y. (2002). Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. Science *296*, 913-916.

Zehe, C., Engling, A., Wegehingel, S., Schafer, T., and Nickel, W. (2006). Cell-surface heparan sulfate proteoglycans are essential components of the unconventional export machinery of FGF-2. Proc Natl Acad Sci U S A *103*, 15479-15484.

ACKNOWLEDGEMENTS

For my mother and my father, in the hope that they would be proud of me.

Foremost, I am deeply grateful to Walter Nickel, who not only gave me the opportunity to work on such an interesting project in his lab, but who always supported me with scientific guidance, personal advice and a lot of encouragement whenever necessary. I learned a lot.

I would like to thank Prof. Dr. Herwig Brunner and the Peter-und-Traudl-Engelhorn-Stiftung for generous funding, as well as Agnes Speck and the Gleichstellungsbüro of the University for support during my maternity leave.

I absolutely owe to Steffen Remmele for a great collaboration: for his brilliant ideas and programming skills and all the time and patience he invested in this project. It was great fun and very illuminating.

I am deeply thankful to all people from the EMBL RNAi screening facility, especially Rainer Pepperkok, Faba Neumann and Jutta Bulkescher for granting their scientific and technical support, access to their equipment and, last but not least, made my time at their lab a funny and fruitful time. Without them, I would have got lost along the way.

I am indebted to Thomas Söllner and Blanche Schwappach for scientific advice during this thesis and for being referees and examiners.

A great thank you goes to all past and present members of the Nickel lab for a pleasant and funny working atmosphere, scientific input and interesting (non-scientific) lunch discussions: Lucía Cespón Torrado, Antje Ebert, Mareike Laußmann, Hans-Michael Müller, Claudia Seelenmeyer, Carolin Stegmayer, Julia Steringer, Koen Temmerman, Stella Tournaviti, Tao Wang, Sabine Wegehingel, Georg Weidmann and Sonja Zacherl. It was great fun working with you.

I would like to especially thank Koen Temmerman not only for his scientific and technical support and a great time in the lab but above all for always being there whenever he was needed and making my time in Heidelberg so much more enjoyable. Thank you, Annabel, for many lunches at EMBL and interesting girl's and mother's discussions.

I would like to appreciate my family: we've been through rough times and I am glad we are sticking together. Special thanks go to Hans and Gerlinde Malsam for their great emergency babysitting service.

Last, but most important, I want to thank Jörg and Theo for their mere presence in my life. They mean everything to me.