INAUGURAL - DISSERTATION

zur

Erlangung der Doktorwürde

der

Naturwissenschaftlich-Mathematischen

Gesamtfakultät

der

Ruprecht-Karls-Universität

Heidelberg

Vorgelegt von

MSc Biophysik Tautvydas Lisauskas

aus Vilnius, Litauen

Tag der mündlichen Prüfung: 31.01.2014

Thema

Application of high-content microscopy to investigate membrane trafficking

Gutachter: Prof. Dr. Jürgen Wolfrum

Prof. Dr. Mike Heilemann

Abbreviations

BFA	Brefeldin A
cDNA	complementary DNA
CFP	cyan fluorescent protein
DNA	deoxyribonucleic acid
dsRNA	double stranded RNA
dSTORM	direct stochastic optical reconstruction microscopy
FRET	fluorescence resonance energy transfer
GAP	GTPase-activating protein
GDP	guanosine diphosphate
GEF	GDP/GTP exchange factor
GFP	green fluorescent protein
GTP	guanosine triphosphate
YFP	yellow fluorescent protein
miRNA	micro RNA
mRNA	messenger RNA; protein coding RNA
NRK	normal rat kidney
PDI	protein disulphide isomerase
PFA	Paraformaldehyde
RAB	protein family name - RAS in brain
RAB GDI	RAB GDP dissociation inhibitor
RAS	protein family name related to an origin of the first observation - rat sarcoma
RISC	RNA silencing complex
RNA	ribonucleic acid
RNAi	RNA interference
siRNA	short interfering RNA
STED	stimulated emission depletion (microscopy)
STORM	stochastic optical reconstruction microscopy

Acknowledgement

Here I would like to acknowledge and express sincere gratitude to a number of people those who one or the other way supported me on the way to this thesis and hopefully will continue to do so on my new journeys.

First of all, I thank my parents and my senior brothers who have been helping me in various situations to find the way. I thank my best friend Simona for her support which can still be felt despite long distance and rare opportunities to meet.

I thank Dr. Ramūnas Valiokas and all members of his lab for introducing me to fascinating world of self-assembled biosensors and for an opportunity to work on my master thesis in his lab. Without Ramūnas input I would not come to Heidelberg.

I thank Dr. Haisen Ta and Dr. Konstantinos Lymperopoulos for friendly support when I most needed it.

I am grateful to PD. Dr. Dirk-Peter Herten and his lab members for collaboration and wise advises. The friendship with Arina Rybina was the most important experience for me in Heidelberg. Numerous scientific discussions on various topics and especially Arina's insistent proposal to apply the model of consequent chemical reactions on my experimental data had substantial impact on my work.

I am thankful to the all of my laboratory colleagues and especially to Andrius Serva and Dr. Manuel Gunkel for strong collaboration and valuable input into my work.

I am sincerely grateful to Dr. Vytautė Starkuvienė-Erfle and Dr. Holger Erfle for inviting me to their labs and granting me an opportunity to have a peek on membrane trafficking and RNA interference fields.

Finally, I thank Prof. Dr. Jürgen Wolfrum for an opportunity to work on my doctoral thesis and for his priceless advises and encouragement to reach for more.

Abstract

Membrane trafficking is an abstraction describing material transport within different organelles of the cell. It is one of the areas of interest in molecular and cell biology. The practical interest for the investigation of membrane trafficking is lying in its relation to an increasing list of diseases, such as cancer. Identification of genes involved in membrane trafficking and their regulation within the cell are in particular focus of interest.

Microscopy based observation of membrane trafficking events in the native or close to native context is the method of choice. It allows relating different stages of the whole process to specific cellular organelles with spatial and temporal resolution. It also has the advantage of better statistics over other population mean based methods, as results are concluded from the observation of a high number of individual cells. Despite the versatility of fluorescence microscopy, the application of any particular technique for cell biology requires more application layers such as fluorophores and labelling strategies, data acquisition and storage, digital image analysis, data interpretation in a biological context, and last but not least – the biological process of interest itself. Improvements of those application layers or even synergy of different microscopy techniques increases information content what can be extracted from a single experiment. Currently we call this approach as high-content microscopy.

During the time of my studies I was working on improving some of application layers as well as creating stronger link between them. This thesis describes my experimental and data analysis work which was done to address some of the specific membrane trafficking related cases in context of high-content microscopy:

a) The organization kinetics of the Golgi complex, which is the major membrane trafficking organelle, and the influence of over-expressed membrane proteins on it. Live cell fluorescence screening microscopy was employed to address this question.
b) Statistical data analysis of a cargo protein secretion screen. Influence of specific RAB GTPases over-expression on collagen-I secretion was evaluated.
c) The development of a comprehensive fluorescence microscopy based platform to investigate messenger RNA (mRNA) and micro RNA (miRNA) interactions in live cells.

Kurzfassung

Der Membrantransport ist eine Abstraktion, welche den Materialtransport innerhalb verschiedener Organellen der Zelle beschreibt. Es ist eines der Interessensgebiete der Molekularbiologie. Das Interesse an der Untersuchung des Membrantransports begründet sich durch die wachsende Liste von Krankheiten, die mit ihm in Beziehung stehen, wie beispielsweise Krebs. Deshalb ist die Identifizierung von Genen, die an Membrantransport und deren Regulierung von größter Wichtigkeit.

Abläufe des Membrantransports in nativem oder nahezu nativem Kontext werden bevorzugt mikroskopisch beobachtet, da hierdurch ein räumlicher und zeitlicher Zusammenhang zwischen verschiedenen Stufen des gesamten Prozesses und spezifischen Zellorganellen hergestellt werden kann. Außerdem kann eine bessere Statistik erzielt werden, als in Methoden, welche auf einem Populationsmittelwert beruhen, da die Ergebnisse auf der direkten Beobachtung einer großen Anzahl individueller Zellen beruhen. Trotz der Vielseitigkeit der Fluoreszenzmikroskopie benötigt die Anwendung einer speziellen Technologie im Bereich der Zellbiologie mehrere zusätzliche Anwendungsebenen. Einige zusätzliche Anwendungsebenen sind: Fluorophore und Markierungsstrategien, Datenakquisition und -speicherung, digitale Bildverarbeitung, Dateninterpretation im biologischen Kontext und - nicht zuletzt – der zu untersuchende biologische Prozess selbst. Verbesserungen dieser Anwendungsschichten oder sogar Synergie der unterschiedlichen Mikroskopie-Techniken erhöht den Informationsgehalt gegenüber den Einzel-Experimenten. Derzeit nennen wir diesen Ansatz als High-Content-Mikroskopie.

Während der Zeit meiner Doktorarbeit arbeitete ich an der Verbesserung der Anwendungsschichten sowie der Etablierung und Verbesserung der Anbindung zwischen den verschiedenen Systemen. Diese Dissertation beschreibt die von mir durchgeführten experimentellen Arbeiten und deren Datenanalyse, um folgende spezifischen Fälle im Zusammenhang mit dem Membrantransport zu untersuchen und auszuwerten:

a) Der Einfluss überexprimierter Membranproteine auf die Organisationskinetik des Golgi Apparates, welcher eines der Hauptorganellen des Membrantransportes ist, durch Lebendzell-Fluoreszenzmikroskopie.
b) Statistische Datenanalyse eines Transportprotein-Sekretions-Screens. Der Einfluss der Überexpression spezifischer RAB GTPasen auf die Kollagen-I-Sekretion wurde bestimmt.
c) Die Entwicklung einer umfassenden Fluoreszenzmikroskopie Plattform zur Untersuchung der Messenger-RNA (mRNA) und Mikro-RNA (miRNA) Interaktionen in lebenden Zellen.

Table of Contents

Introduction

This thesis is focused on describing my work that had been done at the laboratories of dr. Vytaute Starkuviene (Screening of cellular networks) and dr. Holger Erfle (ViroQuant-Cell Networks RNAi Screening facility). The main interest of my study was investigation of membrane trafficking by application of fluorescence microscopy techniques. As stated in both laboratory names, a screening is the main approach of study. The screening approach is defined as being a test or examination to discover if there is anything wrong with someone[1] or the testing of a person or group of people for the presence of a disease or other condition[2]. The main task of the screening approach is to find conditions of interest via testing the effect of wide range of different conditions rather than describe them in great details.

To perform a biological screening one has to take control on the experimental conditions, establish evaluation method and then make the parallelization of the experiment to increase the throughput of the data acquisition and minimize effects of the unknown (or not controlled) conditions. Throughput could be increased as well by increasing speed of sequential read-out, but in case of biological experiments, there time scale of the experiment including preparation stage is in the range of days, it is not suitable.

Membrane trafficking is very wide abstraction and one could not investigate all the questions related to it in frame of three year study. My work was directed toward development of fluorescence microscopy based biological applications and analysis approaches of some specific questions as evaluation of gene importance on secretion of collagen-I or maintenance of the Golgi complex, and evaluation of post-transcriptional gene expression regulation by micro RNAs.

Generally speaking, the work described in this thesis was carried out by me while I was working on three joint collaboration research projects related to the topics of membrane trafficking:

Project 1 - **Enhanced-throughput screening microscopy development**. Main task of the project was to improve data acquisition throughput of current screening microscope. My objective was to develop an application for fast screening microscope, naming: to develop experimental

routine; perform experiment; develop data analysis end evaluation routine. Investigation of the organization kinetics of the Golgi complex was chosen due to its major role in membrane trafficking as well as practical aspects like not complicated and reproducible triggering of morphology changes by a number of chemicals.

Project 2 - **Role of RAB proteins in collagen secretion model system**. Main task of the project was to identify members of the RAB family proteins involved in the collagen-I secretion pathway. My objectives in the project were to analyse experimental data and develop evaluation routine.

Project 3 - **Integrated High-content fluorescence microscopy platform to study miRNA targets in live cells**. Main task of the project is to develop integrated fluorescence microscopy platform suitable to perform high-throughput screening, high resolution and super resolution (above diffraction limit) imaging of biological samples. My objective in the project was to develop experiment routine suitable for imaging on the integrated microscopy platform; investigate mRNA labelling techniques; investigate siRNA/miRNA labelling techniques; develop fluorescence based method to evaluate RNAi efficiency. The interest in miRNAs is due to their role in regulation of gene expression, thus interfering or potentially controlling various cellular processes like membrane trafficking.

All named projects were linked towards improvement of application of fluorescence microscopy techniques for investigation of biological questions. Once membrane trafficking is in concern, a microscopy based observation in the native or close to native context is the method of choice. It allows observe different stages of the whole process, to specific cellular organelles with spatial and temporal resolution. It also has the advantage of better statistics over other population mean based methods, as results are concluded from the observation of a high number of individual cells.

Introduction into molecular biology of the cell

The molecular biology of the cell is a growing set of concepts and knowledge being developed while studying various cells and organisms with the focus kept on molecular structures and their function in cellular processes. Currently it is accepted to date the beginning of the molecular biology of the cell as a discipline with proposal of the model for deoxyribonucleic (DNA) formation into dual helix by J.D. Watson and F.H.C. Crick in 1953[3]. Molecular biology of the cell is built on so called doctrine of the triad, developed and published by F.H.C. Crick in 1958[4,5], central dogma of which postulates the direction of information transfer, there information is understood as sequence of the DNA, ribonucleic acid (RNA) or amino acid residues. Once information got into the sequence of amino acid residues it cannot be transferred any more. While central dogma was supposed to model information transfer directionality, for long time the function of DNA or RNA was understood just as intermediate state for protein translation and protein non-coding DNA or RNA were implied as "junk". Yet, in recent years, functional activities of non-coding DNAs and RNAs have been discovered. An example of non-coding RNA function is post-transcriptional protein translation regulation by a class of small RNA molecules, known as micro RNAs.

The cell is basic structural unit of any living organism. Organisms can be classified as unicellular or multicellular; for example human body consists of about 10^{14} cells[6]. Dimensions of most of animal cells are in the range from 10 to 100 µm, it makes visual investigation of single cells impossible without tools such as microscopes are. In my study I have had used eukaryotic cells so everywhere will be spoken about model of eukaryotic cells, unless stated differently.

In Figure 1 a simple model representation of trafficking related organelles of the eukaryotic cells is shown. The eukaryotic cell is a three dimensional enclosed space containing intracellular organelles, functional compartments formed by intracellular membranes and separate from cytosol, which are responsible for one or more functions required for cell functioning. Membranes that comprise intracellular organelles contain specialized domains.

Each organelle must have mechanism for importing, and incorporating into the organelle, the specific proteins that make the organelle unique.

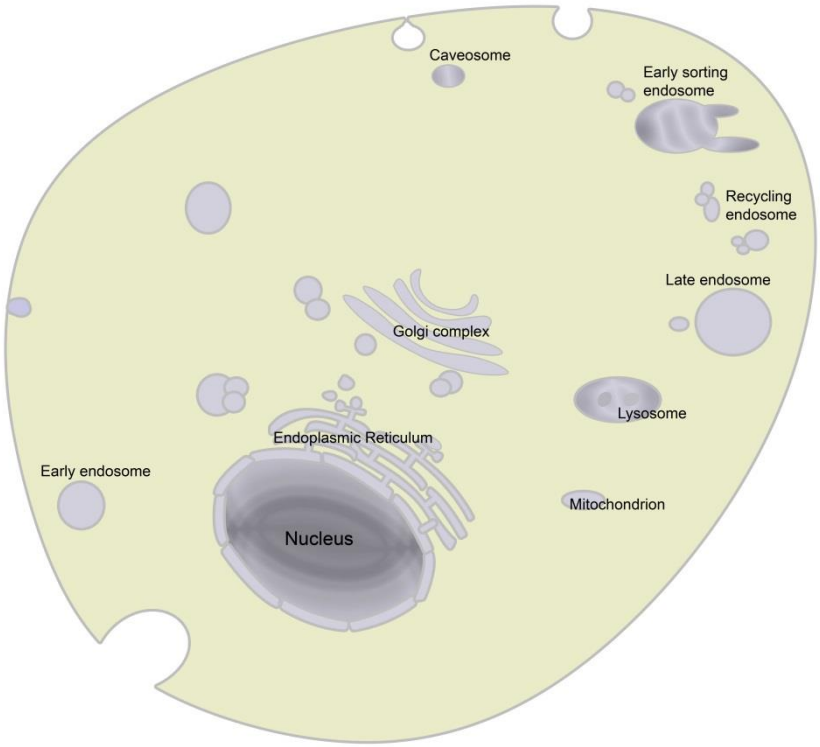


Figure 1. A simple model representation of trafficking related organelles of the eukaryotic cell

The major intracellular compartments of the eukaryotic cell are the plasma membrane, the cytosol, endoplasmic reticulum (ER), Golgi apparatus (complex), nucleus, mitochondrion, endosomes, and lysosomes. The plasma membrane is working as cell boundaries and as site for import and export of various cargoes. The cytoplasm consists of the cytosol and cytoplasmic organelles suspended in it. The cytosol constitutes around half of the total volume of the cell, and is the site of protein synthesis and degradation. Large part of the total area of membrane in eukaryotic cell encloses networked spaces of the endoplasmic reticulum. ER, according its specialized domains, is divided to the rough, smooth endoplasmic reticulum and nuclear envelope. The rough ER contains bound ribosomes on the cytosolic side, these synthetize both

soluble and membrane proteins, which later are destined for secretion to the cell exterior or for other organelles. The Golgi apparatus (or Golgi complex) is mostly responsible for sorting and modifying of various cargo molecules. The nucleus contains the genome and is the principal site of DNA and RNA synthesis. Mitochondrion generates most of ATP that cells use to drive various reactions. Endosomes are responsible for import of material from plasma membrane and transportation to destination organelles, such as lysosomes. Lysosomes contain digestive enzymes that degrade defunct intracellular organelles, as well as macromolecules and particles taken in from outside the cell by endocytosis. It is accepted that all proteins intended for secretion to cell exterior after being synthetized on the rough ER must be transported through ER to Golgi complex, early secretory pathway, and then through secretory bodies to plasma membrane, that briefly could be called as late secretory pathway[7].

DNA and RNA are polymer molecules containing small building blocks – nucleotides. A nucleotide is a molecule made up of a nitrogen-containing ring compound linked to five carbon sugar, which in turn carries one or more phosphate groups. The nitrogen-containing rings are generally referred as bases. The five-carbon sugar can be either ribose or deoxiribose. Nucleotides containing ribose are known as ribonucleotides, and those containing deoxiribose as deoxiribonucleotides. There are five nucleotides: adenine (A), cytosine (C), guanine (G), thymine (T), and uracil (U). Each nucleotide named after base it contains. RNA usually contains of A, C, G an U ribonucleotides, DNA – A, C, G, T. The ability of the bases in different nucleic acid molecules to recognize and pair with each other by hydrogen-bonding is usually referred as base-pairing or Watson-Crick base pairing - C makes hydrogen bonds with G, A with T or U.

Genetic information stored in DNA is transcribed to RNA, which may carry many functions; one of those is to be as a template for protein translation. RNA which is used for protein translation is called messenger RNA (mRNA)[6]. In eukaryotic cells production of mature mRNA, the one ready to be used as template for protein translation, requires post translational processing. First, RNA transcribed from DNA has non-coding sequences, called introns. Introns are removed from RNA by a RNA-protein complex, spliceosome, stitched coding sequences, the exons, form mRNA. Second, 5’ end of RNA polymer is modified by addition of 7-Methylguanosine, termed as RNA cap. RNA cap is required for mRNA recognition by the

ribosome and for protection from RNases, RNA digesting enzymes. Third modification is polyadenylation, the covalent linkage of adenylyl polymer on the 3' end of RNA polymer; this protects mRNA from degradation by exonucleases and is important for export of the mRNA from the nucleus and protein translation. Every mRNA contains three logical elements: 5' untranslated region (5' UTR), coding sequence, and 3' UTR.

After transcription and modifications of the mRNA in a nucleus, mRNA is exported to the cytosol there translation of proteins occurs. Exact localization of specific mRNA, in respect to other cytosolic organelles, is still under question. Concepts of molecular biology suggest that translation of mRNA into proteins can occur via free, membrane not bound, ribosomes, at the same time reporting no structural difference of membrane bound (rough ER) ribosomes to the ones found free in cytosol[6]. There is existing idea that mRNA coding membrane or secretory proteins are attracted to the ER membranes in a signal specific manner, by direct interaction or via attracting other factors called signal recognition particles (SRP). Report of experimental study in HeLa cells suggested existence of other, signal sequence independent mechanisms for mRNA localization to the ER, concluding localization of all mRNA to be at the ER[8].

Distribution of different proteins within the cell is not uniform, proteins localise to specific organelles, according to their function, or they are transported via secretory membranes to extracellular space. Usually, protein-membrane interactions are governed by a short amino acid sequence at the C end of a polypeptide, termed as signal sequence. Protein localisation usually is assigned to specific organelle via co-localisation with another protein which localisation is known in biochemical or microscopy based experiments. Biochemical approach is limited to cell lysate therefore it is cell population average describing strong interacting partners and knowledge on localisation within the cell is lost. Microscopy based localisation, however, can give single cell deviations but no information on interaction strength and is limited by resolution of microscopy technique used for imaging.

Membrane trafficking

The transfer of material in cells, between different secretory organelles, is mediated by carrier vesicles or tubules that continuously bud from one membrane and fuse with the other. Each vesicle transport event can be divided into essential steps: budding, transport, tethering, and fusion.

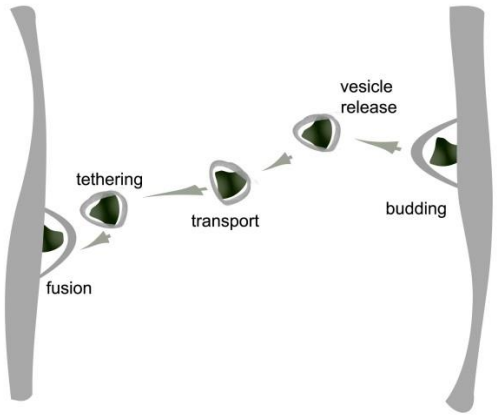


Figure 2. The essential steps in vesicle transport

On the donor membrane cargo molecules are recognized and selected from the donor compartment by being sorted into a vesicle or a larger intermediate that then detaches from the donor compartment. This event is called vesicle budding. Vesicle budding is mediated by protein coats, which are dynamic structures that cycle on and off membranes. Coat proteins are recruited from cytosol onto donor membranes by active GTP-bound forms of the Arf1/Sar1 family small GTPases[9–11] and then deform flat membranes into round buds, that leads to the release of coated vesicles. Clathrin was the first coat protein identified, clathrin coated vesicles are mainly derived from the plasma membrane or trans-Golgi network end are transported to endosomes[12]. Cop I and Cop II are example of other essential coat proteins which are involved in vesicle transport between ER and Golgi complex[13].

After budding, vesicles are transported to their final destination by diffusion or by motor-mediated transport along a cytoskeletal track (microtubules or actin; molecular motors like kinesin, dynein, and myosin are involved in this process)[14].

Tethering is a term used to describe the initial interaction between a vesicle and its target membrane. It precedes the pairing of transmembrane soluble N-ethylmaleimide sensitive factor attachment protein receptor (SNAREs; a family of membrane proteins) on opposing membranes, an event that leads to membrane fusion[11]. Another family of proteins together with SNAREs involved in specific tethering of cargo vesicles is Rab GTPases[14,15].

The last step in vesicle-mediated transport is the fusion of the vesicle with its target membrane. Fusion is thought to occur by the pairing of SNAREs. A SNARE on a transport vesicle (v-SNARE) pairs with its cognate SNARE-binding partner (t-SNARE) on the appropriate target membrane. However, the specific role of the SNAREs in membrane fusion remains to be determined.

In general, vesicle transport is well controlled process by various specially separated factors, thus different cargo vesicles from donor membrane are capable of reaching different destination membranes. Controlled and selective transportation of cargo vesicles is usually referred as molecular sorting; first, the ability to selectively include or exclude individual membrane and content proteins during the formation of transport vesicles, and second, the ability to segregate the vesicular container from its cargo after vesicle fusion. This segregation permits the recycling or salvage of essential components involved in vesicle formation and targeting, which must be returned to the donor compartment to allow continued transport[13].

GTPases

There is a number of different families of GTPases, which take role in cellular processes such as protein synthesis, transmembrane signalling, cell differentiation and proliferation, regulation of cell shape and migration, guidance of vesicular traffic within cells by employing basic cycle of GTP binding and hydrolysis[16,17].

All GTPase switches go through the same cycle of reactions. Binding and hydrolysis of GTP drive transitions among three conformational states: GDP-bound, ‘empty’, and GTP-bound (Figure 3).

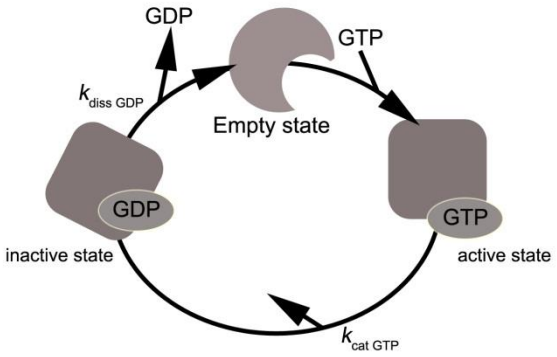


Figure 3. The basic GTPase cycle. Adapted from[16]

The switching function of each individual GTPase in the cell is determined by the differing abilities of conformational states to interact with specific macromolecules[16]. For most GTPases, the fraction of protein molecules in the active state (GTP-bound) depends on the relative rates of two reactions: dissociation of GDP from GDP-bound form and hydrolysis of bound GTP. These rates are characterized, respectively, by two fractional rate constants, $k_{diss\ GDP}$ and $k_{cat\ GTP}$. The ratio between GTP-bound and GDP-bound forms is:

$$\frac{GTPase\ GTP}{GTPase\ GDP} = \frac{k_{diss\ GDP}}{k_{cat\ GTP}} \qquad [1]$$

This equation depends on a number of assumptions, most of which are valid most of the time – the concentration of GTP must not be limiting, for example GTP must bind rapidly to the empty site on a GTPase that has just released GDP. For many GTPases, specific protein ligands regulate $k_{diss\ GDP}$ and $k_{cat\ GTP}$ constants. Those that increase $k_{diss\ GDP}$ are called guanine nucleotide exchange proteins or release proteins (GNRPs; guanine exchange factors (GEFs)), whereas GTPase activating proteins (GAPs) increase $k_{cat\ GTP}$ [16,18].

Rab GTPases

The RAS (the name for the protein family originated from organism of the first observation, which had happened in rat sarcoma) superfamily of monomeric GTP-binding proteins, or small GTPases, comprises a large family of regulatory molecules that collectively

regulate diverse and critical cellular processes in eukaryotes. For example, RAS proteins control cellular signalling pathways responsible for growth, migration, adhesion, cytoskeletal integrity, survival and differentiation[17]. Originally, RAS genes were identified as oncogenes in rat sarcoma in beginning of the second part of 20th century[19].

The members of the RAS superfamily are in the range of 20 to 29 kDa, share sequence homologies and common motifs[17]. The RAB (RAS in brain) and ARF GTPases are members of two of the RAS-related subfamilies that function in regulating vesicle trafficking, starting from regulating the formation of vesicles on donor membranes and directing trafficking specificity to and facilitating vesicle docking on target membranes. RABs are also known for playing essential role in signalling, cell differentiation and proliferation. Over 70 human RAB and Rablike members of the RAS superfamily have been identified[18,20]. RAB proteins are present on all compartments of the endomembrane system. Each RAB protein is found at a particular stage of a membrane transport pathway[21].

In most cases examined, sequence conservation among yeast, mammalian, and plant RAB GTPases has correlated with similar localization on discrete subcellular compartments of the endomembrane trafficking systems in these diverse organisms[18]. The localization of particular RAB family member to a specific compartment is governed by the amino acid sequence at its carboxyl termini[22], cysteine motive (CXXX, CXC, CC, CCXX, CCXXX where X is any other amino acid) overcomes post translational modification with one or two highly hydrophobic geranylgeranyl groups[23]. Post translational modification requires recognition of newly translated RAB by another protein, RAB escort protein, which presents RAB to the geranylgeranyl transferase.

Despite differences in amino acid sequence the RAB family shares common higher order structure (Figure 4 Ribbon drawing of Rab3a): Rab GTPases consists of a six stranded β sheet, comprising five parallel strands and one antiparallel one, surrounded by five α helices[24].

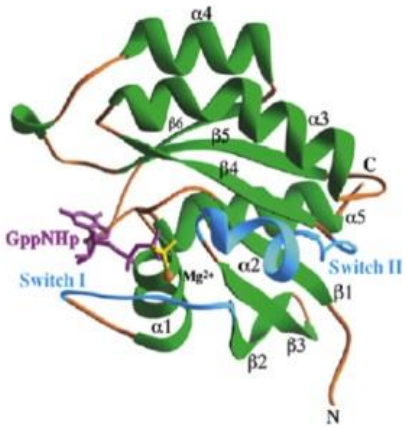


Figure 4. Ribbon drawing of Rab3a[24]

RAB GTPases utilize guanine nucleotide exchange and GTP hydrolysis to switch between active (GTP-bound) and inactive (GDP-bound) conformations. The molecular switch function of a small GTPase is carried out by cycling between its active GTP-bound form and inactive GDP-bound form. Switching from the inactive to the active state is accomplished by replacing bound GDP with GTP, this process requires GEF. After performing functions in their active form, GTPases are inactivated by hydrolysing GTP to GDP, which is accelerated by the GAP. Once inactivated, small GTPases detach from the membrane and are kept in the GDP-bound inactive state until the next round of the GTPase activation cycle begins. While SAR/ARF members do not require specific factors for this process, dissociation of most RAB GTPases from membranes is mediated by a conserved protein family, the RAB GDP dissociation inhibitor (RAB GDI). RAB GDI binds only to GDP-bound RABs and keeps them in the GDP-bound state by inhibiting GDP release.

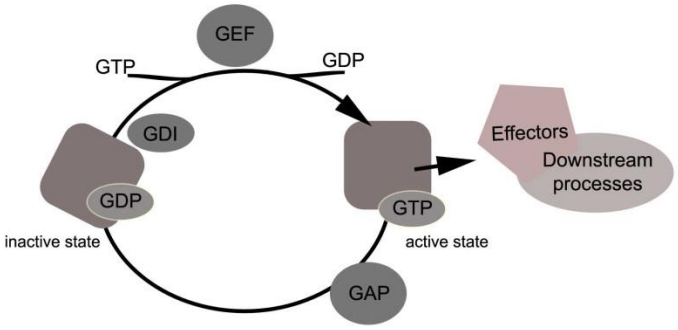


Figure 5. The Rab GTPase cycle between an active GTP- bound and inactive GDP-bound state. GEFs activate GTPases, which in turn interact with specific effectors to mediate downstream processes. GAPs stimulate the inherent GTPase activity of these small G proteins, accelerating the inactivation of their regulatory activity.

In-vitro experiments on RAB 5 mutant revealed that GTP cycle of RAB proteins differs that of other GTPases and indicate that GTP hydrolysis acts as a timer that determines the frequency of membrane tethering and fusion events. GTP hydrolysis is not conditional on membrane fusion as it was thought before but rather serves to sustain a dynamic equilibrium between GTP and GDP-bound forms. Continuous switching of RAB 5 between GDP to GTP-bound forms might act as recruitment regulator of Rabaptin-5 (RAB 5 interaction partner)[25].

RAB regulatory proteins (GEFs, GAPs and GDIs) are phosphorylated in response to stress and growth factor signalling, thereby enhancing or diminishing RAB activity and resulting in up- or down-regulation of constitutive and regulated trafficking[20].

The Example of a cargo protein. The Collagen

The collagens are large family of structural proteins that form a diverse range of highly organized supra-molecular assemblies in the extracellular matrix and are abundant in all animals. In humans, collagens comprises one third of the total protein, accounts for three-quarters of the dry weight of skin, and are the most prevalent component of the extracellular matrix[26]. The superfamily now includes more than 20 collagen types with altogether at least 38 distinct polypeptide chains, and more than 15 additional proteins that have collagen-like domains[27].

There are many classes of collagenous structures in the extracellular matrix, including fibrils, networks, and trans-membrane collagenous domains. According structural conformation the superfamily can be divided into several families based on these assemblies and other features. All collagens also contain non-collagenous domains, and many of these have important functions that are distinct from those of the collagen domains[27].

Various defects of collagen supra-molecules formation in extracellular matrix leads to number of known diseases, for example: after injury healing involves the formation of scar and fibrous tissue which largely consists of collagen fibrils and if fibrous material accumulates in excessive amounts that impairs the normal functioning of an affected organ. The normal formation of collagen might be interfered by mutations in both triple-helical and non-triple-helical domains of various collagens or by mutations of other biomolecules involved in the collagen biogenesis[26,27].

All collagen molecules consist of three α polypeptide chains that are wrapped around each other into a triple helix. In some collagen types, all three α chains of the molecule are identical, while in other types the molecule contains two or even three different α chains. In each of the polypeptide chains every third amino acid is glycine, and thus the sequence of an α chain in a collagen domain can be expressed as $(Gly\text{-}X\text{-}Y)_n$, there X and Y represent any other amino acid than glycine and n is number of repetitions in the chain, the repeat motif is continuous for approximately 1000 amino acids.

The collagen biosynthesis involves a number of post-translational modifications and requires at least nine ER-resident enzymes and molecular chaperones[28]. Five collagen specific enzymes are required: three collagen hydroxylases, two collagen glycosyltransferases[27]. The fibril-forming collagens are first synthesized as procollagen molecules that have propeptide extensions at both their N- and C terminal ends. After cleavage of the signal peptides number of modifications occurs: hydroxylation of certain proline and lysine residues to 4-hydroxyproline, 3-hydroxyproline and hydroxylysine, glycosylation of some of the hydroxylysine residues to galactosylhydroxylysine and glucosylgalactosyl-hydroxylysine, glycosylation of certain asparagine residues in one or both of the propeptides, association of the C propeptides through a process directed by their structures, and formation of intrachain and interchain disulphide bonds. Hydroxylation of proline residues, and interaction with the molecular chaperone HSP47 are unique collagen biosynthetic events. Protein disulphide isomerase (PDI) assists protein folding in the ER by catalysing the formation, reduction and isomerisation of disulphide bonds. As a subunit of the enzyme prolyl 4-hydroxylase, PDI is also essential for proline hydroxylation, the key post-translational modification underpinning helix formation. The main function of PDI in prolyl 4-hydroxylase seems to be to keep the highly insoluble a subunits in a catalytically active, non-aggregated state. Prolyl 4-hydroxylase requires ferrous ions, 2-oxoglutarate, O_2, and ascorbate[27].

When disulphide bond formation is completely prevented by incubation of fibroblasts with the reducing agent dithiothreitol, type I collagen does not assemble and is not secreted[28], that could be called primary quality control in collagen secretion pathway.

After the C propeptides have become associated and about 100 4-hydroxyproline residues have been formed in each of the pro α chains, a nucleus of the triple helix forms in the C-terminal region, and the triple helix is then propagated towards the N terminus in a zipper-like fashion[29,30]. The procollagen molecules are transported from the endoplasmic reticulum across the Golgi stacks without ever leaving the lumen of the Golgi cisternae. During this transport, the molecules begin to aggregate laterally, and increased condensation of the aggregates results in the formation of granules ready for secretion[31]. The extracellular steps in biosynthesis include cleavage of the N and C propeptides[32], assembly of the collagen

molecules into fibrils by nucleation and propagation[29,33,34], and formation of covalent cross-links[35]. The processing and assembly of other collagens feature the same steps as for the fibril-forming collagens, but with certain exceptions. Many collagen molecules contain N and/or C-terminal non-collagenous domains that are not cleaved and are hence not called propeptides.

Procollagen molecules formed into triple helices are known as tropocollagen, assemble in a complex, hierarchical manner that ultimately leads to the macroscopic fibers and networks observed in tissue, bone, and basement membranes. Fibrils are primarily composed of type I collagen. Procollagen of type I collagen have the unique property of actually being unstable at body temperature. The trick is that collagen fibrillogenesis, process of tropocollagen monomers assembling into mature fibrils, has a stabilizing effect on triple helices. Moreover, the assembly of macromolecular structures is essential to enable collagen to support stress in three dimensions[36].

Extracellular modifications require two specific proteinases to cleave the N and C propeptides[32] and one specific oxidase to convert certain lysine and hydroxylysine residues to their reactive aldehyde derivatives that then participate in cross-link formation[35].

Upon entering the secretory pathway, procollagen undergoes a secondary quality control with not properly folded molecules being degraded in the endosomal–lysosomal system[37]. The processing and assembly of other collagens feature the same steps as for the fibril forming collagens with certain exceptions. Many collagen molecules contain N and/or C-terminal non collagen domains that are not cleaved[27].

Revealing of the RNA interference

Initially, there was an approach to increase gene production rate by introducing transgene DNA into plants, which resulted unexpected silencing of the gene[38]. Results of that experiment led to the new era of gene function analysis experiments in living organisms as introduction of complementary RNA to silence gene of interest, later dubbed as anti-sense technology. Anti-sense experiments saved a lot of time comparing to mutagenesis methods used before. Gene silencing by introduction of high amounts of endogenous RNA had short

activity window, as concentration of anti-sense molecules diminish rapidly as consequence of RNA degradation and cell division. However, some experiments were showing unexpected efficiency on gene silencing over number of populations and not limited by small number of anti-sense molecules per cell. At that time anti-sense molecules used to be produced using biotechnology rather than using synthetic synthesis and hypothesis was carried out that during production of anti-sense molecules some fraction of molecules were double stranded. Indeed, experiments with double-stranded RNA molecules showed highly increased efficiency of gene down-regulation in comparison to single-stranded molecules[39]. That was first step into revealing cell internal post-transcriptional gene expression regulation mechanism, now called as RNA interference (RNAi).

RNAi is a concept of gene expression silencing by protein complex, which is guided by the small RNA molecule complementary pairing to mRNA.

Small non-coding RNAs

Three major classes of functional small non-coding RNAs (~21-31 nucleotides long) have been found in mammals: microRNAs (miRNAs), endogenous small interfering RNAs (endo-siRNA), Piwi-interacting RNAs (piRNAs). These classes differ in their biogenesis. Here I will focus on siRNA and miRNA as main players in RNAi pathway.

In general, siRNA is a response to foreign or invasive nucleic acids such as viruses, transposons, and transgenes. siRNAs are ~ 21-23 nucleotides long double-stranded RNA with 2 nucleotides overhangs on 3' ends derived from long double stranded RNAs (dsRNAs) which are directly processed by Dicer, RNase III family enzyme, along dsRNA to produce multiple siRNAs[40]. Dicer can cleave dsRNA in a sequence-independent manner and this includes dsRNA with noncanonical base-pairings and minor duplex deviations. Specific cleavage at this site is determined by the Dicer PAZ domain that contains a hydrophobic protein fold which recognises the two nucleotide overhang at the 3'-end. This locks the dsRNA into position and explains why Dicer generally prefers to have at least one 3'-termini of its dsRNA substrate with two nucleotide overhangs. However, Dicer can accept other RNA substrates including those with

blunt ends and 5'-overhangs. Once dsRNA is locked by PAZ domain and positioned within Dicer, the duplex lines up along the edges of both RNase III domains. Here, the two RNase III catalytic sites cleave the dsRNA to leave siRNA with two nucleotide overhangs at both 3'-termini. The larger RNase IIIa domain cleaves the dsRNA to leave the strand containing the 5'-phosphate, and the RNase IIIb domain independently cleaves the opposing strand two nucleotides further along to leave the two nucleotide 3'-overhang containing a terminal hydroxy group. The siRNA product length is determined by the distance between the PAZ binding domain and the active sites of both RNase III domains; the active sites of each RNase III domain harbours two magnesium atoms which are essential for phosphodiester hydrolysis[41]. Both siRNA strands usually are perfectly complementary.

miRNAs are initially transcribed by RNA polymerase II from independent genes or represent introns of protein-coding genes as a long RNAs containing hairpins (pri-miRNAs). Drosha, RNase III family enzyme, cleaves hairpins resulting ~70 nt long pre-miRNAs, which are then released to the cytoplasm. Dicer cleaves pre-miRNAs into short 18-25 nt double stranded miRNAs[42]. Complementarity of miRNA strands usually is not perfect, making different conformations between strands.

Critical feature of double stranded small RNAs is a phosphorylation of 5’ ends, as it is required for recognition by Dicer[43,44]. In human cells phosphorylation of 5’ ends occurs naturally upon entry to the cytoplasm[45].

Synthetic siRNA molecules as technique are used in various biological assays as a tool for down-regulation of gene expression. In contrast to miRNAs, siRNAs are designed to keep high complementarity to target mRNA, main target of siRNA is defined by a design. Reliable prediction of endogenous miRNA targets is problematic task as miRNAs usually have no highly complementary target genes; bioinformatics approaches that are based on known rules derived from biochemical experiments cover just a fraction of possible arrangements; bioinformatics approaches that are based on unattended machine learning also fail (personal correspondence with Andrius Serva; AG Starkuviene). Additionally, interaction of short RNAs with their target mRNA is not always leading to suppression of gene expression, upon binding on 5’UTR end of mRNA, miRNA might enhance translation rate[46].

RNAi machinery

RNAi is an efficient and complex mechanism, which involves number of proteins, activity steps, and short double-stranded RNA molecules. Main player is so called RNA induced silencing complex (RISC). According to literature, RISC variates in different organisms but the key factor is an argonaute protein (historically named as slicer) which contains an RNase-H type fold. RNase-H is responsible for cleaving the phosphodiester bond of RNA in a RNA/RNA or RNA/DNA duplex. All argonaute proteins investigated so far contain three principal domains: the N-terminal, PAZ and the PIWI domain[47]. Approximately the first 150 amino acids of argonaute constitute the N-terminal region which is joined to PAZ by a β-sheet stalk. The PAZ domain of argonaute proteins is structurally and functionally similar to the PAZ domain of Dicer, containing a hydrophobic protein fold that specifically recognises siRNA containing two-nucleotide overhangs and responsible for binding 3' end of single stranded short RNA[48]. The PIWI domain can be split into two sub-domains known as the A (Mid/A) and B (PIWI/B) sub-domains and has two distinctive roles. The Mid/A sub-domain is the primary region of the 5'-phosphate binding pocket whereas the PIWI/B subdomain contains the active site of Slicer, the key characteristic of which is the RNase-H like fold that is responsible for mRNA cleavage[48].

Not all argonautes show endoribonuclease activity and are thus active as Slicer. Humans, for example, have eight homologues, the Ago subfamily (hAgo1–4) and the Piwi subfamily (HIWI, HIWI2, HILI and PIWIL3)[49], but only hAgo2 has shown Slicer type activity[50]. The activity of the argonautes can partly be attributed to the catalytic site of the protein (which is contained within the PIWI/B sub-domain), with an Asp-Asp-X amino acid trio of active site residues essential for Slicer activity (where X = Asp, Glu, His or Lys)[51]. Even though some argonautes have the complete trio of active site residues, they remain inactive (hAgo3, for example) which suggests there are other unknown factors that are involved in determining whether an Argonaute protein is an active Slicer or not[50].

Formation of active RISC requires not only Argonaute protein and small RNA but also many cofactors and formation steps. Initial step is formation of RISC loading complex (RLC), which is composed from Dicer, Argonaute, TRBP (HIV-1 transactivating response RNA-binding

protein), and small dsRNA. Co-factor TRBP is essential for efficient transfer of short RNA from Dicer to Argonaute2, it acts as stabilizing agent of Dicer and Argonaute interaction[52]. After short RNA is transferred into Argonaute, RISC is assembled, yet it is not clear if single or double strand small RNA molecule is transferred as human Dicer has ATP-dependent helicase domain on its N-terminal which could unwind short RNA, but some experiments showed an efficient ATP independent loading of the RISC. It is known that the RNA strand whose 5'-end is less stably paired with its complement within the small RNA duplex is chosen as the guide strand[53,54]. In year 2005 two RISC loading mechanisms were proposed:

1) Passenger strand cleavage mechanism[55,56]: RISC works as an endoribonuclease by guide-strand directed hybridisation and cleavage of mRNA, it was suggested that the siRNA duplex is analogous to the guide strand/mRNA duplex and the passenger strand could act as the first target for RISC. This would result in the passenger strand cleavage within the inactive RISC complex. This was proven using antisense RISC inhibitors that block hybridisation to target mRNA. Their failure to stop the production of passenger strand cleavage products showed that passenger strand degradation preceded siRNA dissociation.
2) Bypass mechanism[55]: hAgo2 is the only active human Slicer, but evidence suggests that the other hAgo homologues are loaded with single-stranded guide RNA strands. Since the other hAgo homologues cannot act as Slicer, they are unable to cleave the passenger strand as required for the passenger strand cleavage mechanism. This means an alternative pathway, which is independent of passenger strand cleavage, must be operational and this is known as the bypass mechanism. It relies on the assumption that some helicase is involved in siRNA unwinding. Helicases require an energy source and previous reports suggested that RISC assembly was ATP dependent, which would account for helicase activity in the bypass mechanism. However, it was observed no such requirement for ATP. Passenger strand cleavage appears to

be the most favoured mechanism for siRNA dissociation, with most evidence pointing towards this.

When the guide strand is incorporated into RISC with its 3'-end attached to the PAZ domain of argonaute and the 5'-phosphate is coordinated into a 5'-binding pocket. The key amino acid residues of the 5'-binding pocket are located primarily in the Mid/A sub-domain of PIWI, but also partly in the PIWI/B sub-domain. The 5'-phosphate binding pocket and 3'-hydroxyl group bound in PAZ seem to act as molecular anchors, holding the guide strand in place and ensuring cleavage of the mRNA target. The nucleotides along the guide strand do not equally contribute to mRNA recognition: the key nucleotides are from 2^{nd} to 8^{th}, known as the seed region, as measured from the 5'-end. Since the 5'-phosphate is bound strongly to the PIWI binding pocket, nucleotide 1 is bent away from the mRNA strand and does not participate in mRNA recognition. The 5'-seed region initially hybridises to the mRNA target in the PIWI domain. This initial affinity between guide and mRNA determines most of the binding energy between the guide and target RNA strands. Although nucleotides 9 to the 3'-terminus are of less importance, the propagation of duplex formation towards the guide strand 3'-end is thought to be required for efficient catalytic cleavage of mRNA. However, evidence suggests that as long as the seed region and the next few bases can locally form an A-type helix, cleavage of the scissile bond still occurs. This mean that active RISC complexes can cleave other mRNAs that only possess partial complementarity and thus leads to off-target effects. Once the guide/mRNA duplex is formed in argonaute, the phosphate group between the 10th and 11th nucleotide is positioned in the PIWI/B sub-domain, directly in the Rnase H type fold[57]. Cleavage then occurs aided by two divalent magnesium ions to leave products with a 5'-phosphate and 3'-hydroxyl group[58]. For miRNAs, the miRNA–guide hybridises to the mRNA in an imperfect manner and the mRNA is not actually cleaved by RISC, instead, the hybridisation inhibits translation or RISC recruits deadenylation factors like poly(A) binding protein (PABP) and the CAF1 and CCR4 deadenylases[59].

Fluorescence microscopy

A light microscope is an optical instrument that uses visible light to produce a magnified image of an object that is projected onto the retina of the eye or onto the photosensitive surface of an imaging device[60]. During its long development the microscopy field have developed number of various techniques applicable for different tasks, despite differences, all microscopes share the same basic optical principles. Two microscope components are of critical importance in forming the image: the objective, which collects by the specimen emitted or diffracted light and forms a magnified real image at what is called the real intermediate image plane near the eyepieces or oculars, and the condenser, which focuses light from the illuminator onto a small area of the specimen. The fundamental difference between the ordinary light and fluorescence microscope is in the way of image formation: normally, the image is formed by the modification of light passing through the specimen; a fluorescent image is formed due to visible light emission from the specimen itself, the illumination beams which excite fluorophores do not contribute to the formation of the image.

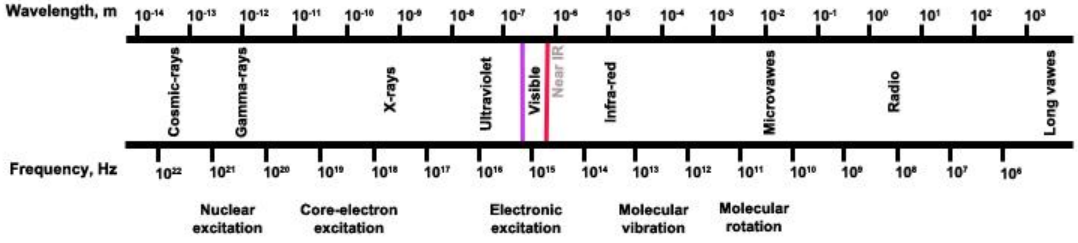


Figure 6 summary of the electromagnetic spectrum and classification of the spectral regions

The light can be described as electromagnetic field, an oscillating electric and magnetic disturbance that spreads as a harmonic wave through space or a particle, called photon. The electromagnetic field is characterized by wavelength (λ), the distance between two neighbouring peaks of the wave, and its frequency (ν), the number of times per second at which its displacement at a fixed point returns to its original value. In Figure 6 the electromagnetic spectrum is summarized, the description and classification of the electromagnetic field according to its frequency and wavelength. Incident light may interact with material objects in different ways: light may be reflected, absorbed or transmitted through

the object, also scattered in all directions. Visible to the human eye part of the electromagnetic spectrum corresponds to wavelengths from about 380 nm to 700 nm and energies of the photons corresponds to excitation energies of outer electronic layer of atoms and molecules[61]. In case of specific visual light absorption by an atom or molecule a number of different relaxation events might occur, principle diagram is depicted in Figure 7: the electronic state of the molecule changes from the ground state to one of many vibrational levels in one of the excited electronic states, once the molecule is in excited state, relaxation to the lowest level of excited state occur via vibrational relaxation process, now exited molecule might relax to the ground electronic state by emitting fluorescence photon or through non-radiative internal conversion or via intersystem crossing transfer excitation energy to triplet excited state, from lowest level of triplet excited state molecule might relax by emitting phosphorescence photon or through non-radiative external conversion.

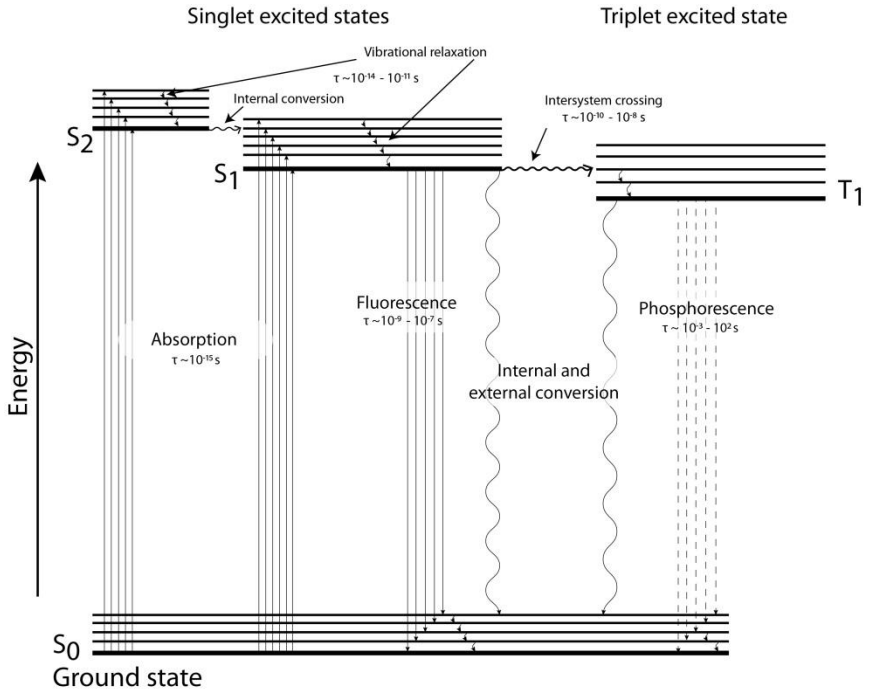


Figure 7 Electronic transition energy level diagram, Jablonski diagram

Ernst Abbé's (1840–1905) theory of image formation in the light microscope, developed while working in the Carl Zeiss optical workshop in Jena, is based on three fundamental actions: diffraction of light by the specimen, collection of diffracted rays by the objective, and

interference of diffracted and non-diffracted rays in the image plane. The key element in the microscope's imaging system is the objective, which determines the precision with which these actions are effected. An objective is described by the numerical aperture (NA), the angle over which the objective can collect diffracted rays from the specimen and the key parameter determining spatial resolution[62].

In the light microscope, the angular aperture is described in terms of the numerical aperture (NA), as:

$$NA = nsin\Theta \qquad [2]$$

where Θ is the half angle of the cone of specimen light accepted by the objective and n is the refractive index of the medium between the lens and the specimen.

For point objects that are self-luminous (e.g. fluorophores), or for nonluminous points that are examined by bright field microscopy in transmitted light where the condenser NA is greater than or equal to the objective NA, the resolving power of the microscope is defined as:

$$d = \frac{0.61\lambda}{NA} \qquad [3]$$

where *d* is the minimum resolved distance in µm, λ is the wavelength in µm, and NA is the numerical aperture of the objective.

Just as diffraction and the wave nature of light determine that the image of a point object is a diffraction disk of finite diameter, so do the same laws determine that the disk has a measurable thickness along the z-axis. Depth of field (*Z*) in the object plane refers to the thickness of the optical section along the z-axis within which objects in the specimen are in focus. For diffraction-limited optics, the wave-optical value of *Z* is given as:

$$Z = n\lambda/NA^2 \qquad [4]$$

where n is the refractive index of the medium between the lens and the object, λ is the wavelength of light in air, and *NA* is the numerical aperture of the objective. Thus, the larger the aperture (the higher the *NA*), the shallower will be the depth of field.

The other existing fluorescence microscopy techniques are capable to resolve spatial features even below diffraction limit, down to 10 nm resolution[63–65]. Basic principle to overcome the diffraction limit for optical resolution is to separate fluorophores in time and/or space, and later construct the whole image of observed area by summarizing determined

fluorophore locations. It can be done by physically controlling excitation volume, as it is realized in STED microscopy, or by photo-physical switching of most fluorophores to non-fluorescent state (triplet state), as it is realized in dSTORM. In such case improvement of spatial resolution comes with expense of time required to acquire microscopy data, computing power and data storage.

The term of screening microscopy is usually referred to an automated wide field microscope, as devices of this type are relatively fast once observation of bigger sample area is in consideration. However, screening microscopy is not suitable to address all questions, as the application of microscopy is always about compromise which usually is about size of observation area, data acquisition speed and resolution. For example, then a conventional wide field microscope is used with a 10x air objective planar resolution is in the range of 800 nm and depth of field is 3 µm, once the same microscope is equipped with 60x oil immersion objective area of view is reduced by a factor of 2×10^2 and planar resolution is 230 nm, depth of field is 420 nm. Here high-content microscopy approach becomes very handy to overcome unavoidable compromise. Once a screening is performed in a wide field mode, imagery can be analysed and a high number of observed cells may be evaluated. Here specific examples of statistical distribution can be selected for investigation with a higher resolution microscopy technique. In this case a statistical distribution of many observed cells comes in one set with high resolution imagery of selected single cells. This combination creates higher information content. That requires for improvements and interconnections of different application layers like digital image analysis, statistical data analysis and even synergy of different microscopy techniques.

Methods of molecular biology

Growth and storage of bacteria cultures

Preparation of Lysogeny broth medium

Lysogeny broth (LB) is a nutritionally rich medium primarily used for the growth of bacteria. It was developed by Giuseppe Bertani and published for the first time in 1951[66]. Typical composition of LB medium is: 10 g/L Tryptone, 5 g/L Yeast Extract, 10 g/L Sodium Chloride (NaCl), deionized water. pH of LB naturally is close to pH 7 but might be affected by yeast extract. Freshly prepared LB has to be autoclaved on liquid cycle for 20 minutes at 1 atm. and cooled down before addition of antibiotic. LB can be stored at room temperature or at 4°C if antibiotic was added.

Preparation of LB selection plates

To normal composition of LB medium add 15 g/L of agar. After autoclaving and cooling down to 55˚C required antibiotic should be added and solution poured into Petri dishes. Plates should be left over night at room temperature to dry before storage at 4˚C.

Preparation of dYT medium

dYT is more rich medium compared to LB, usually used then bacterial cells have to overcome higher stress. Typical composition of dYT medium is: 16 g/L Tryptone, 10 g/L Yeast Extract, 5 g/L NaCl, deionized water. pH has to be adjusted by addition of NaOH. Freshly prepared dYT has to be autoclaved on liquid cycle for 20 minutes at 1 atm.

Growth and storage of Escherichia coli

Bacteria were grown on the LB agar plates and in the liquid LB media at 37 °C. For the selective conditions appropriate antibiotics were added. For long-time storage overnight bacteria cultures were mixed 50:50 with sterile 30% glycerol and placed into -80 °C.

Preparation of chemically competent cells of Escherichia coli

Inoculate 5 ml of dYT medium with a glycerol culture of the particular strain, shake overnight at 37 °C, dilute 1:50 in a fresh dYT medium and grow cells at 37 °C to an OD_{600} of 0,6. After centrifugation at 4000 rpm and 4 °C for 20 min, the cells has to be re-suspended in 20 ml 100 mM ice-cold $CaCl_2$ medium and incubated 20-24 h on ice. Collect the cells by centrifugation; re-suspended the pellet in 3 ml of 100 mM ice-cold $CaCl_2$ and incubated for 30 min on ice. Aliquot cells and add DMSO 3,5% of total volume, mix and stored at -70 ˚C.

Transformation of the chemically competent cells of Escherichia coli

In molecular biology a transformation is the genetic alteration of a cell resulting from the direct uptake, incorporation and expression of exogenous genetic material from its surroundings and taken up through the cell membrane. For the transformation, 10-300 ng of DNA should be added to 200 µl aliquot of the cell suspension, and then cells should be incubated on ice for 30 min, following by 3 min heat shock at 37 ˚C. Then mix the cells with 0,8 ml LB medium and incubate for 40-60 min at 37 ˚C. Then plate 50-100 µl of cell suspension on the selective LB agar plates.

Plasmid DNA preparation

The basic steps of DNA isolation are disruption of the cellular structure to create a lysate, separation of the soluble plasmid DNA from cell debris and other insoluble material and purification of the DNA from soluble proteins and other nucleic acids. As there are many methods to extract and purify plasmid DNA from cell cultures, here briefly described sample preparation method used for this work[67]. Primary consideration then preparing plasmid DNA is contamination of the sample with chromosomal DNA. SDS-alkaline denaturation (sodium dodecyl sulfate) of DNA technique exploits the difference in denaturation and renaturation characteristics of covalently closed circular plasmid DNA and chromosomal DNA fragments. After plasmid and chromosomal DNA are efficiently denatured, rapid neutralization with a high-salt buffer such as potassium acetate in the presence of SDS has two effects that contribute to the overall effectiveness of the method. First, rapid neutralization causes the chromosomal DNA to base-pair in an intrastrand manner, forming an insoluble aggregate that precipitates out of solution. The covalently closed nature of the circular plasmid DNA promotes interstrand rehybridization, allowing the plasmid to remain in solution. Second, the potassium salt of SDS is insoluble, so the protein and detergent precipitate and aggregate, which assists in the entrapment of the high molecular weight chromosomal DNA. Separation of soluble and insoluble material is accomplished by a clearing method: filtration or centrifugation[68]. The soluble plasmid DNA is ready to be further purified.

I have used two methods for soluble plasmid purification: binding of the plasmid to silica in the presence of high concentrations of chaotropic salts[69] and differential precipitation of plasmid DNA from ethanol solutions[70]. Then silica is used the plasmid DNA can be isolated by virtue of its ability to bind silica in the presence of high concentrations of chaotropic salts. These salts are then removed with an alcohol-based wash and the DNA eluted in a low-ionic-strength solution such as water.

Determination of DNA concentration by UV-spectroscopy

In molecular biology, quantitation of nucleic acids is commonly performed to determine the average concentrations of DNA or RNA present in a mixture, as well as their purity. There are several methods to establish the concentration of a solution of nucleic acids, here UV spectrophotometric quantification method is described.

Beer–Lambert–Bouguer law states: the absorbance of a beam of collimated monochromatic radiation in a homogeneous isotropic medium is proportional to the absorption path length, *l*, and to the concentration, *c*, or — in the gas phase — to the pressure of the absorbing species. The law can be expressed as:

$$A = log_{10}\left(\frac{P_\lambda^0}{P_\lambda}\right) = \varepsilon cl \text{ or } P_\lambda = P_\lambda^0 10^{-\varepsilon cl} \qquad [5]$$

where the proportionality constant, ε, is called the molar absorption (extinction) coefficient, *A* usually referred as optical density (OD). At a wavelength of 260 nm, the average extinction coefficient for double-stranded DNA is 0.020 $(\mu g/ml)^{-1}cm^{-1}$, for single-stranded DNA it is 0.027 $(\mu g/ml)^{-1}cm^{-1}$, for single-stranded RNA it is 0.025 $(\mu g/ml)^{-1}$ cm^{-1}. Thus OD_{260} of 1 corresponds to a concentration of 50 µg/ml for double-stranded DNA.

Purity of DNA sample by proteins may be estimated by ratio OD_{260}/OD_{280}, pure DNA ratio should be between 1.8 and 2.0.

Growth and storage of adherent eukaryotic cells

All HeLa cells were grown in medium composed of: DMEM, 10% FBS, 1% L-Glu, 1% Pen/Step (see Chemicals). NRK-GalT-CFP cells were grown in medium composed of: DMEM, 5% FBS, 1% NEAA, 1% L-Glu, 1% Pen/Step. NIH 3T3 cells were grown in medium composed of: DMEM, 10% NCS, 1% L-Glu, 1% Pen/Step. Cells were grown at +37 ˚C and 5% CO_2 conditions. Cells were passaged periodically, 3-4 times per week, with dilution factor of 1:3~5.

For transfection experiments growth medium without antibiotic was used. For long storage, cells used to be collected by centrifugation at 1000 RPM for 5 min and pellet resuspended in growth medium containing 10% of DMSO.

Preparation of stably transfected eukaryotic cell line

In general, stable cell line means that population of cells in long term is expressing gene that is introduced to cells intentionally by researcher. There are many different approaches for establishing stable cell lines, depending on the type of expression one is interested in and the construct that is being incorporated. This protocol is specific for establishment of cell lines that constitutively express fluorescent protein tagged genes:

1. Generation of DNA vector with construct of interest (plasmid) and resistance gene (for example: kanamycin resistance for selection in bacteria and neomycin for selection in mammalian cells).
2. Transfection of cells with DNA vector
3. Day after transfection passage cells (high dilution ratio; 1:20) and start selection treatment (G418; concentration has to be selected before, advisable concentration is represented by reduction of cell population by 50% after three days of incubation and complete death of population after 7 days of incubation with drug)
4. Cells have to be grown under drug treatment, changing media each 2-3 days. Cells which did not incorporate plasmid DNA into genomic DNA should start suffering and in 10-14 days massive cell death should be observed.

5. After some time survived cells will be growing in clump-colonies. Colonies of interest has to be selected according fluorescence signal, as there will be colonies not expressing expected amount of endogenous gene. Pick-up of positive cell colony might be done by scratching them of the plate using pipet tip, it is recommended to rinse the dish with PBS and then add warm PBS containing 5% trypsin-EDTA.
6. Picked-up colony has to be transferred into new 24-well plate. With every consequent passage cell may be transferred to bigger dish.
7. If cell population is not homogeneous (according fluorescence signal), it is possible to use fluorescence activated cell sorting (FACS).

Protocol for fluorescence image based assay to over-expression effect on collagen secretion

Plate cells one day before transfection at $0.5x10^4$ cells per well concentrations on 96 well plates (Ibidi). Transfect cells with transfection mixture composed of 100ng of plasmid with 0.8µl Enhancer and 2.5µl Effectene. After incubation with transfection mixture for 6 hours wash cells once with growth medium and leave in new growth medium for another 16 hours and 15 min. Perform chase experiment by incubating cells with ascorbic acid (1:100) and cycloheximide (1:1000) for 1 h 45 min. After chase cells must be fixed (incubation in 3% paraformaldehyde mixture for 20 min). 0,1% Triton-X-100 mixture and 5 min incubation has to be used for cell membrane permeabilization. Antibody labelling of collagen-I should be performed by sequentially incubating cells with primary and secondary antibodies for 30 min at room temperature, including three times washing of the cells by 1xPBS solution for intermediate steps. Cell nuclei have to be labelled with appropriate dye, for example - Hoechst 33342 (1:15000). Acquire fluorescence microscopy images for cell nuclei, over-expressed GFP fusion clones and collagen-I.

Sample preparation for Solid phase transfection

Sample preparation for solid phase transfection in a form of small contact printed arrays on glass substrates was developed and reported previously[71,72]. The same preparation procedure of solid phase transfection is used for spin-coating of cell growth multi-well plates except the dilution of transfection mixture (step 4) 1:10 with dH_2O.

Preparation and fabrication of siRNA (gene loss of function) cell arrays

Preparation of the transfection solution for cell array:

1. 0.4M OptiMEM/sucrose: dissolve 1,37 g Sucrose in OptiMEM, total volume of mixture should be 10 mL.
2. Gelatine preparation: add 0.2 g gelatine in 100 mL water (see Note 2), place to water bath preheated to 56°C for 20 minutes. Cool down gelatine to room temperature (RT), then filter with sterile filter (0,45µm).
3. Add 1 µL of fibronectin to 100 µL gelatine.
4. Transfection mixture: mix 4,0 µL OptiMEM/sucrose with 3,0 µL Lipofectamine2000 and incubate for 15 minutes at RT. Add 2,0 µL cDNA (concentration of 500 ng/µL) and incubate for 15 minutes at RT. Add 7,0 µL fibronectin/gelatin mixture. Spin shortly in centrifuge at 1000 rpm.

Preparation and fabrication of cDNA over-expression cell arrays

Preparation of the transfection solution for cell array:

1. 0.4M OptiMEM/sucrose: dissolve 1,37 g Sucrose in OptiMEM, total volume of mixture should be 10 mL.
2. Gelatine preparation: add 0.2 g gelatine in 100 mL water (see Note 2), place to water bath preheated to 56°C for 20 minutes. Cool down gelatine to room temperature (RT), then filter with sterile filter (0,45µm).
3. Add 1 µL of fibronectin to 100 µL gelatine.

4. Transfection mixture: mix 3,0 µL OptiMEM/sucrose with 3,5 µL Lipofectamine2000 and 5 µL siRNA (30mM; 400ng/µL) incubate for 30 minutes at RT. Add 7,25 µL fibronectin/gelatin mixture. Spin shortly in centrifuge at 1000 rpm.

Chemicals and materials

Cell culture reagents:

DMEM - Dulbecco's Modified Eagle Medium GIBCO/Invitrogen Dulbecco's Modified Eagle Medium (1X) Cat. No. 31885-023

FBS - Foetal Bovine Serum Standard Quality PAA Laboratories GmbH Cat. No. A15-101

NCS – Newborn Calf Serum, Biochrom AG, Cat. No. S0125

NEAA – Non-Essential Amino Acids (100X) GIBCO/Invitrogen 11140-035

Pen/Strep – Penicillin Streptomycin GIBCO/Invitrogen 15140-122

L-Glu – L-Glutamine 200mM (100X) GIBCO/Invitrogen 25030-024

Trypsin-EDTA - GIBCO/Invitrogen Trypsin, 0.05% (1X) with EDTA, Cat. No. 25300-054

MEM - Minimum Essential Medium Eagle, Sigma-Aldrich M3024

Cell culture vessels and filters:

Petri Dish – Petri Dish (94x16 mm) CELLSTAR®, physical surface treatment, sterile, Greiner bio-one, No 633171

1 chamber labtek - Lab-Tek™ Chambered Coverglass, Thermo Fisher Scientific, No155361

8 chamber labtek - Lab-Tek™ Chambered Coverglass, Thermo Fisher Scientific, No155411

8 chamber µ-Slide - µ-Slide 8 well, Ibidi, No80826

96 well plate - µ-Plate 96 well, Ibidi, No 89626

MILLIPORE Steritop-GP Millipore Express PLUS (PES) 0.22µm, Millipore Corporation, Billerica, MA. SCGP T01 RE

Transfection reagents:

Effectene – Effectene Transfection Reagent, Qiagen, Cat. No 301425

Lipofectamine - Lipofectamine™ 2000 Transfection Reagent, Invitrogen Cat. No. 11668-019

Opti-MEM - Opti-MEM® I Reduced Serum Medium (1X), liquid - with GlutaMAX™ I, Invitrogen Cat. No. 51985-026

Cellular dyes:

Hoechst – Bisbenzimide Hoechst 33342, Sigma-Aldrich B2261

Polyclonal antibody against mouse collagen-I - Millipore, Temecula, CA, Cat. No. AB765P

Anti-GM130 - BD Biosciences (monoclonal mouse) 610823

Anti-GFP - Abcam (polyclonal rabbit) 6556

Alexa Fluor® 532 Goat Anti-Rabbit IgG (H+L) Cat. No. A-11009

Alexa Fluor® 647 Goat Anti-Rabbit IgG (H+L) Cat. No. A21244

Alexa Fluor® 647 Goat Anti-Mouse IgG (H+L) Cat. No. A21236

Other reagents:

Brefeldin-A - Brefeldin A, Eupenicillium brefeldianum, Calbiochem/Merck Chemicals Cat. No 203729

Nocodazole - R17934, Methyl-[5-(2-thienylcarbonyl)-1H-benzimidazol-2-yl]-carbamate, Calbiochem/Merck Chemicals Cat. No 487928

Cycloheximide - Cycloheximide solution, Sigma-Aldrich C 4859

HEPES - N-2-Hydroxyethyl piperazine-N'-2-ethane sulphonic acid, ROTH Art. 9105.2 EG-Nr. 2309079

Sodium chloride – AnalaR NORMAPUR analytical reagent, VWR Prolabo 27810.259

Hydro chloric acid - AnalaR NORMAPUR analytical reagent, VWR Prolabo 20252.290

Poly-L-lysine - Sigma-Aldrich P8920

Pure water (dH_2O) prepared by TKA 08.2207 UV-TOC/UF

TetraSpeck™ microspheres (0.1 µm) – Invitrogen T-7279

Small RNAs:

siRNAs or miRNAs were purchased from Ambion, Qiagen and Dharmacon.

Ambion short double stranded RNA molecules: MGAT1 siRNA ID 11481; DLST siRNA ID 111044; SNRPG siRNA ID 44849; ALDH4A1 siRNA ID 13635; AEBP2 siRNA ID 39787; CPB1 siRNA ID 104003; GABRQ siRNA ID 118966; GPR6 siRNA ID 42020; HBB siRNA ID 45933; ERBB3 siRNA ID 184; PCDHAC2 siRNA ID 27411; Silencer NC (negative control) AM4635; miR-17 AM12412.

Dharmacon short RNA molecules: miRIDIAN miR mimics miR30b (MIMAT0000420); miRIDIAN miR mimics miR30b-dy647 custom made; miRIDIAN miR hairpin inhibitor anti-miR30b (MIMAT0000420).

Qiagen siRNAs: AllStars Negative Control siRNA Cat #1027280; GBF1 SI01707678.

Modified λ_{N22} aptamers DNA vectors

Modified λ_{N22} aptamers DNA vectors were designed by Andrius Serva and prepared by Susanne Reusing:

TBC1D2_mRFP_boxB (ENST00000375066);

LDLR_mRFP_boxB (ENST00000558518);

mRFP_LDLR-3'UTR_boxB(ENST00000558518);

mRFP_TBC1D2-3'UTR_boxB (ENST00000375066);

mRFP_miR30b_boxB;

Computational Methods

Wide field fluorescence microscopy image analysis (Scan^R software)

An evaluation of a biological experiment is a complex task, not always result can be expressed in binary system as yes or no answer, numeric and quantitative evaluation is required. Fluorescence microscopy due to digital image acquisition is natively a numeric data collection method thus already simplifying evaluation process. Depending on a biological question, different data features might be extracted from recorded images. For screening experiments image analysis and interpretation is rather simplified, as the task usually is to observe population as large as possible and draw a yes or no conclusion. Numeric evaluation of cellular processes recorded on fluorescence microscopy images first of all requires cell based image analysis and data extraction; many different biologically relevant features might be extracted and exported as object based numeric data sets. For example, it could be the number of "bright" spots or their size. When fluorescence intensity signal is used as a reporter for the amount of molecule of interest image intensity values are integrated in the area of the cell, having in mind that cells differ in size, usually mean intensity values are used, integrated value divided by area of the object.

Here is described a method for evaluation of fluorescence microscopy images using Olympus Scan^R automated Image and data analysis software package (version 1.3.0.3) then fluorescence intensity is the main feature of interest. There are open source image analysis software packages available that could perform similar or even more intelligent tasks, but as images are acquired on Olympus Scan^R screening microscope, using mentioned package is simply more convenient and straight forward.

Image processing and analysis routine is segregated into main steps: uneven background correction, image segmentation in order to detect objects of interest, usually it is cell nuclei, then determination of area of interest (mask) and feature extraction. In Figure 8 an example of image analysis is shown: A - original cell nuclei fluorescence microscopy image (Hoechst 33342) with red rectangular shapes marking detected objects, B - software generated

image representing found objects, C and D example masks for fluorescence intensity integration of other fluorescence channels.

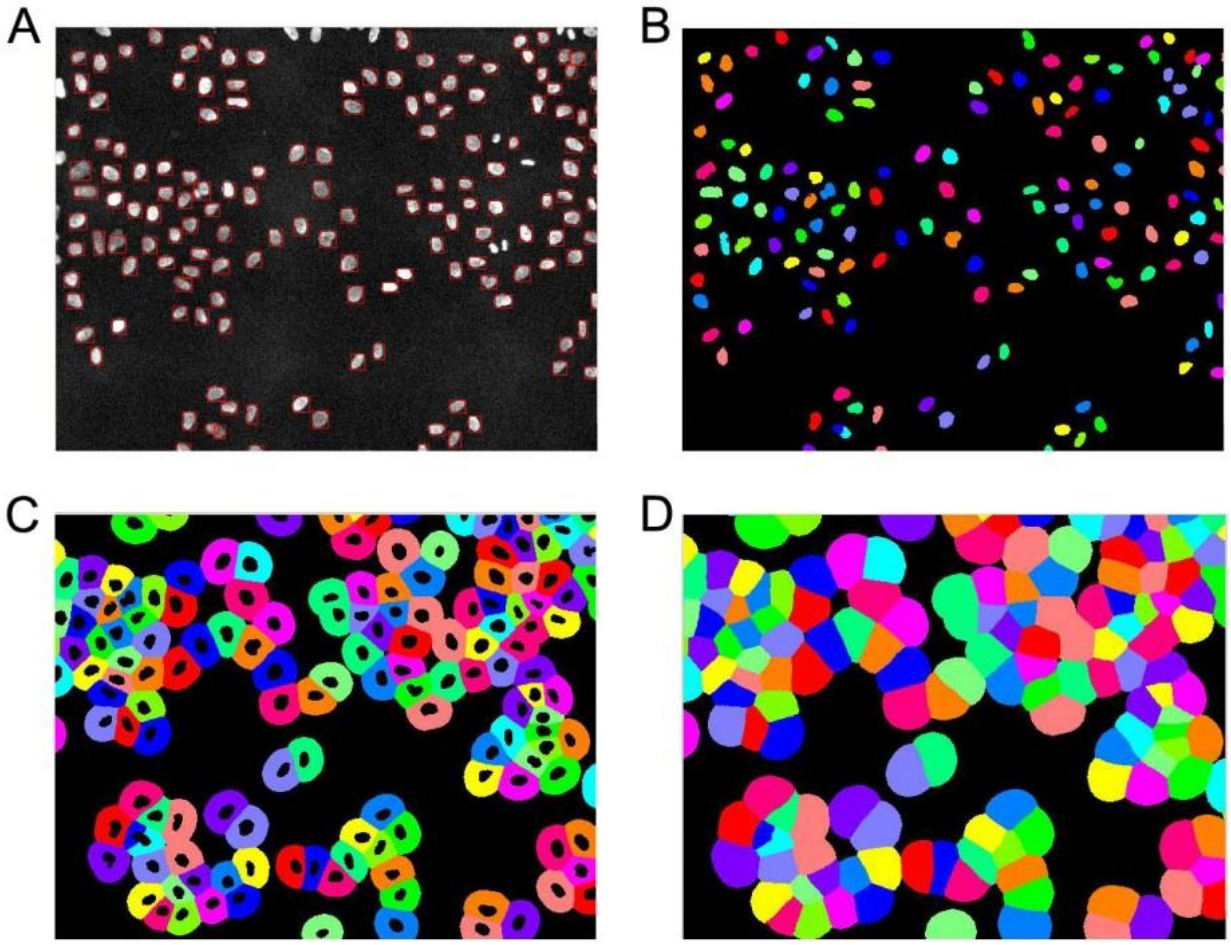


Figure 8 Example of object based image analysis (Olympus Scan^{R} analysis package). A) Original fluorescence image of Hoechst 33342 staining cell nuclei, B) Recognized objects based on fluorescence intensity features of Hoechst 33342 fluorescence signal, C and D examples of masks for object base evaluation of other fluorescence channels.

First of all images are processed by correcting uneven background using "rolling ball" algorithm: a local background value is determined for every pixel by averaging over a large ball around the pixel, this value is hereafter subtracted from the original image. The radius should be set to at least the size of the largest object that is not part of the background.

The objects are recognized using intensity threshold method, where pixels with intensities above a predefined threshold are united to one individual object. Here it is possible to apply object size mask, this allows detecting only objects within specified size (area) window.

Mask for each found object is created by drawing a new shape perpendicularly from the edge of the main object in a desired distance. Depending on a task, the mask may be created as donut. If two objects are too close and their masks overlap new masks are created with consensus border in the middle between the objects. Later this mask is used to perform measurements in selected fluorescence channels.

After image processing and object based measurements of microscopy images, numeric data is exported as table in ACSII format.

Wide field fluorescence microscopy image analysis for live cell Golgi kinetics evaluation

An automated image analysis routine was developed in collaboration with dr. Petr Matula (AG Rohr) and described in details here[73].

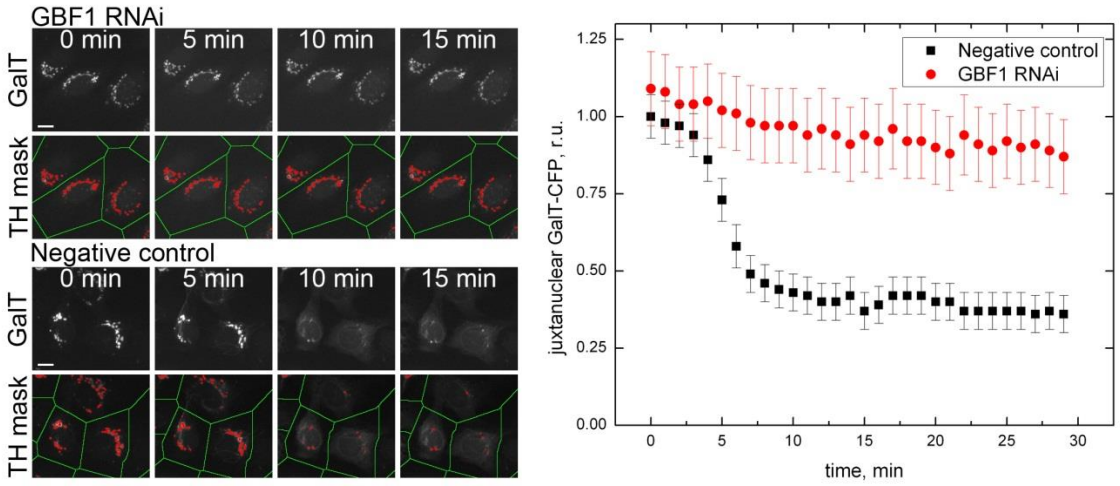


Figure 9 Quantification of the Golgi phenotype from fluorescence microscopy images. Examples of image analysis approach to quantify BFA-induced redistribution of Golgi-resident proteins. NRK-GalT-CFP cells were transfected with siRNA targeting GBF-1 and the negative control, and were incubated for 48h. After addition of 5 µg/mL of BFA together with 0.1 mg/mL of cycloheximide, images were acquired every minute and processed as described in[73]. A) TH mask is top-hat mask, green lines indicate segmented influence zones of individual cells, and red colour marks GalT-CFP positive structures. Scale bar=10µm. B) Quantification of BFA-induced redistribution of GalT-CFP. Relative units on y-axis represent the population average of GalT-CFP top-hat value in relation to first time point of control population. At each time-point, more than 50 cells were analysed. Error bars indicate standard deviation of the population.

A developed image analysis approach was applied to time series of fluorescence microscopy images, each time point containing three different fluorescence channels. Image analysis was performed in four steps:

(i) Detection and tracking of cell nuclei in the Hoechst channel
(ii) Determination of influence zones of detected cell nuclei at each time-point in the GalT-CFP channel
(iii) Segmentation and quantification of GalT-CFP structures in the GalT-CFP channel
(iv) Classification of transfected and non-transfected cells based on the information in the YFP-tagged protein channel.

For the detection and tracking of cell nuclei in the first step, the images of the Hoechst channel were smoothed using a Gaussian filter with a large standard deviation sigma and regional maxima in the blurred images were detected. Cell nuclei centres were detected by cluster analysis of the projected regional maxima along the time dimension. In the second step, influence zones of the detected cell nucleus centres were determined at each time-point by a marker-controlled watershed transform. Detected cell nucleus centres were used as markers and the influence zones were determined by combining information in the GalT-CFP channel and the distance from the markers. An example of computed influence zones is presented in Figure 9 (represented by green lines). To quantify the assembled GalT-specific structures in the third step, small bright structures in the GalT-CFP channel were segmented by applying a threshold to the result of an opening of a morphological top-hat transform of the GalT-CFP channel. The threshold was determined automatically by the triangle method. For all images and for all time-points the same threshold value was used. In Figure 9, the segmentation result is shown by the red colour in the overlay images. For each cell nucleus, the number of segmented pixels (corresponding to the area of the bright structures) was determined and denoted as its top-hat value. In the fourth step, a cell is classified as transfected if either its nucleus or the majority of segmented GalT-CFP structures co-localize with segmented YFP-positive structures. Images of the YFP-tagged marker were segmented by threshold followed by morphological opening and closing. The threshold was set automatically by the triangle method applied to the histogram of the logarithm of the image intensity values. The classification

precision of transfected and non-transfected cells of 92.8% was defined as TP/(TP + FP) and the recall precision of 94.4% was defined as TP/(TP + FN), where TP, FP and FN denote true positives, false positives and false negatives, respectively. The top-hat transformation was selected over other quantification measures as it gave largest dynamic range between assembled and disassembled Golgi.

The image analysis software program was written in MATLAB (Mathworks), using the MATLAB IMAGE processing toolbox, and the DIPIMAGE toolbox (TU Delft).

RNA secondary structure prediction

The RNA is not always playing a role of intermediate in information transfer process from DNA sequence to proteins. For many genes RNA is the final product, and like proteins, many of these RNAs fold into three-dimensional structures that have structural, catalytic, and regulatory roles in the cell[6], prediction of secondary and tertiary structure of RNA sequence is not only important to understand its function but as well for construction of well-defined 3D objects in bottom-up approach of nanotechnology[74].

Contrary to DNA, there two complementary sequences form stable complex, RNA transcripts are single stranded and secondary structure is determined by intramolecular base pairing. As secondary structure of RNA is governed by intramolecular base pairing formed secondary structure of the RNA polymer has some bases paired while others remain free and thus forming loops in the molecule.

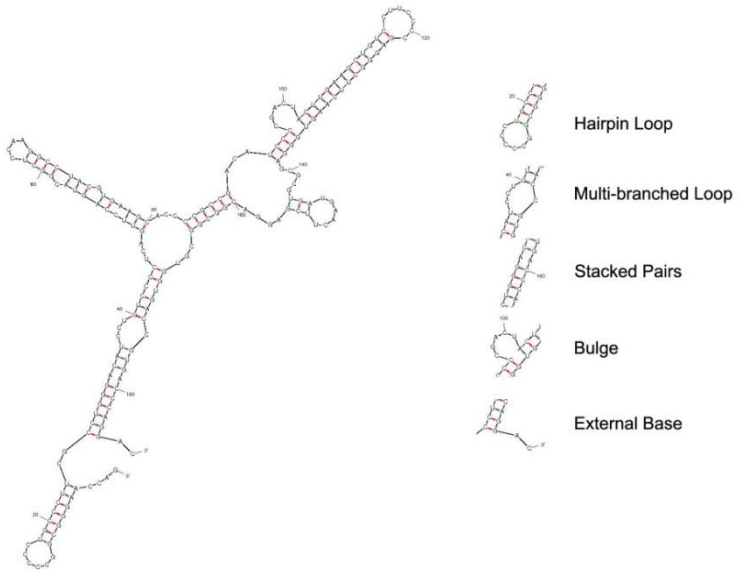


Figure 10 Illustration of loop formation in secondary RNA structure. The black lines indicate the backbone, and the color, red and blue, lines indicate paired bases. Illustration prepared by using Mfold predicted secondary RNA structure.

Formation of RNA secondary structure leads to formation of energetically favourable helical regions and energetically unfavourable loop regions. Such energetics leads to a vast combinatorial set of helix and loop arrangements of an individual RNA sequence[75]. Figure 10 illustrates stereotypical elements that form complete RNA secondary structure, such as stacked base paired stems, hairpin loops, bulges, multi branched loops, and external bases. Another possible and usually excluded form prediction algorithms structural element of secondary RNA structure is pseudo knot. An example of simple pseudo knot would be hairpin loop nucleotides base paired to other part of whole RNA[76].

Correct and accurate prediction of the secondary structure of a single RNA molecule is a difficult task. Basic algorithms suffer both from the simplifications of the underlying models as well as from the uncertainty concerning the exact values of the energy parameters. These problems lead to typical prediction accuracies (measured as the fraction of correctly predicted base pairs) of between 45 and 70%[77].

Mfold was the first prediction approach that modelled nucleic acids as a thermodynamic system assuming that the correct structure is the one with lowest Gibbs free energy or one of

the suboptimal structures. The free energy parameters used were derived from experimental thermodynamic studies of RNA and DNA[74].

The name ”mfold” describes a number of related software applications available for the prediction of the secondary structure of single stranded nucleic acids, letter “m” refers to “multiple”. The objective of mfold web service is to provide access to RNA and DNA folding and hybridization software to the scientific community[78]. The ‘mfold’ software for RNA folding was developed in the late 1980s and became available as Web service in the autumn of 1995, hosted by Washington University School of Medicine[78,79]. The core algorithm predicts a minimum free energy, ΔG, as well as minimum free energies for foldings that must contain any particular base pair. Any base pair, r_i-r_j, between the i^{th} nucleotide and the j^{th} nucleotide that is contained in a folding no more than δδG from the minimum, is plotted in a triangular plot called the “energy dot plot”. The base pair r_i-r_j is plotted in row i and column j of this matrix. The free energy increment, δδG, is chosen a priori by the user, who selects a “percent suboptimality”, P. From this, δδG is computed to be P/ 100 |ΔG|. Base pairs within this free energy increment are chosen either automatically, or else by the user, and foldings that contain the chosen base pair are computed. They have minimum free energy conditional on containing the chosen base pair.

The UNAFold software package is an integrated collection of programs that simulate folding, hybridization and melting pathways for one or two single-stranded nucleic acid sequences. The name is derived from “Unified Nucleic Acid Folding”. Folding (secondary structure) prediction for single-stranded RNA or DNA combines free energy minimization, partition function calculations and stochastic sampling. The UNAFold software predicts nucleic acid foldings, hybridizations and melting profiles using energy methods and a general computational technique known as dynamic programming. Early software for RNA folding predicted minimum free energy foldings only. Such methods were unreliable in the sense that many different foldings, with free energies close to the computed minimum, could exist. Although constraints deduced from experiments or phylogenetic analyses could be applied to reduce uncertainty, a method to compute a variety of close to optimal foldings was needed. UNAFold encompasses all three methods by computing minimum and suboptimal foldings in

the mfold style, and full partition functions that allow both exact base pair computations as well as stochastic sampling. It is important to emphasize that the computations used by the mfold and UNAFold software are based on a number of assumptions.

Basics of applicable statistics

Statistics is the practice or science of collecting and analysing numerical data in large quantities, especially for the purpose of inferring proportions in a whole from those in a representative sample[80]. For data analysis it is crucially important to determine the distribution of acquired data as the same statistical operators for different statistical distributions may have different meaning. There is a lot of literature available covering various statistics applications. As my aim here is to describe very basic concepts, current section is mainly based on later resource[81].

Random quantity distribution

Let say, we are observing an experiment, random quantity, and for each observation we assign a value. To describe our random quantity we need to describe range there values of our random quantity fall, but just the range is not enough. We need to have description of how often we observe one value and how often the other.

Relation between random quantity value and its probability is called random quantity distribution.

Random quantity X, x_i is its value and p_i corresponding probability. Then discrete random quantity distribution may be visualized graphically, shown in a table or expressed analytically.

It is clear that

$$\sum_{i=1}^{n} p_i = 1 \qquad [6]$$

n stands here for a number of possible values of random quantity X. Equation describes distribution normalization condition.

Distribution completely describes statistical characteristics of discrete random quantity; it might be described in other ways, but they will be equal to distribution.

For continues quantities a distribution describing discrete random quantity is useless, thus another form describing distribution is required. Universal characteristic, suitable for discrete and continues quantities, is called distribution function. Function *F(x)* is called a

distribution function (or integral probability distribution), if it describes corresponding probability P, that random quantity X will acquire values lower than x:

$$F(x) = P(X < x) \qquad [7]$$

Probability for random quantity to fall in the range between x_1 and x_2 is equal to difference distribution function at those points:

$$P(x_1 \leq X < x_2) = F(x_2) - F(x_1) \qquad [8]$$

this quality is valid for both discrete and continues distributions.

Another usually used distribution characteristics is probability density. Probability density (function) *w*(*x*) multiplied by small interval Δ*x* is equal to probability of random quantity to fall in interval between *x* and *x*+Δ*x.*

$$w(x) = \frac{dF(x)}{dx} \qquad [9]$$

Probability density function has these qualities:

- It is non-negative as for any *x* values, *F*(*x*) is not decaying
  - Probability that continues random quantity will fall into interval between x_1 and x_2 is equal

$$P(x_1 \leq X < x_2) = \int_{x_1}^{x_2} w(x)dx \qquad [10]$$

- Probability density integral in infinite interval is equal to one:

$$\int_{-\infty}^{\infty} w(x)dx = 1 \qquad [11]$$

this equation is called as probability density normalization condition.

Sometimes random quantity is described without using distribution or probability density functions but rather using other numeric characteristics of the distribution (also called moments). Simplest and most common characteristics are statistical mean and dispersion of the distribution.

Statistical mean of the random quantity is a value around which values of arithmetic mean are grouping, when number of repetitions is suitably large.

$$M(X) = \sum_{i=1}^{n} x_i p(x_i) \qquad [12]$$

$$M(X) = \int_{-\infty}^{\infty} x w(x)dx \qquad [13]$$

Some qualities of the mean:

- Random quantity multiplied by constant mean will be equal to random quantity's mean multiplied by the same constant
- Random quantity's sum or difference theoretical mean is equal to their sum or difference
- Mean of product of independent random quantity's is equal to product of their means

In statistics random quantity dispersion is used to describe spread of random quantity. Dispersion for discreet random quantity is expressed as:

$$\sigma_x^2 = \sum_{i=1}^{n}(x_i - m_x)^2 p_i \qquad [14]$$

And for continues random quantity expressed as integral:

$$\sigma_x^2 = \int_{-\infty}^{\infty}(x - m_x)^2 w(x)dx \qquad [15]$$

Another frequently used characteristic is square root of dispersion; it is called as standard deviation.

Generalized distribution characteristic is called moment. Moments are segregated in two classes: primary moments and central moments. Primary moments are expressed as:

$$m_k = M(X^k) = \sum_{i=1}^{n} x_i^k p_i \qquad [16]$$

$$m_k = M(X^k) = \int_{-\infty}^{\infty} x^k\, w(x)dx \qquad [17]$$

Central moments are expressed as:

$$\mu_k = M\left(\overline{X}^k\right) = \sum_{i=1}^{n}(x_i - m_x)^k p_i \qquad [18]$$

$$\mu_k = M\left(\overline{X}^k\right) = \int_{-\infty}^{\infty}(x - m_x)^k w(x)dx \qquad [19]$$

According to equations random quantity mean is first order primary moment and dispersion is second order central moment.

Normal (Gauss) distribution

Normal (Gauss) distribution is very common and is used in many areas. According to central limit theorem - at given certain conditions, the mean of a sufficiently large number of independent random variables, each with finite mean and variance, will be approximately normally distributed.

Probability density for "normally" distributed random quantity is expressed as:

$$w(x) = Cexp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \qquad [20]$$

m, σ and C are parameters describing normal distribution. Distribution parameter C could be found by applying normalization condition and solving integral; then $is\ equal\ to$:

$$C = \frac{1}{\sigma\sqrt{2\pi}} \qquad [21]$$

Now normal distribution could be written as:

$$w(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \qquad [22]$$

Parameter *m* represents the mean, and σ^2 - the dispersion of the normal distribution; if one has observed mean and knows its dispersion (or standard deviation), distribution density can be reconstructed.

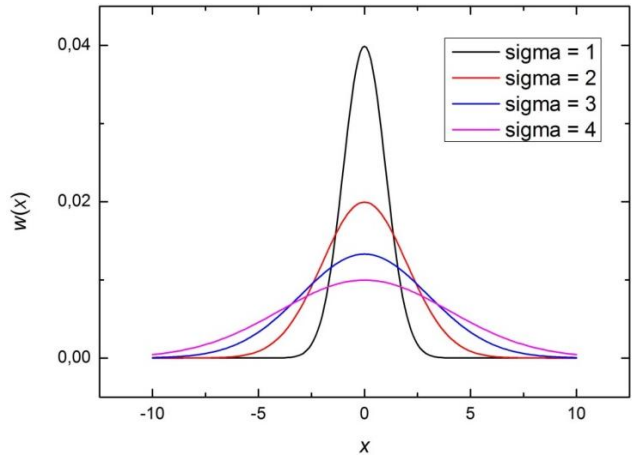


Figure 11 Normal distribution probability density function dependency on standard deviation.

In Figure 11 represents normal distribution probability density function dependency on standard deviation, when mean is set to zero. Normal distribution probability density function is symmetric; probability to observe random quantity in a distance $\pm\Delta x$ from *m* is equal.

It is important to note that for normal distribution the mode, the most probable value, the mean and the median, numerical value separating the distribution in halves, are equal to the parameter m.

Gamma distribution

A gamma distribution is a general type of statistical distribution that arises naturally in processes for which the waiting times between Poisson distributed (independent events described by discrete probability distribution) events are relevant[82]. The gamma distribution is a continuous probability distribution described by two parameters representing shape (*k*) and scale (*Θ*).

Gamma distribution probability density function for all *k*>0, *Θ* >0, and *X*>0:

$$w(x) = \frac{1}{\theta^k \Gamma(\alpha)} x^{k-1} \mathrm{e}^{-\frac{x}{\theta}} \qquad [23]$$

where $\Gamma(k) = (k-1)!$

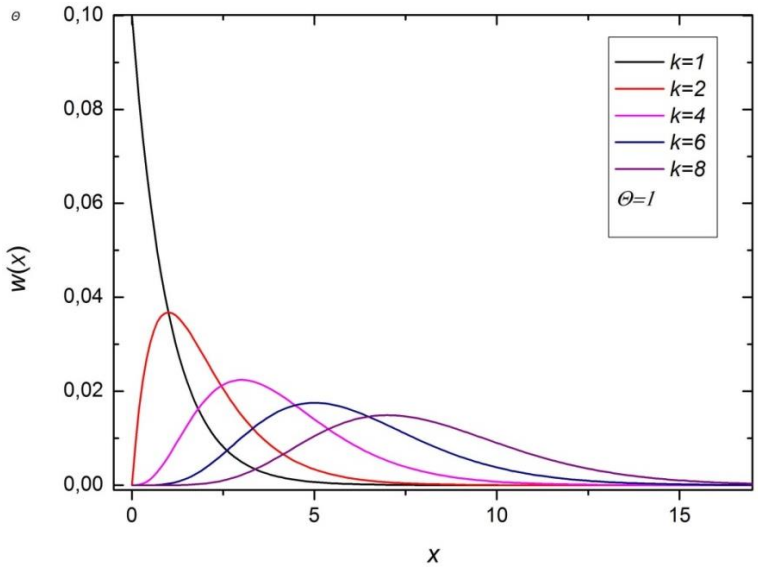


Figure 12 Gamma distribution probability density function dependency on shape parameter *k, when scale parameter Θ is fixed and equals to one*.

Figure 12 represents gamma distribution density function dependency on shape parameter *k*, where with high values of α gamma distribution is close to a normal distribution.

The mean and dispersion of gamma distributed quantity can be expressed by distribution parameters:

$$M(X) = k\theta \qquad [24]$$

$$\sigma^2 = k\theta^2 \qquad [25]$$

Distribution mode is equal to $(k-1)\theta$, for k>1.

Unlike the normal distribution, the mode, the mean and the median in gamma distribution are not equal. An application of normal distribution moments on the data which is distributed according to gamma distribution would lead to false conclusions.

Evaluation of the influence of over-expressed membrane proteins on the organization kinetics of the Golgi complex

Initial background for the development of an assay

Various biological processes have different time interval and period; some processes produce detectable phenotypes in time scale of days, then other processes are relatively fast and for observation require sampling rate of minutes or even seconds. When screening microscopy is selected as a tool of investigation, the throughput of the system is one of the main factors for designing the experiment as it limits the number of conditions being tested and sampling rate. The purpose for a live cell observation of a membrane trafficking event instead of observation of fixed time points before and after the event was supported by the idea that it should be possible to identify effectors which could not be revealed by other biochemistry techniques for example due to the different interaction mechanism.

Throughput limiting factors are: velocity of sample positioning, time required for autofocus procedure, number of different fluorescence channels (excitation-emission ranges), as it is necessary to switch between different filters, exposure time, time and resources required for computer to generate digital image and store it. By our collaborative partners (AG Fisher, Uni-Mannheim) many of these factors were thought to be reduced by redesigning Olympus Scan^R fluorescence screening microscope; introducing parallel excitation and image acquisition of different fluorophores; optimizing algorithm of sample positioning and autofocus procedure. It was decided that a biological process has to be evaluated on the experimental screening microscope as a test of made improvements. Initial conditions for biological experiment were suggested by the theoretical hardware limitations: up to 96 different positions with sampling rate of one minute, and up to three fluorescence channels, with autofocusing once at the start of an experiment.

The investigation of the Golgi complex organization was chosen due to a number of factors such as its major role in membrane trafficking, relatively easy visualization, reliable induction of morphology changes by a number of different chemicals in a time frame of 5 to 20 minutes.

The Golgi complex and the basis for assay development

The Golgi complex fulfils multiple functions that are essential for growth, homeostasis and division of eukaryotic cells. The Golgi complex merges and coordinates secretory, endocytic and recycling protein transport pathways. Newly synthesized proteins and lipids are modified by abundance of Golgi-resident enzymes. Over the last decades the morphology of the Golgi complex has been investigated in great detail by different approaches[83–94]. A number of cell biological studies revealed that the structure of the Golgi complex is highly dynamic and is crucially dependent on the balance of material flow[95–97]. A comprehensive list of proteins residing at and functionally related to the Golgi complex is being currently compiled[98–102]. Research on basic Golgi functions has profited much from application of various drugs that interfere with maintenance of the Golgi structure and/or with its function in protein transport. Despite a wealth of knowledge on the function and structure of the Golgi complex, the information was largely obtained by studies in fixed cells. As a result, factors and mechanisms underlying dynamics of this organelle are little understood.

First observations of the Brefeldin-A (BFA) effect on the organization of the Golgi complex were reported by Toshiyuki Fujiwara *et al*. in 1988[103] and seconded by Jennifer Lippincott-Schwartz *et al.* in 1989[104]. In 1997 the first live cell microscopy experiment studying BFA treatment effect on Golgi redistribution to the ER was reported by Noah Sciaky *et al*.[105], there confocal microscopy was used to observe disappearance of bright Golgi compartments with evaluated halftime of ~4min. A fixed cell experiment was reported by Heling Pan *et al*.[106] in 2008, this study is worth to mention due to an application of automated fluorescence microscopy image analysis approach.

NRK (normal rat kidney) cells stably transfected with the Golgi enzyme GalT, tagged with CFP (NRK-GalT-CFP), were chosen for the development of the approach as the efficient redistribution of GalT-CFP to the ER after the addition of BFA and reassembly after BFA wash-out was previously reported for these cells in a number of studies[107–109]. The goal was to measure GalT-CFP dynamics on a single cell basis in an automated assay in living cells. An over-expression of YFP (yellow fluorescent protein) tagged ER and/or the Golgi complex resident proteins was selected as gene amount altering method.

Experiment set-up

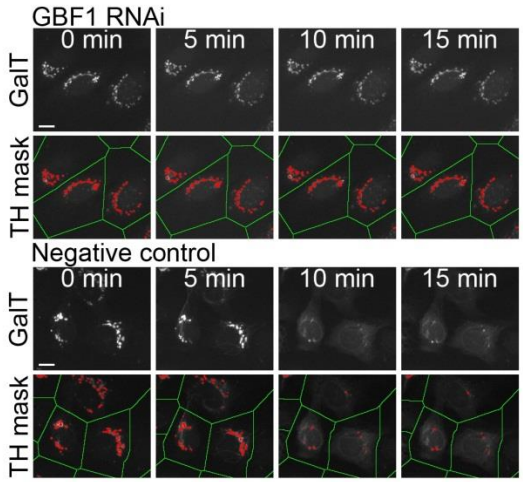
All experiments were performed on standard Olympus Scan^R screening microscope (based on inverted Olympus IX 81), which allowed to acquire 10 different positions in three fluorescence channels (Hoechst 33342, excitation- emission max at λ - 350/461 nm, for labelling cell nuclei; CFP, excitation- emission max at λ - 435/475 nm, for trans-Golgi resident GalT tagged with cyan fluorescent protein; YFP, excitation- emission max at λ – 515/530 nm , for over-expressed test proteins tagged with yellow fluorescent protein) within one minute. Experiments were performed in 8 well microscopy plates, Nunc Lab-Tec or Ibidi µ–slide 8.

Cells used to be plated 48h before imaging experiment; 24h before experiment cells used to be transfected with plasmid encoding protein of interest. 15min before experiment cells used to be stained by Hoechst 33342. All experiments were performed under Cycloheximide treatment (0,1 mg/ml; 0,355 mM) to block translation of new proteins.

BFA treatment (5 µg/ml; 17,8 µM) experiments were performed on the microscope in environmentally controlled conditions (37°C) using imaging media, imaging two wells at a time with one minute sampling rate. BFA wash-out experiments cells used to be treated with BFA for 30 min without imaging, then BFA used to be washed-out by exchanging medium for three times. Standard growth medium was used instead of imaging medium to keep cells vital for longer interval of time. 5 images per well used to be imaged with autofocus procedure for each time point, with sampling interval of 10 min for 3h.

Data analysis and interpretation

A



B

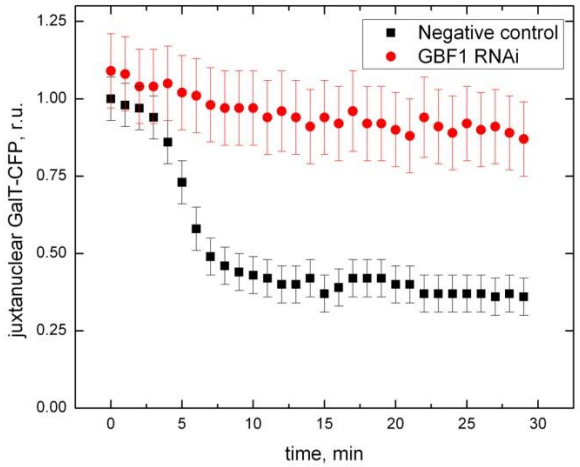


Figure 13 Evaluation of the Golgi phenotype from fluorescence microscopy images. Examples of image analysis approach to quantify BFA-induced redistribution of Golgi-resident proteins. NRK-GalT-CFP cells were transfected with siRNA targeting GBF1 and the negative control, and were incubated for 48h. After addition of 5 µg/mL of BFA together with 0.1 mg/mL of cycloheximide, images were acquired every minute. A) TH mask is top-hat mask, green lines indicate segmented influence zones of individual cells, and red colour marks measured GalT-CFP structures. Scale bar=10µm. B) Evaluation of BFA-induced redistribution of GalT-CFP. Relative units on y-axis represent the population average of GalT-CFP top-hat value in relation to first time point of control population. At each time-point, more than 50 cells were analysed. Error bars indicate standard error of the mean for observed population.

An automated image analysis routine was developed in collaboration with AG Rohr (Dr. P. Matula) (see **Figure 13.A**), which may be segregated into four major steps: 1) Detection and tracking of single cell nuclei stained with Hoecht 33342, 2) Partitioning of the images into disjoint regions, each encompassing a cell nucleus, called influence zones, 3) Quantification of GalT-CFP structures defined by top-hat transformation for each influence zone, 4) Classification of cells expressing YFP-tagged candidate proteins versus non-expressing cells based on information in the YFP channel[73]. Evaluation of the algorithm for the detection and tracking of cell nuclei in time based on two image time series with around 100 cell nuclei per image was performed and resulted, that in total 98.1% of the cell nuclei were correctly detected and tracked. The algorithm failed only for cell nuclei that were very close to each other, for example, multinucleate cells. Evaluation of the second part of our approach, namely image

partitioning into influence zones, was based on 7 images from different experiments with 104 cell nuclei on average. 98% of the cells GalT-CFP specific structures were correctly assigned to the nucleus. In **Figure 13.B** an example of the Golgi area quantification for two different biological conditions is represented, RNA interference and down regulation of GBF1, primary known BFA target, resulted a little to none response of the cell population upon the treatment by BFA, while the control population shows short latency and rapid reduction of the Golgi area.

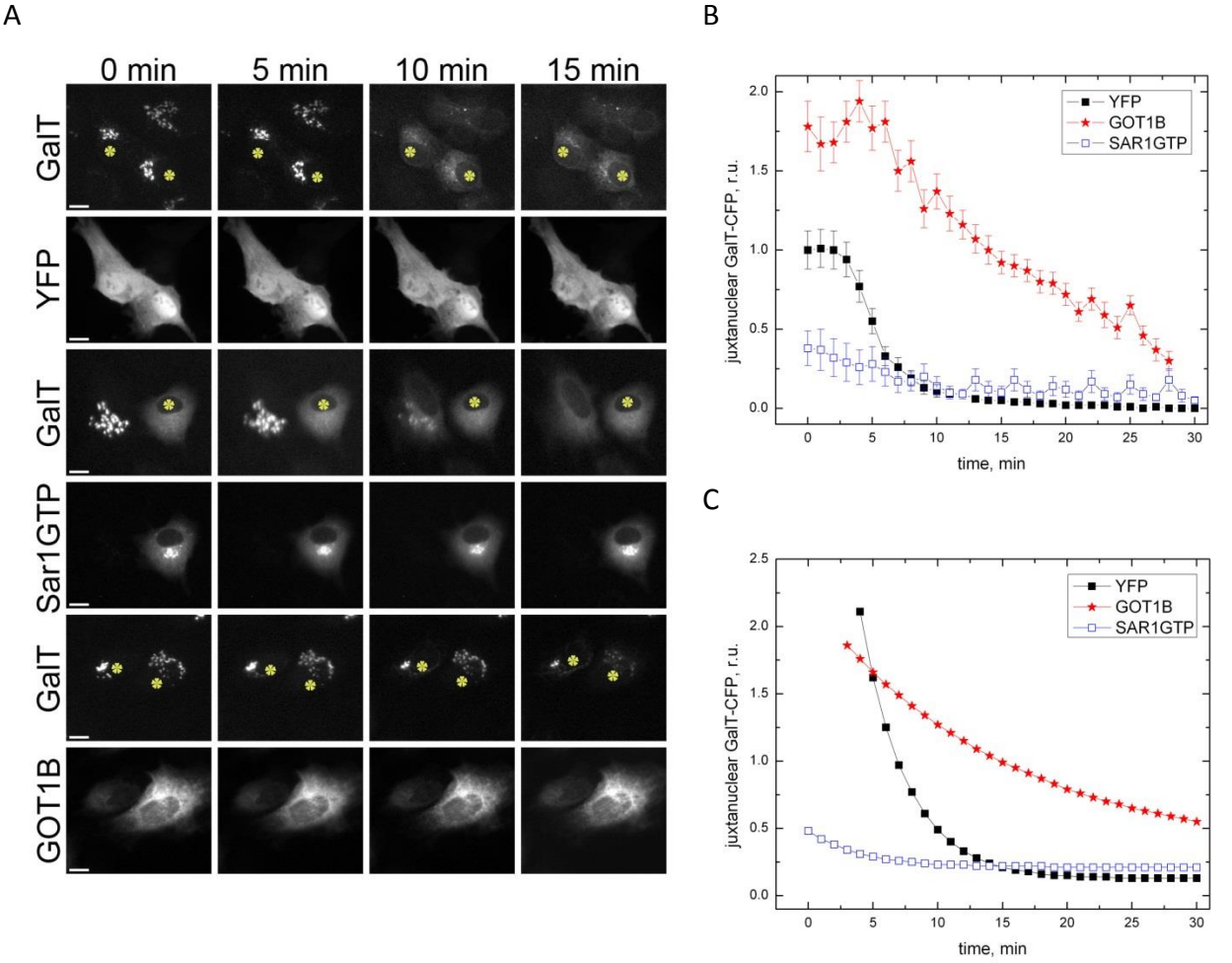


Figure 14 Over-expression of YFP tagged proteins can suppress GalT-CFP redistribution upon the addition of BFA. A. Examples of GalT-CFP measured area response to BFA addition under conditions of over-expression of YFP, SAR1GTP and GOT1B. Yellow asterisks indicate expressing cells, scale bar = 10µm. B. Quantification of BFA induced GalT-CFP responses. Relative units on Y axis represent average top-hat value of a juxtanuclear GalT-CFP, error bars represent standard error of the mean. C. Experimental data were fitted to the exponential decay function; fitting parameters were collected, averaged, and used to determine the average decay curves.

Data from all BFA treatment experiments essentially had the same trend – short delay of 3-5 min, which could be called "BFA lag time", with no significant changes in the Golgi area (Top-Hat values) and following it near exponential decrease of GalT-CFP area representing the Golgi complex. We could not accurately measure BFA lag for all proteins in one experiment. Therefore, we excluded calculations of the lag time with the candidate proteins from further analysis. Soluble YFP expression (always imaged as the first clone in all experiments) revealed a BFA lag time of 3.25±0.9 min. That fits well with the BFA lag time previously reported, ~ 4 min, for GalT-GFP in HeLa cells[107]. For the evaluation of BFA induced redistribution of GalT-CFP, the average top-hat value for each cell expressing protein-YFP at each time point was calculated and normalised to the average top-hat value of the time point 0 min of cells expressing soluble YFP (relative units of top-hat). In the case of down-regulating GBF-1 by RNAi (**Figure 13**) the average top-hat value of cell treated with siRNA was normalized to the average top-hat value of the time point 0 min of cells transfected with the negative control. The data was plotted against time and used for fitting to an exponential decay function ($A(t) = y_0 + A_0 e^{-\frac{t}{t_1}}$), where A_0 is the initial amount of GalT-CFP specific fluorescence, t_1 is the exponential time constant, y_0 is the offset value, *A(t)* denotes the amount of GalT-CFP specific fluorescence at time *(t)*. The fitting parameters (y_0, t_1, A_0) were collected and averaged for the corresponding experiments. As seen in **Figure 14.A** over-expression of YFP leads to virtually no changes (upper panel), over-expression of SAR1GTP induces relocation of GalT-CFP to the ER prior the addition of BFA (middle panel), over-expression of GOT1B hinders GalT-CFP relocation to the ER when BFA is added (lower panel). In **Figure 14.B** quantification of BFA induced GalT-CFP responses are plotted; relative units on Y axis represent average top-hat value of the GalT-CFP structures for observed population, error bars represent standard error of the mean for observed population. Top-hat values were normalised to the average top-hat value of cell expressing YFP at 0 min. In **Figure 14.C** average exponential decay functions for each over-expressed protein derived from collected fitting parameters are presented.

Data from all BFA-washout experiments had similar trend – relatively fast growth reaching maximum value and consequent slow decay. The average top-hat values for all cells expressing the protein-YFP at each time point were calculated. Then the average top-hat value

of the time point 0 min was subtracted from each data point, assuming that the starting point for assembly is the same for every candidate protein. This allowed us to compare assembly rates of different experiments directly.

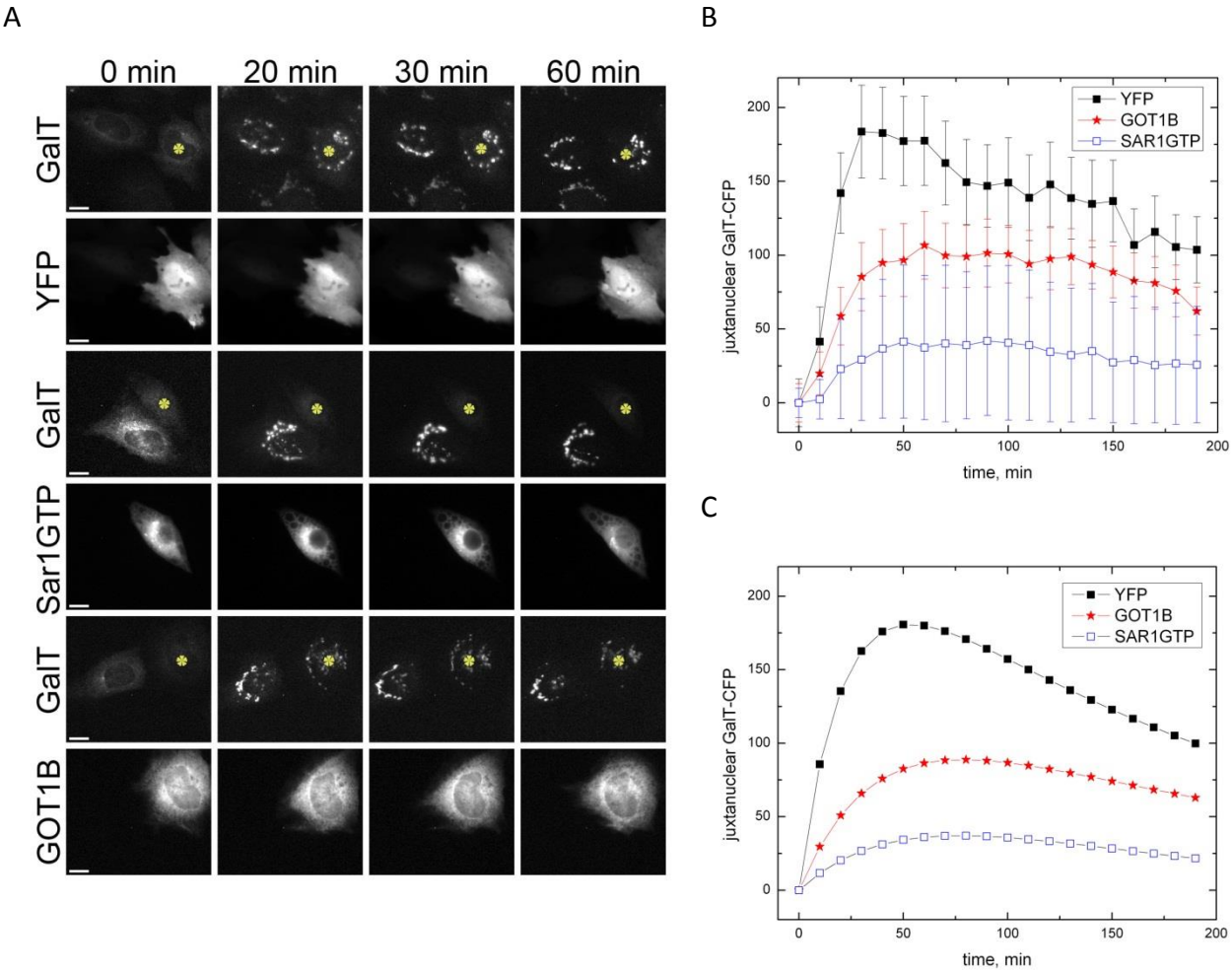


Figure 15 Over-expression of YFP tagged proteins alters GalT-CFP reassembly upon the wash-out of BFA. A. Examples of GalT-CFP response to BFA wash-out under conditions of over-expression of YFP, SAR1GTP, GOT1B. Yellow asterisks indicate protein expressing cells, scale bar = 10µm. B. Quantification of GalT-CFP responses to BFA wash-out. The curves represent experimental data. Y axis shows the GalT-CFP expressed as averaged top-hat values. Error bars represent standard error of the mean. C. Experimental data were fitted to the model, describing consecutive elementary reactions, fitting parameters were collected, averaged and used to determine time of the top-hat maximum of the GalT-CFP structures after BFA wash-out as well as to determine average curves.

The data was plotted against time and fitted to the model describing the amount of the intermediate product in consecutive two steps elementary reactions[61], where k_a describes

formation rate of intermediate product, k_b describes decrease rate of intermediate product, A_0 is the amount of initial product, I is GalT-CFP area measured by Top-Hat:

$$I = \frac{k_a}{k_b - k_a}(e^{-k_a t} - e^{-k_b t})A_0 \text{ [26]}$$

The fitting constants k_a and k_b were collected and averaged for every clone. Positions of GalT-CFP fluorescence intensity maxima were determined by finding extrema of the fitting function using the collected fitting constants.

In **Figure 15.A** it is shown that over-expression of YFP leads to virtually no changes (upper panel) over-expression of SAR1GTP nearly completely blocks relocation of GalT-CFP to the Golgi complex (middle panel), over-expression of GOT1B slows-down GalT-CFP relocation to the Golgi complex when BFA is removed (lower panel). In **Figure 15.B** a quantification of GalT-CFP structures after BFA wash-out is plotted; units on y axis represent average top-hat value of the GalT-CFP structures for observed population, error bars represent standard error of the mean for observed population. In **Figure 15.C** an average curves based on data fitting to the model of consecutive chemical reactions are presented; fitting parameters for each over-expressed protein were collected, averaged, and used to determine the average curves and time when top-hat value reaches its maximum.

Conclusion

We developed a fluorescence microscopy based approach that allows evaluation of regulators of ER to Golgi trafficking and/ or maintenance of the Golgi complex integrity in living cells; our experiment is the first where Golgi relocation kinetics was approached in live cells in over-expression conditions of fluorescently tagged proteins. The approach is based on the measurement of the Golgi area (defined by a morphological Top-Hat transformation) labelled by enzyme GalT (beta1,4-galactosyltransferase I) tagged with CFP in conditions of Golgi to ER redistribution and reassembly induced by addition and wash-out of Brefeldin-A (BFA). Results of the proof of principle screen identified a number of proteins that have significant effect on the change of the rate of GalT-CFP redistribution and reassembly when over-expressed; and the majority of them were missed in functional analysis of fixed cells[110].

Statistical data analysis of RAB GTPases over-expression on collagen-I secretion

The collagen biosynthesis involves a number of post-translational modifications and requires at least nine ER-resident enzymes and molecular chaperones[28]. Five collagen specific enzymes are required: three collagen hydroxylases, two collagen glycosyltransferases[27]. The fibril-forming collagens are first synthesized as procollagen molecules that have propeptide extensions at both their N- and C terminal ends. After cleavage of the signal peptides number of modifications occurs: hydroxylation of certain proline and lysine residues to 4-hydroxyproline, 3-hydroxyproline and hydroxylysine, glycosylation of some of the hydroxylysine residues to galactosylhydroxylysine and glucosylgalactosyl-hydroxylysine, glycosylation of certain asparagine residues in one or both of the propeptides, association of the C propeptides through a process directed by their structures, and formation of intrachain and interchain disulphide bonds. Protein disulphide isomerase (PDI) assists protein folding in the ER by catalysing the formation, reduction and isomerisation of disulphide bonds. As a subunit of the enzyme prolyl 4-hydroxylase, PDI is also essential for proline hydroxylation, the key post-translational modification underpinning helix formation. The main function of PDI in prolyl 4-hydroxylase seems to be to keep the highly insoluble a subunits in a catalytically active, non-aggregated state. Prolyl 4-hydroxylase requires ferrous ions, 2-oxoglutarate, O_2, and ascorbate[27]. Strong PDI functional dependency on ascorbate makes the procollagen-I very good secretory cargo for the well-controlled experimentation, as in the cells grown without ascorbate procollagen-I is trapped in the ER[37].

A family of RAB GTPases is known as controlling membrane trafficking. Despite differences in amino acid sequence the RAB family shares common higher order structure: RAB GTPases consists of a six stranded β sheet, comprising five parallel strands and one antiparallel one, surrounded by five α helices[24]. RAB GTPases utilize guanine nucleotide exchange and GTP hydrolysis to switch between active (GTP-bound) and inactive (GDP-bound) conformations. The molecular switch function of a small GTPase is carried out by cycling between its active GTP-bound form and inactive GDP-bound form. Switching from the inactive to active state is accomplished by replacing bound GDP with GTP, which requires GEF (GDP/GTP exchange

factor). After performing functions in their active form, GTPases are inactivated by hydrolysing GTP to GDP, which is accelerated by the GAP (GTPase-activating protein). Once inactivated, small GTPases detach from the membrane and are kept in the GDP-bound inactive state until the next round of the GTPase activation cycle begins. Dissociation of most RAB GTPases from membranes is mediated by a conserved protein family, the RAB GDP dissociation inhibitor (RAB GDI) which binds only to GDP-bound RABs and keeps them in the GDP-bound state by inhibiting GDP release.

Our goal was to identify the RAB family members having a role in the trafficking of a procollagen-I. For this task we used library of human GFP tagged RAB GTPases (kind gift from B. Gaud, Institute Currie, Paris) and NIH-3T3 cell line as it is known for high native procollagen-I expression levels. In total, we tested 178 different clones. Experiments were performed as follows: 1) cells are plated on multi-well cell culture plate, 2) transfected with according cDNA vector 24h before trafficking experiment, 3) cells are incubated (chased) with ascorbate and cycloheximide for 105 min before fixation with paraformaldehyde, 4) last step is immuno staining and microscopy image acquisition using 10x air objective. Microscopy images were analysed on a single cell basis using Olympus Scan^R automated image analysis software. It should be noted that we actually measured procollagen-I retained in cells rather than secreted one, due to this it is better to describe the experimentation as procollagen-I retention assay.

Data analysis and interpretation

While there is no standard way to interpret biological screening data, data normalization according standard deviation of the negative (neutral) control is the mostly used method to interpret RNAi (RNA interference) screens. For our experimental data such population normalization was a not sufficient method. Two main factors supporting this claim are: first, the probability to observe cell over-expressing fluorescently tagged protein decreases exponentially towards higher mean fluorescence values of over-expressed protein and such probability varies from one clone to the other, second, the distribution of procollagen-I mean fluorescence values in a population is not normal but is in good agreement with gamma distribution. In **Figure 16.A** average distributions for the procollagen-I values in control populations are shown, average distributions are result of averaged parameters derived from

more than 30 independent experiments, for each experiment procollagen-I values were scaled according mean value of the negative control populations (CFP).

A B

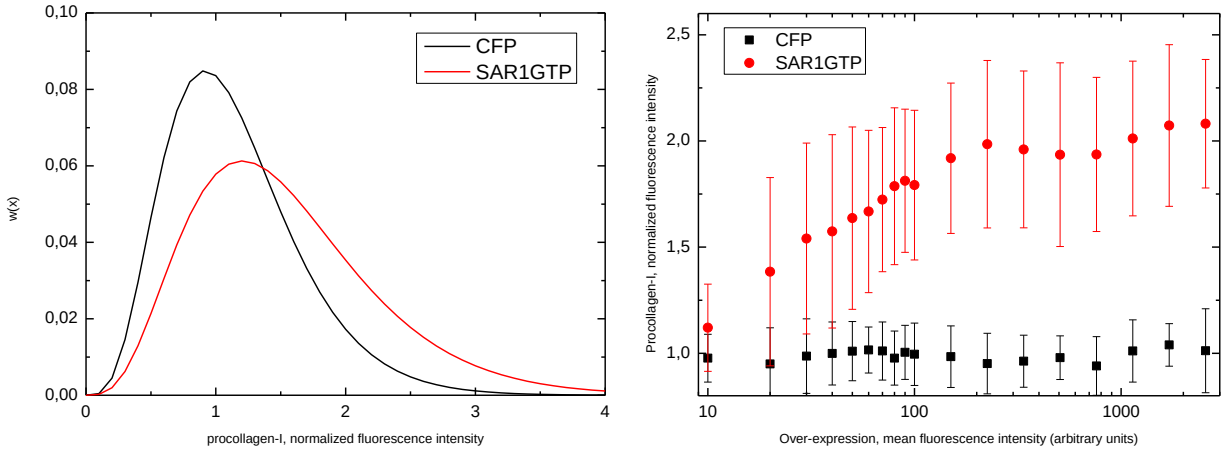


Figure 16 Cell based distribution of procollagen-I values in control populations; CFP (cyan fluorescent protein) is assumed as having no influence on procollagen-I secretion. SAR1GTP is a strong effector of ER to Golgi trafficking. A) Average Gamma distributions of the control populations; average distribution parameters collected from more than 30 individual experiments B) Sub-population based analysis of over-expression influence on procollagen retention in cells. Error bars represent standard deviation of more than 30 individual experiments.

Due to these reasons, another data interpretation approach had to be developed. As over-expressed clones were reported by fussed GFP, it was possible to segregate observed populations according mean fluorescence values of over-expressed proteins. Mean fluorescence values of over-expressed clones were lying in the interval from 0 to 3000. Interval from 0 to 100 was segregated into 10 equidistant ranges. Cells with values over 100 were segregated into 9 geometrically increasing ranges as cell numbers were getting low; a minimal threshold of 20 cells per range was used. That allowed plotting of retained procollagen-I in dependence on relative amount of over-expressed protein. That confirmed the assumption that over-expression of negative control (CFP) has little to none effect on procollagen retention in the cells (**Figure 16.B**). In contrast, over-expression of Sar1GTP clone had statistically not significant increase of retained procollagen-I at low levels (background) of over-expressed protein and then fast growth till procollagen-I saturation is reached at about twice the value of the negative control population. Additionally, data scaling according the procollagen-I mean value of negative control populations for separate experiments was used, that allowed to

reduce day to day variations of the acquired data (evaluated as a value of standard deviation) from about 70% of the value to 10% when scaling was used.

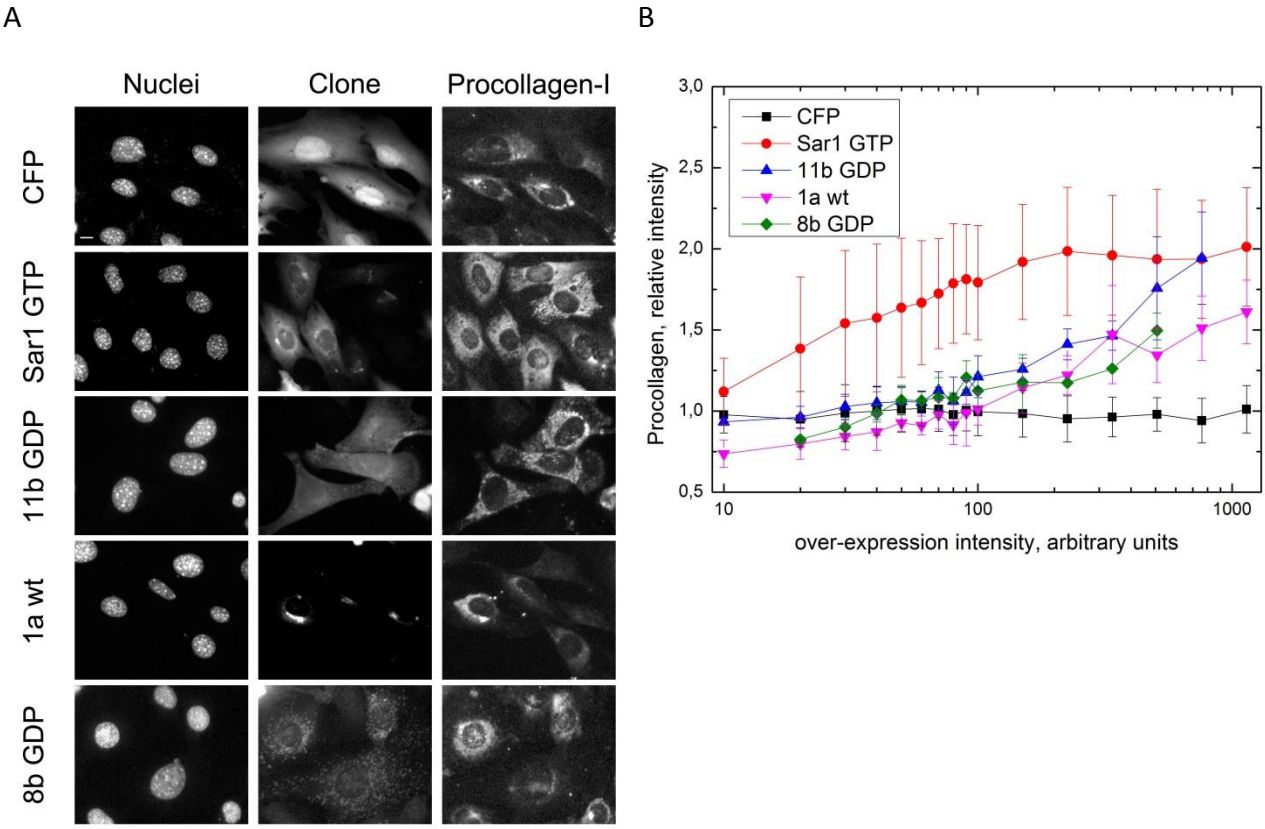


Figure 17 an example of over-expressed protein effect on procollagen-I retention. Image intensity values were scaled for better perception. Scale bar=10µm. A) CFP represents negative (neutral) control; residual amount of procollagen-I can be seen in ER and presumably in Golgi complex. Sar1 GTP (dominant active clone) acts as positive control; procollagen-I is trapped in ER. To judge about procollagen-I entrapment locations for RAB 11b GDP (dominant inactive clone), RAB 1a wt (wild type) and RAB 8b GDP additional co-localization experiments required. B) Sub-population based analysis of over-expression influence on procollagen retention in cells. Error bars represent standard deviation of individual experiments.

Figure 17.A. represents example of actual screening imagery; rows represent set of images for different clone, columns represent fluorescence images for cell nuclei (Hoechst 33342), clones tagged by CFP and procollagen-I immuno stained by Alexa Fluor647. CFP represents negative (neutral) control; residual amount of procollagen-I can be seen in ER and presumably in Golgi complex. Sar1 GTP (dominant active clone) acts as positive control; procollagen-I is trapped in ER. To judge about procollagen-I entrapment locations for RAB 11b GDP (dominant inactive clone), RAB 1a wt (wild type) and RAB 8b GDP additional co-localization experiments required using fluorescent markers of other organelles and using higher resolution imaging. **Figure 17.B**

represents retained procollagen-I in dependence on relative amount of over-expressed protein for selected example proteins.

Conclusion

A data analysis approach was developed to evaluate the effect of RAB GTPases on the secretion of procollagen-I. The difference of the retained procollagen-I fluorescence signal in sub-populations sorted according RAB GTPase over-expression intensity value showed to be the best “hit” calling parameter for our experimental data. 43 out of 178 molecules had shown statistically significant effect. Our screen reproduced results of functionally described RAB GTPases as well as brought not yet described proteins into the picture of procollagen-I trafficking. Further analysis is required to describe the effect and function of the “hit” proteins on the trafficking of procollagen-I.

The development of a fluorescence microscopy based experiment to investigate mRNA and miRNA interactions in live cells

Initial background for the development of an assay

Currently available prediction algorithms of miRNA (micro RNA) interaction with mRNA (messenger RNA) in the RNAi (RNA interference) pathway have low accuracy, and at the time of project start available experimental techniques were limited to cell lysate and capable just to report cell population average of endogenous mRNA amount. Time course experiments on microArray suggested that primary targets of miRNA reliably could be identified just at early time points (12h), with incubation time increasing (over 24h) the transcriptome changes could no longer be explained using sequence alignment (personal correspondence with Andrius Serva; formerly AG Starkuviene). For biology community it is still common to perform RNAi experiments for 48-72h, thus results gathered in such a way could no longer be used for confirmation of miRNA-mRNA interactions.

In the project proposal for "Baden Württemberg Stifftung" funding a collaborative initiative suggested to address miRNA-mRNA interaction (RNAi) question by combining knowledge derived from biological trafficking screen, bioinformatics, and development of integrated high content fluorescence microscopy platform, capable of producing images in wide field screening, confocal and super resolution (below diffraction limit) modes.

The whole project could be partitioned into three general parts: first, selection of small subset of potential miRNA-mRNA interaction partners by combining results of trafficking screen (produced by A. Serva) and bioinformatics analysis (dr. U. Rost); second, development of experiment strategy, this includes labelling strategies for mRNA and short RNAs (siRNAs and miRNAs), data acquisition and analysis, and third, development of the link between wide field, confocal, and dSTORM microscopes (dr. M. Gunkel, B. Flottmann).

Main biological questions concerning development of the live cell capable strategy had been raised:

1) Where is mRNA localized in the cell and how to label it specifically for the microscopy needs?
2) It was clear, what there is no technology available for labelling of endogenous miRNAs, question was how to label synthetic miRNAs and how to test possible efficiency reduction which might occur?
3) Transfection seemed to be only available high throughput capable technique for probe delivery, how long does it take for cells to be transfected?
4) When and where RNAi occurs?

Initial idea was built on assumption that if we are able to label RNAi players – we can image miRNA-mRNA interaction as spatial co-localization of fluorophores and thus confirm or deny the interaction.

A common practice for miRNA activity evaluation is a “luciferase assay”, which basically is a DNA vector containing the firefly luciferase gene (mRNA) with fused nucleotide sequence of interest; initially this sequence used to be completely complementary to the miRNA of study or taken from a real gene as potential target site, later, for luciferase assays it has been started to use longer sequence inserts up to complete 3’UTR of the gene. The miRNA activity is evaluated by measuring luminescence photon yield of the cell lysate; cells are transfected with DNA vector and miRNA of interest, after incubation for 24h (or longer) the cell lysate is collected. Using this method an average value is collected for the cell population, but there is no information left about population size itself or deviation of the value.

As in natural conditions components of the RNAi pathway, like Argonaute family proteins, mRNA and miRNA are not fluorescent, a method for labelling each component has to be chosen; preferably biological activity has to be retained. That requires a method to evaluate biological activity as well. Argonaute family proteins were fluorescently tagged by fusion with GFP or other GFP derivatives suitable for super-resolution (below diffraction limit) microscopy by collaboration partners (AG Grimm). The question remained how to label mRNA and miRNA.

Labelling of the siRNA or miRNA (endogenous molecules)

At the time of project start there were no commercially available short RNAs with fluorescently labelled guide strand, we had to order custom made product. Initially we thought to test labelled siRNA or miRNA activity by observation of a known cellular phenotype. A Järve *et al.*[111] reported an application of FRET technique on siRNA, there 3'end of the guide and 5'end of the passenger strands were labelled, that allowed observing siRNA duplex integrity in the cells. By author evaluation the dual labelling of the siRNA reduced RNAi efficiency by 25% comparing to a non-labelled siRNAs. In our case, we were interested in visualizing single stranded siRNAs - the state after incorporation into the RISC. We took the sequence of a siRNA which is producing clear cellular phenotype (against gene PLK1; once PLK1 is down-regulated normal cell mitosis is interrupted) and labelled 3' end of the guide strand with Alexa Fluor647 and 5' end of the guide strand with BHQ3 quencher, Alexa Fluor647 was chosen due to its known applicability for dSTORM microscopy technique, in such constellation siRNA had to be non-fluorescent while double stranded, as BHQ3 is quenching Alexa Fluor647, and fluorescent when passenger strand is separated. The approach had not achieved expectations: first, penetration of the cellular phenotype was reduced dramatically, and second, RNA melting measurement shown that we can't discriminate between double and single stranded molecules. It was clear that profound research is necessary and as it was not the scope of our main topic, we reduced complexity to single labelled siRNAs.

Labelling of the mRNA in living cells

Two general concepts for specific fluorescence labelling of mRNA in cell cultures were developed so far. First strategy is based on Watson-Crick base pairing, and evolved from in situ hybridization methodologies (ISH). Fluorescence ISH became known as FISH; later FISH limitations were crossed by new derived methods as molecular beacons and smart probes. The second concept is based on aptamers, oligonucleic acid or peptide macromolecules capable to bind target molecule, in this case - polypeptide recognizing specific secondary structure of RNA. For broader and more detail review number of reviews could be referred to, to name just a few Sanjay Tyagi[112],Timothy T Weil *et. al.*[113], Konstantinos Lymperopoulos *et. al.*[114]. While ISH method is capable of detecting endogenous mRNAs, probe design, in-vitro validation and

delivery to live cells are still not at the level for high-throughput applications. In contrast, RNA binding aptamers derived from bacteriophage MS2 or bacteriophage λ in over-expression vector systems are capable of detection just exogenous mRNA, but due to compatibility with cationic lipid transfection methods are more applicable for high-throughput and live cell experimentation, as well probe design is less complicated. Another advantage of aptamers vectors over ISH is possibility not only to observe mRNA, but as well to observe product of mRNA translation. Due to these considerations a bacteriophage λ based aptamers[115] were chosen as mRNA labelling method, in Figure 18 a schematic representation is presented: A) system is built of two DNA vectors, first vector encodes mRNA with specific RNA sequence (boxB), second DNA vector encodes reporter construct, which should be stably integrated into cell culture; B) schematic of detection approach – reporter construct reports mRNA in cytosol (3x GFP), mRNA encoded protein location within the cell is reported by monomeric red fluorescent protein (mRFP); C) as miRNAs primarily target 3'UTR of a gene we redesigned the mRNA construct.

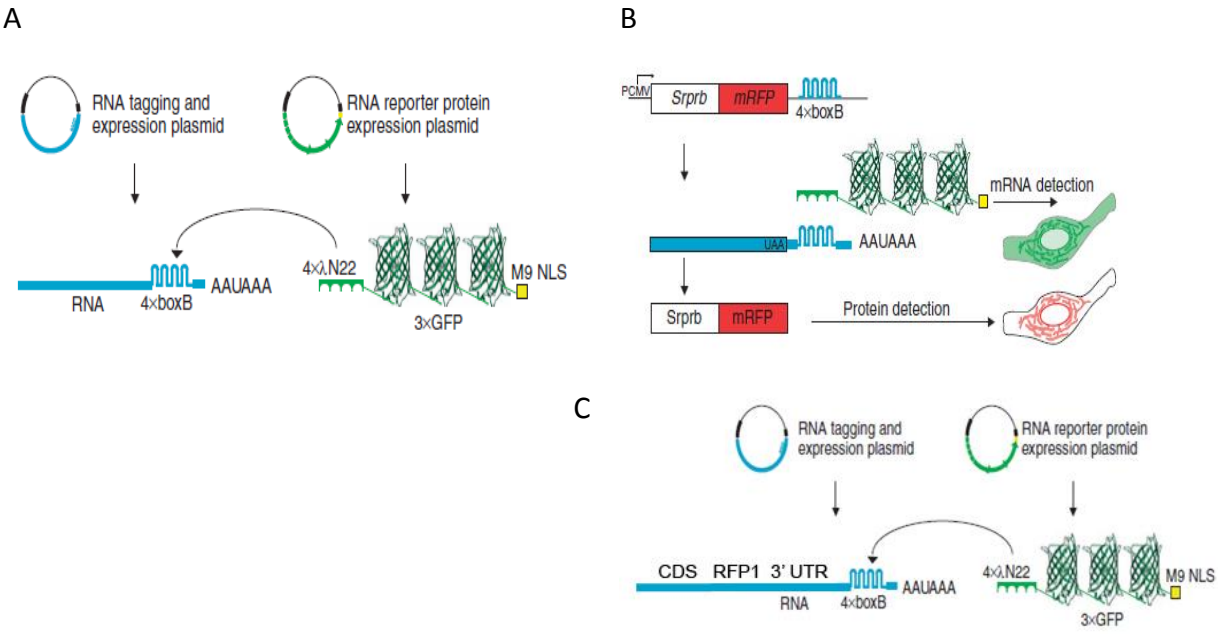


Figure 18 Schematic representation of bacteriophage λ based aptamers system to fluorescently label mRNA in live cells; figures adopted from [115].

This method allows not only visualize mRNA in live cells but also observe translated protein. That combines ideas of different experimental techniques: for example, microarray chip reports detected mRNA amount, then western blot qualitatively reports protein amount, application of

bacteriophage λ aptamers based system allows to visualize mRNA and translated protein in live cells at the same time. In Figure 19 an example of live cell microscopy based imagery is presented: transmission light images to visualize cell boundaries, GFP images to visualize mRNA in cytozol (mRNA not bound reporter localizes in cell nucleus), RFP reports translated protein, DY647 reports labelled miRNA 30b. Presented cell overcomes mitosis at 12^{th} hour of observation; at 13^{th} hour two cells can be observed, GFP fluorescence signal is distributed over full volume of the cells; at 14h increase of RFP fluorescence signal can be observed. GFP fluorescence fignal representing mRNA in cytosol is transient and requires sensitive detection device.

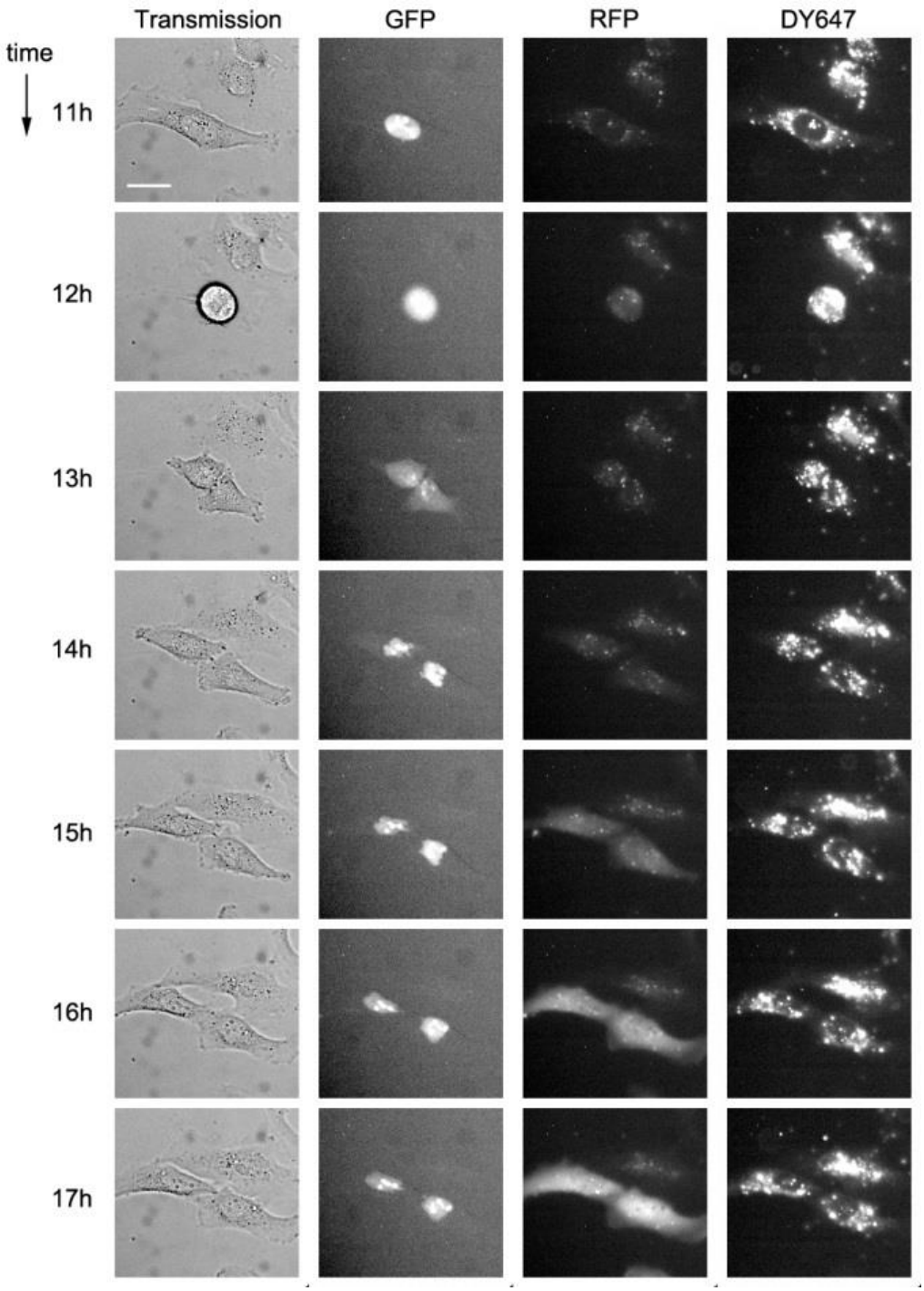


Figure 19 An example of live cell experiment imagry: transmission light images to visualize cell boundaries, GFP images to visualize actin (human ACTB) mRNA in cytozol (mRNA not bound reporter localizes in cell nucleus), RFP reports translated protein, DY647 reports labelled miRNA 30b. At 12th hour of observation cell overcomes mitossis. Images acquired using 20x magnification objective and 2 by 2 camera binning. Scale bar = 10 µm.

Fluorescence microscopy observed cell transfection

Labelled control siRNAs (producing no cellular phenotype change) have been used for long time to evaluate siRNA uptake; usually such evaluation is done at the same time with other siRNAs producing phenotype changes, meaning 24-72h after addition of transfection reagents to cell culture. We applied λ_{N22} aptamers vector system in combination with labelled siRNAs to address transfection process itself, as it was not known when cells are transfected, and when transfected DNA is being transcribed, except common practice to plan imaging 24-48h after transfection. Usually researchers start counting incubation time 4-6h after transfection reagents were applied, while it does not change experiment itself, it creates wrong impression that transfection takes relatively long time. Our results suggest that transfection reagents are uptaken by cells (HeLa) in first 5-30min of incubation, fluorescent probes mainly show dispersed cytosolic distribution which in later time points aggregates to bright fluorescent spots. mRNA encoded in transfected DNA is starts to be observable already in 1-3h after if cells are overcoming mitosis process. mRNA might be visualized in cells which did not overcome mitosis in time of observations as well, usually mRNA fluorescence signal level is lower and product fluorescence signal rises slower, unfortunately numeric image analysis approach has not been developed till now and this conclusion is made from visual observations of acquired fluorescence microscopy images.

Evaluation of RNAi efficiency from fluorescence imagery

Visualization of the mRNA requires temporal sampling and sensitive detection while exposure dose is kept as low as possible; it is very demanding experiment with low throughput. In other hand the observation of product protein fluorescence does not require high temporal sampling rate and may be done in fixed cell cultures, thus the throughput of the experiment can be increased. For a start, the decision was made to observe cells 24h after transfection, as it is the same time as used for “luciferase assays” and corresponds to observation that our cultured cells overcomes mitosis in a frame of such period. We made a number of constructs based on the λ_{N22} aptamers described above, for example RFP_LDLR-3’UTR_boxB, RFP_TBC1D2-3’UTR_boxB, RFP_Mir30b_boxB (complementary binding site for miR30b). In conditions of co-transfection with miR17 the RFP_LDLR-3’UTR_boxB and RFP_TBC1D2-3’UTR_boxB constructs

have shown statistically significant reduction of population mean fluorescence intensity as compared to the same constructs co-transfected with negative (neutral) control siRNA (see Figure 20), in agreement to earlier "luciferase assay results acquired by A.Serva (AG Starkuviene). The same approach was used to test if fluorescently labelled miRNA retains its functionality. miR30b, miR30b-dy647 and anti-miR30b oligonucleotides were purchased from Dharmacon (Thermo Fisher Scientific, UK). As seen in Figure 20, in co-transfection conditions of the RFP_Mir30b_boxB construct, correspondingly with mir30b and mir30b-dy647, showed significant reduction of population mean fluorescence intensity as compared to the same construct co-transfected with negative control siRNA; labelled miRNA performed less efficiently in gene silencing than its non-labelled counterpart, however its activity was evaluated as reasonably good to be comparable.

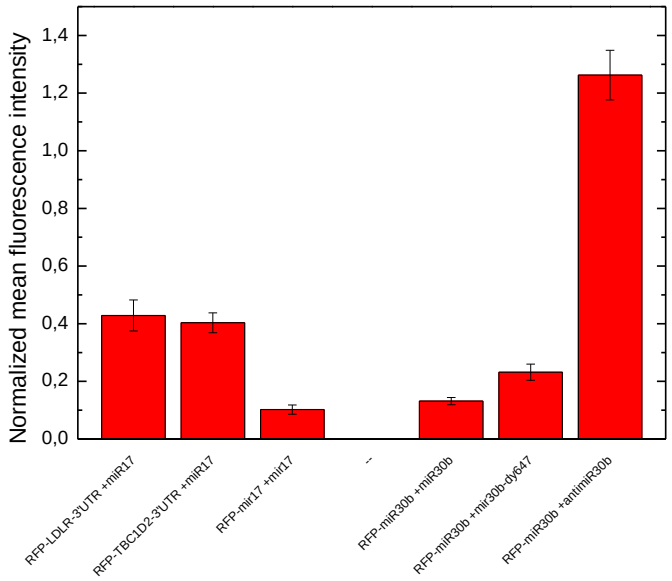


Figure 20 Fluorescence microscopy based evaluation of RNAi activity; LDLR and TBC1D2 genes silencing by miR17 and fluorescent miRNA labelling influence on miRNA activity. Fluorescence imagry was analysed as described in method part of this thesis and population mean calculated. Error bars represent standard error of the mean.

The miR30b is natively expressed in HeLa cells, co-transfection of the RFP_Mir30b_boxB construct with anti-miR30b, specially designed RNA molecule with complementary sequence to miR30b, showed increase of population fluorescence mean intensity comparing to the same construct co-transfected with negative control siRNA. These results show that such microscopy

based approach could be used not only for RNAi evaluation of endogenous molecules, but as well to evaluate effect of native miRNAs on gene expression rate.

The co-transfection experiments have shown applicability of the λ_{N22} aptamers vector system for RNAi evaluation. The question was if this system could be useful for systematic experimentation: as mRNA would be always the same an influence of binding site location (5'UTR, CDS, 3'UTR) could be evaluated as well as secondary structure of the mRNA.

Our collaboration partner (AG Kummer) compared at the time available miRNA-mRNA interaction prediction algorithms to explain microarray experimental results; the conclusion was that available miRNA-mRNA interaction prediction algorithms fail in accurate prediction and that miRNA down-regulated mRNA are better explain using "seed" sequence alignment (6 nucleotide length sequence; 2-7th nucleotide from 5' end of a miRNA) algorithm. Seed sequence alignment of known human miRNA library[116] against TBC1D2 nucleotide sequence resulted a high number of miRNAs, more than 500 out of 1900; then seed sequence length increased to 8 nucleotides 9 miRNA found potentially targeting 5'UTR, 64 miRNAs coding sequence and 25 miRNAs 3'UTR. The question was if we really can simply use seed sequence alignment for miRNA-mRNA interaction prediction. As miRNAs are still relatively expensive reagents, it was decided to test if siRNAs selected by the seed sequence length criteria would indeed result effective RNAi. For this reason seed sequences of Silencer® siRNA library (Ambion) were aligned with RFP-TBC1D2_3'UTR_boxB construct and 11 siRNAs from the ones having match (having seed length from 9 to 14 nucleotides) were randomly selected for the experiment. It was expected that the seed length should have a direct correlation with RNAi activity.

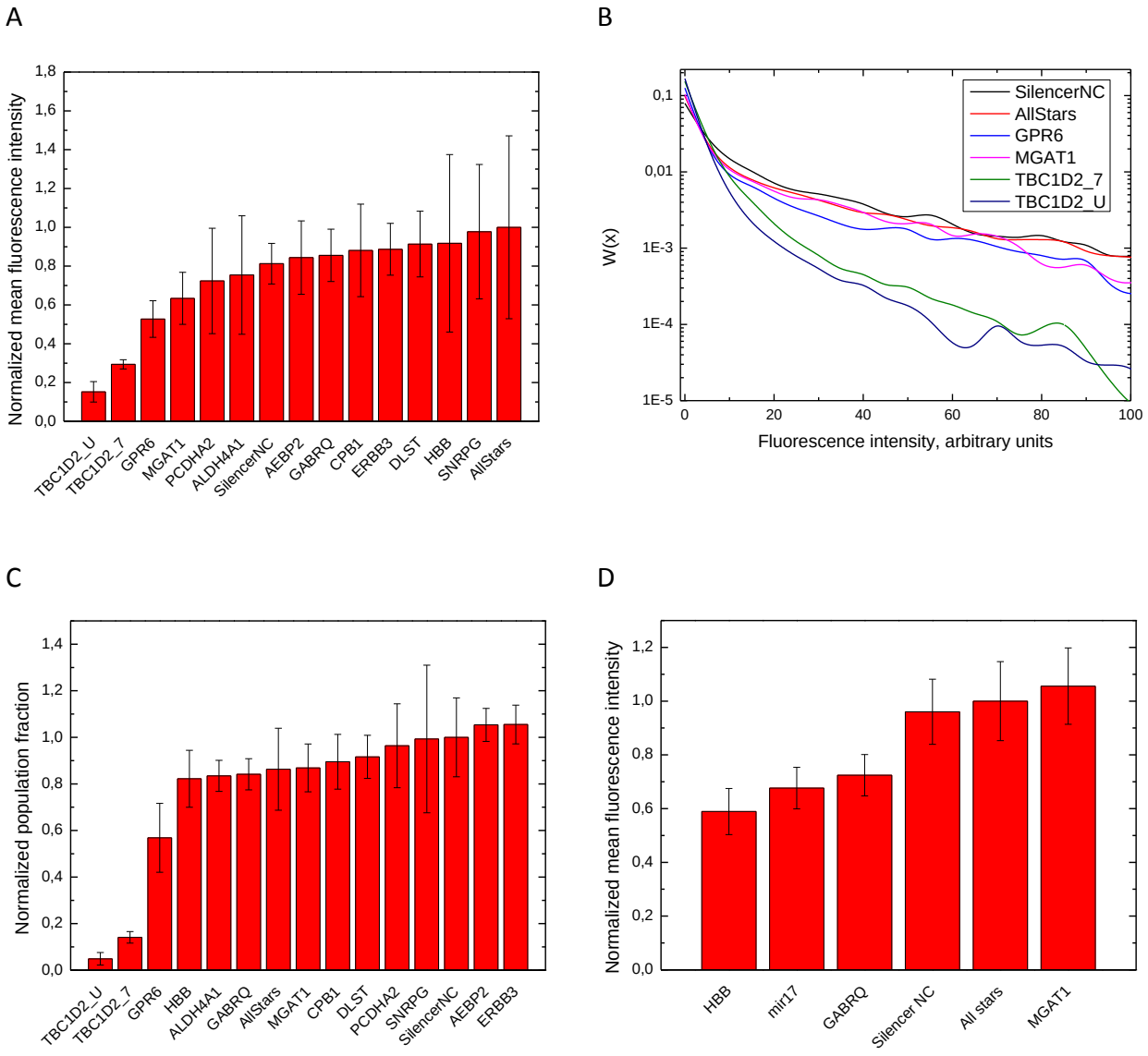


Figure 21 Evaluation of gene expression efficiency. A) Population mean based evaluation of RFP-TBC1D2_3'UTR_boxB expression changes due to co-transfection with corresponding siRNA; B) Population distribution; C) Population distribution based evaluation of RFP-TBC1D2_3'UTR_boxB expression changes due to co-transfection with corresponding siRNA; D) Population mean based evaluation of RFP-LDLR_3'UTR_boxB expression changes due to co-transfection with corresponding siRNA. Error bars represent standard deviation.

Transfection experiments were performed independently 5 times. Two negative control siRNAs were used (AllStars (Qiagen); Silencer NC(Ambion)) and two positive control siRNAs, being fully complementary to the target sites (TBC1D2_U, TBC1D2_7(Qiagen)). It is difficult to judge negative control siRNAs as companies producing them usually keep the sequence

information in secret, therefore we don't know sequence of AllStars negative control siRNA. Silencer NC siRNA sequence is known, and it has 5nt of the seed sequence match to our construct. It became clear that population mean based evaluation is difficult due to high variation of control population mean values and is not conclusive unless the interaction is strong. Using such evaluation method just two siRNAs (see Figure 21.A) could be evaluated as having significant effect on gene expression: GPR6 with 8nt seed sequence match with our gene, and MGAT1 having 13 nt seed sequence match.

In contrast to population average based evaluation, population distribution based analysis (Figure 21.B) demonstrated to be less sensitive to day to day variations. Distribution based analysis may report many different parameters derived from distribution or cumulative distribution functions which could be used to score effect of one or the other siRNA. In Figure 21.C an approach to evaluate RNAi is presented: population fractions in the range where average distribution function of the negative controls intersects with positive controls to the max of positive control were compared, comparison of cumulative distribution functions would give very similar result. Just one siRNA, GPR6, can be evaluated as having significant effect on gene expression.

Thermodynamic analysis of siRNAs and their target sites (using mfold/Unafold software) suggested that secondary structure of the mRNA is an important factor for an effective RNAi occurrence; single stranded count parameter was related to total number of computed structures and presented for predicted target sites in Table 1. Prediction of the secondary structure for RFP-LDLR_3'UTR_boxB construct suggested that binding sites for some of the siRNAs have different conformation (probability to form single stranded structure) in contrast to the RFP-TBC1D2_3'UTR_boxB construct, in case if secondary structure of the mRNA is really a factor for efficient RNAi an experiment should confirm this. In Figure 21.D results of RFP-LDLR_3'UTR_boxB co-transfection experiment with corresponding siRNA are presented, the experiment confirmed hypothesis, that siRNA activity correlates with predicted conformation of the mRNA. While these results have shown that mRNA secondary structure has to be accounted for prediction of miRNA-mRNA interaction, gathered data were not sufficient for building a quantitative interaction prediction model. A large scale experiment had to be set-up.

Construct	siRNA	ΔG full length	ΔG of the seed sequence	siRNA guide strand nucleotide number (5'->3') and corresponding probability of the mRNA to be single stranded																				
				21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
TBC1D2 3'UTR	MGAT1	-17,1	-7,4	0,1	0,1	0	0	0	0,1	0	0	0	0	0	0,9	0,9	0,9	0,9	1	0	0	0	0,1	0
TBC1D2 3'UTR	MGAT1	-23,1	-10,3	0,4	0,7	0,7	0,7	0,6	0,5	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1
TBC1D2 3'UTR	DLST	-21,8	-6,6	0,1	0,1	0,4	0,3	0,8	0,4	0,3	0,3	0,6	0,4	0,4	0,4	0	0	0	0	0	0	0	0	0
TBC1D2 3'UTR	SNRPG	-21	-6,9	0,1	0,9	0,9	0,9	0,1	0,1	0,1	0,1	0,9	0,9	0,9	0,9	0	0,1	0,1	0,1	0,8	0,8	0,1	0,1	0,7
TBC1D2 3'UTR	ALDH4A1	-14,9	-11	0,8	0,8	0,8	0	0	0,1	0,1	0,1	0,8	0	0	0,6	0,9	0,4	0,1	0,1	0,3	0,6	0,7	0,6	0,6
TBC1D2 3'UTR	ALDH4A1	-16,8	-11	0,7	0,6	0,6	0,5	0,5	1	0,3	0	0	1	0,1	0	0	0	0	0	0,9	0,9	0,9	0,9	0,9	0,9
TBC1D2 3'UTR	ALDH4A1	-20,2	-11	0,7	0,2	0	0,1	0,1	0,8	0,9	1	0,9	0,7	0,7	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,4	0,8	0,8
TBC1D2 3'UTR	AEBP2	-17,9	-5,9	0,9	0,9	0,7	0,7	0,8	0,8	0,1	0,1	0	0,1	0,2	0,2	0,9	0,8	0,8	0,8	0,8	0,9	0,9	1	0,9
TBC1D2 3'UTR	CPB1	-15,7	-8,7	0,1	0,1	0,7	0,7	0,8	1	1	1	0,3	0,3	1	1	0	0	0	0,3	0,9	1	1	1	0,9
TBC1D2 3'UTR	CPB1	-9,9	-7,8	0	0	0,1	0,1	0,9	0,9	0,9	0,1	0,1	0,1	0,1	0,9	0,9	0,9	0,9	0	0,1	0,1	0,1	0,8	0,8
TBC1D2 3'UTR	GABRQ	-17,1	-6,7	0	0	0	0,9	0,9	0,9	0,9	1	0	0	0	0,1	0	0,1	0,1	0,1	0,9	0,3	0	0,1	0,6
TBC1D2 3'UTR	GABRQ	-14,8	-6,7	0,9	1	0,9	1	1	0,9	0,9	0,9	0,1	0	0	0,1	0,2	0,1	0,8	0	0	0	0,1	1	0,3
TBC1D2 3'UTR	GPR6	-17,5	-7,6	0,3	0,2	0,2	0,2	0,1	0,1	0	0	0,9	0,9	0,1	0,1	0,1	0,2	0,1	0,6	0,8	0,7	0,3	0,3	0,3
TBC1D2 3'UTR	HBB	-15,6	-6,9	0,3	0,2	0,3	0,4	0,3	0,3	0,3	0,3	0,4	0,6	0,7	0,5	0,7	0	0,1	0,8	0,1	0,1	0,1	0,8	0,9
TBC1D2 3'UTR	ERBB3	-12,5	-9	0,2	0,2	0,2	0,2	0,2	0,9	0,2	0,2	0,8	0	0	1	1	0,8	0	0	0	0	0,8	1	1
TBC1D2 3'UTR	ERBB3	-16	-9,9	0	0,9	0,9	0	0	0,1	0,1	0,1	1	0,1	0,1	0	0	0	0,1	0	0	0	0	0	0,9
TBC1D2 3'UTR	ERBB3	-14,3	-10,1	0,1	0,8	0	0	0	0,1	1	0,3	0,3	0,2	0,2	0,9	0,8	0,8	0,8	0,1	0,1	0,1	0,2	0,9	0,2
TBC1D2 3'UTR	PCDHAC2	-9,9	-7,6	0,9	0	0	0,1	0,1	0,1	1	0,1	0,1	0	0	0	0,1	0	0	0	0	0	0,9	0,9	0,9
TBC1D2 3'UTR	PCDHAC2	-10,9	-7,8	0,1	0,7	0,7	0,8	1	1	1	0,3	0,3	1	1	0	0	0	0,3	0,9	1	1	1	0,9	0,9
TBC1D2 3'UTR	PCDHAC2	-10,4	-7,9	0	0,1	0,1	0,9	0,9	0,9	0,1	0,1	0,1	0,1	0,9	0,9	0,9	0,9	0	0,1	0,1	0,1	0,8	0,8	0,1
TBC1D2 3'UTR	PCDHAC2	-8,7	-7,6	0	0	0	0,1	1	0,3	0,3	0,2	0,2	0,9	0,8	0,8	0,8	0,1	0,1	0,1	0,2	0,9	0,2	0,1	0,1
TBC1D2 3'UTR	TBC1D2_U	-33,4	-7	0,4	0,6	0,7	0,5	0,7	0	0,1	0,8	0,1	0,1	0,1	0,8	0,9	0	0	0,9	0,1	0	0,9	0,9	0,9
TBC1D2 3'UTR	TBC1D2_7	-30,2	-5,8	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0,4	0,6
LDLR 3'UTR	MGAT1	-17,1	-7,4	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0
LDLR 3'UTR	MGAT1	-23,1	-11,8	0,2	1	1	1	0,8	0,8	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1
LDLR 3'UTR	miR17	-10,5	-7,9	0,5	0,4	0,1	0,1	0	0	0,4	0,3	0,4	0,7	0,4	0,5	0,4	0,3	0,4	0,9	0,9	0,8	0,8	0,9	0,5
LDLR 3'UTR	GABRQ	-17,1	-9,3	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0,9
LDLR 3'UTR	GABRQ	-16,4	-7,8	0,3	0,4	0,7	0,4	0,5	0,4	0,3	0,4	0,9	0,9	0,8	0,8	0,9	0,5	0,5	0,5	1	0,8	0,8	0,8	0,8
LDLR 3'UTR	HBB	-15,3	-7,9	0,5	0,4	0,1	0,1	0	0	0,4	0,3	0,4	0,7	0,4	0,5	0,4	0,3	0,4	0,9	0,9	0,8	0,8	0,9	0,5

Table 1 some of thermodynamic parameters computed by mfold and Unafold software for mRNA target sites.

Solid phase transfection

For large scale experiments the application of a solid phase transfection seemed to be the technique of choice. Initial experiments showed the potential - the results were comparable to liquid transfection experiment, but as the over-expression efficiency showed to be much more reduced, we decided to test different transfection reagents and optimize solid phase transfection in order to perform conclusive large scale experiment. We tested a number of transfection reagents in order to select the one most effective for co-transfection (cDNA and siRNA/miRNA). We found that in contrast to usually used "Lipofectamine 2000" (Invitrogen, Life Technologies), "peqfect siRNA"(peqlabs) produces much higher protein expression level (Figure 22; Figure 24 A and B) and fluorescence signal significantly higher than background is presented in more than 65% of the observed population, when for lipofectamine transfected cells this parameter is just in the range of 35%. "peqfect siRNA" transfection reagent is less toxic as judged from observed cell numbers (Figure 23) ~4000 cells transfected by lipofectamine 2000 and ~5500.

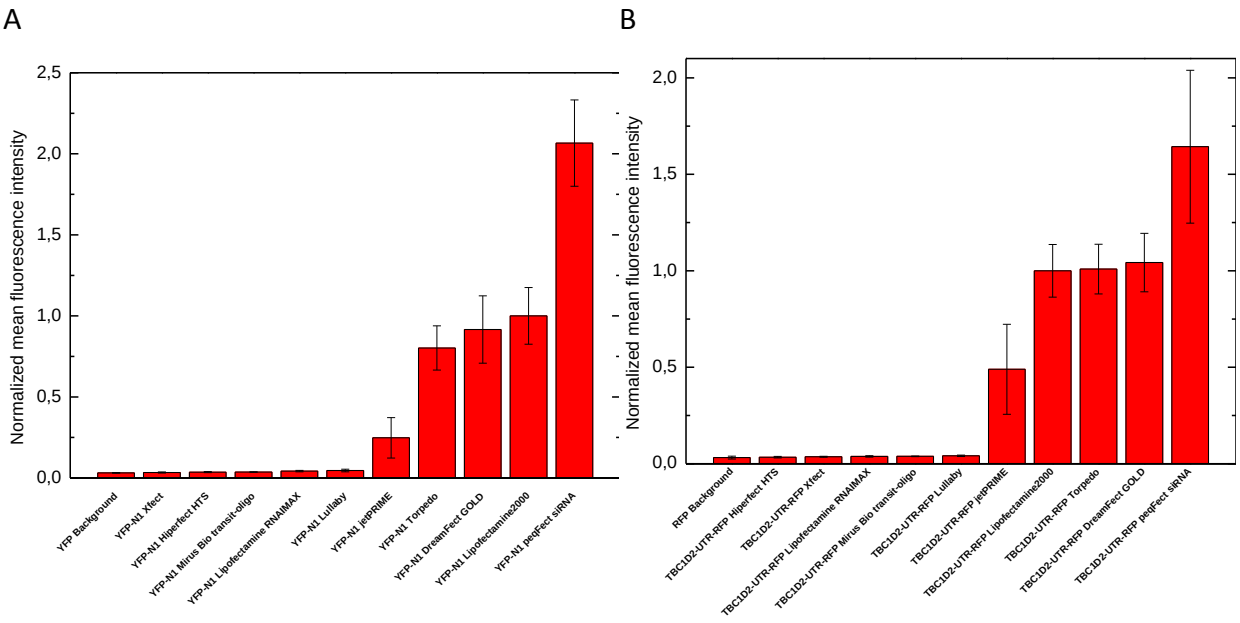


Figure 22 Evaluation of DNA plasmid solid phase transfection. A) Cell population average of over-expressed YFP protein fluorescence mean intensity values; values are related to cell population transfected using Lipofectamine2000. B) Cell population average of over-expressed RFP(TBC1D2 3'UTR construct) protein fluorescence mean intensity values; values are related to cell population transfected using Lipofectamine2000. Error bar represents standard deviation of individual observations.

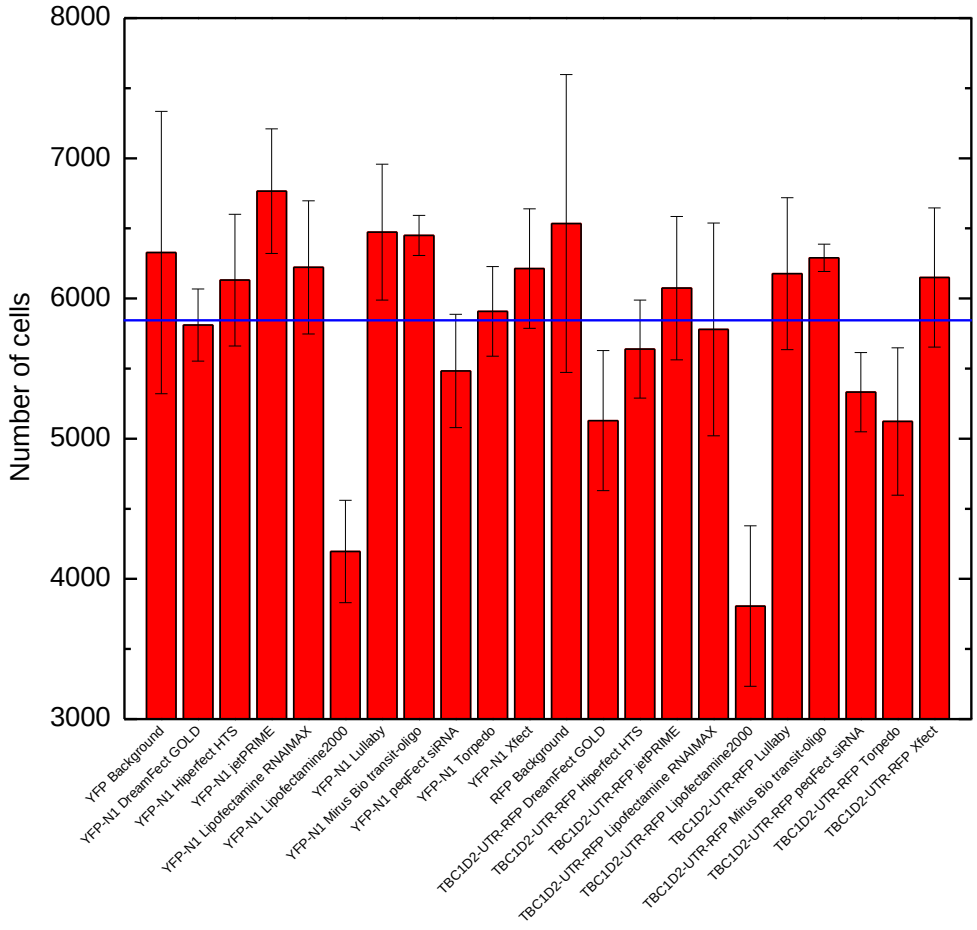


Figure 23 Cell number variations for observed populations due to different transfection reagent. Blue line represents mean of all observed populations. YFP Background and RFP Background were populations grown without any transfection reagent.

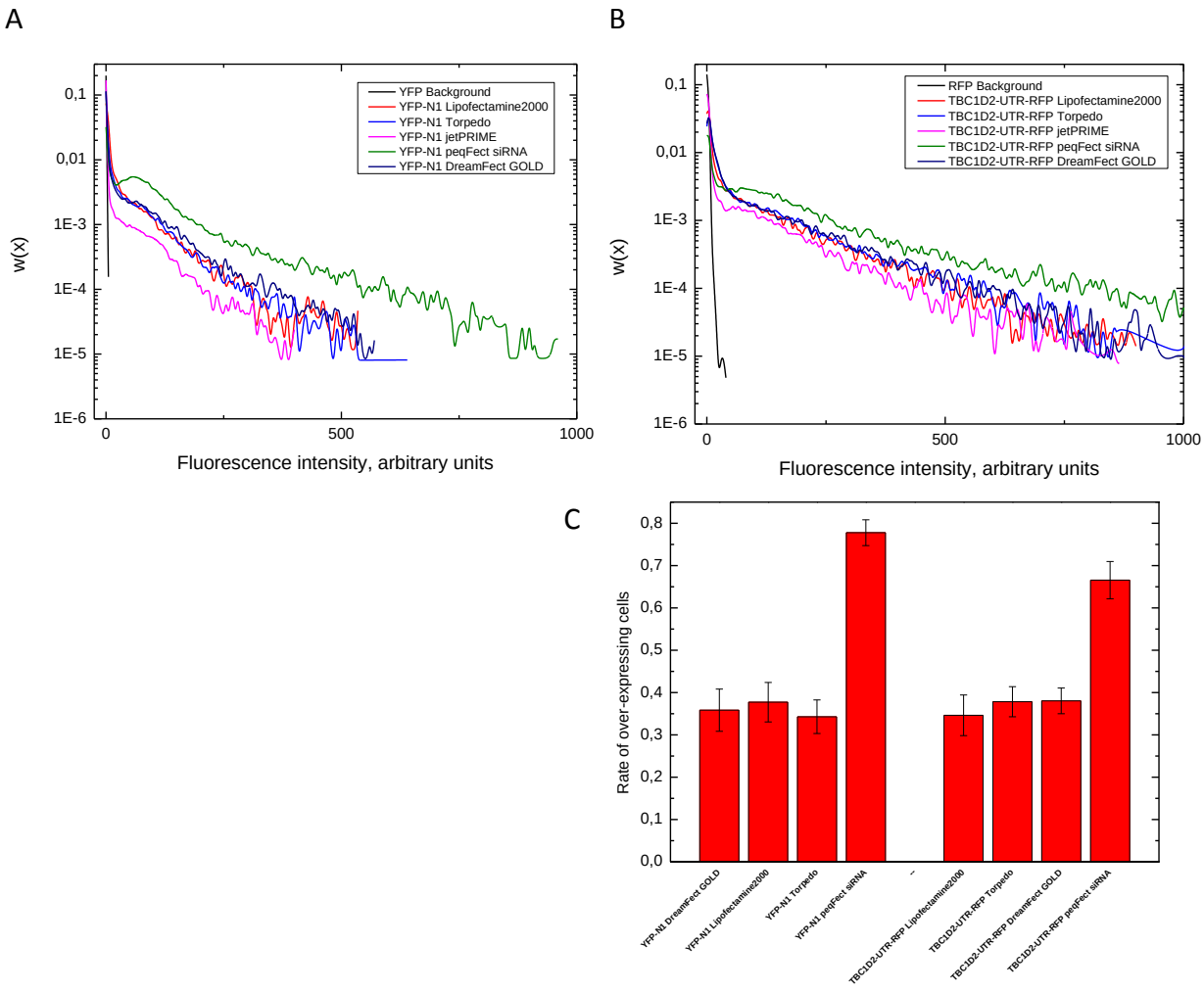


Figure 24 evaluation of DNA solid phase transfection efficiency. A) Probability density for YFP over-expression using different transfection reagents. B) Probability density for overexpression of RFP(TBC1D2 3'UTR construct) using different transfection reagents. C) Rate of expressing cells in observed population when particular transfection reagent is used.

Gathered experience on solid phase co-transfection of DNA and siRNA, as well as data analysis and interpretation, will be used for a large scale miRNA-mRNA interaction evaluation screen for which 104 miRNAs potentially targeting human TBC1D2 gene, selecting miRNAs according seed sequence criteria, were acquired.

Correlative fluorescence microscopy

In the scope of microscopy based evaluation of miRNA-mRNA interactions project there was a task to build an integrated microscopy platform which would combine features of wide field screening fluorescence microscopy, confocal microscopy and super-resolution microscopy (stochastic optical reconstruction microscopy; STORM). At the current stage of research, the STORM microscopy technique showed to be not necessary for the evaluation of miRNA-mRNA interaction experiments.

Sub-diffraction fluorescence microscopy techniques as STORM or STED (Stimulated emission depletion) have requirement of fine spatial structure for imaging process control as otherwise it is difficult to judge if technique was applied correctly and the expected spatial resolution was reached. We decided that such integrated microscopy platform, a correlative fluorescence microscopy, could be useful to address morphological changes of the Golgi complex, especially the co-localization or co-existence of different Golgi compartments. Main reasons to go with Golgi complex investigation were: first, the structural organization of the Golgi complex is quite well known; second, using various drugs or gene down-regulation Golgi complex phenotype can be changed dramatically, this feature of this organelle is essential to test the range of detectable morphology changes; third, Golgi complex is essential organelle for membrane trafficking processes and it was observed earlier (A. Serva; AG Starkuviene) that some miRNAs have significant effect on the organization of the Golgi complex.

Labelling of cellular organelles is usually performed by over-expression of fluorescently tagged organelle residing proteins or by imuno-staining via fluorescently tagged antibodies. In our case there was a requirement for fluorophores to be suitable for all microscopy techniques and especially to photo-physically switch in the same buffer conditions for STORM imaging. A pair of suitable dye was found by Benjamin Flottmann: Alexa Fluor 532 and Alexa Fluor 647. For proof of principle experiments we choose NRK cells stably expressing GalT-CFP, a *cis*-Golgi marker (the same cell line was used for the Golgi complex organization kinetics evaluation experiments). We used anti-GFP antibodies to co-stain *cis*-Golgi membranes by Alexa Fluor 532 and GM-130 antibodies to stain *trans*-Golgi membranes by Alexa Fluor 647.

For experiments µ-Slide 8-well slides with pre-coating by poly-L-lysine (15 min before cell plating) were used. Cells were treated by BFA or nocodazole at concentrations of 5

µg/mL and 1 µg/mL, respectively, for 20 min. in presence of 0.1 mg/mL of cycloheximide to abolish protein synthesis. For gene down regulation experiments cells were incubated for 30h past transfection with corresponding siRNA. Cells were fixed by incubation in 3% PFA for 20 min. TetraSpeck™ microspheres (0.1 µm) were applied and incubated for 15 min, and cells were permeabilized by incubating with 0.1% Triton-X-100 for 5 min. Fluorescence coordinates of TetraSpeck™ microspheres were used for planar re-alignment of reconstructed STORM images.

In Figure 25 set of different Golgi complex morphological phenotypes acquired by confocal microscopy is presented. Cells transfected with negative control siRNA represent normal Golgi phenotype; Down-regulation of GBF1 induced dispersal and fragmentation of Golgi stacks; cell treatment with nocodazole, microtubule depolymerising agent, fragmented Golgi stacks; treatment with BFA relocated *cis*-Golgi marker to the ER, *trans*-Golgi marker relocated to punctuate structures, most likely – ER exit sites.

We used sample based co-ordinate system to relocate sample on different microscopes, basically it were fixed geometrical objects observable in full scale trough the microscope optical system. A schematic example of the workflow is depicted in Figure 26: once sample coordinates are set, sample is imaged using 10x wide field microscopy mode, from those images coordinates of all individual cells are generated, here selection of candidates for higher resolution imaging can be selected manually or using computer algorithm (in development by Dr. Ullrich Koete; AG Hamprecht); sample based coordinates are used for transferring the sample onto different microscopes and relocating selected cell automatically; high magnification images presented to demonstrate resolving power of corresponding microscopy technique; overlap of fluorescence signal in separate channels is represented by yellow, as seen in wide field image - yellow is heavily presented while in dSTORM image presence of yellow is vanishingly low.

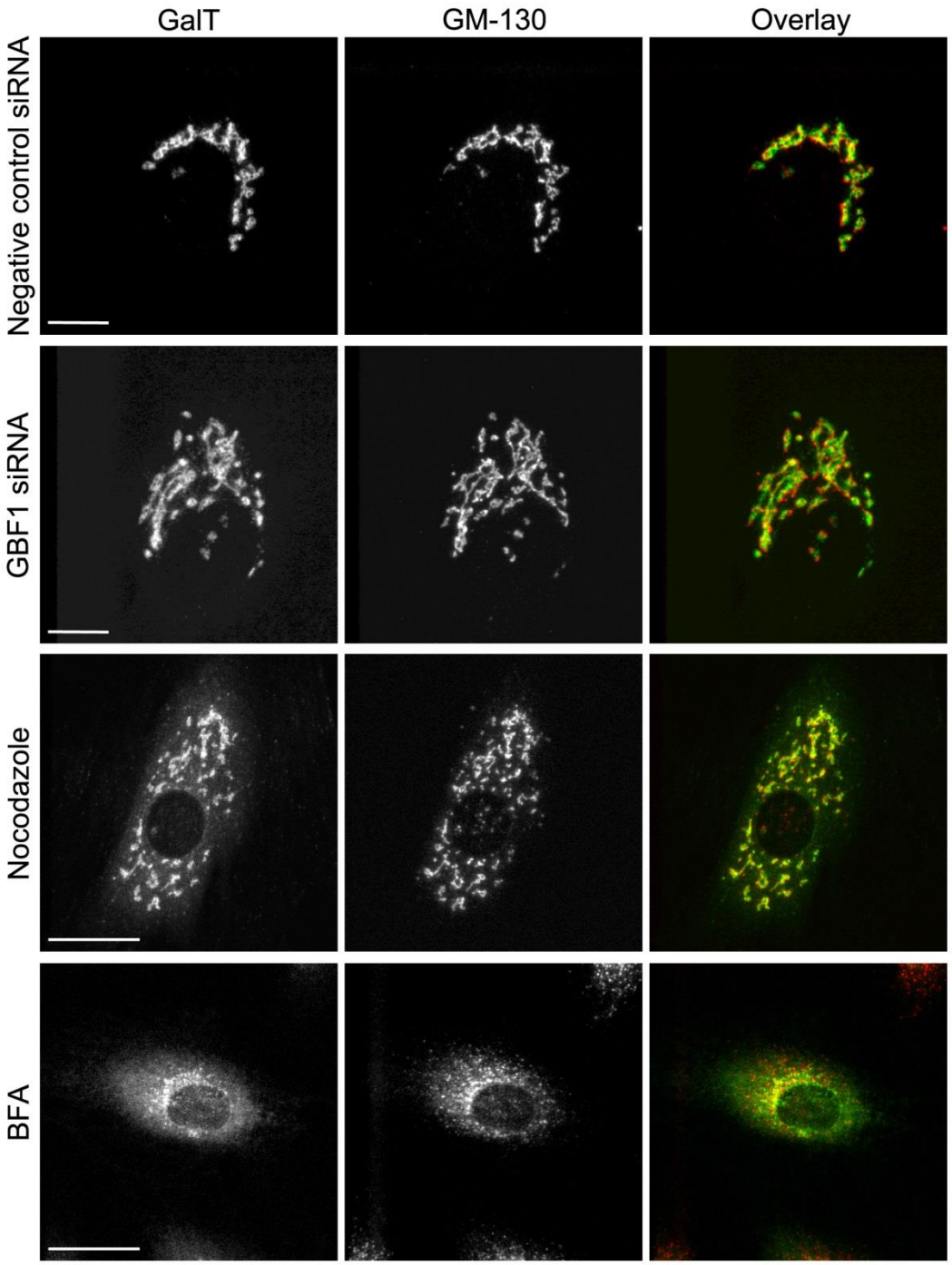


Figure 25 Examples of drug or gene down-regulation induced Golgi phenotype changes acquired by confocal microscopy. After transfection cells were incubated for 30h before fixation. Cell were treated by Nocodazole and BFA for 20min. GalT-CFP is labelled via anti-GFP antibody and Alexa Fluor 532; GM130 Alexa Fluor 647. Scale bar = 10 µm.

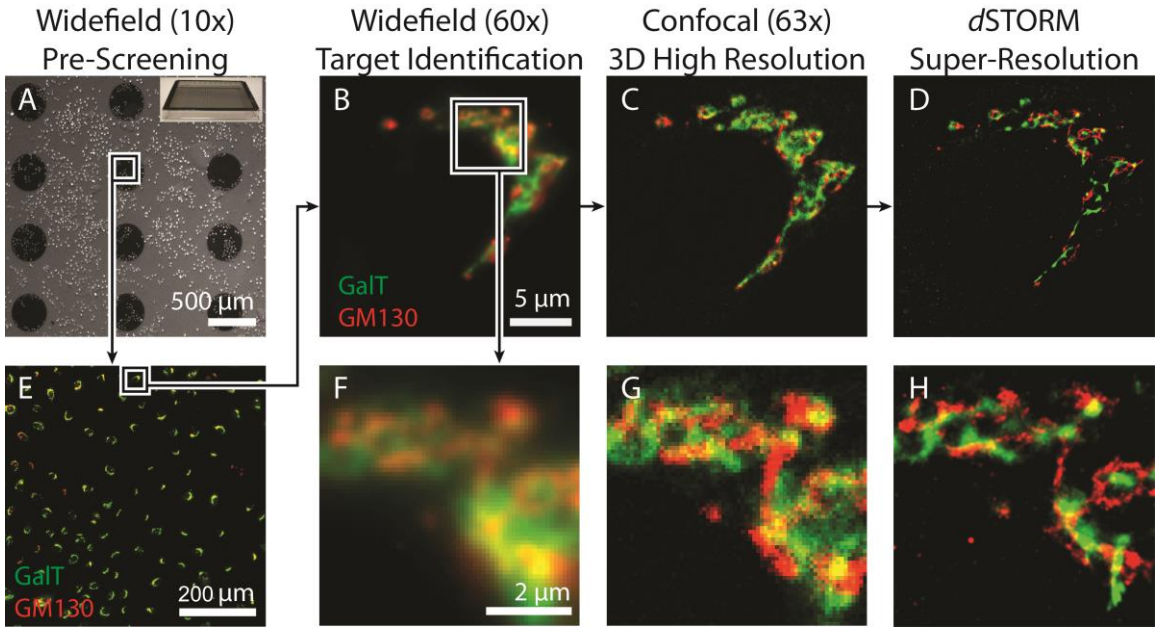


Figure 26 Imaging routine: Fluorescence correlative microscopy – high content data acquisition. The sample is observed in 10x magnification wide field mode and coordinates of individual cells are generated, then sample is transferred over higher resolution microscopes, cells are automatically relocated and imaged. High magnification images presented to demonstrate resolving power of corresponding microscopy technique. Fluorescence images artificially colored: GalT – green, GM130 – red.

Further numeric evaluation of fluorescence imagery acquired on correlative fluorescence microscopy platform was performed by Benjamin Flottmann and currently is being submitted in a publication[117].

In cell biology it is important to know relation of different cellular structures, their overlap in space usually is interpreted as direct interaction, fluorescence microscopy is very helpful for this task, unfortunately the limitations of particular microscopy technique is often forgotten when the conclusion is made. It is known that *cis* and *trans* membranes of the Golgi complex are related and in general belong to the same organelle, interpretation of wide field fluorescence microscopy images would allow to think that those membranes mix, however judging data from dSTORM images such hypothesis cannot be confirmed.

Conclusions

We have been investigating possibility to apply fluorescence microscopy based method to investigate miRNA-mRNA interactions in living cells. The RNAi is relatively fast process; our current instrumentation is not suitable for high-throughput live cell

observations of mRNA and protein translation kinetics, especially due to requirement for low photon yield sensitive camera. Nerveless an application of λ_{N22} aptamers vector system allowed us to observe cell transfection process and its consequence – mRNA and translated protein distribution in the cytosol in low scale experiment. Cell transfection takes place in time frame of 5-30min after transfection reagents are applied, once cell overcome mitosis new daughter cells show mRNA reporter distribution in the cytoplasm and consequent appearance of translated protein reporting fluorescence. mRNA distribution and protein translation process may be observed anytime in the frame of 20h past transfection reagent addition. RNAi efficiency estimation by evaluation of translated protein amount by the means of chemoluminescence ("luciferase assay") is accepted in a field, it is reporting fluorescence population average without even indicating population size. Fluorescence microscopy based evaluation allows observing whole population by investigating single units; such method allows not only conclude the mean value but as well to observe distribution.

Techniques for labelling of endogenous mRNA even being developed for quite long time at best are applicable for observation of single cell scale. Biological aptamers based labelling technique is not applicable for observation of endogenous mRNA and is more artificial but due to compatibility with live cell experiments and high-throughput approaches was more suitable for our needs. As RNAi consequence is change of gene production, fluorescence intensity based evaluation of gene amount seems to be suitable way to interpret RNAi efficiency.

Discussion

The cell biology is vast set of knowledge. Much of this knowledge was generated by simple visual inspection and verbal description. It is common to hear explanation that biology is too complex for application of quantitative approaches. In many cases it is due to yet not existing quantitative methods. While in many cases qualitative analysis would be sufficient, recent trends in research blindly require numeric expression without much care if it represents qualitative or quantitative measurement.

During the period of time I spent in the labs of Dr. Vytaute Starkuviene and Dr. Holger Erfle I had a chance to investigate several cases of the membrane trafficking events as well as two versatile techniques used to investigate gene function, naming over-expression of GFP tagged clones and RNA interference. Main results of my work are: first, developed fluorescence microscopy based assay to investigate live cell Golgi complex dynamics, second, developed data analysis and interpretation approach for collagen-I secretion assay, third, developed fluorescence microscopy based High-throughput approach to investigate gene silencing efficiency by short double stranded RNAs in live or fixed cells.

The proof of principle screen on factors potentially influencing the Golgi complex dynamics revealed a number of proteins which were previously not reported as influencing Golgi complex organization. The developed live cell assay showed to be useful technique for identification of Golgi complex organization influencing genes.

Fluorescence microscopy based observation of gene over-expressing cell population has advantage over other cell lysate based techniques as it allows addressing cell distribution, for example, probability to observe transfected cell with a level of expressed exogenous protein decay according exponential decay function toward high expression levels as judged from fluorescence intensity values. In this case population mean value might not reveal influence of over-expressed gene on specific cellular process. Parametric sorting of cells was helpful to address influence of RAB GTPases on collagen-I secretion in cell culture experiments. Similar distribution based analysis was helpful for optimisation and evaluation of solid phase transfection protocols as well as interpretation of RNA interference efficiency for short double stranded RNA molecules having partial complementarity to target mRNA.

Summarizing said above, during my study I developed and/or performed couple of different biological experiments and by applying fluorescence microscopy and numeric data analysis approaches extracted higher content of meaningful data comparing to the methods used before. For example, fluorescence intensity based evaluation of cellular processes is basically qualitative approach, an application of statistical analysis methods and quantitative description of qualitative data distribution is critically important in order to detect and describe influence of factors being under investigation.

Summary

The molecular biology of the cell is a growing set of concepts and knowledge being developed while studying various cells and organisms with the focus kept on molecular structures and their function in cellular processes. An accepted model of the mammalian cell suggests that the cell is a membrane based system where all processes are driven by proteins; proteins and their specific interactions render specific properties for membranes, which in consequence form individual functional units within the cell. Membrane trafficking is an abstraction describing material transport within different organelles of the cell and is one of the areas of interest in molecular biology. A practical interest for the investigation of membrane trafficking is lying in its relation to a number of serious diseases. Membrane trafficking is organized as multi-stage cargo delivery from donor to destination membranes via vesicle formation, transportation, and fusion processes. Genes involved in membrane trafficking and their regulation within the cell are in particular focus of interest.

Microscopy based observation of membrane trafficking events in the native or close to native context such as fixed or live cell cultures is the method of choice as it allows relating different stages of the whole process to specific cellular organelles in space and time domains and has the advantage of better statistics over population mean based methods, as results are concluded from the observation of a high number of individual cells. Fluorescence microscopy allows addressing questions like how many copies of a specific protein are in the cell or where they are located in relation to other cellular organelles, in some cases the time of the protein translation can be addressed. Despite the versatility of fluorescence microscopy, the application of any particular technique for cell biology requires more application layers. To name just a few of important additional application layers: fluorophores and labelling strategies, data acquisition and storage, digital image analysis, data interpretation in a biological context, and last but not least – the biological process of interest itself.

This thesis describes my experimental and data analysis work which was done to address some of specific membrane trafficking related cases:

a) The influence of over-expressed membrane proteins on the organization kinetics of the Golgi complex, which is a major membrane trafficking organelle. Live cell

fluorescence screening microscopy was employed to address this question[73,110,118].

b) Statistical data analysis of a cargo protein secretion screen. Influence of specific RAB GTPases over-expression on collagen-I secretion was evaluated[119].
c) The development of a fluorescence microscopy based experiment to investigate messenger RNA (mRNA) and micro RNA (miRNA) interactions in live cells.

Evaluation of the influence of over-expressed membrane proteins on the organization kinetics of the Golgi complex

The purpose for a live cell observation of a membrane trafficking event was supported by the idea that it should be possible to identify effectors which could not be revealed by other biochemistry techniques for example due to the different interaction mechanism. The investigation of the Golgi complex organization was chosen due to its major role in membrane trafficking, relatively easy visualization, and reliable induction of morphology changes by a number of different chemicals.

Over the last decades the morphology of the Golgi complex has been investigated in great detail by different approaches[83–94]. A number of cell biology studies revealed that the structure of the Golgi complex is highly dynamic and is crucially dependent on the balance of material flow[95–97]. A comprehensive list of proteins residing at and functionally related to the Golgi complex is being currently compiled[98–102]. Despite a wealth of knowledge on the function and structure of the Golgi complex, the information was largely obtained by studies in fixed cells. First observations of the Brefeldin-A (BFA, a fungal metabolite) effect on the organization of the Golgi complex were reported by Toshiyuki Fujiwara *et al*. in 1988[103] and seconded by Jennifer Lippincott-Schwartz *et al.* in 1989[104]. In 1997 the first live cell microscopy experiment studying the BFA treatment effect on Golgi redistribution to the Endoplasmic Reticulum (ER) was reported by Noah Sciaky *et al*[105], there confocal microscopy was used to observe disappearance of bright spatial structures. A fixed cell experiment was reported in 2008 by Heling Pan *et al*.[106], this study is worthy to mention due to the application of automated fluorescence microscopy image analysis approach.

In context of the biological question it was decided to develop a highly automated experiment setup: automated sample preparation, automated image acquisition and image

analysis routines. In our experiment setup the throughput limiting factor was image acquisition velocity.

For the experiment normal rat kidney cells (NRK) stably transfected with the Golgi enzyme GalT-CFP (beta1,4-galactosyltransferase I tagged by cyan fluorescent protein) were chosen. Efficient redistribution of GalT-CFP to the ER after the addition of BFA and reassembly after BFA wash-out was previously reported for this type of cells in a number of studies[107–109]. The goal was to evaluate GalT-CFP dynamics on a single cell basis in living cells. An over-expression of ER and/or the Golgi complex resident proteins tagged by yellow fluorescent protein (YFP) was selected as gene amount altering method.

An automated image analysis routine was developed in collaboration with AG Rohr (Dr. P. Matula), which may be segregated into four major steps: 1) Detection and tracking of single cell nuclei stained with Hoechst 33342, 2) Partitioning of the images into disjoint regions, each encompassing a cell nucleus, called influence zones, 3) Quantification of GalT-CFP structures defined by Top-Hat transformation, the method to extract small spatial features from uneven background, for each influence zone, 4) Classification of cells expressing YFP-tagged candidate proteins versus non-expressing cells based on information in the YFP channel.

Data from all BFA treatment experiments essentially had the same trend – short delay of 3-5 min, which could be called “BFA lag time”, with no significant changes in the Golgi area (Top-Hat values) and following it near exponential decrease of the GalT-CFP area representing the Golgi complex. Soluble YFP expression revealed a BFA lag time of 3.25±0.9 min. That fits well with the BFA lag time previously reported, ~ 4 min, for GalT-GFP in HeLa cells[107]. Data from all BFA-washout experiments had a similar trend – relatively fast growth reaching a maximum value and consequent slow decay. The data were plotted against time and fitted to the model describing the amount of the intermediate product in consecutive two steps elementary reactions[61]. The fitting parameters were collected and averaged for every clone. Positions of GalT-CFP fluorescence intensity maxima were determined by finding extrema of the fitting function using the collected fitting parameters.

Our experiment is the first where Golgi relocation kinetics was approached in live cells in over-expression conditions of fluorescently tagged proteins. Results of the proof of principle screen identified a number of proteins that have a significant effect on the change

of the rate of GalT-CFP redistribution and reassembly when over-expressed; and the majority of them were missed in functional analysis of fixed cells.

Additionally to live cell experiments, fixed time point experiments were carried out to investigate morphology of the Golgi complex in greater details as there are evidences that functional activity is related to its spatial organization. Results supported the claim that live cell experiments and kinetic data analysis reveals more information on the biological process. In many cases we could not recognize significant differences of the Golgi complex morphology in fixed time point microscopy images. That could be concluded as: first, there are no morphological differences, second, wide field and confocal resolution is not sufficient to address the question. For collaborative project we had a task to build an integrated microscopy platform which would combine features of wide field, confocal and super-resolution (stochastic optical reconstruction microscopy; STORM) fluorescence microscopy techniques. To test our microscopy platform we required a biological sample. We decided that investigation of the Golgi complex morphology, especially the co-localization or co-existence of different Golgi compartments on such integrated microscopy platform, a correlative fluorescence microscopy, could be of great interest. The results of the proof of principle experiments confirmed the lack of resolving power of the wide field and confocal microscopy techniques for addressing co-organization of the Golgi complex compartments, where STORM is capable to resolve fine structure[117].

Statistical data analysis of a cargo protein secretion screen

The main aim of this work was to identify which members of the RAB GTPase family are involved in the collagen-I secretion pathway. The RAB GTPase family is known for controlling various stages of membrane trafficking. RAB GTPases utilize guanosine nucleotide exchange and guanosine tri-phosphate (GTP) hydrolysis to switch between the active (GTP-bound) and the inactive (guanosine di-phosphate; GDP-bound) conformations. Switching from the inactive to active state requires the GDP/GTP exchange factor (GEF). After performing functions in their active form, GTPases are inactivated by hydrolysing GTP to GDP, which is accelerated by the GTPase-activating protein (GAP). Once inactivated, small GTPases detach from the membrane and are kept in the GDP-bound inactive state until the next round of the GTPase activation cycle begins. Dissociation of most RAB GTPases from membranes is mediated by a conserved protein family, the RAB GDP dissociation inhibitor

(RAB GDI) which binds only to GDP-bound RABs and keeps them in the GDP-bound state by inhibiting GDP release.

In our study collagen-I is the cargo of membrane trafficking. Collagen biosynthesis involves a number of post-translational modifications and requires at least nine ER-resident enzymes and molecular chaperones[28,27]. Protein disulphide isomerase (PDI) assists protein folding in the ER by catalysing the formation, reduction and isomerisation of disulphide bonds. As a subunit of the enzyme prolyl 4-hydroxylase, PDI is also essential for proline hydroxylation, the key post-translational modification underpinning helix formation. The main function of PDI in prolyl 4-hydroxylase seems to be to keep the highly insoluble α subunits in a catalytically active, non-aggregated state. Prolyl 4-hydroxylase requires ferrous ions, 2-oxoglutarate, O_2, and ascorbate[27]. Strong PDI functional dependency on ascorbate makes procollagen-I a very good secretory cargo for well-controlled experimentation, as in the cells grown without ascorbate procollagen-I is trapped in the ER[37].

We used a library of human GFP tagged RAB GTPases (kind gift from B. Gaud, Institute Currie, Paris) and NIH-3T3 cell line as it is known for high native procollagen-I expression levels. Experiments were performed as follows: 1) cells are plated on multi-well cell culture plate, 2) transfected with according cDNA vector 24h before the trafficking experiment, 3) cells are incubated with ascorbate and cycloheximide for 105min before fixation with paraformaldehyde, 4) the last step is immuno staining and microscopy image acquisition. Microscopy images were analysed on a single cell basis using the Olympus Scan^R automated image analysis software. It should be noted that we actually measured procollagen-I retained in cells rather than the secreted one, due to this it is better to describe the experimentation as procollagen-I retention assay.

Data normalization according to the standard deviation of the negative (neutral) control is one of the mostly used methods to interpret RNAi (RNA interference) screens. For our experimental data such population normalization was not a sufficient method. Two main factors supporting this claim are: first, the probability to observe the cell over-expressing specific amount of fluorescently tagged protein decreases exponentially towards high mean fluorescence values of over-expressed protein and such probability varies from one clone to the other, second, the distribution of procollagen-I mean fluorescence values in a population is not a normal but is in good agreement with gamma distribution. Due to these reasons, another data interpretation approach had to be developed. As over-

expressed clones were reported by the GFP tag, it was possible to segregate the observed populations according to the mean fluorescence values of the over-expressed proteins. That allowed plotting of the procollagen-I retained in cells in dependence on the amount of the over-expressed protein. The assumption that over-expression of the negative control (CFP) has little to none effect on procollagen retention in the cells was confirmed. In contrast, low level of over-expression of the Sar1GTP had a statistically not significant increase of retained procollagen-I and then a fast growth till procollagen-I saturation is reached at about twice the amount of the negative control population. Additionally, data scaling according to the procollagen-I mean value of the negative control populations for separate experiments was used. That allowed reducing day to day variations of the acquired data (evaluated as standard deviation) from about 70% of the value to 10% when scaling was used.

The difference of the retained procollagen-I fluorescence signal in sub-populations sorted according RAB GTPase over-expression intensity value showed to be the best “hit” calling parameter for our experimental data. Our screen reproduced the results of functionally described RAB GTPases as well as brought not yet described proteins into the picture of procollagen-I trafficking. Further extended analysis is required to describe the effect and function of the “hit” proteins on trafficking of procollagen-I.

An application of over-expression of fluorescently tagged RAB GTPases in context of vesicular trafficking or other cell cycle related functions was published in numerous studies, to name just a few: Nahuel Romero *et al*.(RAB1b)[120], Alexander K. Haas *et al*.(RAB1 and RAB43)[121], Laura Cogli *et al*.(RAB7a)[122]. The recent publication by Noushin Nabavi *et al*.[123] reported investigation of RAB GTPase mRNA up-regulation in osteoblast cells while stimulated by ascorbic acid and correlating RAB GTPase mRNA up-regulation to collagen production; in this publication over-expression of fluorescently tagged RAB GTPases also was used together with numeric evaluation of fluorescence imagery, however results were expressed just as population average and conducted from a relatively low population (~50 cells) in contrast to our experiments where populations over 3000 cells were analysed. Our experiment is the first systematic approach to investigate the role of RAB GTPases in the procollagen secretion pathway, as we tested all currently known RAB members, most of them in form of wild type, dominant active and dominant inactive clones.

The developed data analysis approach for an evaluation of over-expressed RAB GTPases effect on collagen secretion could be applied for various different biological cases and it was already applied to evaluate RNAi rescue experiments[124].

The development of a fluorescence microscopy based experiment to investigate mRNA and miRNA interactions in live cells

Micro RNAs (miRNAs) are known for interfering with the messenger RNA (mRNA) translation process through complex cellular mechanism known as RNA interference (RNAi). Currently available computational prediction algorithms of interacting miRNA-mRNA pairs lack of accuracy. Direct observation of miRNA-mRNA interaction is hardly feasible so interaction evaluation is based on indirect measurements of mRNA or translation product amount in cell lysate. Our aim was to develop a high-throughput fluorescence microscopy based experiment to evaluate RNAi activity in single live cells.

Our developed fluorescence microscopy based approach allows a systematic numeric evaluation of RNAi in context of miRNA-mRNA interactions in live and fixed cell cultures in high-throughput manner. Additionally, our approach allows tracking of mRNA localization in live cells. We adopted the mRNA labelling technique described by Nathalie Daigle and Jan Ellenberg[115], a λ_{N22} aptamers based DNA vector system. Initially this system was designed for visualization of protein coding mRNA and product protein tagged with RFP. In order to study miRNA-mRNA interactions we had to redesign the system by introducing 3'UTRs of the genes of interest (for example:TBC1D2), which were selected from previous protein secretion studies (Andrius Serva, AG Starkuviene).

DNA vector based approaches with fluorescence or firefly luciferase readouts studying miRNA activity were known for a while, for example, an application of a DNA vector based fluorescent miRNA activity reporter in developing vertebrate embryos was reported by Davide De Pietri Tonelli *et al*. in 2006[125], this approach was designed to monitor the dynamics of specific miRNA activity in individual live cells of the developing organism. Stefan Ludwig Ameres *et al*.[126] in 2007 presented luciferase assay results that correlated mRNA secondary structure to gene silencing activity by siRNAs. Another DNA vector based fluorescent RNAi activity reporter approach was reported by Joseph A. Grendell *et al*.[127], the authors used DNA vector encoding GFP protein and a number of siRNAs targeting different sites of the gene. A fluorescence read out was performed using a plate reader, thus giving an integrated value for the whole population. RNAi activity

correlation to secondary structure of mRNA was reported, target sites with single stranded regions were prone to show higher silencing activity.

The seed sequence (2-7th nucleotide (nt) from 5' end of a miRNA) alignment algorithm was proposed by our collaborative partners (AG Kummer) as reliable approach for miRNA-mRNA pair identification. Seed sequence alignment of a human miRNA library[116] against TBC1D2 resulted in a high number of miRNAs, more than 500 out of 1900. As miRNAs are still relatively expensive reagents, it was decided to test if siRNAs selected by the seed sequence length criteria would indeed result in an effective RNAi. For this reason seed sequences of the Silencer® siRNA library (Ambion) were aligned with our DNA construct and 11 siRNAs from the ones having a match (seed length from 9 to 14 nt) were randomly selected for the experiment. It was expected that the seed length should have a direct correlation with RNAi activity. Just two siRNAs generated a statistically significant reduction of the reporter translation. Opposite to the expectation, the most effective siRNA had the shortest seed sequence. Thermodynamic analysis (mfold software) of siRNAs and their target sites suggested that secondary structure of the mRNA is an important factor for an effective RNAi occurrence. As gathered data was not sufficient to build a quantitative interaction prediction model, a large scale experiment to investigate miRNA-mRNA interactions was developed and will be executed in the nearest future.

Zusammenfassung

Die Molekularbiologie der Zelle ist eine anwachsende Sammlung von Konzepten und Erkenntnissen, welche sich durch das Studium verschiedenster Zellen und Organismen ständig weiterentwickelt. Das Hauptaugenmerk liegt hierbei auf molekularen Strukturen und deren Funktion in zellulären Prozessen. Ein akzeptiertes Modell von Säugerzellen geht davon aus, dass die Zelle ein membranbasiertes System ist, in welchem alle Prozesse durch Proteine in Gang gesetzt werden. Proteine und deren spezifische Interaktionen bestimmen spezifische Eigenschaften von Membranen, welche daraufhin individuelle funktionale Einheiten innerhalb der Zelle ausbilden. Mebrantransport, eines der Interessensgebiete der Molekularbiologie, ist eine Abstraktion, welche den Materialtransport innerhalb verschiedener Organellen der Zelle beschreibt. Bei der Untersuchung des Membrantransports ist dessen Zusammenhang mit einer Vielzahl schwerer Erkrankungen von besonderem Interesse. Der Membrantransport ist ein mehrstufiger Materialtransportprozess von der Donor- zur Zielmembran, wobei Vesikel ausgebildet und sowohl Transport- als auch Fusionsprozesse eingesetzt werden. Von besonderem Interesse sind Gene, welche in den Membrantransport und dessen Regulierung innerhalb der Zelle involviert sind.

Abläufe des Membrantransports in nativem oder nahezu nativem Kontext, wie etwa in fixierten oder lebendigen Zellkulturen, werden bevorzugt mikroskopisch beobachtet da hierdurch ein räumlicher und zeitlicher Zusammenhang zwischen verschiedenen Stufen des gesamten Prozesses und spezifischen Zellorganellen hergestellt werden kann. Außerdem kann eine bessere Statistik als in Methoden, welche auf einem Populationsmittelwert beruhen erzielt werden, da die Ergebnisse auf der direkten Beobachtung einer großen Anzahl individueller Zellen beruhen. Durch die Fluoreszenzmikroskopie können verschiedene Fragestellungen adressiert werden, beispielsweise wie viele Kopien eines spezifischen Proteins vorliegen, wo diese innerhalb der Zelle lokalisiert sind oder, in manchen Fällen, zu welchem Zeitpunkt diese translatiert werden. Trotz der Vielseitigkeit der Fluoreszenzmikroskopie benötigt die Anwendung jeder speziellen Technologie im Bereich der Zellbiologie mehrere zusätzliche Anwendungsebenen. Einige zusätzliche Anwendungsebenen sind: Fluorophore und Markierungsstrategien, Datenakquisition und –

speicherung, digitale Bildverarbeitung, Dateninterpretation im biologischen Kontext und – nicht zuletzt – der zu untersuchende biologische Prozess selbst.

Diese Dissertation beschreibt die von mir durchgeführten experimentellen Arbeiten und die zugehörige Datenanalyse, um folgende spezifischen Fälle im Zusammenhang mit dem Membrantransport zu untersuchen und auszuwerten:

a) Der Einfluss überexprimierter Membranproteine auf die Organisationskinetik des Golgi Apparates, welcher eines der Hauptorganellen des Membranstransportes ist. Hierzu wurde Lebendzell-Fluoreszenzmikroskopie eingesetzt[73,110,118].
b) Statistische Datenanalyse eines Transportprotein-Sekretions-Screens. Der Einfluss der Überexpression spezifischer RAB GTPasen auf die Kollagen-I-Sekretion wurde bestimmt[119].
c) Die Entwicklung Fluoreszenzmikroskopie basierter Methoden, um die Interaktionen von Boten-RNA (*messenger RNA*, mRNA) und microRNA (miRNA) in lebenden Zellen zu untersuchen.

Bestimmung des Einflusses überexprimierter Membranproteine auf die Organisationskinetik des Golgi Apparates

Zur Identifikation von Effektoren, welche nicht durch andere biochemische Verfahren entdeckt werden können, beispielsweise durch einen unterschiedlichen Interaktionsmechanismus, wurde eine Lebendzellmessung des Membrantransports durchgeführt. Die Untersuchung der Organisation des Golgi Apparates wurde aufgrund dessen entscheidender Rolle beim Membrantransport, dessen relativ leichter Visualisierbarkeit und der leichten Induzierung morphologischer Veränderungen durch mehrere verschiedene Chemikalien gewählt.

Die Morphologie des Golgi Apparates wurde bereits mit unterschiedlichen Methoden ausführlich untersucht[83–94]. Mehrere biologische Studien zeigen, dass die Struktur des Golgi Apparates hochdynamisch und stark abhängig vom Gleichgewicht des Materieflusses ist[95–97]. Eine umfassende Auflistung von Proteinen, welche am Golgi Apparat angesiedelt sind und mit diesem in funktionalem Zusammenhang stehen, wird momentan zusammengestellt[98–102]. Die Mehrzahl der vielseitigen Erkenntnisse über Struktur und Funktion des Golgi Apparates wurden jedoch aus Experimenten mit fixierten

Zellen gewonnen. Erste Berichte zu den Untersuchungen des Effektes von Brefeldin A (BFA, ein Pilzmetabolit) auf den Golgi Apparat kamen 1988 von Toshiyuki Fujiwara *et al*.[103], gefolgt von Jennifer Lippincott-Schwartz *et al.* 1989[104]. 1997 berichteten Noah Sciaky *et al*.[105] von ersten Lebendzellexperimenten, in welchen der BFA-Effekt auf die Umverteilung vom Golgi Apparat zum Endoplasmatischen Retikulum (ER) untersucht wurde. Hierbei wurde das Verschwinden heller räumlicher Strukturen mittels Konfokalmikroskopie untersucht. Heling Pan *et al*.[106] beschrieben 2008 ein Experiment mit fixierten Zellen, welches nennenswert ist, da hier die Fluoreszenzmikroskopie automatisiert angewandt wurde.

Im Kontext der biologischen Fragestellung wurde beschlossen, einen hochautomatisierten Experimentaufbau zu entwickeln: automatisierte Probenvorbereitung, automatische Bildaufnahme und Bildauswerteroutinen. In unserem Aufbau war der den Durchsatz limitierende Faktor die Geschwindigkeit der Bildaufnahme.

Für das Experiment wurden stabil transfizierte normale Rattennieren (*normal rat kidney cells*, NRK) mit dem Golgi-Enzym GalT-CFP (Beta1,4-Galactosyltransferase I markiert mit dem cyan fluoreszenten Protein) ausgewählt. Die effiziente Umverteilung von GalT-CFP zum ER nach der Zugabe von BFA und die Wiederherstellung nach dem Auswaschen von BFA werden in einigen Studien beschrieben[107–109]. Ziel war es, die GalT-CFP-Dynamik auf der Basis einzelner Zellen in lebenden Zellen zu bestimmen. Als Methode zur Veränderung der Genanzahl wurde die Überexprimierung von dem ER und / oder dem Golgi Apparat zugehöriger Proteine gewählt.

In Zusammenarbeit mit der AG Rohr (Dr. P. Matula) wurde eine automatische Bildverarbeitungsroutine entwickelt, welche in vier Hauptschritte unterteilt werden kann:
1) Erkennung und Verfolgung von, mit Hoechst 33342 markierten, Zellkernen, 2) Aufteilung der Bilder in disjunkte Regionen, Einflusszonen genannt, von denen jede einen Zellkern umgibt, 3) Quantifizierung der durch Top-Hat-Transformation definierten GalT-CFP-Strukturen in jeder Einflusszone; hierdurch können kleine räumliche Eigenschaften von einem ungleichmäßigen Hintergrund getrennt werden, 4) Klassifizierung der Zellen in YFP-markierte Zielproteine exprimierende und nichtexprimiernde Zellen aufgrund der Informationen aus dem YFP-Farbkanal.

Die Daten aller Experimente mit BFA-Behandlung zeigten die gleiche Tendenz – eine kurze Verzögerung von 3-5 Min, welche „BFA-Latenzzeit“ genannt werden kann, bei keiner

signifikanten Änderung innerhalb der Golgi Fläche (Top-Hat-Werte), gefolgt von einem nahezu exponentiellen Abfall der den Golgi Apparat repräsentierenden Fläche von GalT-CFP. Die Expression von löslichem YFP ergab eine BFA-Latenzzeit von 3.25±0.9 Min, was gut mit der zuvor berichteten BFA-Latenzzeit von ~4 Min für GalT-GFP in HeLa-Zellen zusammenpasst[107]. Die Daten aller BFA-Auswaschexperimente hatten eine ähnliche Tendenz – ein relativ zügiges Wachstum bis zu einem Maximalwert und darauffolgend ein langsamer Abfall. Die Daten wurden gegen die Zeit aufgetragen und hieran ein Modell, welches die Menge des Zwischenproduktes in aufeinanderfolgenden Zweischrittelementarreakktionen beschreibt, angepasst[61]. Die Fitparameter wurden für jeden Klon zusammengefasst und gemittelt. Die Positionen der GalT-CFP-Intensitätsmaxima wurden durch Auffinden der Extrema der Fitfunktion unter Verwendung der gesammelten Fitparameter bestimmt.

Unser Experiment ist das erste, in welchem die Golgi-Relokationskinetik in lebenden Zellen unter Überexpressionsbedingungen von fluoreszenzmarkierten Proteinen untersucht wurde. In Proof-of-Principle-Screens wurden mehrere Proteine identifiziert, welche bei Überexprimierung einen signifikanten Einfluss auf die Rate der GalT-CFP Umverteilung und Wiederherstellung hatten. Die Mehrzahl dieser Proteine konnten nicht durch funktionale Analysen mit fixierten Zellen ermittelt werden.

Zusätzlich zu den Lebendzellmessungen wurden Experimente mit fixem Zeitpunkt durchgeführt, um die Morphologie des Golgi Apparates genauer zu untersuchen, da es Hinweise auf einen Zusammenhang zwischen funktionaler Aktivität und räumlicher Anordnung gibt. Die Ergebnisse unterstützen die These, dass Lebendzellexperimente und die Analyse kinetischer Daten mehr Informationen über den biologischen Prozess liefern. In vielen Fällen konnten wir in Mikroskopaufnahmen mit fixem Zeitpunkt keine signifikanten Änderungen in der Morphologie des Golgi Apparates erkennen. Mögliche Schlussfolgerungen hiervon sind: Erstens, es gibt keine morphologischen Unterschiede, zweitens, die Auflösung von Weitfeld- und Konfokalmikroskopie ist nicht ausreichend, um diese Frage zu beantworten.

In einem Kollaborationsprojekt hatten wir die Aufgabe, eine integrierte Mikroskopplattform zu entwickeln, welche die Eigenschaften der Fluoreszenzmikroskopietechniken Weitfeldmikroskopie, Konfokalmikroskopie und Superauflösung (*stochastic optical reconstruction microscopy*; STORM) kombiniert. Um

unsere Mikroskopplattform zu entwicheln benötigten wir eine biologische Probe. Wir entschieden uns für die Untersuchung der Morphologie des Golgi Apparates, insbesondere der Kolokalisation oder Koexistenz verschiedener Golgi-Kompartimente. Die Ergebnisse der Proof-of-Principle-Experimente bestätigen, dass das Auflösungsvermögen von Weitfeld- und Konfokalmikroskopie nicht ausreicht, um die Koorganisation des Golgi Apparates zu untersuchen, während STORM in der Lage ist, auch feine Strukturen aufzulösen[117].

Statistische Datenanalyse eines Cargoprotein-Sekretionsscreens

Das Hauptziel dieser Arbeit war es, zu identifizieren welche Mitglieder der RAB GTPase-Familie im Kollagen-I Sekretionspfad involviert sind. Es ist bekannt, dass die RAB GTPase-Familie verschiedene Stufen des Membrantransportes kontrolliert. RAB GTPasen verwenden Guanosinnukleotidaustausch und Guanosintriphosphathydrolyse, um zwischen der aktiven (GTP-gebundenen) und der inaktiven (Guanosindiphosphat, GDP-gebundenen) Konfiguration umzuschalten. Beim Wechsel von der inaktiven zur aktiven Konformation wird der GDP/GTP Austauschfaktor (GEF) benötigt. Nach der Ausführung von Funktionen in ihrer aktiven Form werden GTPasen durch Hydrolyse von GTP zu GDP deaktiviert, was durch das GTPase-aktivierende Protein (GAP) beschleunigt wird. Sind sie einmal inaktiv, lösen sich GTPasen von der Membran ab und werden in dem GDP-gebundenen inaktiven Zustand gehalten bis der nächste Durchlauf des GTPase-Aktivierungszyklus beginnt. Die Dissoziation der meisten RAB GTPasen von der Membran wird durch eine konservierte Proteinfamilie vermittelt, die RAB GDP Dissozationsinhibitoren (RAB GDI), welche nur an GDP-gebundene RABs binden und diese durch Inhibition der GDP-Freigabe im GDB-gebundenen Zustand halten.

In unserer Studie ist die Fracht beim Membrantransport Kollagen-I. Die Biosynthese von Kollagen beinhaltet einige posttranslationale Modifizierungen und benötigt mindestens neun im ER angelagerte Enzyme und molekulare Chaperone[28,27]. Protein-Disulphid-Isomerase (PDI) unterstützt die Proteinfaltung im ER durch Katalyse der Bildung, Reduktion und Isomerisierung von Disulfidbrücken. Als eine Untereinheit des Enzyms Proxyl 4-Hydroxylase ist PDI essentiell für die Hydroxylierung von Prolin, die posttranslationale Schlüsselmodifikation welche die Helix-Formation unterstützt. Die Hauptfunktion von PDI in der Prolyl 4-Hydroxylase scheint es zu sein, die hochgradig unlöslichen α Untereinheiten in einem katalytisch aktiven, nichtaggregierten Zustand zu halten. Prolyl 4-Hydroxylase benötigt Eisenionen, 2-Oxoglutarat, O_2, und Ascorbat[27]. Eine starke funktionale

Abhängigkeit von PDI von Ascorbat macht Prokollagen-I zu einer geeigneten Sekretionsfracht für gut kontrollierte Untersuchungen, da in Zellen, welche ohne Ascorbat aufgezogen werden, Prokollagen-I im ER eingeschlossen bleibt[37].

Wir verwendeten eine Bibliothek menschlicher GFP-markierter RAB GTPasen (freundlicherweise zur Verfügung gestellt von B. Gaud, Institute Currie, Paris) und die Zellinie NIH-3T3, da diese für hohe native Prokollagen-I Expressionslevel bekannt ist. Den Experiment wurde folgendermaßen durchgeführt: 1) die Zellen werden auf Multiwell-Kulturplatten aufgebracht und 2) mit dem entsprechenden cDNA-Vektor 24 Std vor dem Transportexperiment transfiziert, 3) die Zellen werden vor der Fixierung mit Paraformaldehyd für 105 Min mit Ascorbat und Cycloheximid inkubiert, 4) der letzte Schritt ist die Immunfärbung und die mikroskopische Bildaufnahme. Die Mikroskopbilder wurden auf Einzelzellbasis mittels der automatischen Analysesoftware des Olympus Scan^R Systems ausgewertet. Es sollte erwähnt werden, dass anstelle des sekretierten, das in der Zelle verbleibende, Prokollagen-I gemessen wurde, daher lässt sich das Experiment besser als Prokollagen-I Retentionsassay beschreiben.

Datennormalisierung aufgrund der Standardabweichung der Negativkontrolle (Neutralkontrolle) ist eine weit verbreitete Methode, um RNAi (RNA Interferenz) Screens zu interpretieren. Im Falle unserer Daten war dies nicht zufriedenstellend. Zwei Hauptgründe hierfür sind: erstens sinkt die Wahrscheinlichkeit zu beobachten, wie eine Zelle eine bestimmte Menge eines markierten Proteins überexprimiert, exponentiell zu hohen Werten für die mittlere Fluoreszenz von überexprimierten Proteinen ab und diese Wahrscheinlichkeit schwankt von einem Klon zum anderen, zweitens sind die Werte der mittleren Prokollagen-I Fluoreszenz in einer Population nicht normalverteilt, sondern lassen sich in guter Näherung durch eine Gammaverteilung beschreiben. Daher musste ein anderer Ansatz zur Dateninterpretation entwickelt werden. Da überexprimierte Klone durch die GFP-Markierung angezeigt wurden war es möglich, die beobachteten Populationen aufgrund der mittleren Fluoreszenzwerte der überexprimierten Proteine aufzuteilen. Hierdurch konnte der Anteil des in den Zellen zurückgehaltenen Prokolagen-I in Abhängigkeit des überexprimierten Proteins aufgetragen werden. Die Annahme, dass eine Überexpression der Negativkontrolle (CFP) einen geringen bis keinen Effekt auf den Rückhalt von Prokollagen in den Zellen hat wurde bestätigt. Im Gegensatz hierzu zeigte ein geringer Level an Sar1GTP Überexpression einen statistisch nicht signifikanten Anstieg an

zurückgehaltenem Prokollagen-I und dann ein schnelles Wachstum bis zur Sättigung von Prokollagen-I bei etwa dem zweifachen Wert der Negativkontrolle. Zusätzlich wurden die Daten für jedes Experiment bezüglich der Prokollagen-I Mittelwerte der Negativkontrolle skaliert. Durch die Skalierung konnten die Variationen der aufgenommen Daten (berechnet als Standardabweichung) von einem Tag zum anderen von 70% auf 10% reduziert werden.

In unserem Experiment war der Unterschied des zurückgehaltenen Prokollagen-I Fluoreszenzsignals in Subpopulationen angeordnet nach den Intensitätswerten der Überexpression der RAB GTPase der beste „Hit"-Parameter. Unser Screen reproduzierte die Ergebnisse funktionell beschriebener RAB GTPasen und konnte zusätzlich Proteine in den Zusammenhang mit Prokollagen-I Transport bringen, für welche dies zuvor noch nicht bekannt war.

Eine Anwendung überexprimierter fluoreszenzmarkierter RAB GTPasen im Kontext von vesikulärem Transport auf andere mit dem Zellzyklus in Verbindung stehende Funktionen wurde in vielzähligen Studien veröffentlicht. Um einige zu nennen: Nahuel Romero *et al*.(RAB1b)[120], Alexander K. Haas *et al*.(RAB1 and RAB43)[121], Laura Cogli *et al*.(RAB7a)[122]. Die aktuelle Publikation von Noushin Nabavi *et al*.[123]beschreibt die Untersuchung von RAB GTPase mRNA Hochregulierung in Osteoblastenzellen bei gleichzeitiger Stimulierung durch Ascorbinsäure und korreliert RAB GTPase mRNA Hochregulierung mit der Kollagen Produktion. In dieser Publikation wurde ebenfalls die Überexpression fluoreszenzmarkierter RAB GTPasen zusammen mit einer numerisch berechneten Fluoreszenzanimation beschrieben, allerdings wurden im Gegensatz zu unserem Experiment, in welchem Populationen von über 3000 Zellen analysiert wurden, die Ergebnisse lediglich als Populationsmittel angegeben und von einer relativ kleinen Population (~50 Zellen) abgeleitet. Unser Experiment ist der erste systematische Ansatz zur Untersuchung der Rolle der RAB GTPasen im Sekretionspfad von Prokollagen, da wir alle derzeit bekannten RAB Mitglieder getestet haben, die meisten in Form des Wildtyps, als dominant aktiven und als dominant inaktiven Klon.

Der hier entwickelte Ansatz zur Analyse des Effektes überexprimierter RAB GTPasen auf die Kollagensekretion kann für viele verschiedene biologische Fragestellungen angepasst werden und wurde bereits zur Auswertung von RNAi-Rettungsexperimenten verwendet[124].

Entwicklung eines Fluoreszenzmikroskop basierten Experiments zur Untersuchung von mRNA und miRNA Interaktionen in lebenden Zellen

MicroRNAs (miRNAs) sind bekannt für ihre Interferenz mit dem Botschafter-RNA (mRNA) Translationskomplex durch den komplexen zellulären Mechanismus der RNA Interferenz (RNAi). Den derzeit verfügbaren computergestützten Vorhersagealgorithmen zur Bestimmung von interagierenden miRNA-mRNA Paaren mangelt es an Genauigkeit. Direkte Beobachtung der miRNA-mRNA Interaktion ist schwerlich möglich, daher basiert die Bestimmung der Interaktion auf indirekten Messungen der mRNA oder der Menge an Translationsprodukten im Lysat der Zellen. Unser Ziel war die Entwicklung eines auf Fluoreszenzmikroskopie basierten Hochdurchsatzexperiments zur Bestimmung der RNAi Aktivität in einzelnen lebenden Zellen.

Unser Ansatz ermöglicht eine systematische numerische Bestimmung der RNAi bezogen auf miRNA-mRNA Interaktionen in lebenden und fixierten Zellkulturen bei hohem Durchsatz. Gleichzeitig kann die mRNA Lokalisation in lebenden Zellen verfolgt werden. Hierfür passten wir die mRNA Markierungstechnologie beschrieben von Nathalie Daigle und Jan Ellenberg[115] an, ein λ_{N22} Aptamer basiertes DNA Vektorsystem. Ursprünglich wurde dies zur Visualisierung proteinkodierender mRNA und deren Produktprotein mittels RFP entwickelt. Um miRNA-mRNA Interaktionen zu untersuchen musste das System umgestaltet werden. Hierzu wurden 3'UTRs der zu untersuchenden Gene (beispielsweise TBC1D2) eingeführt, welche aufgrund vorhergehender Sekretionsstudien (Andrius Serva, AG Starkuviene) ausgewählt wurden.

DNA Vektor basierte Ansätze zur Untersuchung der miRNA Aktivität, welche als Indikatoren Fluoreszenz oder Glühwürmchen Luziferase verwenden, sind schon seit längerem beschrieben. Ein Beispiel eines derartigen Systems mit fluoreszentem Reporter in sich entwickelnden Wirbeltierembryos wurde 2006 von Davide De Pietri Tonelli *et al.*[125] veröffentlicht. Das Ziel war, die Dynamik spezifischer miRNA Aktivität in den lebenden Zellen sich entwickelnder Organismen zu überwachen. Stefan Ludwig Ameres *et al*.[126] präsentierten 2007 Ergebnisse eines Luziferase Assays, welcher die sekundäre mRNA Struktur mit der Stillegung von Genen korrelierte. Ein weiterer Ansatz mit einem DNA Vektor basierten fluoreszenten RNAi Aktivitätsreporter wurde von Joseph A. Grendell *et al*.[127] beschrieben, die Autoren verwendeten einen GFP kodierenden DNA Vektor und mehrere siRNAs, welche auf verschiedene Sequenzen dieser DNA abzielten. Die

aufsummierten Fluoreszenzwerte der jeweiligen Populationen wurden simultan mit einem Plattenlesegerät ausgelesen. Es konnte ein Zusammenhang der RNAi Aktivität mit der sekundären Struktur festgestellt werden, Zielsequenzen mit Einzelstrang-Regionen wiesen eine erhöhte Stillegungsrate auf.

Ein Algorithmus zur Anordnung der Keimsequenz (2-7. Nukleotid (nt) vom 5‘ Ende der miRNA) wurde von unseren Kollaborationspartnern (AG Kummer) als verlässlicher Ansatz zur Auffindung von miRNA-mRNA Paaren vorgeschlagen. Die Keimsequenzanordnung einer Bibliothek menschlicher miRNA[116] gegen TBC1D2 führte zu einer großen Anzahl miRNAs, mehr als 500 von insgesamt 1900. Da miRNAs immer noch relativ kostspielige Reagenzien sind wurde entschieden zu testen, ob nach der Länge der Keimsequenz ausgewählte siRNAs tatsächlich zu einer effektiven RNAi führen. Zu diesem Zweck wurden Keimsequenzen der Silencer® siRNA Bibliothek (Ambion) mit unserem DNA Konstrukt abgeglichen und von den übereinstimmenden (Länge der Keimsequenz 9-14 nt) 11 siRNAs zufällig für das Experiment ausgewählt. Es wurde ein direkter Zusammenhang zwischen der Länge der Keimsequenz und der RNAi Aktivität erwartet. Nur zwei siRNAs induzierten einen signifikanten Rückgang bei der Translation des Reporters. Im Gegensatz zur Erwartung hatte die effizienteste siRNA die kürzeste Keimsequenz. Thermodynamische Analysen (mfold Software) der siRNAs und ihrer Zielregionen legten nahe, dass die sekundäre Struktur der mRNA einen wichtigen Faktor für eine effektive RNAi darstellt. Da die aufgenommenen Daten nicht ausreichten, um ein quantitatives Vorhersagemodell für die Interaktion zu erstellen, wurde ein großangelegten Experiment zur Untersuchung der miRNA-mRNA Interaktion entwickelt, welches in der nahen Zukunft durchgeführt wird.

List of publications arising from my work described in this thesis

1. Erfle H, **Lisauskas T**, Claas C, Reymann J, Starkuviene V (2011) Cell arrays for the measurement of organelle dynamics in living cells. Methods in molecular biology 706: 73–81.
2. Matula P, **Lisauskas T**, Starkuviene V, Rohr K (2011) Quantification of Golgi Complex Assembly and Disassembly in Live Cell Fluorescence Microscopy Images. Proceedings of Microscopic Image Analysis with Applications in Biology (MIAAB) Workshop, Heidelberg: 1–5.
3. **Lisauskas T**, Matula P, Claas C, Reusing S, Wiemann S, et al. (2012) Live-cell assays to identify regulators of ER-to-Golgi trafficking. Traffic 13: 416–432.
4. Serva A, Knapp B, Tsai Y-T, Claas C, **Lisauskas T**, et al. (2012) miR-17-5p Regulates Endocytic Trafficking through Targeting TBC1D2/Armus. PloS one 7: e52555.

Publications in preparation or submitted at the time of writing this thesis:

1. Flottmann B*, Gunkel M*, **Lisauskas T***, Heilemann M, Starkuviene V, Reymann J and Erfle H. Correlative light microscopy for high-content screening. (2013) (BioTechniques, **Accepted**).
2. Tsai Y-T*, **Lisauskas T***, Claas C*, Knapp B, Reusing S, Kaderali L, Erfle H, Goud B, Starkuviene V. Comparative knock-down and knock-in screens of Rab GTPases identifies Rab40c as a novel regulator of biosynthetic trafficking. (2013) (In preparation).
3. Eskova A, Knapp B, Matelska D, **Lisauskas T**, Pepperkok R, Russell R, Eils R, Kaderali L, Erfle H, and Starkuviene V. 1KIF15 emerges as a novel regulator of integrin endocytic trafficking. (2013) (Submitted).

*Authors contributed equally

Comments on my participation

The research projects that I was working on and describe in this thesis were built as collaborative initiatives and could not be realized without involvement of many other people, so here I would like to state my contribution.

My contribution for High-throughput screening microscopy project which in my thesis stands as “Evaluation of the influence of over-expressed membrane proteins on the organization kinetics of the Golgi complex” was realized in sample preparation, cell culture, experiment execution and fluorescence microscopy image derived data analysis. In frame of this project an automated image analysis routine was developed in collaboration and executed by Dr. Petr Matula (AG Rorh).

My contribution for “Statistical data analysis of RAB GTPases over-expression on collagen-I secretion” was in development of High-throughput oriented fluorescence microscopy derived data analysis. The experiment was developed by Dr. Vytaute Starkuviene and executed by Susanne Reusing.

My contribution for the project “High-throughput and super resolution fluorescence microscopy platform to study miRNA targets in live cells” which in my thesis stands as “The development of a fluorescence microscopy based experiment to investigate mRNA and miRNA interactions in live cells” was in development of High-throughput capable experiment and data analysis approach, that includes investigation of available mRNA labelling techniques, evaluation of siRNA/miRNA activity reduction due to fluorescent labelling, bioinformatics based selection of siRNAs for the experiments. Modifications of the λ_{N22} aptamers based system were designed by Andrius Serva and executed by Susanne Reusing. Experiment for the correlative fluorescence microscopy approach was designed and executed by me; wide field and confocal imagery acquired in collaboration with Dr. Manuel Gunkel; dSTORM images acquired by Benjamin Flottmann. Samples for solid phase transfection were prepared in collaboration with Ruben Bulkescher.

References

1. Cambridge Dictionaries Online. (1999). at <http://dictionary.cambridge.org>

2. Oxford Dictionaries Online. at <http://oxforddictionaries.com>

3. Kadziauskas, J. *Biochemistry basics*. 648 (Vilnius University press, 2012).

4. Crick, F. H. On protein synthesis. *Symposia of the Society for Experimental Biology* **12,** 138–63 (1958).

5. Crick, F. Central dogma of molecular biology. *Nature* **227,** 561–3 (1970).

6. Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, P. W. *Molecular biology of the cell*. 1392 (Garland Science, 2008).

7. Lodish H, Berk A, Zipursky SL, et al. *Molecular Cell Biology*. (2000). at <http://www.ncbi.nlm.nih.gov/books/NBK21471/>

8. Pyhtila, B. *et al.* Signal sequence- and translation-independent mRNA localization to the endoplasmic reticulum. *Spring* 445–453 (2008). doi:/10.1261/rna.721108

9. Springer, S., Spang, A. & Schekman, R. A primer on vesicle budding. *Cell* **97,** 145–8 (1999).

10. Sztul, E. & Lupashin, V. Role of tethering factors in secretory membrane traffic. *American journal of physiology. Cell physiology* **290,** C11–26 (2006).

11. Aridor, M. & Traub, L. M. Cargo selection in vesicular transport: the making and breaking of a coat. *Traffic (Copenhagen, Denmark)* **3,** 537–46 (2002).

12. Owen, D. J., Collins, B. M. & Evans, P. R. Adaptors for clathrin coats: structure and function. *Annual review of cell and developmental biology* **20,** 153–91 (2004).

13. Mellman, I. & Warren, G. The road taken: past and future foundations of membrane traffic. *Cell* **100,** 99–112 (2000).

14. Cai, H., Reinisch, K. & Ferro-Novick, S. Coats, tethers, Rabs, and SNAREs work together to mediate the intracellular destination of a transport vesicle. *Developmental cell* **12,** 671–82 (2007).

15. Pfeffer, S. R. Rab GTPases: specifying and deciphering organelle identity and function. *Trends in cell biology* **11,** 487–91 (2001).

16. Bourne, H. R., Sanders, D. A. & McCormick, F. The GTPase superfamily: a conserved switch for diverse cell functions. *Nature* **348,** 125–32 (1990).

17. Rajalingam, K., Schreck, R., Rapp, U. R. & Albert, S. Ras oncogenes and their downstream targets. *Biochimica et biophysica acta* **1773,** 1177–95 (2007).

18. Nielsen, E., Cheung, A. Y. & Ueda, T. The regulatory RAB and ARF GTPases for vesicular trafficking. *Plant physiology* **147,** 1516–26 (2008).

19. Barbacid, M. ras genes. *Annual review of biochemistry* **56,** 779–827 (1987).

20. Schwartz, S. L., Cao, C., Pylypenko, O., Rak, A. & Wandinger-Ness, A. Rab GTPases at a glance. *Journal of cell science* **120,** 3905–10 (2007).

21. Novick, P. & Brennwald, P. Friends and family: the role of the Rab GTPases in vesicular traffic. *Cell* **75,** 597–601 (1993).

22. Chavrier, P. *et al.* Hypervariable C-terminal domain of rab proteins acts as a targeting signal. *Nature* **353,** 769–72 (1991).

23. Anant, J. S. *et al.* Mechanism of Rab geranylgeranylation: formation of the catalytic ternary complex. *Biochemistry* **37,** 12559–68 (1998).

24. Stenmark, H. & Olkkonen, V. M. The Rab GTPase family. *Genome biology* **2,** REVIEWS3007 (2001).

25. Rybin, V. *et al.* GTPase activity of Rab5 acts as a timer for endocytic membrane fusion. *Nature* **383,** 266–9 (1996).

26. Shoulders, M. D. & Raines, R. T. Collagen structure and stability. *Annual review of biochemistry* **78,** 929–58 (2009).

27. Myllyharju, J. & Kivirikko, K. I. Collagens and collagen-related diseases. *Annals of medicine* **33,** 7–21 (2001).

28. Lamandé, S. R. & Bateman, J. F. Procollagen folding and assembly: the role of endoplasmic reticulum enzymes and molecular chaperones. *Seminars in cell & developmental biology* **10,** 455–64 (1999).

29. Prockop, D. J. & Kivirikko, K. I. Collagens: molecular biology, diseases, and potentials for therapy. *Annual review of biochemistry* **64,** 403–34 (1995).

30. McLaughlin, S. H. & Bulleid, N. J. Molecular recognition in procollagen chain assembly. *Matrix biology : journal of the International Society for Matrix Biology* **16,** 369–77 (1998).

31. Bonfanti, L. *et al.* Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. *Cell* **95,** 993–1003 (1998).

32. Prockop, D. J., Sieron, A. L. & Li, S. W. Procollagen N-proteinase and procollagen C-proteinase. Two unusual metalloproteinases that are essential for procollagen

processing probably have important roles in development and cell signaling. *Matrix biology : journal of the International Society for Matrix Biology* **16,** 399–408 (1998).

33. Kadler, K. E., Holmes, D. F., Trotter, J. a & Chapman, J. a. Collagen fibril formation. *The Biochemical journal* **316 (Pt 1,** 1–11 (1996).

34. Graham, H. K., Holmes, D. F., Watson, R. B. & Kadler, K. E. Identification of collagen fibril fusion during vertebrate tendon morphogenesis. The process relies on unipolar fibrils and is regulated by collagen-proteoglycan interaction. *Journal of molecular biology* **295,** 891–902 (2000).

35. Smith-Mungo, L. I. & Kagan, H. M. Lysyl oxidase: properties, regulation and multiple functions in biology. *Matrix biology : journal of the International Society for Matrix Biology* **16,** 387–98 (1998).

36. Buehler, M. J. Nature designs tough collagen: explaining the nanostructure of collagen fibrils. *Proceedings of the National Academy of Sciences of the United States of America* **103,** 12285–90 (2006).

37. Starkuviene, V. & Pepperkok, R. Differential requirements for ts-O45-G and procollagen biosynthetic transport. *Traffic (Copenhagen, Denmark)* **8,** 1035–51 (2007).

38. Napoli, C., Lemieux, C. & Jorgensen, R. Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *The Plant cell* **2,** 279–289 (1990).

39. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391,** 806–11 (1998).

40. Filipowicz, W. RNAi: the nuts and bolts of the RISC machine. *Cell* **122,** 17–20 (2005).

41. Macrae, I. J. *et al.* Structural basis for double-stranded RNA processing by Dicer. *Science (New York, N.Y.)* **311,** 195–8 (2006).

42. Krol, J., Loedige, I. & Filipowicz, W. The widespread regulation of microRNA biogenesis, function and decay. *Nature reviews. Genetics* **11,** 597–610 (2010).

43. Nykänen, a, Haley, B. & Zamore, P. D. ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **107,** 309–21 (2001).

44. Pellino, J. L., Jaskiewicz, L., Filipowicz, W. & Sontheimer, E. J. ATP modulates siRNA interactions with an endogenous human Dicer complex. 1719–1724 (2005). doi:10.1261/rna.2102805.stabilize

45. Weitzer, S. & Martinez, J. The human RNA kinase hClp1 is active on 3’ transfer RNA exons and short interfering RNAs. *Nature* **447,** 222–6 (2007).

46. Ørom, U. A., Nielsen, F. C. & Lund, A. H. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Molecular cell* **30,** 460–71 (2008).

47. Gaynor, J. W., Campbell, B. J. & Cosstick, R. RNA interference: a chemist's perspective. *Chemical Society reviews* **39,** 4169–84 (2010).

48. Song, J.-J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science (New York, N.Y.)* **305,** 1434–7 (2004).

49. Kim, V. N., Han, J. & Siomi, M. C. Biogenesis of small RNAs in animals. *Nature reviews. Molecular cell biology* **10,** 126–39 (2009).

50. Liu, J. *et al.* Argonaute2 is the catalytic engine of mammalian RNAi. *Science (New York, N.Y.)* **305,** 1437–41 (2004).

51. Tolia, N. H. & Joshua-Tor, L. Slicer and the argonautes. *Nature chemical biology* **3,** 36–43 (2007).

52. Wang, H. *et al.* Structural insights into RNA processing by the human RISC-loading complex. *Nature structural & molecular biology* **16,** 1148–53 (2009).

53. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115,** 209–16 (2003).

54. Schwarz, D. S. *et al.* Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115,** 199–208 (2003).

55. Matranga, C., Tomari, Y., Shin, C., Bartel, D. P. & Zamore, P. D. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **123,** 607–20 (2005).

56. Preall, J. B. & Sontheimer, E. J. RNAi: RISC gets loaded. *Cell* **123,** 543–5 (2005).

57. Elbashir, S. M., Martinez, J., Patkaniowska, a, Lendeckel, W. & Tuschl, T. Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *The EMBO journal* **20,** 6877–88 (2001).

58. Parker, J. S. & Barford, D. Argonaute: A scaffold for the function of short regulatory RNAs. *Trends in biochemical sciences* **31,** 622–30 (2006).

59. Fabian, M. R. *et al.* Mammalian miRNA RISC recruits CAF1 and PABP to affect PABP-dependent deadenylation. *Molecular cell* **35,** 868–80 (2009).

60. *Optical Fluorescence Microscopy*. (Springer Berlin Heidelberg, 2011). doi:DOI 10.1007/978-3-642-15175-0

61. Atkins, P. & Paula, J. de. *Atkin's physical chemistry*. 1064 (Oxford University Press, 2006).

62. Murphy, D. B. & Davidson, M. W. *Fundamentals of Light Microscopy and Electronic Imaging*. (John Wiley & Sons, Inc., 2012). doi:10.1002/9781118382905

63. Hell, S. W. Far-field optical nanoscopy. *Science (New York, N.Y.)* **316,** 1153–8 (2007).

64. Heilemann, M. Fluorescence microscopy beyond the diffraction limit. *Journal of biotechnology* **149,** 243–51 (2010).

65. Galbraith, C. G. & Galbraith, J. a. Super-resolution microscopy at a glance. *Journal of cell science* **124,** 1607–11 (2011).

66. Bertani, G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. *Journal of bacteriology* **62,** 293–300 (1951).

67. DNA purification. *Promega Protocols & Applications Guide* (2012). at <https://www.promega.com/resources/product-guides-and-selectors/protocols-and-applications-guide/dna-purification/>

68. Birnboim, H. C. & Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic acids research* **7,** 1513–23 (1979).

69. Boom, R. *et al.* Rapid and simple method for purification of nucleic acids. *Journal of clinical microbiology* **28,** 495–503 (1990).

70. Wilcockson, J. The differential precipition of nucleic acids and proteins from aqueous solutions by ethanol. *Analytical biochemistry* **66,** 64–8 (1975).

71. Erfle, H. *et al.* Reverse transfection on cell arrays for high content screening microscopy. *Nature protocols* **2,** 392–9 (2007).

72. Erfle, H. *et al.* Work flow for multiplexing siRNA assays by solid-phase reverse transfection in multiwell plates. *Journal of biomolecular screening* **13,** 575–80 (2008).

73. Matula, P., Lisauskas, T., Starkuviene, V. & Rohr, K. Quantification of Golgi Complex Assembly and Disassembly in Live Cell Fluorescence Microscopy Images. *Proceedings of Microscopic Image Analysis with Applications in Biology (MIAAB) Workshop, Heidelberg* 1–5 (2011).

74. Andersen, E. S. Prediction and design of DNA and RNA structures. *New biotechnology* **27,** 184–93 (2010).

75. Wuchty, S., Fontana, W., Hofacker, I. L. & Schuster, P. Complete suboptimal folding of RNA and the stability of secondary structures. *Biopolymers* **49,** 145–65 (1999).

76. Rivas, E. & Eddy, S. R. A dynamic programming algorithm for RNA structure prediction including pseudoknots. *Journal of molecular biology* **285,** 2053–68 (1999).

77. Bernhart, S. H. & Hofacker, I. L. From consensus structure prediction to RNA gene finding. *Briefings in functional genomics & proteomics* **8,** 461–71 (2009).

78. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* **31,** 3406–15 (2003).

79. Zuker, M. On finding all suboptimal foldings of an RNA molecule. *Science (New York, N.Y.)* **244,** 48–52 (1989).

80. Dictionary of British and World English. *Oxford Dictionaries* at <http://oxforddictionaries.com>

81. Palenskis, V. & Maknys, K. *Random processes; educational book*. 176 (Vilnius University, Department of Radio Physics, 2008).

82. Weisstein, E. W. Gamma distribution. *MathWorld - A Wolfram Web Resource* at <http://mathworld.wolfram.com/GammaDistribution.html>

83. Perinetti, G. *et al.* Correlation of 4Pi and electron microscopy to study transport through single Golgi stacks in living cells with super resolution. *Traffic (Copenhagen, Denmark)* **10,** 379–91 (2009).

84. Bouchet-Marquis, C., Starkuviene, V. & Grabenbauer, M. Golgi apparatus studied in vitreous sections. *Journal of microscopy* **230,** 308–16 (2008).

85. Koga, D. & Ushiki, T. Three-dimensional ultrastructure of the Golgi apparatus in different cells: high-resolution scanning electron microscopy of osmium-macerated tissues. *Archives of histology and cytology* **69,** 357–74 (2006).

86. Grabenbauer, M. *et al.* Correlative microscopy and electron tomography of GFP through photooxidation. *Nature methods* **2,** 857–62 (2005).

87. Marsh, B. J., Volkmann, N., McIntosh, J. R. & Howell, K. E. Direct continuities between cisternae at different levels of the Golgi complex in glucose-stimulated mouse islet beta cells. *Proceedings of the National Academy of Sciences of the United States of America* **101,** 5565–70 (2004).

88. Egner, A., Verrier, S., Goroshkov, A., Söling, H.-D. & Hell, S. W. 4Pi-microscopy of the Golgi apparatus in live mammalian cells. *Journal of structural biology* **147,** 70–6 (2004).

89. Marsh, B. J., Mastronarde, D. N., Buttle, K. F., Howell, K. E. & McIntosh, J. R. Organellar relationships in the Golgi region of the pancreatic beta cell line, HIT-T15, visualized by high resolution electron tomography. *Proceedings of the National Academy of Sciences of the United States of America* **98,** 2399–406 (2001).

90. Ladinsky, M. S., Mastronarde, D. N., McIntosh, J. R., Howell, K. E. & Staehelin, L. A. Golgi structure in three dimensions: functional insights from the normal rat kidney cell. *The Journal of cell biology* **144,** 1135–49 (1999).

91. Suzaki, E. & Kataoka, K. Three-dimensional visualization of the Golgi apparatus: observation of Brunner's gland cells by a confocal laser scanning microscope. *Journal of structural biology* **128,** 131–8 (1999).

92. Rambourg, A. & Clermont, Y. Three-dimensional electron microscopy: structure of the Golgi apparatus. *European journal of cell biology* **51,** 189–200 (1990).

93. Beams, H. W. & Kessel, R. G. The Golgi apparatus: structure and function. *International review of cytology* **23,** 209–76 (1968).

94. Mollenhauer, H. H. an Observation on the Functioning of the Golgi Apparatus. *The Journal of Cell Biology* **17,** 222–225 (1963).

95. De Matteis, M. A., Luini, A. & Matteis, M. A. De. Exiting the Golgi complex. *Nature reviews. Molecular cell biology* **9,** 273–84 (2008).

96. Storrie, B. Maintenance of Golgi apparatus structure in the face of continuous protein recycling to the endoplasmic reticulum: making ends meet. *International review of cytology* **244,** 69–94 (2005).

97. Lee, M. C. S., Miller, E. A., Goldberg, J., Orci, L. & Schekman, R. Bi-directional protein transport between the ER and Golgi. *Annual review of cell and developmental biology* **20,** 87–123 (2004).

98. Kondylis, V., Tang, Y., Fuchs, F., Boutros, M. & Rabouille, C. Identification of ER proteins involved in the functional organisation of the early secretory pathway in Drosophila cells by a targeted RNAi screen. *PloS one* **6,** e17173 (2011).

99. Wendler, F. *et al.* A genome-wide RNA interference screen identifies two novel components of the metazoan secretory pathway. *The EMBO journal* **29,** 304–14 (2010).

100. Boulaflous, A., Faso, C. & Brandizzi, F. Deciphering the Golgi apparatus: from imaging to genes. *Traffic (Copenhagen, Denmark)* **9,** 1613–7 (2008).

101. Bard, F. *et al.* Functional genomics reveals genes involved in protein secretion and Golgi organization. *Nature* **439,** 604–7 (2006).

102. Gilchrist, A. *et al.* Quantitative proteomics analysis of the secretory pathway. *Cell* **127,** 1265–81 (2006).

103. Fujiwara, T., Oda, K., Yokota, S., Takatsuki, A. & Ikehara, Y. Brefeldin A causes disassembly of the Golgi complex and accumulation of secretory proteins in the endoplasmic reticulum. *The Journal of biological chemistry* **263,** 18545–52 (1988).

104. Lippincott-Schwartz, J., Yuan, L. C., Bonifacino, J. S. & Klausner, R. D. Rapid redistribution of Golgi proteins into the ER in cells treated with brefeldin A: evidence for membrane cycling from Golgi to ER. *Cell* **56,** 801–13 (1989).

105. Sciaky, N. *et al.* Golgi tubule traffic and the effects of brefeldin A visualized in living cells. *The Journal of cell biology* **139,** 1137–55 (1997).

106. Pan, H. *et al.* A novel small molecule regulator of guanine nucleotide exchange activity of the ADP-ribosylation factor and golgi membrane trafficking. *The Journal of biological chemistry* **283,** 31087–96 (2008).

107. Sciaky, N. *et al.* Golgi tubule traffic and the effects of brefeldin A visualized in living cells. *The Journal of cell biology* **139,** 1137–55 (1997).

108. Nelson, K. K., Holmer, M. & Lemmon, S. K. SCD5, a suppressor of clathrin deficiency, encodes a novel protein with a late secretory function in yeast. *Molecular biology of the cell* **7,** 245–60 (1996).

109. Ward, T. H., Polishchuk, R. S., Caplan, S., Hirschberg, K. & Lippincott-Schwartz, J. Maintenance of Golgi structure and function depends on the integrity of ER export. *The Journal of cell biology* **155,** 557–70 (2001).

110. Lisauskas, T. *et al.* Live-cell assays to identify regulators of ER-to-Golgi trafficking. *Traffic* **13,** 416–32 (2012).

111. Järve, A. *et al.* Surveillance of siRNA integrity by FRET imaging. *Nucleic acids research* **35,** e124 (2007).

112. Tyagi, S. Imaging intracellular RNA distribution and dynamics in living cells. *Public Health* **6,** (2009).

113. Weil, T. T., Parton, R. M. & Davis, I. Making the message clear: visualizing mRNA localization. *Trends in cell biology* **20,** 380–390 (2010).

114. Lymperopoulos, K., Kiel, A., Seefeld, A., Stöhr, K. & Herten, D.-P. Fluorescent probes and delivery methods for single-molecule experiments. *Chemphyschem : a European journal of chemical physics and physical chemistry* **11,** 43–53 (2010).

115. Daigle, N. & Ellenberg, J. LambdaN-GFP: an RNA reporter system for live-cell imaging. *Nature methods* **4,** 633–6 (2007).

116. miRBase RELEASE 19. *miRBase: the microRNA database* (2012). at <http://www.mirbase.org/>

117. Flottmann, B. *et al.* Correlative light microscopy for high-content screening. (2013). **Accepted,** BioTechniques.

118. Erfle, H., Lisauskas, T., Claas, C., Reymann, J. & Starkuviene, V. Cell arrays for the measurement of organelle dynamics in living cells. *Methods in molecular biology* **706,** 73–81 (2011).

119. Tsai, Y.-T. *et al.* Comparative knock-down and knock-in screens of Rab GTPases identifies Rab40c as a novel regulator of biosynthetic trafficking. (2013). **In preparation**

120. Romero, N. *et al.* Rab1b overexpression modifies Golgi size and gene expression in HeLa cells and modulates the thyrotrophin response in thyroid cells in culture. *Molecular biology of the cell* **24,** 617–32 (2013).

121. Haas, A. K. *et al.* Analysis of GTPase-activating proteins: Rab1 and Rab43 are key Rabs required to maintain a functional Golgi complex in human cells. *Journal of cell science* **120,** 2997–3010 (2007).

122. Cogli, L., Progida, C., Bramato, R. & Bucci, C. Vimentin phosphorylation and assembly are regulated by the small GTPase Rab7a. *Biochimica et biophysica acta* (2013). doi:10.1016/j.bbamcr.2013.02.024

123. Nabavi, N., Pustylnik, S. & Harrison, R. E. Rab GTPase mediated procollagen trafficking in ascorbic acid stimulated osteoblasts. *PloS one* **7,** e46265 (2012).

124. Serva, A. *et al.* miR-17-5p Regulates Endocytic Trafficking through Targeting TBC1D2/Armus. *PloS one* **7,** e52555 (2012).

125. De Pietri Tonelli, D. *et al.* Single-cell detection of microRNAs in developing vertebrate embryos after acute administration of a dual-fluorescence reporter/sensor plasmid. *BioTechniques* **41,** 727–32 (2006).

126. Ameres, S. L., Martinez, J. & Schroeder, R. Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* **130,** 101–12 (2007).

127. Gredell, J. a, Berger, A. K. & Walton, S. P. Impact of target mRNA structure on siRNA silencing efficiency: A large-scale study. *Biotechnology and bioengineering* **100,** 744–55 (2008).

Hiermit erkläre ich an Eides Statt, dass ich die vorliegende Arbeit selbstständig und ohne unerlaubte Hilfsmittel durchgeführt habe.

Heidelberg, den 11.11.2013 .....

Eidesstattliche Versicherung gemäß § 8 der Promotionsordnung der Naturwissenschaftlich-Mathematischen Gesamtfakultät der Universität Heidelberg

1. Bei der eingereichten Dissertation zu dem Thema

Application of high-content microscopy to investigate membrane trafficking.

handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich ~~wie folgt~~/bislang nicht[1)] an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

Titel der Arbeit:_____

Hochschule und Jahr:_____

Art der Prüfungs- oder Qualifikationsleistung:_____

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Heidelberg, 11.11.2013

Ort und Datum

Unterschrift

[1)] Nicht Zutreffendes streichen. Bei Bejahung sind anzugeben: der Titel der andernorts vorgelegten Arbeit, die Hochschule, das Jahr der Vorlage und die Art der Prüfungs- oder Qualifikationsleistung.