Single molecule tracking of Myosin VI in the nucleus of living cells and increasing the toolbox for future studies

Dissertation zur Erlangung des Doktorgrades Dr.rer.nat. der Fakultät für Naturwissenschaften der Universität Ulm

Vorgelegt von:
Andreas Große-Berkenbusch
aus Backnang

2020
Amtierender Dekan der Fakultät:
Prof. Dr. Thorsten Bernhardt

Gutachter:
Prof. Dr. Christof Gebhardt
PD Dr. Thomas Gronemeyer

Tag der Promotion: 26.03.2021

Version April 18, 2021

© 2020 Andreas Große-Berkenbusch
Satz: PDF-LATEX 2ε
Copyright notification

Partial results of this thesis were previously published in the following journal:

**Myosin VI moves on nuclear actin filaments and supports long-range chromatin rearrangements**

*Andreas Große-Berkenbusch, Johannes Hettich, Timo Kuhn, Natalia Fili, Alexander W. Cook, Yukti Hari-Gupta, Anja Palmer, Lisa Streit, Peter J.I. Ellis, Christopher P. Toseland, J. Christof M. Gebhardt*

bioRxiv 2020.04.03.023614
Contents

5 Methods 31
5.1 Biochemical methods .................................................. 31
  5.1.1 Cloning ......................................................... 31
  5.1.2 Western Blot ..................................................... 33
  5.1.3 EMSA ............................................................. 35
5.2 Cell culture .............................................................. 36
  5.2.1 Generation of stable cell lines .................................. 36
  5.2.2 Fluorescent Activated Cell Sorting - FACS .................... 37
5.3 Halo-Tag labeling ...................................................... 39
5.4 Spinning disc confocal microscopy ................................... 40
5.5 Single molecule fluorescence microscopy ......................... 40
  5.5.1 Interlaced Time-lapse Microscopy (ITM) ....................... 42
  5.5.2 Time-lapse microscopy .......................................... 43
  5.5.3 GRID ............................................................ 44
  5.5.4 Measurement of directed motion .............................. 45
  5.5.5 SPLIT .......................................................... 45
5.6 CRISPR/Cas ............................................................ 47

6 Fluorescent labeling of proteins in living cells 53
6.1 Selective chemical labeling of purified proteins ................. 58
6.2 Protein and organic dye delivery across cell membranes ....... 60
  6.2.1 Interim results .................................................. 63
  6.2.2 MscL membrane channels ...................................... 64
  6.2.3 His-Tag labeling ................................................ 70
  6.2.4 TMP-Tag labeling .............................................. 79
  6.2.5 Discussion and outlook ........................................ 85

7 Results 91
7.1 Nuclear myosin VI ..................................................... 91
  7.1.1 ITM .............................................................. 93
  7.1.2 Time lapse microscopy .......................................... 93
  7.1.3 Motion analysis ................................................. 95
7.2 Investigation of MVI motion ......................................... 102
7.3 Nuclear actin .......................................................... 107
7.4 MVI moves on nuclear actin filaments .............................. 111
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5 MVI supports long range chromatin rearrangements</td>
<td>117</td>
</tr>
<tr>
<td>7.6 Summary</td>
<td>121</td>
</tr>
<tr>
<td>8 Discussion</td>
<td>125</td>
</tr>
<tr>
<td>A Appendix</td>
<td>135</td>
</tr>
<tr>
<td>References</td>
<td>137</td>
</tr>
<tr>
<td>Acknowledgment</td>
<td>163</td>
</tr>
</tbody>
</table>
1 Abstract

In the past decades, it has been shown that the 3D organization of the chromatin is essential for many biological processes. Chromosome conformation capture methods were able to reveal previously unknown layers of chromatin organizations. These were also shown to dynamically change during the cell cycle or upon different cell stimuli, and they play essential roles in various regulatory mechanisms. Advances in this field led to the identification of different molecular motors, which play specific roles in transcription, gene repositioning, whole chromosome translocation and DNA damage repair. However, most underlying molecular mechanism are still unclear. Single molecule microscopy offers the possibility to study biological systems from a new perspective. I used live cell single particle tracking to study the role of nuclear myosin VI (MVI) in the context of RNA polymerase II-dependent transcription. I found that, besides static chromatin interactions, MVI performs ATPase dependent directed motion. This newly identified process requires the full length motor and might be mediated by cooperating monomeric motors. Additional, actin perturbation and two-color experiments showed that this motion occurs on nuclear actin filaments. In cooperation with the Toseland group chromosome, paint experiments complemented my single molecule studies and showed that MVI is required for transcription-dependent long-range chromatin rearrangements. Overall I propose that the found motion process is directly involved in this chromatin reorganization mechanism. Further studies on this biological system and other nuclear processes will require orthogonal fluorescent labeling systems, and the advances in genome editing present an elegant way to reduce perturbations of the sensitive biological system. I used proteins that are involved in the formation of the transcription preinitiation complex as a model system to establish additional single molecule fluorescent labeling methods, and I added the genome editing system CRISPR/Cas, for direct endogenous labeling of proteins, to the available toolbox.
2 Overview

The nucleus of a cell is similar to a densely packed anthill. It is a crowded environment, consisting mainly of the DNA divided into the different chromosomes and associated proteins forming the chromatin. The nucleus is the site for DNA transcription. This complex mechanism involves numerous individual proteins and is tightly regulated by transcription factors and other regulatory proteins. All those proteins may be described as ants, each one executing its given task. On top of the locally defined regulation by transcription factors, there is an additional level of regulation, the 3D organization of the chromatin.

This work's primary focus is based on single molecule fluorescence microscopy, tracking of proteins in the nucleus of living cells, and the required biochemical methods surrounding this thematic.

Light sheet microscopy enables the tracking of single fluorescent-labeled proteins in living cells. Like slicing through an anthill full of ants, only a single plane of the whole system can be studied per time without creating background noise coming from beyond and above the selected plane. Considering the sectioning of an anthill, that will result in a single plane that is still heavily crowded by ants. Going one step further by selectively labeling a few individuals and following only those labeled ants makes the task a lot easier. Experimentally, this first step can be done by live cell compatible single protein labeling systems, like the well-established Halo-Tag system. However, to study complex multi-protein system like the transcription machinery, there is a need for different labeling strategies to differentiate them, like the ants—for example, male and female ants at the same time.

One layer of transcription regulation is the three-dimensional nuclear organization of the chromatin. On a basic level, the nucleus can be divided into chromosome territories and is locally structured into small substructures. The introduction of
chromatin capture methods revealed that the chromatin is organized into smaller substructures required for the regulation of gene transcription. This organizational layer is mainly regulated by the CCCTC-binding factor (CTCF) and the motor protein cohesin. However, it has been found that cellular triggers like hormone stimulation or cell stress can trigger long-range chromatin reorganizations that might end up in the formation of nuclear hubs, where transcription or DNA repair can be executed with higher efficiencies. The biological mechanisms of this process are poorly understood and was the major goal of this thesis. The study of nuclear myosin VI (MVI) revealed its role as an active player in this regulation process by being involved in actin dependent chromosome reorganization processes.

Gene transcription is regulated on multiple levels. One layer is the direct regulation of the general transcription factors by recruiting the transcription machinery. This formation of the transcription preinitiation complex (PIC) is a tightly regulated and complex process that is not yet fully understood. The previously introduced goal of establishing an orthogonal protein labeling system did go hand in hand with the kinetic investigations of the general transcription factors the TATA binding protein (TBP) and the transcription factor IIB (TFIIB), which are both involved in the PIC formation and are two key proteins that were used on a single molecule level, to shed light on the timings of the PIC formation.

Another goal of this thesis was the establishment and characterization of orthogonal fluorescent labeling systems. Up until now, like the study of additionally applied and labeled ants to the studied colony, labeled proteins were added to the system on top of the endogenous ones. Secondly, to take the protein labeling to the next level and reduce artifacts generated by protein overexpression, a CRISPR/Cas mediated gene-editing system was established. This method can in the future be used to maintain the sensitive cellular environment and to label the endogenous protein population directly.
3 Introduction

3.1 Molecular motors

Molecular motors are proteins that transform chemical energy into mechanical work. In biological systems, there is a diverse set of molecular motors. Cytoskeletal motors can be divided into three different families, namely myosins, kinesins, and dyneins. Other known proteins acting as molecular motors include polymerases and rotary motors like the bacteria flagella (Cooke and Francisco, 2001). Myosins and kinesins share one common ancestor, the RasGTPases, whereas dyneins are a member of ATPases (Satir, 2009). All cytoskeletal motors share distinct characteristic features. They use ATP as an energy source to enforce movement. The generated energy is directed by the use of cytoskeletal polymers. Myosin motors use actin networks, whereas kinesins and dyneins move on microtubules (Gross et al., 2007). Structural analysis revealed that all motor proteins consist of a highly conserved motor domain and a tail domain that includes a binding site for other proteins through which, for example, cargo elements can bind (Satir, 2009).

Depending on the relevant task, motor proteins convert energy in three different ways. First, if the motor is anchored, the underlying polymer moves, whereas if the polymer is anchored, the motor moves. If both systems are anchored, the energy can be used to stretch elastic elements (Satir, 2009). Two systems form the underlying cytoskeletal polymer network, necessary for motor actions.

In this work myosin VI, a motor protein from the myosin family that uses actin polymers as a structural network was studied. For completeness, also kinesins and dyneins that move along microtubules are briefly introduced. Figure 3.1 shows the different molecular motors in their correct cytoskeletal context.
3 Introduction

Figure 3.1: (a) Myosin motors move along actin filaments, whereas (b) dynein and kinesin motors move along microtubules.

Microtubules are the underlying framework for kinesins and dyneins. Microtubules form a very dynamic hollow structured polymer with two distinct ends. This polarity arises from its substructure that is formed by a head to tail assembly of tubulin dimers. These dimers form protofilaments that assemble around a hollow core (Cooper, 2000). The directionality is mirrored in the behaviour of the two molecular motor proteins, kinesin and dynein, as they move on microtubules. Most kinesins move toward the plus end of microtubules, whereas dyneins walk in the opposite direction, towards the minus end. The bidirectional behaviour of both motors gives rise to a multitude of biological functions. Through ATP hydrolysis, kinesins and dyneins move vesicles, chromosomes, and other cargo through the cell (Cooke and Francisco, 2001).

3.1.1 Structural and mechanical insights into myosins

Myosin motors either move along actin filaments or serve as an anchor by fixing structures or proteins to actin. Myosin motors belong to a large superfamily of motors that can be further divided into 15 classes (Foth et al., 2006; Sellers, 2000). The first reported myosin was discovered in 1864 in muscle extracts (Kühne, 1864). The muscle myosin II (MII), also called conventional myosin, forms filamentous structures. Two MII molecules can dimerize via their α-helical tail domain. The formed rods bind other dimers in a bipolar manner, forming myosin filaments that can bundle actin filaments (Alexander and Johnson, 1965). Following myosin I’s discovery, all later discovered myosins were numbered according to the date of their discovery (Soldati et al., 1999). In general, all myosins share the same three features: a motor domain, a neck domain, and a tail domain (Sellers, 2000).
3.1 Molecular motors

Motor domain

The highly conserved motor head is an 80 kDa domain that is almost exclusively found at the N-terminus and which is composed of several important features, including the actin-binding site (Goodson and Spudich, 1993). The motor head also contains the nucleotide-binding pocket, where ATPase hydrolysis takes place. ATP hydrolysis fuels the conformational change of myosins, which is translated into movement on or of actin, or actin-associated proteins (Houdusse and Sweeney, 2016). Despite the motor domain’s conserved nature, myosins show significant differences in their ATPase cycle rate and their actin-binding time. The latter is also known as the duty ratio (Dillon and Murphy, 1982). The variation of the duty ratio gives rise to a diversity of cellular functions. Starting with the simple idea that the duty ratio \( r \) is defined by the fraction of time that the motor spends in its actin bound state (see Formula 3.1) (O’Connell et al., 2007). In combination with the knowledge that ATP and ADP+Pi bound myosins are weak actin binders, and the ADP bound conformation is tightly associated with actin, one already ends up with several possible regulatory modes. The whole ATPase cycle of myosins is shown and described in figure 3.2.

\[
    r = \frac{t_{on}}{t_{on} + t_{off}} = \frac{t_{on}}{t_{total}} \tag{3.1}
\]

The smooth muscle myosin MII shows high ATP and ADP+Pi affinities, making phosphate release the rate-limiting step in its duty cycle. As described above, both states have low actin affinities, which decreases the on rate and thereby the duty ratio, which makes MII a low duty ratio myosin. Furthermore, the duty ratio of MII was experimentally found to be 0.04 (Lymn and Taylor, 1971). For the muscle movement of MII, this biochemical behaviour is favourable since many individual dimers move corporately along a single actin filament (Lymn and Taylor, 1971).

In contrast, in the case of a single motor dimer executing long-distance transports, a high duty ratio is a prerequisite. The duty ratio of purified zipped MVI dimers was measured to be 0.8. Similar to MV, MVI shows a high ADP binding affinity, making
3 Introduction

Figure 3.2: The ATPase cycle of myosins. (a) ATP bound myosin with low actin affinity. (b) Myosin in ADP+Pi-bound state in pre-power stroke position binds actin filaments. (c) Upon phosphate release, myosin’s affinity to actin increases, and the motor facilitates a working stroke. (d) Exchange of ADP to ATP decreases its actin affinity. (d-b) During its unbound state the recovery stroke takes place and the myosin folds back into its original conformation.

ADP release the rate-limiting step (De la Cruz et al., 2001). This ensures that one head of a dimer always stays bound, preventing complete dissociation of the motor and its cargo from actin filaments. Contrary to MII, where a higher motor density leads to higher total velocity, MVI motors show higher velocities at lower densities, strengthening the idea of tightly and long bound motor heads blocking binding sites and thereby slowing down trailing motors (Lister et al., 2004; A. D. Mehta et al., 1999).

Another term used to describe myosin motion is processivity. The processivity goes hand in hand with the duty ratio. A processive motor moves over long distances without detaching from an actin filament to facilitate its directed transport function. This requires a long and coordinated actin bound state of one motor head per time. Non-surprisingly, MV and MVI are known as processive motors (Sakamoto et al., 2000, Rock et al., 2001), whereas MII with a low duty ratio is a non-processive motor (Uyeda et al., 1990).
3.1 Molecular motors

Neck domain

The neck domain is directly adjacent to the motor domain. It consists of a variable number of isoleucine-glutamine (IQ) repeats with the sequence motive IQXXXR-GXXXR (Cheney and Mooseker, 1992). The neck domain folds into a proline devoid \( \alpha \)-helical structure (Bähler and Rhoads, 2002), that serves as a lever arm to amplify ATPase dependent changes occurring in the motor domain. The repeat number of IQ motives determines the length of the lever arm and directly influences the step size of the motor (Fili and Toseland, 2020; Purcell et al., 2002). Despite its function as a mechanical amplifier, the neck domain serves as a binding site for various proteins summarized under the term myosin light chains, whereas the core domains motor, neck, and tail of the motor are termed heavy chains (Heissler and Sellers, 2014). The first identified binding partners were proteins that bound conventional myosins. The identified essential light chains are crucial for the structural integrity of MII, whereas additional binders modulate its function (Sellers et al., 1980). Later, the unconventional binding partner calmodulin (CaM) was identified (Coluccio et al., 1990). CaM is a \( \text{Ca}^{2+} \) binding protein that is highly conserved and acts as a regulatory second messenger (Chin and Means, 2000). The study of various myosins has shown that binding of \( \text{Ca}^{2+} \) to CaM can either strengthen or weaken its actin binding affinity (Heissler and Sellers, 2014).

Tail domain

The tail domain is the most versatile region among the myosin superfamily and is highly myosin specific. The range of elements goes from short or long coiled coils and helical regions, used as lever arm extensions or dimerization domains, and phospholipid-binding domains to various protein-interacting domains. In addition, alternative splicing is used to increase the diverse functionality of a single myosin motor. In summary, the tail domain defines the oligomerization state, the intracellular localization, and the cargo binding specificity (Fili and Toseland, 2020).
Power stroke

As depicted in figure 3.2, the ATPase cycle of myosins includes four distinct states. Overall, the force by myosins is generated via ATP hydrolysis. As described earlier, ATP hydrolysis takes place in the absence of actin. However, in the unbound pre-powerstroke state, the ADP+Pi is trapped inside its binding pocket (Sweeney and Houdusse, 2010). Upon binding to actin, structural changes are initiated, and electrostatic interactions holding the phosphate are weakened. Through the opening of the phosphate-binding pocket, Pi can escape through the so-called “back door” mechanism. This release comes with additional structural changes, allowing the closure of the actin-binding cleft and initiating a large movement and the swing of the lever arm (powerstroke) (Muretta et al., 2015). In a second step, the rigor conformation, ADP binding contacts are lost, and ADP can be released (Coureux et al., 2003). Rebinding of ATP comes with a reopening of the actin-binding cleft, which terminates the power stroke and releases the motor from the actin filament (Houdusse and Sweeney, 2016).

3.1.2 Myosins in the nucleus

Phylogenetic analysis revealed 15 different myosin classes. Among those, up until now, seven were found to be present inside the nucleus (Caridi et al., 2018; Sarshad and Percipalle, 2014). The most extensively studied cytoplasmic myosin MII has also been found in the nucleus. However, in contrast to the cytoplasmic pool that forms filamentous structures, no such filaments were found in the nucleus (Rodgers, 2005). MII was identified as a core transcription factor in the RNA-PolII transcription complex (Lindsay and McCaffrey, 2009). The most recent study describes an ICAM-1 specific gene regulation mechanism and proposes a direct interaction of MII with RNA-PolII. A proposed model suggests that MII might slide DNA along the transcription complex (Percipalle, 2013).

MV has been shown to be relevant for RNA-PolII transcription of rRNA in nucleoli and RNA processing, as it was also found in nuclear speckles (Lindsay and McCaffrey, 2009; Pranchevicius et al., 2008). As a processive motor, it might also directly be involved in nuclear virus particle trafficking (K. L. Roberts and Baines, 2011), and together with MI, MV is involved in the relocalization of heterochromatin breaks.
3.1 Molecular motors

to the nuclear periphery (Caridi et al., 2018).
The most extensively studied nuclear motor is MI. It was first described in 1997 (Nowak et al., 1997). Many studies show the involvement of MI, also in context of nuclear actin filaments in different processes such as RNA-PolII dependent transcription (Fomproix and Percipalle, 2004; Hofmann et al., 2006; Sarshad and Percipalle, 2014) and chromosome movement and repositioning after hormone stimulation (Chuang et al., 2006; Q. Hu et al., 2008; I. S. Mehta et al., 2010). Apart from the ability of the MI tail to directly bind DNA (De Lanerolle et al., 2005), the direct working mechanism of MI remains elusive.

MVI is the second most extensively studied nuclear myosin. As it is the subject of this study, it will be described in more detail below (see chapter 3.1.3). Furthermore, MX, MXVI, and MXVII were also identified in the nucleus. However, their roles are still not yet clear (Cameron et al., 2007; Salamon et al., 2003; Woolner et al., 2008). Until now, nuclear myosins were found to play roles in various processes. However, all studies so far lack detailed mechanisms of the direct role of these motors in their specific context.

3.1.3 Unconventional myosin VI

Myosin VI (MVI) is an unconventional 150 kDa motor protein, that was first identified in Drosophila melanogaster in 1992 (Kellerman and Miller, 1992), however later found to be expressed in a wide range of species from Caenorhabditis elegans to humans (Mooseker and Cheney, 1995). While, all other known myosins move towards the plus end of actin filaments (barbed end), MVI was found to head towards the minus end (pointed end) (Wells et al., 1999). This unique directionality opens a wide spectrum of biological functions. MVI is involved in cell migration (R. Roberts et al., 2004), endocytosis (Arden et al., 2007; Aschenbrenner et al., 2003; Buss et al., 2001; R. Roberts et al., 2004), exocytosis (Tomatis et al., 2013), membrane anchoring of stereocilia (Cramer, 2000; Self et al., 1999; G. Spudich et al., 2007), cytokinesis (Arden et al., 2007; Hecht et al., 2019) and transcription (Cook et al., 2018; Fili et al., 2020; Fili et al., 2017; Loikkanen et al., 2009; Vreugde et al., 2006). The structure of MVI consists of the highly conserved N-terminal motor domain, which is responsible for its high duty ratio (De la Cruz et al., 2001). The neck do-
3 Introduction

Figure 3.3: Schematic representation of myosin VI. The bar graph shows the different domains, in respect to their size. The structural representation contains the specific structural annotations of the different domains.

main, that follows the motor domain consists of two subunits, one being a unique insert that is responsible for the reverse directionality, and the other being an IQ motive, a binding site for CaM (Wells et al., 1999). The large tail domain includes a three-helix bundle, a single \( \alpha \)-helix (SAH), and a C-terminal cargo binding domain (Knight et al., 2005; Mukherjea et al., 2009) (Figure 3.3). Each domain plays a distinct role in the functionality of the motor. Apart from the conserved and general functionalities of the individual domains described above, experiments gave specific insights into its mechanisms.

MVI moves on actin filaments towards the minus end, which is a unique feature among myosins (Wells et al., 1999). However, the neck domain can explain this directionality. With a single IQ-motive and an additional insert, the neck domain binds the Ca\(^{2+} \) responsive second messenger CaM. Structural analysis revealed that this binding involves four hydrophobic interactions that reposition the lever arm, during the power stroke, in the opposite direction (Bahloul et al., 2004). Additionally, CaM binding stiffens the lever arm (Heissler and Sellers, 2014), whereas an increase in Ca\(^{2+} \) leads to the unbinding of CaM, thereby decreasing the stiffness (J. Li et al., 2016). In order to facilitate MVI motion, high stiffness is thought to be needed, which requires low Ca\(^{2+} \) levels, making apo-CaM the responsible subunit for stiffening the lever arm (Bahloul et al., 2004). However, prior to transport at low Ca\(^{2+} \) levels, a change in Ca\(^{2+} \) from low to high levels might facilitate cargo binding (Batters et al., 2016).
3.1 Molecular motors

MVI an irregular motor that questions the lever arm hypothesis
First MVI analysis of the tail domain revealed a coiled-coil domain, similar to MII and MV, suggesting an analogous dimeric stepping mechanism. Initial *in vitro* studies on MVI, assessing this question, were performed using an artificial dimer, in which the whole C-terminal CBD was substituted by a leucine zipper motif. These experiments revealed a "hand over hand" stepping mechanism with a step size of 36 nm, which did not fit the classical swinging lever arm theory. Additionally, MVI was isolated from cell extracts as a monomer (Lister et al., 2004), and the predicted coiled-coil domain formed by the tail domain was identified to be separated into a proximal tail forming a three α-helix motive (Mukherjea et al., 2009) followed by the medial tail presenting a stable single α-helix (SAH) (Knight et al., 2005; Lister et al., 2004; Spink et al., 2008). These findings questioned the initially predicted stepping mechanism as well as the required dimer formation of MVI.

Follow-up experiments using a bead-bound MVI monomer showed that diffusion-limited monomers can walk processively on actin filaments with a step size of 40 nm. However, the motion was limited to a very short distance below 100 nm (Iwaki et al., 2006). Another study found that dimerization occurs locally in regions with high MVI concentrations, resulting in processive 1-2 µm hand-over-hand motion, with a step size of 30 nm. However, the molecular mechanism facilitating the monomer to dimer formation remained elusive (H. Park et al., 2006). Additionally, a theoretical model based on multiple dimers moving nanospheres processively on an actin network suggests that MVI monomers can exert similar motions if they are present at high local concentrations (Sivaramakrishnan and Spudich, 2009).

The C-terminal tail of MVI with its alternative splicing site might play an essential role in dimer formation. There are two possible splice sites, one N-terminal to the cargo binding site and one within the CBD, resulting in a large insert (LI) and a small insert (SI) isoform, respectively (Buss et al., 2001). The resulting isoforms, including the non-insert (NI), are subject to differential intracellular localization. The LI localizes to the cytosol, with targets at the cell periphery, being deployed in clathrin-mediated endocytosis (Buss et al., 2001). The NI localizes to cytosol as well as to the nucleus but is there occluded from the nucleoli, playing a role in transcription (Fili et al., 2017). The SI was found to localize to the cytosol, where it regulates the
Recent experiments showed that MVI binding partners regulate the localization and function of the motor. It adopts, in its inactive monomeric state, a backfolded conformation with a stroke size of 18 nm (Lister et al., 2004), and in the same conformation, individual protein binding sites at the CBD are blocked (Wollscheid et al., 2016). Several adapter proteins of MVI binding the CBD were identified. Among those, GiPC and DAB2 play an important role, as they are found to dimerize on endocytotic vesicles (Lou et al., 2001; Xu et al., 1998). Interestingly, MVI was in this context also found to dimerize indirectly via DAB2 fragments (Phichith et al., 2009; C. Yu et al., 2009). NDP52 represents another CBD binding partner, that was found to form homodimers (Fili et al., 2017).

There are two opposing protein binding sites at the CBD, one containing a WWY and the other an RRL core motive. DAB2 was found to bind the WWY motive (G. Spudich et al., 2007), whereas NDP52 binds the RRL motive. In this context, the different isoforms with their specific cellular localization play an important role (Fili et al., 2017). Interestingly the LI, with its additional α-helix, blocks the RRL binding site, thereby impeding NDP52 binding (Morriswood et al., 2007). This makes the LI-isoform insensitive and the nuclear NI-isoform prone to the binding of NDP52. Overall, both MVI isoforms can thereby form direct or indirect dimers in the presence of DAB2, whereas NDP52 only facilitates the dimerization of the NI isoform.

FRET studies showed that the dimerization is preceded by an unfolding of the backfolded state (Fili et al., 2017). Staying with the classical hand over hand stepping mechanism, unfolding and either direct or indirect dimerization of MVI is preceded by protein binding of DAB2 or NDP52. Additionally, unfolding is also relevant for the anchor function of MVI (Altman et al., 2004). Taken together, these studies suggest a cofactor dependent direct or in-direct dimerization of MVI and a mechanistic separation by the expression of the different splice variants.
3.1 Molecular motors

Nuclear functions of myosin VI
After the first report of MVI in 1992, it took 18 years until MVI was identified and described in the nucleus. Preceding this finding, overexpression of MVI was observed in ovarian, breast, and prostate cancer cells, suggesting a role in cell proliferation (Yoshida et al., 2004). In 2006 MVI was identified in the nucleus, where it colocalizes with PolII and sites of active transcription (Vreugde et al., 2006). Interestingly upregulation of MVI expression has also been observed in the context of DNA damage (Jung et al., 2006). In contrast, the depletion of MVI by lentiviral mediated knockdown decreased cell motility and proliferation (Puri et al., 2010; H. Wang et al., 2015; H. Wang et al., 2016). Recently a more direct role has been identified, linking MVI to active gene transcription. qPCR experiments revealed that AR and ER induced gene transcription is regulated by active MVI, whereas the global gene transcription seemed to be unaffected by knockdown of MVI (Fili et al., 2017; Loikkanen et al., 2009). Interestingly the WWY motive of MVI has been found to directly bind chromatin (Fili et al., 2017). Linking these findings with the above described binding partners, it is not surprising that the overexpression of the tumor suppressor DAB2 has a negative effect on target gene transcription (He et al., 2001). Additionally, it has been found that inhibition by TIP, a specific MVI ATPase inhibitor (Heissler et al., 2012), exhibits a similar negative effect on transcription, suggesting an active role of MVI in target gene expression (Cook et al., 2018; Fili et al., 2020). Moreover, it was also shown that the ER related gene transcription is actively regulated by its nuclear binding partner NDP52, which facilitates its unfolding and dimerization (dos Santos et al., 2020; Fili et al., 2017). However, the direct molecular mechanism of MVI in gene expression is not clear yet, and further investigations were the subject of this study.
3 Introduction

3.2 Nuclear actin

Nuclear actin was initially described in 1963. Here actin was found in the nuclear extract of calf thymus cell (Ohnishi et al., 1963). A few years later, filamentous actin rods were described in the nucleus of *Triturus viridescens* oocytes (Lane, 1969). However, due to a sceptical research community and lacking visualization techniques, the role of nuclear actin and the presence of filamentous structures was, for many years, controversially discussed. Recently new visualization techniques based on nanobodies gave new insights into nuclear actin structures (Plessner et al., 2015). Different structures were observed to form upon various cell stimuli and stressors. For example, DMSO, heat and UV, and other stimuli like magnesium, serum, and the ionophore A23187 induce nuclear actin structures (Kelpsche and Tootle, 2018).

3.2.1 Biochemistry

Cytosolic actin dynamics are well characterized. There are three main actin isoforms, the skeletal $\alpha$-actin and the non-muscle and muscle $\beta$- and $\gamma$-actin, that mainly vary in their N-terminus (Herman, 1993). Actin exists in two different states, monomeric (G-actin) and filamentous (F-actin) actin. Actin polymerization starts with the formation of small aggregates. These aggregates can bind additional monomers at either side (barbed and pointed end), whereas the barbed end shows a higher affinity towards another monomer. These dynamics ultimately result in the so-called treadmilling of F-actin, with the barbed end being the growing end and the pointed end being the shrinking end (Petty and Cassidy, 2005). Actin monomers can bind and hydrolyze ATP. Although ATP binding is not required for filament formation, the bound nucleotide determines its affinity to bind to F-actin. Actin monomers have a higher binding affinity while in the ATP bound state. However, the ATPase activity of actin increases while being in its bound state. This gives rise to an additional level of F-actin dynamics (Petty and Cassidy, 2005). The described characteristics and a variety of actin-binding proteins define the dynamics of F-actin. Mainly, profilin together with other factors like the ARP2/3 complex are responsible for filament formation (Witke, 2004), whereas cofilin depolymerizes F-actin structures (Pendleton et al., 2003).
3.2 Nuclear actin

Figure 3.4: Regulation of nuclear actin concentration. Actin/cofilin (1:1) complexes are translocated to the nucleus via the nuclear core complexes (NPC), while bound to importin 9. The export of actin takes place in an actin/profilin (1:1) complex, while bound to exportin 6. In the nucleus actin can adopt several states ranging from monomers to actin structures, like small polymers and actin filaments.

Nuclear actin also comes as G- and F-actin. However, in contrast to the cytosolic fraction, it must be imported into the nucleus. With its globular shape and a size of 42 kDa (Kabsch et al., 1990), actin has to be actively transported into the nucleus via the nuclear core complexes (NPC). In the case of actin, this import is facilitated by Importin 9 (Dopie et al., 2012). Importin 9 is part of the importin-β family and is highly conserved among higher eukaryotes (Jakel, 2002). Since actin lacks a nuclear localization signal (NLS), it is recruited in complex with cofilin. The nuclear localization is facilitated by the NLS of cofilin, which lays opposite its actin-binding pocket (Dopie et al., 2012). Exportin 6, which is also part of the importin-β family, facilitates actin's nuclear export. In line with its import, also the export takes place while bound in complex to another protein. During the export actin is bound to profilin (Witke, 2004)(Figure 3.4). Both mechanisms highlight the tightly controlled nuclear actin concentration that might be important for its dynamic functions. Compared to the regulation of cytosolic actin structures, so far, limited knowledge could be gained on nuclear actin filament formation. A recent study found that nuclear actin structures are involved in DNA double-strand repair and described that
the actin nucleator complex ARP2/3 and the F-actin capping protein are associated with DNA repair sites. Similar to the cytosol, Wiskott–Aldrich syndrome protein (WASP) that activates the ARP2/3 has been found to colocalize with the repair sites (Schrank et al., 2018a). These findings suggest a similar F-actin regulation as found in the cytosol. However, further studies on nuclear actin structures have to reveal its exact regulatory and mechanistic functions.

### 3.2.2 Visualization

A long-standing question was whether different visualization techniques artificially induce and stabilize actin filaments inside the nucleus. The most widely used F-actin labeling method, using fluorescently labeled phalloidin, was in most experimental setups unable to label nuclear actin structures (Lengsfeld et al., 1974; Vandekerckhove et al., 1985). Only a few labs reported nuclear phalloidin labeled filaments (Baarlink et al., 2013; Nishida et al., 1987). These findings supported the concept that nuclear structures are distinct from cytoplasmic ones. This theory was also supported by studies that used different anti-actin-antibodies for visualization of nuclear actin structures. While those antibodies detected different pools and structures of nuclear actin, they were not able to label cytoplasmic F-actin structures (Gonsior et al., 1999; Schoenenberger et al., 2005; Wineland et al., 2018). While the question of G-actin’s presence in the nucleus was solved early on, only recent studies with two GFP tagged antibodies, namely LifeAct and nuclear actin chromobodies (nAc), were able to visualize nuclear F-actin structures in living cells. The latter is now an established and accepted labeling technique that is thought to have minimal influence on actin dynamics. Several studies were preformed that revealed short-lived and dynamic nuclear actin filaments upon serum stimulation (Plessner et al., 2015), during cell spreading (Plessner et al., 2015), and during cell division (Baarlink et al., 2017). However, one has to keep in mind, while this labeling technique can visualize dynamic and stable forms of actin filaments, it is not able to stain actin/cofilin polymers that were found in earlier studies by LifeAct staining (Plessner et al., 2015).

In the last years, many groups reported and visualized different nuclear actin structures ranging from monomers and small polymers to dynamic and stable actin filaments, which are also called actin rods. Additionally, much biochemical work inves-
tigating these structures’ potential functions already revealed various biochemical functions associated with nuclear actin. The manipulation of the nuclear actin import and export mechanism, as well as several actin mutations, mostly S14C, which stabilizes filaments, and R62D, which blocks filament formation, were used to assess its functionality (Posern et al., 2002).

### 3.2.3 Functions

Indirect FRAP analysis of the nuclear actin pool suggests that roughly half of the present actin pool is in its monomeric state while the other half is in different bound and polymeric/filamentous states (McDonald et al., 2006).

Most studies link nuclear actin structures to gene transcription. Direct anti-actin-antibody labeling showed increased binding of actin at transcriptionally active regions (Percipalle, 2002; Percipalle et al., 2001). Additionally, nuclear actin was found to be indirectly involved in transcription by affecting chromatin remodeling complexes, like the Swi/SNF (Schubert et al., 2013) and the INO80 complex (Kapoor et al., 2013). It also directly binds and regulates transcription factors. For example, a key player that is affected by nuclear actin is MRTF-A (Myocardin-related transcription factor A), a coactivator of the serum response factor (SRF). Here, actin was found to be required for the correct localization of MRTF-A during transcription activation (Vartiainen et al., 2007). It was also shown that actin interacts with the ER in the nucleus after Estrogen receptor (ER) activation (Ambrosino et al., 2010).

A more direct effect has been shown by copurification experiments, where actin was found to bind PolII directly (Egly et al., 1984; Hofmann et al., 2004a) and is responsible for the recruitment of the transcription elongation factor P-TEFb (Qi et al., 2011). An active mechanism facilitated by nuclear actin was proposed upon finding that PolII clusters formation after serum stimulation is actin-dependent (Wei et al., 2020a). While there are many studies linking actin and PolII, there are also studies on other RNA polymerases. Direct binding of actin to PolI (Fomproix and Percipalle, 2004) and PolIII (P. Hu et al., 2004) have been shown as well.

However, the role of nuclear actin is not limited to transcription or its regulation. Two recent publications showed that nuclear actin is directly involved in DNA double-strand repair (Schrank et al., 2018a) and the relocalization of the repair loci towards
3 Introduction

the nuclear laminar (Caridi et al., 2018). Similar to the involvement in this active transport mechanism, nuclear actin also plays a role in long-distance transport of chromatin (Chuang et al., 2006) and the repositioning of active gene loci (Dundr et al., 2007). Force generation of nuclear actin filaments is also relevant for the baculovirus infection (Ohkawa and Welch, 2018). However, whether these active mechanisms are directly facilitated by actin filaments or by the help of nuclear myosins remains a part of future studies.
3.3 Chromatin organization

In the last years, the advances in biochemical methods to study the nucleus and the chromatin led to a more in depth understanding of the chromosome architecture and its associated function and dynamics. The human haploid genome has a size of roughly 3000 Mbp and has a length of over one meter. However, it still fits inside a cellular nucleus with a volume of roughly 200 µm$^3$ (Misteli, 2008). The chromatin has to be packed and organized in a defined manner in order to exhibit its regular function. From the lowest level of compaction, the DNA is packed tightly around nucleosomes, which is described in the "beads on a string" model. These strings assemble into chromatin fibers and are further packed into loop structures (Figure 3.5(a)) (Chien and Van Der Heijden, 2014).

But how are those chromatin fibers spatially organized, and how does this concept allow for proper biological dynamics and control?

The 3D organization of the chromatin into chromosomal territories was already suggested in 1885 (Rabl, 1885). Over the years, as more and more biochemical methods such as FISH found their way into the labs, the proposed concept could be visually verified (Boveri, 1909; Cremer and Cremer, 2010). The model proposing distinct chromosome territories was further validated by the advance in chromosome conformation capture methods (3C). 3C methods provided another way to study the 3D conformation of the chromatin. 3C is a high throughput method that is able to estimate chromatin contact frequencies of different gene loci. However, one has to keep in mind that the resulting data is an average generated by a pool of cells (Dekker et al., 2002). A few years ago, HiC methods pioneered by Liebermann-Aiden could improve the resolution limit of chromatin conformation capture methods, which led to the proposal and the confirmation that the chromatin is arranged into different nuclear compartments that are named A- and B- compartments. Additionally, the so-called topological associated domains (TADs) were identified. TADs are defined as microcompartments, which play a direct role in transcription (Lieberman-Aiden et al., 2009). However, all C-methods are indirect ones, and it became apparent that results obtained by chromosome conformation capture methods need to be verified by an alternative and complementary approach like FISH (Williamson et al., 2014).

Apart from chromosomal territories (Figure 3.3(b)), early observations already found
3 Introduction

Figure 3.5: Chromatin organization inside the nucleus. (a) Double-stranded DNA wraps around histones forming the nucleosome. The nucleosomal organization of the DNA is described by the "beads on a string" model. Higher-order structures are formed by loop formation of the chromatin. (b) Inside the nucleus, chromosomes are localized to distinct territories. (c) Chromatin is separated into active A- and inactive B-compartments. (d) At shorter length scales, DNA is organized into TADs, which also include DNA loops formed by cohesin and CTCF (red), which bring promoter and enhancer regions together.

that the chromatin has more and less condensed regions, with active and inactive sites of transcription, respectively (Monneron and Bernhard, 1969; Verschure et al., 1999). These open and closed chromatin regions were found to have a strong overlap with the structurally, by HiC, identified active A- and inactive B- compartments (Figure 3.3(c)) (Lieberman-Aiden et al., 2009). Other nuclear membrane-free compartments like the nucleoli, the nuclear lamina, splicing speckles, cajal bodies, and the recently identified phase-separated particles, which include transcription hubs, present additional layers of structural organization (Kempfer and Pombo, 2020). The improved resolution of C-methods led to the identification of self-associated chromatin domains, called TADs (topological associated domains) (Dixon et al.,
3.3 Chromatin organization

TADs harbor co-regulated promotor and enhancer regions and consist of smaller looping substructures. They are essential for the positive and negative regulation of gene transcription (Figure 3.3(d)). The formation of DNA loops in these subcompartments is described by the so-called "loop extrusion model." The rod-shaped molecular motor cohesin actively pulls one DNA strand with respect to the other strand through its internal structure. This process is terminated by two convergently oriented CTCF (CCCTC-binding factor) molecules, which are bound to conserved CTCF binding sites, flanking the whole DNA loop (Fudenberg et al., 2016; Ganji et al., 2018). The formed promoter and enhancer loops are essential regulators for gene expression. Artificial changes of TADs by manipulating its boundaries can lead to transcriptional changes and disease phenotypes (Franke et al., 2016). Although TADs regulate the transcription of many genes, the general hierarchical chromatin structure is unaffected by the depletion of cohesin, and many genes are not solely regulated by the formation of TADs (Seitan et al., 2013).

The chromatin is organized into distinct and dynamic structures responsible for the regulation of gene expression. Ongoing research already revealed several concepts and mechanisms for the three-dimensional organization of the genome. However, apart from these, many identified processes like chromatin reorganization upon transcription stimulation or the dynamics involved in DNA damage repair are found to be regulated by other factors like nuclear actin and myosins.
4 Materials

4.1 Cell lines

Bacterial cells

XL1-Blue  
One Shot Stbl3

Stratagene, USA  
Life Technology, Darmstadt

Mammalian cells

Lenti X 293T  
HeLa CCL-2  
MCF7 HTB-22  
NIH 3T3

Clontech Laboratories Inc. USA  
ATCC - Cell Culture Collection, USA  
ATCC - Cell Culture Collection, USA  
Sigma Aldrich, Steinheim

4.2 Antibodies

• γ-tubulin, ab11316, abcam, UK
• Myosin VI, M0691, Sigma, Germany
• mouse alkaline phosphatase, A24527 Invitrogen, US
4 Materials

4.3 Kits

- QIAprep Spin Miniprep Kit
- QIAGEN Plasmid Plus Midi Kit
- QIAGEN RNeasy Plus Mini Kit
- Invitrogen SuperScript VILO cDNA Synthesis Kit
- Qiagen DNeasy Blood and Tissue Kit
- QIAquick PCR Purification Kit
- QIAquick Gel Extraction Kit
- NEB Monarch DNA Gel Extraction Kit

4.4 Media and Buffers

Buffers

<table>
<thead>
<tr>
<th>Solution</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ringer Solution</td>
<td>10 mM Hepes pH 7.4</td>
</tr>
<tr>
<td></td>
<td>145 mM NaCl</td>
</tr>
<tr>
<td></td>
<td>5 mM KCl</td>
</tr>
<tr>
<td></td>
<td>2 mM CaCl$_2$</td>
</tr>
<tr>
<td></td>
<td>1 mM MgCl$_2$</td>
</tr>
<tr>
<td></td>
<td>10 mM Glucose</td>
</tr>
<tr>
<td>Delivery Solution</td>
<td>10 mM Hepes pH 7.3</td>
</tr>
<tr>
<td></td>
<td>140 mM KOH</td>
</tr>
<tr>
<td></td>
<td>100 mM Aspartic Acid</td>
</tr>
<tr>
<td></td>
<td>12 mM NaOH</td>
</tr>
<tr>
<td></td>
<td>4 mM HCl</td>
</tr>
<tr>
<td></td>
<td>1 mM MgCl$_2$</td>
</tr>
<tr>
<td></td>
<td>2 mM ATP</td>
</tr>
</tbody>
</table>
### 4.4 Media and Buffers

<table>
<thead>
<tr>
<th>Buffer Type</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>WB Transfer Buffer</td>
<td>25 mM Tris Base pH 8.3&lt;br&gt;150 mM Glycine&lt;br&gt;100 mM Aspartat&lt;br&gt;10 % (v/v) Ethanol</td>
</tr>
<tr>
<td>WB TBS Buffer</td>
<td>10 mM Tris HCl pH 7.5&lt;br&gt;150 mM NaCl</td>
</tr>
<tr>
<td>WB TBS-T Buffer</td>
<td>20 mM Tris HCl pH 7.5&lt;br&gt;500 mM NaCl&lt;br&gt;5% (v/v) Tween 20&lt;br&gt;0.2% Triton X-100</td>
</tr>
<tr>
<td>WB Blocking Buffer</td>
<td>5 % (w/v) Milk Powder in TBS-T</td>
</tr>
<tr>
<td>TBP TFIIB storage buffer</td>
<td>10 mM Hepes pH 7.3&lt;br&gt;150 mM NaCl&lt;br&gt;2 mM DTT</td>
</tr>
<tr>
<td>EMSA Assembly buffer</td>
<td>50 mM Hepes pH 7.5&lt;br&gt;40 mM (NH₄)₂SO₄&lt;br&gt;10 μM ZnCl₂&lt;br&gt;5% (v/v) Glycerol&lt;br&gt;10 mM DTT (fresh)</td>
</tr>
</tbody>
</table>

### Cell culture media

- DPBS: Gibco Thermo Fischer, USA
- DMEM: Gibco Thermo Fischer, USA
- MEM: Gibco Thermo Fischer, USA
- OptiMEM: Gibco Thermo Fischer, USA
4.5 Primers

hTBP-XbaI-fw  
aatggtTCTAGAatggatcagaacaacagcgc

hTBP-AscI-rv  
ttaccaGGCGCGCttacgtgtctttgatcccttttag

hTFIIB-XbaI-fw  
aatggtTCTAGAatggcgtctaccagccg

hTFIIB-AscI-rv  
ttaccaGGCGCGCttatagctgtgtagttgttgccttg

Spt4-XbaI-fw  
aatggtTCTAGAATGGCCCTGGAGACG

Spt4-AscI-rv  
accattGGCGCGCCCTAGGTCTTTTATAGCTGTGCTCT

Spt5-XbaI-fw  
aatggtTCTAGAAGTGGCTACACAAGATCCC

Spt5-AscI-rv  
accattGGCGCGCCCTAGGCTTCCAGGAGCTTCC

RNGTT2-XbaI-fw  
aatggtTCTAGAATGGCTACAACAAGATCCC

RNGTT2-AscI-rv  
ttaccaGGCGCGCCttacgtgtctttgatcccttttag

Spt5-Paci-fw  
aatggtATGTCGGACAGCGAGGA

Spt5-AscI-rv  
accattGGCGCGCCCTAGGCTTCCAGGAGCT

10His-Halo-fw  
aatggtgaattaactcatcatacatcatacatcatacatcacaaccat cacGCA-

10His-Halo-rv  
accattTCTAGAGTTATCGCTCTGAGCTAAGTAGCT

hIIB-10His-fw  
attacaaGCATGCGCCCACTGAGCTGAGATGTACAG

hIIB-10His-rv  
attacaaGCATGCGCCCACTGAGCTGAGATGTACAG

Kozak-MscL-HindIII-fw  
attacaaGCATGCGCCCACTGAGCTGAGATGTACAG

Kozak-MscL-SalI-fw  
attacaaGCATGCGCCCACTGAGCTGAGATGTACAG

MscL-SalI-rv  
ttaccaTCTAGAAGATGCGCTTCTCCTCCTCCT

Xpo6-fw-III  
attacaaGCATGCGCCCACTGAGCTGAGATGTACAG

Xpo6-rv  
attacaaGCATGCGCCCACTGAGCTGAGATGTACAG
### 4.6 EMSA Oligos

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP-EcoRI-NLS-fw</td>
<td>ggtgaattcatgcggaaaagaagcgcgtgtgATGGACAAA</td>
</tr>
<tr>
<td>SNAP-Linker-Xbal-Ascl-rv</td>
<td>aatggcgccctTTATCTAGAGTTATCGCTCTGAAAGTA</td>
</tr>
<tr>
<td>NLS-TMP-EcoRI-fw</td>
<td>aatgtGAATTCatgcCAGAAGAAGCGGAAGGTC</td>
</tr>
<tr>
<td>TMP-Xbal-rv</td>
<td>ttaccaTCTAGAACCCGTACCCGACCCACC</td>
</tr>
<tr>
<td>pnAcGFP-Ascl-rv</td>
<td>ttaccaGGCGCACCTTACCTTCCCGTTTTTTTTTG</td>
</tr>
<tr>
<td>pnAcGFP-EcoRI-fw</td>
<td>aatgtGAATTCATGGCTAGGTGCAGGCAC</td>
</tr>
<tr>
<td>seq-pICE-fw</td>
<td>GGTGGGAGGTTCTATATAAG</td>
</tr>
<tr>
<td>seq-MyoVI-fw</td>
<td>CCGAGGCTGCACTAGA</td>
</tr>
<tr>
<td>seq-MyoVI-fw-II</td>
<td>CATTACGGCCATTTTGGC</td>
</tr>
<tr>
<td>seq-MyoVI-rv</td>
<td>GCCTCACTTCACGTCT</td>
</tr>
<tr>
<td>seq-XPO6-II-rv</td>
<td>ACATCTGACTGATGGAT</td>
</tr>
<tr>
<td>seq-XPO6-rv-III</td>
<td>CTGCTCACCGCGAT</td>
</tr>
<tr>
<td>seq-ICE-rv</td>
<td>TCAATGAACTCATGTCGA</td>
</tr>
<tr>
<td>Seq-nAcinGFP-fw</td>
<td>CGGAGATCAAGTCTTCATCTG</td>
</tr>
<tr>
<td>Seq-nAcinGFP-rv</td>
<td>GTAGCAGAGGTTGGTC</td>
</tr>
<tr>
<td>seqIRESrv</td>
<td>TAACATTATGACAAACGCAC</td>
</tr>
</tbody>
</table>

### 4.7 Plasmids

<table>
<thead>
<tr>
<th>Name</th>
<th>Reference</th>
<th>Backbone</th>
</tr>
</thead>
<tbody>
<tr>
<td>LVTO-Halo-MyoVI-wt</td>
<td>this work</td>
<td>Addgene # 19766</td>
</tr>
<tr>
<td>LVTO-Halo-MyoVI-zip</td>
<td>this work</td>
<td>Addgene # 19766</td>
</tr>
<tr>
<td>LVTO-Halo-MyoVI-CBD</td>
<td>this work</td>
<td>Addgene # 19766</td>
</tr>
<tr>
<td>LVTO-Halo-MyoVI-motor</td>
<td>this work</td>
<td>Addgene # 19766</td>
</tr>
</tbody>
</table>
### 4 Materials

<table>
<thead>
<tr>
<th>Construct</th>
<th>Source</th>
<th>Addgene #</th>
</tr>
</thead>
<tbody>
<tr>
<td>LVTO-Halo-MyoVI-SAH</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-Halo-MyoVI-RRL</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-Halo-MyoVI-RRL-NLS</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-Halo-MyoVI-WWY</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-Halo-MyoVI-WWY-NLS</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-Halo-Ypet-TFIIB</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-Halo-Ypet-RNGTT</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-Halo-Ypet-Spt4</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-10His-TBP</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-10His-TFIIB</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-10His-Halo-TBP</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-10His-Halo-TFIIB</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-Halo-TBP</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-Halo-H4</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-10His-Halo-H4</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-Halo-TBP</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-Halo-TFIIB</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-eDHFR/TMP-Stop</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-eDHFR/TMP-TBP</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>LVTO-nAC-GFP</td>
<td>this work</td>
<td>19766</td>
</tr>
<tr>
<td>pLentiCMV-Tet3G-MscL-Kozak</td>
<td>this work</td>
<td>17452</td>
</tr>
<tr>
<td>pIRES-eGFP-XPO6</td>
<td>this work</td>
<td>Doerner et al., 2012</td>
</tr>
<tr>
<td>pIRES-eGFP-MscL-G26C</td>
<td>this work</td>
<td>Doerner et al., 2012</td>
</tr>
<tr>
<td>pNLS-actin-S14C-YFP</td>
<td>this work</td>
<td>Toseland Lab</td>
</tr>
<tr>
<td>pNLS-actin-R62D-YFP</td>
<td>this work</td>
<td>Toseland Lab</td>
</tr>
<tr>
<td>pSNAP-H2B</td>
<td>this work</td>
<td>NEB # N9186</td>
</tr>
<tr>
<td>pnAC-GFP</td>
<td>this work</td>
<td>Chromotek # acg-n</td>
</tr>
<tr>
<td>pMD2.G</td>
<td>this work</td>
<td>Addgene # 12259</td>
</tr>
<tr>
<td>psPAX2</td>
<td>this work</td>
<td>Addgene # 12260</td>
</tr>
</tbody>
</table>
5 Methods

5.1 Biochemical methods

5.1.1 Cloning

All constructs were cloned using standard molecular cloning methods. Complementary DNA (cDNA) was recovered from HeLa cells with the QIAGEN RNeasy Plus Mini Kit and the Invitrogen SuperScript VILO cDNA Synthesis Kit, according to the manufacturer’s protocol. Genomic DNA (gDNA) was isolated using the Qiagen DNeasy Blood and Tissue Kit. The standard cloning procedure includes the following steps:

**PCR mix:**

- Q5 Buffer 5x 20 µl
- dNTPs 10 mM each 2 µl
- fw Primer 10 µM 6 µl
- rv Primer 10 µM 6 µl
- Q5 Polymerase 2 u/µl 1 µl
- template 1-10 ng
- ddH2O fill up to 100 µl
5 Methods

**Standard PCR:**

1. 98°C 1-5 min
2. 98°C 10 s
3. 55-60°C 30 s
4. 72°C 60 s per 1000 b
go to 2 35x
5. 72°C 5 min
6. 10°C forever

**Touchdown PCR:**

Touchdown PCR was used for the amplification of fragments from mammalian gDNA This PCR method generates PCR products with a higher specificity. Additionally, the primers used were blasted at NCBI.nlm.nih.gov.

1. 98°C 1-5 min
2. 98°C 10 s
3. 70°C (-0.5°C per cycle) 30 s
4. 72°C 60 s per 1000 b
5. go to 2 20x
6. 98°C 10 s
7. 55-60°C 30 s
8. 72°C 60 s per 1000 b
9. go to 6 20x
10. 72°C 5 min
11. 10°C forever

**Restriction digest:**

Restriction digest was performed using the appropriate restriction enzyme and the recommended buffer conditions. For standard procedures, 2 μg plasmid and 1.5 μg insert DNA was digested for 3h, at the recommended temperature. Plasmids were dephosphorylated using rSAP alkaline phosphatase for 1h at 37°C. Non-phosphorylated inserts, created by annealing of two complementary oligo strands,
were phosphorlyated using T4 polynucleotide kinase. According to the manufacturer's protocol, digested plasmids were purified using the QIAquick Gel Extraction Kit or the Monarch DNA Gel Extraction Kit. Digested inserts were purified using the QIAquick PCR Purification Kit according to the manufacturer's protocol.

**Ligation, Transformation and Selection:**
Ligation was done using the QuickLigase with 50 ng of target plasmid and a molar ratio of plasmid to insert of 1:1, 1:3 or 1:5 (plasmid:insert). Ligation products were transformed into chemical competent E.coli strains using heat shock, followed by antibiotic selection on LB plates. Standard plasmids were amplified at 37°C using X11 blue or Dh5α cells. Lentiviral target plasmids contain long terminal repeats (LTRs) that are prone to homologous recombination. Stbl3 and ccdB cells show reduced homologous recombination frequencies of the LTRs. Therefore, plasmids for the production of lentiviral targets were amplified in Stbl3 or ccdB cells at 33°C. Finally, plasmids were purified using QIAPrep Spin Miniprep Kit or QIAGEN Plasmid Plus Midi Kit according to the manufacturer's protocol. Successful cloning was verified by sequencing.

Used primers and created plasmids are listed in chapter 4.5 and 4.7. Cloning of the MyosinVI constructs was performed by the Toseland group (Grosse-Berkenbusch et al., 2020).

**5.1.2 Western Blot**
Western Blot is a technique to detect specific proteins in a cellular sample. The protein sample can be derived from different sources, and the protein of interest is ultimately detected via antibodies. A standard western blot procedure includes sample preparation, gel electrophoresis, blotting, and detection. Western blot analysis was used to compare endogenous and ectopic expression levels of myosinVI.
Sample preparation
Protein samples were derived from a stable polyclonal cell line, expressing Halo tagged MVI. Cells were grown on a 10 cm dish to a confluency of 80%. Cells were trypsinized, and the cell pellet was washed twice with cold PBS. Afterwards, cells were resuspended and incubated for 30 min on ice in 400 µl cold RIPA lysis buffer, containing 2 mg protease inhibitor. Lysed cells were subjected to three freeze (liquid nitrogen) and thaw cycles (37°C). Finally, the protein mix was separated by centrifugation at maximal speed in a table-top centrifuge for 30 min at 4°C. The supernatant was collected, and the protein concentration was determined by BCA assay. For the BCA assay, the protein lysate was diluted in a final volume of 10 µl at a ratio of 1:4 in water. Eight different protein standards were prepared to final concentrations ranging from 0 to 1.5 mg/ml BSA. 2.5 µl lysis buffer from the sample preparation was included in each standard, to correct for the added protease inhibitor. Protein samples and standards, as duplicates, were mixed with 200 µl BCA working reagent and further incubated at 37°C for 30 min. Absorption measurements were performed, and protein concentrations were derived by fitting a linear equation to the values derived from the protein standards.

Blot and Detection
Prepared protein samples were loaded on a 10% tris-glycine polyacrylamide gel. After gel separation, the protein was blotted onto a nitrocellulose membrane. For better transfer of large proteins like myosinVI, the transfer buffer contained 10% Ethanol. Protein detection was done by primary and secondary antibody staining and visualization with alkaline phosphatase. The antibodies were used in the following dilutions:

Primary antibody

- γ-tubulin (1:200,000)
- Myosin VI (1:500)

Secondary antibody

- mouse alkaline phosphatase (1:10,000)
5.1.3 EMSA

The electromobility shift assay (EMSA) is typically used to detect DNA-protein interactions. The goal of the EMSA is to separate DNA bound to proteins and free non-complexed DNA on a native polyacrylamide gel. The migration speed of the sample is determined by the charge and size of the sample. Free DNA migrates faster than DNA-protein complexes. An EMSA preparation consists of several steps. First, the purified protein and the labeled DNA is incubated at physiological conditions to stabilize the complex. Second, the complex is loaded onto a non-denaturing polyacrylamide gel (4-16% Bis-tris PAGE) that facilitates the separation of free from complexed DNA. Finally, the labeled molecules (DNA and potentially labeled proteins) are detected (Holden and Tacon, 2011). A schematic EMSA is shown in figure 5.1.

Figure 5.1: EMSA. Migration of free DNA compared to DNA bound in different protein complexes. DNA and complexed proteins can be detected via fluorescence.

EMSAs were done to test the functionality of labeled TBP and TFIIB. The DNA sequence used contained a TATA-binding site and was labeled with TMR (GGTTTCGA CTTAAAGGTATAAATGCGCCGTG-TAMRA, according to Moore et al., 1999). In brief, the proteins were incubated with DNA in assembly buffer for 25 min at RT. In the meantime, a prerun of the native PAGE was performed to remove debris and clean the gel pockets (TBE, 120V, 30 min, 4°C). The incubated samples were stored on ice and successively loaded on the gel. The gel ran for 40 min at 120V and 4°C. More details about specific experimental conditions can be found in chapter 6.1 and 6.2.3.
5 Methods

5.2 Cell culture

All used cell lines, namely HeLa, MCF7, HEK, and NIH-3T3, were grown in regular DMEM or MEM (MCF7), supplemented with 10% FBS, 1% Glutamax, 1% nonessential amino acids. DMEM used for HeLa, HEK, and NH3T3 cell lines were additionally supplemented with 1% Sodium Pyruvate. Cells were split regularly and kept in a cell incubator at 37°C and at 5% CO₂. Monthly mycoplasma tests were performed to test for possible contamination.

5.2.1 Generation of stable cell lines

Stable cell lines were created by using a lentiviral transfection system. For safety issues, the genes necessary for virus production are encoded on three different plasmids, ensuring that the produced virus cannot replicate on its own. For this 2nd generation lentiviral system, a target plasmid, a packaging plasmid (pMD2.G), and an envelope plasmid (psPAX2) were used. The gene of interest was cloned into the target plasmid, where it was flanked by the required long terminal repeat sequences (LTRs). The gene of interest and the flanking LTRs are later integrated into the host genome. HEK cells (LentiX-293T) were used for the virus production (Figure 5.2).

![Diagram of stable cell line generation](image)

Figure 5.2: Generation of stable cell lines. 1. LentiX-293T cells are transfected with the lentiviral plasmids. 2. Target cells are infected with virus particles. 3. The transduced and stable cells are selected and stored for further use.
5.2 Cell culture

Virus Production
LentiX-293T cells were grown on a 10 cm dish to a confluency of 80%. After exchanging the medium, cells were transfected with 10 µg target plasmid, 2.5 µg packaging plasmid, and 7.5 µg envelope plasmid, by using jetPrime, following the manufacturer’s instruction. Cells were further incubated for 24h. After successful virus production, the medium containing the virus particles was filtered through a 0.45 µm membrane filter, and the obtained virus solution was stored at -80°C.

Cell Transduction
Target cells were seeded on a 6 well plate containing 3 ml of medium. In general, one day prior transduction, three wells were seeded with 30,000 cells per well. On the day of transduction, after exchanging medium, freshly produced or thawed virus solution was applied dropwise onto the cells. Since the virus titer was not determined, three different amounts of virus solution, ranging from 0.5 ml to 3 ml, depending on the construct, were tested. For a first trial, 0.5 ml, 1ml and 1.5ml were used. Cells were incubated for 72h. Successful virus transduction was verified either by antibiotic selection or by fluorescence microscopy followed by FACS. Stable cell lines were stored in liquid nitrogen.

5.2.2 Fluorescent Activated Cell Sorting - FACS

FACS experiments were conducted at the Core Facility Cytometry (Ulm University) with the Core Unit staff’s help. Cell sorts were performed with a BD FACSaria III. FACS measurements were performed using a BDLSR Fortessa.

Cell sorts
Fluorescent activated cell sorting was used for the selection of stable cell lines. After viral transduction, a polyclonal cell line, with variable expression levels was obtained. FACS was used to select a pool of cells with a more homogeneous expression level and to dismiss non transduced cells. Cells were grown on a 10 cm dish to a confluency of 80%. Cells expressing a Halo-Tag construct were labeled with 2.5 µM of TMR-Halo ligand according to the manufacturer’s protocol. After labeling, cells were trypsinized and resuspended in 200 µl PBS, supplemented with 10% FBS (10^6 cells/ml). After labeling, cells were kept on
ice in a light-proof container. Fluorescent positive cells were sorted into DMEM supplemented with 5% Penicillin-Streptomycin, amplified on a 10 cm dish, and stored in liquid nitrogen. Single-cell sorts for the selection of monoclonal, genome-edited cell lines were performed accordingly onto 96 well plates.

**FACS measurements**

Single molecule experiments are performed at a low labeling density. FACS experiments were done to quantify the Halo-Tag labeling efficiency at different dye concentrations.

Sorted cells were labeled with TMR-Halo ligand according to the manufacturer’s protocol. Different dye concentrations, ranging from 50 pM up to 5 μM were used to assess the minimal and maximal labeling efficiency and the resulting fluorescent signal of the expressed Halo-fusion construct. The cells were prepared as described above (Cell sorts), and 50,000 cells were measured for each dye concentration. The fluorescence signal of the labeled cells was recorded, and the mean intensity values were determined by Gaussian fitting.
5.3 Halo-Tag labeling

The Halo-Tag (Los et al., 2008) is the only commonly used tag for nuclear single molecule tracking experiments (Presman et al., 2017). One of the goals of this thesis has been the establishment of an orthogonal tagging system for dual-color single molecule tracking experiments, inside the nucleus of living cells, in unison with the HaloTag. More details about this topic can be found in chapter 6.

Procedure

Cell lines expressing HaloTag fusion constructs were labeled with organic dyes for single molecule tracking experiments. First, cells were seeded on heat-able Delta-T imaging dishes. 100,000 cells were seeded for imaging on the next day, and 80,000 cells were seeded for imaging two days later. In the case of transient transfection, 100,000 cells were seeded on a Delta-T dish, followed, on the following day, by transient transfection of 2 μg of plasmid with jetPrime. Cells were then imaged on the third day, 24-36h after plasmid transfection.

The labeling procedure of proteins tagged with the HaloTag varied depending on the used dye and the tagged protein:

For SiR- and TMR-HaloTag ligands, Halo-MVI expressing cells were incubated with 31 pM or 25 pM dye. Cells expressing Halo-TBP and Halo-TFIIB fusion constructs were labeled with 3.1 pM or 2.5 pM Halo-ligands, respectively. Cells were incubated for 15 min in 1 ml of dye solution. For protein labeling with the PA-JF-647 dye, cells were initially incubated with 100 nM dye for 60 min. After the different labeling processes, the cells were washed once with DPBS, followed by an incubation period of 30 min, to wash out unbound dye molecules. Before imaging, cells were subjected to three short washing steps with DPBS. Finally, the cells were imaged in OptiMEM.
5 Methods

5.4 Spinning disc confocal microscopy

In conventional fluorescence microscopy, the sample is illuminated over a wide area, and the emitted light is collected from all sample planes. This microscopy technique comes with a high out of focus excitation and acquires images with a very low signal to noise ratio. Additionally, when imaging live samples, the large area of illumination presents a source for cellular phototoxicity that is particularly harmful while imaging live-cell samples. In contrast, spinning disc confocal microscopy has a high axial resolution. It uses a set of lenses and pinholes on two rotating discs to focus the excitation beam and eliminates out-of-focus emission. This reduces phototoxicity and improves the SNR. Moving the sample in z gives the opportunity to create a 3D image (Nakano, 2002).

The used setup is built around a commercial corpus (Axio Observer D1, Zeiss) and equipped with four different lasers lines (405 nm, 473 nm, 532 nm, and 638 nm), that can be controlled by an acousto optic tunable filter (AOTF), a spinning disc, a flexible emission filter unit and an emCCD camera (Andor iXon DV887). The user addresses the microscope through the Andor Solis software, and images are acquired with the Andor iQ software.

5.5 Single molecule fluorescence microscopy

Small molecules, like proteins, build up the working machinery of the smallest viable subunit, a cell. The inside of a cell is a densely crowded environment that is hard to assess. One way to study individual processes is through classical protein biochemistry, where proteins are extracted from there biological background and studied in vitro in a defined and controlled reaction environment. However, these systems lack the specific cellular context with all its protein interactions, various regulatory pathways, different control mechanisms, and its specific environment. Live cell single molecule microscopy can be used to study individual proteins in their natural environment (Gebhardt et al., 2013; Tokunaga et al., 2008).

Fluorescence is the emission of light from a fluorophore that has absorbed light. Fluorescence is a three-step process that can be explained using the Jablonski
5.5 Single molecule fluorescence microscopy

Figure 5.3: (a) Jablonski Diagram showing the excitation of an electron from the ground state ($S_0$) to the excited state $S_{1/2}$ and the resulting relaxation possibilities.
(b) Schematic representation of the HILO geometry for live cell single molecule experiments and its comparison to Epi illumination.

diagram (Figure 5.3). First, an electron is excited from the ground state $S_0$ into the singlet $S_1$ or $S_2$ state, by absorption of a photon. In the $S_{1/2}$ state energy is lost, for example, as heat to the environment. This leads to a relaxed $S_1$ state. The lifetime of the excited singlet state is in the nanoseconds range. Finally, the electron returns to the ground state by emitting fluorescence light. However, it is also possible to return to the $S_0$ via nonradiative processes (n.r.) (Figure 5.3). Each state has several different vibrational levels. While the electron loses energy in the $S_{1/2}$ state, it drops to a lower vibrational level. Therefore, the emitted light is of lower energy, shifting the emission spectra, with respect to the excitation spectra, to higher wavelengths. This is called the Stokes shift. The transition process, from one state to another state, with the highest probability, gives the excitation or emission maxima. A third possibility for an electron to drop back to the ground state is an indirect one via the excited triplet state. This is called intersystem crossing. During the transition, the electron is subjected to a spin-flip. The spin-flip occurs with a very low probability, making the triplet state a relatively long-lived state (Fili and Toseland, 2014).

Light sheet microscopy is one way to track single proteins in living cells. The most commonly used technique is highly inclined and laminated optical sheet (HILO) microscopy (Tokunaga et al., 2008). In HILO microscopy, a thin light sheet is created using the refraction properties of light passing through media with different indices of refraction. In contrast to total internal refraction microscopy TIRF (Ambrose, 1956), where the incoming laser beam is completely reflected at the boundary of air to glass, the incoming angle in HILO is close to the critical angle creating a thin light sheet. This sheet can excite a small fraction of the cellular object placed onto the
5 Methods

glass coverslip. HILO microscopy uses a similar principle like TIRF microscopy, and the excitation volume is reduced, thereby reducing the out of focus signal. In TIRF microscopy, the area of excitation starting at the coverslip is limited to roughly 100 nm into the sample, making it a useful tool to study cellular membranes. HILO microscopy extends this range dramatically. The created highly inclined and laminated sheet reaches deeply into cells (Gebhardt et al., 2013). The signal to noise ratio is increased drastically, compared to epi microscopy. Compared to spinning disc confocal microscopy, HILO microscopy offers a higher acquisition time because there is no need for scanning the sample. This gain, however, comes with a decrease in axial resolution.

All necessary ingredients are now part of the game. With fluorescence as the read-out, fluorescently labeled proteins are readily available in live-cell model systems and can be detected at a single molecule level with HILO microscopy. The used setup was built by Matthias Reisser around a commercial corpus (TiE, Nikon) and equipped with five different lasers lines (405 nm, 488 nm, 514 nm, 561 nm, and 640 nm), that can be controlled by an acousto optic tunable filter (AOTFnC-400.650) and an emCCD camera (iXon Ultra DU 897U, Andor). All laser lines were collimated, and the excitation beam was confined by a pinhole (4.25 nm diameter) and focused on the back focal plane of a high NA objective (100x, 1.45 NA Plan Apo, Nikon). The user addresses the microscope through NIS Elements (Nikon NIS-Elements Version 4.40.00) (Reisser, 2019). Live cell single molecule experiments were all performed using highly inclined laminated optical sheet microscopy (Takunaga et al., 2008). Stable or transiently transfected cell lines, expressing tagged protein fusion constructs, were labeled at a low labeling density, ensuring single particle tracking by a nearest neighbor algorithm. The cells were imaged for up to 90 min, in a temperature-controlled environment HEPES buffered system at 37°C (Delta-T dishes).

5.5.1 Interlaced Time-lapse Microscopy (ITM)

Interlaced Time-lapse Microscopy (ITM) is another way to circumvent photobleaching during single molecule experiments. ITM enables to differentiate between three
5.5 Single molecule fluorescence microscopy

Figure 5.4: Schematic representation of Interlaced Time-lapse Microscopy (ITM).

different populations of molecules: diffusive, short bound, and long bound ones (Reisser et al., 2018). The labeled cells are excited sequentially for two consecutive frames, followed by a distinct dark time. The dark time is chosen such, that short bound molecules statistically do not survive the dark time, whereas long bound molecules do. In summary, molecules that are detected once are categorized as diffusive ones. Molecules detected in two consecutive frames are categorized as short bound ones, and molecules that survive at least one dark time are categorized as long bound ones.

For ITM, the cells were prepared, as described in chapter 5.3. SiR-labeled HeLa HaloTag-MVI cells were imaged with an illumination time of 10 ms to reduce blurring and achieve a high two-dimensional resolution of diffusing molecules. The camera exposure time was set to 50 ms, and the dark time was chosen to be 2s (Figure 5.4). The used tracking algorithm is based on a nearest neighbor algorithm, and a molecule was considered to be bound if it did not leave a radius of 160 nm in two consecutive frames. The signal to noise threshold was set to 1.5 (= SNR).

5.5.2 Time-lapse microscopy

Time-lapse microscopy is an illumination scheme that allows for the direct correction of photobleaching (Gebhardt et al., 2013). Another method would be the interference of the bleaching rate by measurement of a statically bound protein like histones (J. Chen et al., 2014). By using time-lapse microscopy, different dark times are introduced between two consecutive frames. Each frame has a fixed camera
integration time of 50 ms, which equals the exposure time. For a complete series of time-lapse measurements, the dark time is increased to a length that ensures the complete acquisition of long binding events without being limited by the photobleaching process.

For measurements of Halo-MVI molecules, continuous movies with an integration time of 50 ms and time-lapse movies with dark times of 100 ms, 200 ms, 400 ms, 800 ms, 1.6 s, 3.2 s, 6.4 s, and 12.8 s were recorded. The used tracking algorithm is based on a nearest neighbor algorithm, and a molecule was considered to be bound if it did not leave a radius of 240 nm in two consecutive frames. For time-lapse microscopy, cells were prepared as described in chapter 5.3. The detected spots were fitted with a single Gaussian function to determine the position of the particle. The signal to noise threshold was set to 3.7 (= SNR), and molecules were allowed to be absent for 1 frame, in order to correct for dynamic photobleaching (Dark Frames = 1). Tracked particles with a track length of two, extracted from continuous movies and time-lapse movies with a dark time of up to 200 ms, contained false-positive connections arising from diffusive molecules. Therefore the minimal track length for these time-lapse conditions was set to three.

5.5.3 GRID

Genuine rate identification (GRID) was used to infer rates and amplitudes of nuclear MVI dissociation processes. Obtained binding time distributions from different time-lapse measurements are fitted with a grid of fixed multiexponential functions with free amplitudes (Reisser et al., 2020).

The measured time-lapse data sets contained binding time distributions with overlapping information that resulted in unstable fitting results. Therefore, only data sets with at least 1000 data points and good signal to noise quality were selected for further analysis. Binding time distributions were inferred from the following time-lapse distributions: 50 ms, 200 ms, 400 ms, 1.6 s, and 6.4 s. The error estimation was done by 500 resampling runs with 80% of the data. The resulting data points, defining one rate cluster, are summarized in the spectral weight (±s.d.). The obtained rate clusters can be presented in two different ways. One way of representation is the event spectrum (Figure 7.4 (c)). The event spectrum displays the relative
frequency of a dissociation event during the observation period. Therefore the amplitude includes information about the number of binding sites and the on rate. The state spectrum arises from weighting the event spectrum with the amplitudes obtained by the event spectrum (Figure 7.4 (b)). The resulting amplitude gives information about the probability to observe, in a snapshot, a molecule in a certain binding state (Reisser et al., 2020).

5.5.4 Measurement of directed motion

The directed motion of MVI molecules was monitored by continuous illumination with a camera integration time of 50 ms, which equaled the exposure time. The cells were prepared as described above in chapter 5.3. The detected spots were fitted with a single Gaussian function to determine the position of the particle. The signal to noise threshold was 0.8 (= SNR), and molecules were allowed to be absent for 1 frame, in order to correct for dynamic photoblinking (Dark Frames = 1). The used tracking algorithm is based on a nearest neighbor algorithm, and a molecule was tracked if it did not leave a radius of 320 nm in two consecutive frames. Overall a track was considered as a run if it lasted for more than 500 ms. The automatic detection of directed movement was done using a recurrence algorithm (SPLIT see chapter 5.5.5).

5.5.5 SPLIT

The initially observed directed MVI motion events were finally automatically identified by the recurrence based algorithm: Speedy partition of Legs in Tracks (SPLIT), that was programmed by Johannes Hettich (Grosse-Berkenbusch et al., 2020). First, localization errors were smoothed by a sliding average of the subsequent x- and y-coordinates. Second, motion events were separated into runs and pauses by setting a threshold defined by the number of points in close proximity (=summation border) and the angle between one step (Figure 5.5), and third, data points from separated track fragments were compressed to correct for false classifications. Finally, tracks were filtered by assuming a diffusion model, in order to remove tracks that could have arisen from anomalous diffusion.
Figure 5.5: The motion of MVI molecules was separated into runs and pauses by a recurrence algorithm. A threshold (red circle) based on the amount of direct neighbors in a defined radius ($\epsilon$) and an angle criterion was used to classify runs and pauses. (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0).

A summation border of five was defined to prevent false classification of circular motion, which would be defined through many points in close proximity. One spot was thereby classified as being directed if it had less than five neighbors in a radius of $\epsilon$ 0.4 $\mu$m. $\epsilon$ was defined by assuming a diffusion coefficient for chromatin-bound molecules of 0.01 $\mu m^2/s$ (Akhtar and Gasser, 2007). Further classification defined a section as run if successive spots varied in its direction by less than 5° or as being non directed if its direction changed by more than 45°. To exclude stochastic diffusive particles with straight sections, the probability that this track could arise from diffusion was calculated using a 2D diffusion model, and the track was excluded if the calculated probability was lower than $10^{-6}$. 

46
5.6 CRISPR/Cas

CRISPR/Cas is an adaptive immune system of prokaryotic cells that targets DNA and RNA of viral intruders (Barrangou et al., 2007). It is based on a specific DNA region that contains the Cas genes and a gene array of clustered regularly interspaced short palindromic repeats (CRISPR). The CRISPR array contains genetic fragments from different viruses and consists of a variable number of repeats separated by unique spacers. The CRISPR/Cas based immune response can be divided into three distinct phases: Adaption, Expression, and Interference. During the adaption phase, pieces of foreign viral DNA or RNA are integrated into the CRISPR locus. These spacers serve later as target recognition elements. In the following expression phase, Cas genes and the CRISPR array are transcribed and built into functional RNA/Protein Cas complexes. In the final interference phase, Cas complexes recognize introduced foreign viral DNA or RNA and degrade it (Hsu et al., 2014).

Genome Editing

Genome editing is an essential tool for live-cell single molecule studies of endogenous proteins (Mahen et al., 2014). The precise induction of a DNA double-strand break is the crucial mechanism for efficient genome editing. The first genome editing approaches used Zinc-Finger Nucleases or TALENS (transcription activator-like effector nucleases). However, both techniques are time-consuming and expensive (Wood et al., 2011). The CRISPR/Cas9 system that was initially derived from the Cas9 nuclease from *Streptococcus Pyogenes* is a fast, efficient, and cheap tool that can be used for genome editing in many living organisms from bacteria to plants (Ran et al., 2013). Cas9 introduces DNA double-strand breaks that lead to apoptosis or are either repaired via the non-homologous end joining (NHEJ) or the homology-directed repair (HDR) pathway. NHEJ is a fast and error-prone process that can result in random insertions or deletions at the target site. This process is used in knock-out experiments to generate premature stop codons (Perez et al., 2008) or to delete large genomic regions by using two cleavage sites (Cong et al., 2013). The HDR pathway uses the homologous information from the sister chromatid to repair the double-strand break. The latter pathway can be used for knock-in experiments. However, the HDR pathway’s efficiency is limited as it competes with
5 Methods

Figure 5.6: Genome editing with CRISPR/Cas. Initially, the assumed workload is mostly only the tip of the iceberg. The final workload of the screening process is directly dependent on the carefully chosen experimental approach. This workload can be reduced by choosing optimal conditions for the individual experimental approach preceding the screening process.

the NHEJ pathway, which has a much higher occurrence rate (Cong et al., 2013). Recent reports of increasing the HDR efficiency for knock-in approaches by inhibiting the NHEJ pathway (Bertolini et al., 2009; Chu et al., 2015; Maruyama et al., 2015; Srivastava et al., 2012) resulted in almost no, or only highly limited increases in HDR repair (Greco et al., 2016; Koch et al., 2018).

For genome editing experiments, an engineered and optimized CRISPR/Cas9 plasmid system was used that contained the Cas9 protein and a synthetic guide RNA (gRNA). The gRNA sequence encoded the information for the target region (crRNA) and the binding site for the Cas protein (tracrRNA) (Cong et al., 2013). The used system was selected due to the described high and efficient genome editing ability of HeLa cells (Koch et al., 2018). Knock-out experiments were done by targeting a single DNA site around the start codon with the Cas9 complex. Knock-in experiments were done by using the Cas9 nickase (Cas9n). The use of Cas9n, in contrast to Cas9, has been shown to increase the HDR efficiency and the chance to generate homozygous cell lines during the integration process of large tags like eGFP or the Halo-Tag (Bothmer et al., 2017; Koch et al., 2018; Miyaoka et al., 2016). The whole genome editing procedure was done according to Koch et al., 2018 and contained the following steps (see also Figure 5.6):
1. design gRNAs
2. test gRNAs
3. Transfection
4. Screening

**gRNA design**
There is a considerable variety of gRNA design tools that use different algorithms to predict "good" gRNAs for the selected Cas protein and its specific PAM motive. The available tools rank the specified guides according to a calculated score, which is mostly defined by the guides' off-target probability. The experimental design of selecting gRNAs was done using at least three different online tools (Sanger, Chopchop, CRISPOR, Atum CRISPR, Cas-OFFinder) to identify around 20 different gRNAs that bind close to the target site. Finally, additional criteria according to Liang et al., 2017 were used to select around 20 gRNAs for further test experiments.

**Test gRNAs**
Most gRNA selection tools rank the guides by their off-target probability. However, many computational approaches reach their limit when it comes to predicting cleavage efficiency and specificity (Liu et al., 2020). An easy and fast way to predict the on-target activity of multiple chosen gRNAs is the T7-Endonuclease I assay (Koch et al., 2018). This assay was used for single gRNAs selected for knock-out experiments and for gRNA pairs selected for HDR facilitated knock-in experiments. In the latter approach, two plasmids, each encoding a single gRNA and the Cas9 nickase, were used for the test.

**Transfection**
The transfection efficiency is an important step that, if optimized, can ultimately reduce the screening process. Since the chosen Cas9 nickase approach relied on the transfection of plasmids rather than purified Cas9 protein and/or synthetic gRNAs, HeLa and MCF7 cells were chosen because they showed a high survival and transfection efficiency for the previously used lipid-based reagents (jetPrime and Lipofectamine). JetPrime was used for the final plasmid transfections.
5 Methods

Screening
The screening process is dependent on the experimental approach. Screening for clonal cell lines that acquired random insertions or deletions via the NHEJ pathway is faster than screening for large protein knock-in cells. Overall the goal is to aim for a monoclonal cell line. This can be done either by limited dilution or by FACS (see chapter 5.2.2). FACS is the faster and more versatile approach. Additionally, sorting single cells into 96 well plates allows the selection of transfected cells that express Cas9-eGFP fusion proteins and/or labeled gRNAs to reduce the following screening time or directly screen for cells with homologous integrated fluorescent tags. Since the survival rate during FACS varies from 50-70 % (HeLa), the following screen should be done roughly 7-10 days after transfection to increase the number of positive cells before the sorting process. Screening for Cas positive cells can be done 3-4 days after transfection (Koch et al., 2018). In general, the FACS procedure should be optimized for high survival rates, since the limited amount of genome editing cells, might already be closer to apoptosis than the non-edited ones.

Validation (Screening) and final notes
The validation of sorted monoclonal cell lines is important, but also the most time-consuming step. The validation time can be reduced by optimizing all preceding experimental steps. Depending on the experimental design, all or several of the following validation steps should be included in the validation process. The following list is in no particular order and further individual validation steps, for example checking the functionality of a tagged or modified protein, or the general cell state (e.g., cell cycle time, morphology) should be included whenever relevant:

1. Junction PCR: A PCR on genomic DNA with primers that bind inside and outside of a recombined site. It can be used to check for the correct integration of an integrated marker and might give first indications about the zygosity.

2. Sanger sequencing: PCR products obtained from a junction PCR should be gel purified, in cases of multiple present products, and send for sequencing to verify the correct integration. Small genomic changes, for example in knock-out experiments, that can not be directly visualized via PCR should be sequenced as well.
(3) Southern Blot: Southern Blot analysis is vital for the validation of endogenous protein tagging as it shows the homozygosity of the integration, and it reveals any off-target integrations of the donor plasmid. It is also sensitive for the detection of genome rearrangements and double integrations.

(4) Western Blot: The validation steps (1-3) confirm the correct genetic manipulation. Western Blot analysis validates the protein expression and shows any changes in the expression level.

(5) Fluorescence microscopy: Live or fixed cell fluorescence microscopy confirms the correct protein localization, and the correct protein localization can indicate correct protein functionality.

Genetic compensation
CRISPR/Cas mediated endogenous protein tagging or the use of protein knock-outs is a general and accepted way for protein studies in living systems (Cattoglio et al., 2019; Chong et al., 2020; Khan et al., 2019; Mahen et al., 2014). However, one has to keep in mind that editing a genomic locus can ultimately result in altering the gene expression by, for example, influencing the promoter region or other regulatory elements (Koch et al., 2018). Genetic compensation is another feature that might arise in protein knock-out systems. It has been shown that two different orthogonal knock-out strategies of the same gene in zebrafish resulted in either a strong phenotype (morpholino), while the other strategy did not result in any morphological alterations (CRISPR/Cas) (Rossi et al., 2015). Further studies revealed that the CRISPR mediated knock-out strategy did result in increased gene expression of a related gene, which is mediated by degraded mRNA, which originated from the edited gene loci. This rescue process compensated the expected phenotype (El-Brolosy et al., 2018).

Once the initial decision for the optimal CRISPR/Cas system that possesses maximal advantages with minimal drawbacks for the edited system has been made, the general establishment of all steps is a relatively cheap and easy process. The ultimate goal is to achieve a gene-edited monoclonal cell line is, which is at least for large protein knock-in experiments, limited by the time-consuming screening process. The establishment of the CRISPR/Cas system was done with the help of my master students Ariel Asuzano, Natalia Deutsch, and Alexandra Junginger.
Fluorescent labeling of proteins in living cells

Fluorescent labeling is a commonly used experimental approach to dissect molecular mechanisms by examining protein structures, locations, and functions at spatiotemporal resolution. In recent years a wide variety of diverse fluorescent labeling methods have been developed for in vivo studies such as protein localization and tracking. The fluorescent labeling of proteins in live cells started with the discovery of the green fluorescent protein (GFP) from Aequorea victoria (Shimomura et al., 1962). Today there are many variations of the original GFP with different excitation and emission spectra, like the blue (BFP) or yellow fluorescent protein (YFP) (Heim et al., 1994). These variations can be used for localization and tracking of different proteins in live cell multicolor experiments. The primary advantage of fluorescent proteins is that the formation of the chromophore is genetically encoded. Thus, GFP can be expressed as a protein fusion construct, that is transferable to a wide range of target organisms (Chalfie et al., 1994). Despite their versatile use, fluorescent proteins do have several limitations, like their large size, that might interfere with normal cellular functions (Dyachok et al., 2006), or their fast photobleaching kinetics, that highly limits the observation period of a single particle (T. S. Chen et al., 2002).

Live cell single-particle tracking, requires minimal interfering labeling systems with high-quality fluorescent labels. First, the protein of interest needs to be labeled directly, either covalent or "almost" covalent, with high affinities in the picomolar range. Additionally, the label should be small and as close as possible at the site of interest, to ensure minimal biological interference and high localization precision. For quantitative experiments, one protein should be labeled with a single emitter (Wombacher et al., 2010). Secondly, single-particle tracking requires cell-permeable fluorescent dyes with improved brightness and a high photostability (Grimm et al.,
Fluorescent labeling of proteins in living cells

The final requirement is the compatibility of the labeling procedure with a living system, and the labeling technique has to exhibit minimal unspecific binding properties (Presman et al., 2017). Initial single and dual-color single molecule tracking experiments were done using fluorescent proteins (Gebhardt et al., 2013; Xie et al., 2008). However, while fluorescent proteins are genetically encoded, direct, and covalent labels that can be easily used in living systems, they show very high photobleaching rates, combined with low photon outputs. These photophysical parameters limit the localization precision and the observation time of the tagged proteins. Another disadvantage is the poorly controllable expression level. This results, even with the use of low expression systems and selected plane illumination, in high protein densities, which additionally limits the localization precision (Barlag et al., 2016).

The development of bright, photostable organic dyes for live-cell imaging (Grimm et al., 2015) in combination with self-labeling enzyme tags (Keppler et al., 2003) or bioorthogonal chemical labeling reactions (Sletten and Bertozzi, 2009), solved many of the above-stated issues, that limit single-particle tracking experiments. Organic dyes, like fluorescent proteins, are available with different excitation and emission maxima, covering the visible light spectrum. Most dyes exhibit good photophysical properties, with high fluorescence quantum yields and high excitation coefficients and low photobleaching and blinking characteristics (Lavis and Raines, 2008). The use of organic dyes also solved dosage issues that arise from low controllable protein expression systems. A well-defined amount of the expressed fusion protein can be labeled by adjusting the dye concentration during the labeling process, giving the user the freedom to adjust the labeling density. That process allows for precise single-particle localization and tracking, at high signal to noise ratios. The labeling density can be varied and adjusted depending on the experimental needs. Recent engineering of organic dyes paved the way to photoactivatable dye molecules. These dyes allow adding an additional layer of flexibility during the labeling process. The labeling density can be directly adjusted during the imaging procedure by UV activation. With the help of this system, a single cell could potentially be imaged over a more extended period of time, since regular activation pulses effectively counter the occurring photobleaching process (Wombacher et al., 2010).
In general, there are two different systems for protein labeling using organic dyes, either by direct bioorthogonal chemical labeling or by labeling via self-labeling enzyme tags. Self-labeling enzyme tags like the SNAP/CLIP (Gautier et al., 2008; Keppler et al., 2003), the Halo (Los et al., 2008), the TMP (Miller et al., 2005), the \(\beta\)-lactamase (BL) (Watanabe et al., 2010) or the Sortase (ST) (Refaei et al., 2011) -Tag are well-established systems that offer the ability for quantitative covalent labeling of proteins in living cells. Self-labeling enzyme tags have similar advantages and disadvantages than fluorescent proteins. They can be genetically encoded and co-expressed as protein fusion constructs, and they have a molecular size of around 20-30 kDa, that might disturb biological processes. However, the ability to label the protein of interest with organic dyes makes them superior to FPs. Recent studies characterized the usability of self-labeling enzyme tags for their usability in nuclear single molecule tracking experiments. It has been shown that the most commonly used tags (SNAP/CLIP/HALO) show unspecific second-long binding events in the nucleus of living cells (Presman et al., 2017). The characterized unspecific binding events occurred, depending on the tag, at different frequencies and ranged from milliseconds to several seconds (Presman et al., 2017). This and other studies left the Halo-Tag as a single commonly accepted option for nuclear single molecule tracking experiments, with commercially available cell-permeable organic dyes. However, many experimental approaches like the study of protein-protein interaction or protein complex formations on a single molecule, require orthogonal labeling techniques, that allow for multicolor fluorescent experiments.

Bioorthogonal direct chemical labeling offers a broad range of available labeling techniques. In general, they can be divided into two approaches—first, live-cell compatible labeling methods, and second, selective chemical labeling of \textit{in vitro} purified proteins. The first approach is mainly limited to cell-permeable dyes, whereas the latter one can, if successful, outcompete the first by its huge variety of available labeling techniques and organic dyes (X. Chen and Wu, 2016). After successfully establishing a labeling protocol, the latter approach offers higher flexibility in terms of dye exchangeability. The established protocol can finally be used to generate a set of proteins with different labels by exchanging the used fluorescent dye. Additionally, this procedure comes with high cellular specificity because the labeled proteins can be purified from the remaining unbound dye molecules by size exclu-
6 Fluorescent labeling of proteins in living cells

However, this approach comes along with a few technical difficulties. First, one needs to have a purified and bioactive protein, and the protein has to be delivered back into the living system, which represents a major experimental issue. Apart from that, chemical tags are very small and therefore impose minimal biological interference, which is highly relevant in the crowded biological environment, especially in studies aiming to investigate large multiprotein complexes. Here, protein based tags like the Halo-Tag, sometimes more than double the environmental space of the protein of interest (Dyachok et al., 2006). Additionally, prior protein engineering makes it possible to selectively define the labeling position, whereas genetically encoded tags are typically added to the N- or C-terminus. The described properties make direct labeling an ideal candidate for single molecule tracking experiments.

Model system: PIC formation and the RNA-PolymeraseII dependent transcription initiation process

Figure 6.1: The sequential assembly pathway of the preinitiation complex (PIC) during RNA-PolymeraseII dependent transcription initiation.

The formation of RNA-PolymeraseII dependent transcription preinitiation complex (PIC) is a complex process that is orchestrated by many different proteins (Hahn,
2004). This system served as a model system to investigate new fluorescent labeling techniques for nuclear single molecule tracking experiments. The formation of this multiprotein complex is not yet fully understood, and multicolor single molecule imaging might shed light on this process (Horn et al., 2016). The core promoter region with its TATA-binding motif and its flanking TFIIB-recognition elements (BRE) serves as a platform for the assembly of the PIC (Lagrange et al., 1998). Two assembly pathways may lead to a fully functional complex formation. The sequential assembly pathway postulates a stepwise recruitment of the general transcription machinery. First, TBP, which is part of the transcription factor TFIID, recognizes the TATA-binding motif on the DNA. After binding, the DNA bends and the transcription factors TFIIA and TFIIB can bind the multiprotein complex. These factors serve as a binding platform for the RNA-PolymeraseII, and the complex formation is completed after the sequential binding of TFIIE, TFIIF, and TFIIH (Figure 6.1)(Thomas and Chiang, 2006). An alternative pathway is the holoenzyme pathway. Here, a preassembled TFIID (TBP+TAFs) complex and an RNA-PolII complex that binds the remaining transcription factors (TFIIA, TFIIB, TFIIE, TFIIF, and TFIH), assemble before promoter binding. Finally both preassembled complexes bind the promotor region in unison (Thomas and Chiang, 2006).

After successful PIC assembly, gene transcription is initiated. TFIIB plays another critical role in this post PIC assembly process. The newly transcribed RNA leaves
the protein complex through the RNA-exit tunnel. However, the exit of this channel is most likely blocked by the B-Finger structure of TFIIB (Bushnell et al., 2004). This structural phenomenon led to the assumption that TFIIB might leave the PIC to facilitate further RNA elongation. Dual-color single molecule tracking and colocalization experiments might reveal the true nature of the PIC formation process and shed light on the time spent from the initial binding of TBP till transition into RNA-PolII elongation state. The last question might be solved with additional studies of several elongation factors like the elongation factor Spt4 or the capping protein RNGTT that both bind the phosphorylated C-terminal RNA-PolII tail after PIC assembly, at the onset of the elongation phase (Figure 6.2) (Henriques et al., 2018). However, a second live-cell compatible, orthogonal, fluorescent labeling method is required to tackle these scientific questions with single molecule fluorescence microscopy.

6.1 Selective chemical labeling of purified proteins

The full range of organic chemistry, for example, condensation reactions through carbonyls or click reactions through azides, is available for selective chemical labeling of proteins (X. Chen and Wu, 2016). However, most techniques require prior protein engineering for locally defined and ratiometric 1:1 labeling. The use of lysines or cysteines is an attractive choice for the site-specific labeling of proteins. Cysteines are especially interesting due to their low abundance in the proteome. Conjugated dye-maleimides are commonly used for the coupling reactions to cysteines. Reduced, accessible thiol groups are required for the conjugation reaction. Reducing agents are used to increase labeling efficiency. However, they need to be removed before dye conjugation, to prevent competition with the thiol groups during the labeling procedure (Kim et al., 2008).

Procedure:
Recombinantly expressed and purified TBP and TFIIB were obtained from the Cramer lab. The functionality of unlabeled and labeled TBP and TFIIB were assessed with an electrophoretic mobility shift assay (EMSA). Initial experiments showed that successful labeling of TBP and TFIIB by labeling accessible lysine residues with NHS-Esters impeded DNA binding of both proteins (data not shown). TFIIB has nine cys-
6.1 Selective chemical labeling of purified proteins

Figure 6.3: Functionality test of labeled TFIIB by using an EMSA. Lane 2, 3, and 5 show DNA shifts induced by binding of TBP (lane 2) or TBP/TFIIB complexes (lane 3 and 5). A large and faint shift is induced by labeled TFIIB (lane 4). Lane 5 shows the incorporation of labeled TFIIB into the DNA protein complex. Free excess DNA can be seen in the lower half of the gel.

Protein residues and crystal structure analysis (PDB 5IY6) showed that three residues are accessible on the protein surface.

TFIIB (65 μM), stored in DTT containing buffer, was labeled in DTT free buffer for 30 min at RT by addition of 150 μM ATTO-655-Maleimide. The labeling reaction was quenched by the addition of 0.5% (v/v) β-Mercaptoethanol, and excess dye was removed by size exclusion chromatography using a Nap5 column. Labeled TFIIB was subjected to an EMSA (chapter 5.1.3). DNA (1 pmol), TBP, and labeled TFIIB were mixed in assembly buffer at a ratio of 1:2:4 (DNA:TBP:TFIIB) and incubated for 35 min at RT. The samples were separated on a native PAGE (4-16%).

Results

Direct labeling of TFIIB with ATTO-655-Maleimide resulted in a loss of 3/4 of initially inserted protein. The labeling of cysteine residues did not impede DNA binding or protein complex formation with TBP (Figure 6.3). The labeled and purified protein was stored in the storage buffer at -20°C. It was further used for protein delivery experiments across the cellular membrane of living HeLa cells (Chapter 6.2). Off-note: TBP also has accessible free cysteine residues; however, the protein precipitated during the labeling process.
6.2 Protein and organic dye delivery across cell membranes

1. Electroporation

Electroporation is a traditional way of transfecting a wide range of different cells, ranging from bacteria to mammalian cells. It achieves high transfection efficiencies in the transport of cargo across the membrane and the number of transfected cells. Electroporation is based on the principle that the cell membrane acts as an electrical capacitor. Subjecting the cellular membrane to high voltage electric fields causes temporary pore formations. This process allows the uptake of small molecules like nucleic acids or the uptake of whole proteins (Shi et al., 2018).

**Procedure**

The used Neon Transfection system (Thermo Fischer) allows for transfection of up to 6 million cells per reaction in a small reaction volume. Apart from the claimed technical superiority, the small transfection volume is particularly interesting, since the used cargos are most often either limited or expensive. 10,000 cells were trypsinized, washed, and resuspended in 100 μl Invitrogen Electroporation Resuspension buffer. The cells were mixed with 1 μl protein solution resulting in final protein concentrations ranging from 1 μM to 1 nM. Finally, the cells were subjected to 2 short electric pulses at 1005 V for 35 ms. The transfected cells were grown on imaging dishes and imaged 6 h after transfection in OptiMEM.

2. Bead loading

Bead loading is a mechanical way to deliver proteins into living cells. Small glass beads are used to rupture the cell membrane mechanically. The molecule or protein of interest can enter the cytosol through the introduced membrane lesions (McNeil and Warder, 1988).

**Procedure:**

One day before protein transfection, 15,000 HeLa cells were seeded on an 8 well Ibidi dish. The cells were washed once with PBS prior to protein bead loading.
Labeled TFIIB was diluted to a final concentration ranging from 1 \( \mu \text{M} \) to 1 nM in 10 \( \mu \text{l} \) HEPES buffer. The protein solution was further dissolved in 90 \( \mu \text{l} \) DMEM and applied to the cells. A small amount of glass beads was added to the cell monolayer and distributed on the surface by tapping the dish 3-6 times. Finally, the cells were incubated for 3h. The cells were washed three times with PBS prior to imaging in OptiMEM.

3. Influx

Influx pinocytic cell-loading reagent (Invitrogen) is based on the osmotic lysis of pinocytic vesicles. First, the cargo is mixed with a hypertonic medium. The cells are incubated with the osmotic buffer reagents, and they take up the cargo via pinocytosis. After transferring the cells into a hypotonic medium, the intracellular vesicles burst and release the trapped protein into the cytosol. It has been shown that the ingested vesicles do not fuse with lysosomes, and neither are lysosomal compartments affected by the procedure (R. D. Park et al., 1988). The Influx of reagent was used as described in the manufacturer’s protocol (Invitrogen).

Procedure

The hypertonic Influx loading reagent was prepared as described in the manufacturer’s protocol and stored at 8 \(^\circ\)C. One day before protein transfection, 15,000 HeLa cells were seeded on an 8 well Ibidi dish. 1 \( \mu \text{M} \) to 1 nM TFIIB was mixed with 40 \( \mu \text{l} \) Influx loading reagent. The protein solution was applied to the cells, followed by an incubation of 10 min. Afterward, the medium was exchanged to 40 \( \mu \text{l} \) hypotonic lysis reagent for 2 min. Finally, the cells were washed with PBS and incubated for another 10 min before imaging in OptiMEM.

4. Pro-Ject

Pierce Protein transfection reagent (Pro-Ject, Thermo) is a cationic lipid-based protein transfection reagent. The cationic reagent complexes with the negatively charged cell surface. The formed complexes fuse directly with the plasma membrane, releasing the cargo, or the cargo is delivered via endosomal uptake, followed by cytosolic release. The protein transfection reagent was used as described in the
manufacter’s protocol.

**Procedure**
The purchased Pro-ject reagent was diluted in 250 µl of methanol, and after shortly vortexing the solution, it was aliquoted into several tubes (1 µl, 2.5 µl, 5 µl). The solution was vacuum dried and stored at -20°C. One day prior to protein transfection, 15,000 HeLa cells were seeded on an 8 well Ibidi dish. The cells were washed with PBS, directly before the addition of the protein solution. Labeled TFIIB was diluted to a final concentration ranging from 5 µM to 5 nM in 10 µl HEPES buffer. The protein solution was added to an epi containing a vacuum dried film of Pro-ject reagent (2.5 µl, 2.5 µl, and 1 µl). After a brief vortexing step (5s) and 15 min incubation, the solution was mixed with 90 µl DMEM and added to the HeLa cells. The cells were incubated for 5h, and they were washed three times with PBS before imaging in OptiMEM.

5. **PULSin**

PULSin (Polyplus) is another chemical reagent to deliver small peptides and proteins into living cells. PULSin contains cationic amphiphile molecules that form non-covalent complexes with the protein of interest. These are internalized via anionic cell adhesion and released into the cytoplasm, where the protein is released. The proteins transfection reagent was used as described in the manufacturer’s protocol.

**Procedure**
One day before protein transfection, 15,000 HeLa cells were seeded on an 8 well Ibidi dish. The cells were washed with PBS directly before the addition of the protein solution. Labeled TFIIB was diluted to a final concentration ranging from 5 µM to 5 nM in 10 µl HEPES buffer. The protein was complexed by the addition of 1.2 µl PULSin. After a brief vortexing step (5s) and 15 min incubation, the solution was mixed with 90 µl DMEM and added to the HeLa cells. The cells were incubated for another 5h, and prior to cell imaging in OptiMEM, cells were washed three times with PBS.
6.2 Protein and organic dye delivery across cell membranes

6.2.1 Interim results

Selective chemical labeling of purified proteins offers a great variety for fluorescent labeling of proteins. Initial labeling experiments of purified TFIIB were promising, and the labeling did not impede with the protein’s functionality (6.1). The second challenge was to deliver the functional protein into living cells. For this, a variety of different delivery methods was tested, ranging from physical transfection methods, like electroporation and bead loading to various chemical transfection systems. Whenever possible, the different cellular transfection methods were optimized, according to the manufacturer’s protocol and the enclosed troubleshooting section, for the cargo delivery into HeLa cells. The control experiments for the cargo delivery with electroporation and PULsin showed promising results (Figure 6.4: Control experiments). The small protein R-Phycocerythrin could be delivered successfully with the lipid-based transfection reagent PULSin and a control plasmid coding for eGFP was successfully delivered via electroporation.

However, TFIIB could not be successfully delivered into HeLa cells (Figure 6.4). Neither of the tested delivery approaches was able to deliver labeled and functional TFIIB across the cellular membrane. The experiments also showed that the main barrier was the plasma membrane and not the nuclear membrane since no cytosolic localization could be detected. Neither the variation of the protein concentration nor changes in the delivery protocols showed any promising effects. Overall all delivery methods had severe effects on the cell viability. The viability rate ranged depending on the delivery method from 70-20% (subjective values determined by eye - not quantified). The cells, transfected with the lipid-based methods, Pro-Ject and PULSin, showed compared with the other methods enclosed strong fluorescent compartments. Those structures might be vesicular compartments enclosing labeled proteins. The bead loaded cells had the least viability rate, and the remaining cells showed strong fluorescent signals similar to cells close to apoptosis.
6 Fluorescent labeling of proteins in living cells

6.2.2 MscL membrane channels

Mechanosensitive channels present a biological way of delivering small cargos across the lipid membrane barrier. Among those, bacterial channels are known for their large pore size, making them attractive for the delivery of larger cargos up to small peptides. Particularly interesting is the membrane channel MscL. The MscL channel has a pore size with a diameter of around 25 Å, and a substitution of a cysteine to a glycine at position 26 enables the gating in a charge induced manner, through the covalent binding of MTSET (2-(Trimethylammonium)ethyl methanethiosulfonate, bromide) (Bartlett et al., 2006). The addition of the reducing agent DTT reverses this effect, thereby closing the membrane channel. It has been shown that the MscL pore complex allows the delivery of fluorescent dyes up to a size of 10 kDa as well as small peptides with a size of 6.5 kDa (phalloidin) (Doerner et al., 2012). During channel opening, the cell viability depends on the duration of MTSET.
6.2 Protein and organic dye delivery across cell membranes

Figure 6.5: Biological principle of MscL (G26C) membrane channels. Pore opening upon binding of MTEST facilitates the delivery of the non-cell-permeable tris-NTA dye.

During the adaptation process, the following parameters were varied and optimized: MscL expression level, time of MTSET treatment and the number of washing steps.

**Procedure**

Initial experiments were performed using cells with transient expression of MscL channels with jetPrime. The plasmid contained an IRES sequence, thereby coexpressing the MscL channels and eGFP as a reporter. For the experiments, 100,000 HeLa cells were seeded on a 3,5 mm Ibidi dish and transiently transfected with jetPrime using 1 µg of plasmid according to the manufacturer’s protocol. 2 µg plasmid decreased the cell viability, whereas 100-500 ng plasmid decreased the delivery efficiency. Transiently transfected cells were used 36 h after transfection, and positive cells were visually identified through the expression of eGFP.

Final experiments were done with a cell line that stably expressed MscL. Lentiviral transduction of HeLa was performed with 1.5 ml lentivirus from pLentiCMVTRET3G-
Puro-DEST-Kozak-MscL-G26C as described in chapter 5.2.1. The transfected cells were initially selected with 4 µg puromycin for one day, followed by growth in DMEM supplemented with 2 µg. Preliminary growth experiments with HeLa cells showed that their lethal puromycin dose is 2 µg. The selected cells were further virally transduced with a plasmid containing the TET transactivator necessary for the induction of the initially transduced TET controlled MscL expression system (pLenti CMV rtTA3 Blast, 1.5 ml virus, Agarwal, 2019). For final experiments, the cellular expression of MscL-G16C was induced by the addition of 250 ng/ml Doxycyclin to the growth medium followed by an incubation time of 16h.

**Opening and closing of membrane channels**

Preparations:

- seed 100,00 cells on a 3.5 mm Ibidi or Delta-T dish
- induce expression with 250 ng/ml Doxycyclin 16h
- dissolve MTSET (fresh)
- dissolve ATP and add to Delivery solution (fresh)

Step by step:

1. 200 µl Ringer solution 1 min
2. 200 µl Delivery solution 1 min
3. 200 µl Delivery solution + 1 mM MTSET + cargo 1.5 - 5 min
4. 200 µl Delivery solution (wash step)
5. 200 µl DMEM 3 min
6. 200 µl Delivery solution + 1 mM DTT 5 min
7. wash with PBS
8. incubate in DMEM 1h or over night
9. imaging in OptiMEM
The non-cell-permeable dye Alexa646-NHS Ester was used as a test cargo to assess the usability of the MscL system. The cells were prepared as described above, and 5 µM of dye was used as cargo. 60 minutes after conducting the cells to the first round of channel opening procedure, the cells were subjected to a second round, to remove the free dye. The incubation time with MTSET was 1.5 min for both opening events.

Figure 6.6: Delivery of fluorescent dyes through MscL channels. (top) HeLa cells transiently expressing MscL channels identified by the expression of eGFP. (bottom) Cellular uptake of Alexa647-NHS Ester. The time points indicate the time after cargo delivery. The second opening was performed 1h after the first opening. The scale bar is 10 µm

**Results**

A general observation was that the cells showed an abnormal roundish morphology, directly and 1h after the opening procedure, for the delivery with and without Alexa647. After 12h, the cell morphology of the control cells (no cargo) and the cells subjected to a second opening procedure to remove the free dye was generally well, whereas the cells treated with Alexa647-NHS showed a poor survival rate (below 10%). The delivery procedure was highly successful, as nearly all cells expressing the MscL channels took up the free dye in varying amounts. The dye persisted inside the cells for at least 12h (Figure 6.6), and the second channel opening resulted in significant removal of Alexa647 from the cells, and complete loss of the remaining dye molecules could be observed 12 h later (Figure 6.6).
Labeling in living cells with the non cell permeable SNAP-Tag ligand

To confirm the labeling capability of dyes delivered via MscL channels, the non-cell-permeable dye BG-Alexa647 was used to label transiently expressed SNAP-CTCF fusion protein according to the protocol described above (2 \( \mu \)g pLenti CMVTRE3G Puro Dest SNAP Tag CTCFwt Flag tag, Agarwal, 2019). 0.1 \( \mu \)M of BG-Alexa647 was used as cargo, and the treatment time with MTSET was 3 min.

![Figure 6.7: Labeling of SNAP-CTCF with non-cell-permeable dye BG-Alexa647. (a-b) HeLa cells transiently expressing SNAP tagged CTCF labeled with BG-Alexa647. BG-Alexa647 was delivered through stably expressed MscL membrane channels. (c-d) Control cells not treated with MTSET. The scale bar is 10 \( \mu \)m](image)

Results

The cells did take up BG-Alexa647 and showed specific staining of SNAP-CTCF inside the nucleus. Therefore, no second opening procedure, to remove unbound excess dye molecules, has been introduced. However this readout is not sensitive enough to detect small amounts of free excess dye molecules. The labeling intensities varied from cell to cell (Figure 6.7 (a-b)). Control cells were incubated with BG-Alexa647, to confirm the non-cell-permeable nature of the SNAP-Tag ligand. The control cells that were not treated with MTSET showed no fluorescent signal (Figure 6.7 (c-d)).
Summary
Initial results on transiently transfected cells proved the usability of the MscL membrane channel system for the delivery of small non-cell-permeable substrates. Bigger cargos such as nanobodies (GFP-VHH-Alexa647, Chromotek) could not be successfully delivered (data not shown). Finally, the MscL system could be stably integrated as an inducible expression system into a HeLa cell line and the adaption of the delivery procedure resulting in successful labeling of SNAP-tag fusion proteins with a non-cell-permeable BG-ligand. Further experiments were done with the non-cell-permeable tris-NTA dyes (Chapter 6.2.3).
6 Fluorescent labeling of proteins in living cells

6.2.3 His-Tag labeling

Labeling of oligo histidine tags presents another system for fluorescent labeling of proteins in living cells. Like the direct labeling of cysteine or lysine residues on protein surfaces, the integration of a small His-tag with 6-10 repeats introduces, with a size of approximately 1 kDa, only small sterical interference. This makes it especially interesting for studies of protein complexes. In contrast to cysteine or lysine labeling, where most often protein engineering is necessary for stoichiometric labeling, the His-tag allows for site-specific labeling with a 1:1 ratio by simple genetic expression of the tag at the N- or C-terminus of the protein. A classical metal chelating agent is nitrilotriacetic acid (NTA). NTA can be complexed with metal ions (e.g., Ni$^{2+}$), and those can form two coordinating bonds to His-tagged protein complexes. NTAs have been widely used in biological applications like protein purification, immobilization, and tethering (Brausemann et al., 2016; Dietrich et al., 1995; E. L. Schmid et al., 1997). Typical metal: NTA complexes can bind with a low affinity of 10 μM, whereas three NTA moieties (tris-NTA) can bind 6 His-tags with high affinities in the nanomolar range (Lata et al., 2005). An additional increase in affinity up to 0.1 nM can be achieved by adding a 10 His-tag to the protein of interest (Lata et al., 2006).

While life cell imaging with tris-NTA dyes of different proteins in the cytosol (TAP1) as well as membrane proteins (Lamin A) and labeling in the nucleus (H2B) has been shown (Kollmannsperger et al., 2016), His-tag labeling comes with several disadvantages. First, tris-NTA-dyes are non-cell-permeable, which makes them unattractive for live-cell imaging. Second, due to the noncovalent nature of the dye bond, there might be a continuous exchange of fluorophores. This can be advantageous for STORM imaging because bleached dyes are continuously exchanged; however, it may hinder single molecule tracking experiments. Third, metal ions in close proximity have a negative effect on the photophysics of the attached fluorophores. This quenching effect of up to 80% can be lessened by an increase in linker length between NTA and the dye. However, it still presents a severe effect for single particle experiments, that rely on high photon output, while working with minimal laser intensities (Lata et al., 2006).
All used tris-NTA-dyes were synthesized according to Lata et al., 2005 by the Tampe group in Frankfurt. Modifications of proteins by adding tags can alter their binding behavior and their functionality. Therefore the protein activity of tris-NTA labeled TBP and TFIIB was assessed in vitro using an EMSA (Chapter 5.1.3).

**Procedure:**
Purified 6His-tagged TBP and TFIIB (Cramer Lab, Göttingen) were labeled with tris-NTA-ATTO655. TBP and TFIIB were incubated in their storage buffers with tris-NTA-ATTO655 in equimolar amounts, for 20 min. Afterward, the labeled proteins were directly used for an EMSA, without an additional purification step (Chapter 5.1.3). TMR labeled dsDNA (1 pmol) was mixed with TBP and TFIIB at a ratio of 1:2:4 (dsDNA:TBP:TFIIB) and incubated in assembly buffer (+10 mM DTT fresh). The DNA protein complex was incubated for 35 min at RT. The samples were separated on a native PAGE (4-16%) for 40 min at 120V and 4°C (pre-run of gel: 30 min, 120V, 4°C).

**Results**
Equimolar labeling of TBP and TFIIB resulted in a very poor labeling efficiency, as indicated by the free dye in figure 6.8. However, the noncovalent nature of the labeling can, in the context of the separation in an electric field, change the reaction equilibrium, by locally withdrawing free dye molecules. Apart from the small labeling efficiency, both proteins form a stable complex with DNA, and detectable amounts of labeled TFIIB are integrated into the TBP/TFIIB complex (Figure 6.8).

**Labeling in living cells**
The first issue, the non-cell-permeability of tris-NTA dyes, was approached by using the MscL membrane channel system (MscL section in chapter 6.2.2). With a size of roughly 1 kDa (Kollmannsperger et al., 2016), tris-NTA dyes present an optimal cargo for this delivery system. HeLa cells transiently expressing N-terminal 10His-Halo constructs (Histone H4, TBP, and TFIIB) were labeled with Halo-TMR ligand, as described in chapter 5.3 and were afterward prepared as described in chapter 6.2.2. Membrane channels were activated with MTSET in the presence of 1 μM of tris-NTA-SiR for 2 min.
6 Fluorescent labeling of proteins in living cells

Figure 6.8: Functionality test of tris-NTA-ATTO655 labeled TBP and TFIIB. The addition of TMR labeled DNA is indicated with green markers, and the addition of tris-NTA-ATTO655 labeled proteins is indicated with red markers. Black markers indicate the addition of unlabeled 6His-tagged proteins. From left to right, the depicted signal of labeled DNA (left), labeled protein (middle), and the composite image (right) are shown. The addition of TBP generates an upshift of free DNA (lane 2), and most of the DNA is complexed in the TBP/TFIIB complex (lane 3). The addition of labeled TBP resulted in no detectable upshift (lane 4, 6). Labeled and unlabeled TFIIB stabilized the protein complex and resulted in a DNA upshift (lane 5, 7). Labeled TFIIB is directly detected in the complex with unlabeled TBP (lane 7). The bottom half shows noncomplexed DNA (lane 1-7) and free dye molecules (lane 4-7).

Results

Halo-Tag labeling of TBP and TFIIB resulted in an expected and confined nuclear labeling, whereas the N-terminally Halo tagged histone H4 could not be labeled (Figure 6.9 (middle row)). Tris-NTA labeling of H4, TBP, and TFIIB succeeded in cells expressing MscL channels. However, a strong cytosolic background signal could be detected for all constructs (Figure 6.9 (bottom row)). The reduction of un-bound dye molecules could be assessed in the future by adding a second channel opening step, as described in chapter 6.2.2. In general, the amount of delivered dye did not directly correlate with the expression level of the MscL channels (e.g., middle lane top image (eGFP=MscL) compared to bottom image (SiR=trisNTA), Figure 6.9).

Single particle tracking with tris-NTA dyes: ATTO655

Functional and successful labeling of proteins inside the nucleus with non-permeable
tris-NTA dyes (Figure 6.8 and 6.9) paved the way for single molecule tracking experiments. The first experiments were conducted with 10His-HaloTFIIB and labeling with 136 nM tris-NTA-ATTO655 (MTSET, 2 min). Addition of a second opening step (MTSET, 1 min) was included to remove unbound dye molecules as successfully shown for Alexa647-NHS (Figure 6.6).

The initial results with tris-NTA-Atto655, revealed strong photoblinking properties of Atto655 (Figure 6.10). This observation led to a change of the dye molecule. All further experiments were performed using tris-NTA-SiR and tris-NTA-Cy3B ligands.
Figure 6.10: Fluorescent intensity analysis of a single 10His-Halo-TFIIB molecule labeled with tris-NTA-Atto655. The intensity timeline indicates multiple photoblinking events of tris-NTA-Atto655.

In general, the second channel opening resulted in a drastic decrease in the cell’s viability in this experimental context. Therefore further test and characterization experiments were done without an additional removal step.

**Single particle tracking with tris-NTA dyes: SiR and Cy3B**

Both tris-NTA-SiR and tris-NTA-Cy3B were extensively tested for the use in live cell single molecule tracking experiments. Stable cells, expressing the MscL channel system, were transiently transfected with either 10His-Halo-TBP or TFIIB fusion constructs (Figure 6.9). For single molecule imaging, cells were sequentially labeled with Halo-TMR, followed by labeling with 1 \( \mu \text{M} \) tris-NTA-SiR/Cy3B (MTSET, 3 min).

Figure 6.11 shows the successful costaining of 10His-Halo-TBP with the Halo-TMR ligand (a) and tris-NTA-SiR ligand (b) in HILO illumination mode. The expression level and the resulting labeling density of the tris-dye varied in both channels from cell to cell (Figure 6.11 (b and c)). Comparing the transfected cells expressing the fusion constructs, with the tris-NTA-SiR labeled control cells showed that a strong unspecific, nuclear signal is present (Figure 6.11 (d)). This unspecific signal was absent in the TMR channel (data not shown). Considering the observed unspecificity, two remaining questions need to be addressed. First, does the unspecific
binding interfere with the single particle tracking analysis, and second, does the specific, but noncovalent bound dye, bind long enough for the planned single particle tracking experiments?

Figure 6.11: HILO imaging of 10His-Halo-TBP. (a) Fluorescence signal of the Halo-TMR ligand. (b) Fluorescence signal of the tris-NTA-SiR ligand. (c) Single molecule labeling of 10His-Halo-TBP. Standard deviation of the fluorescent signal of tris-NTA-SiR (300 frames). (d) Standard deviation of the fluorescent signal of control cells stained with tris-NTA-SiR ligand (300 frames). The nucleus is indicated with dashed lines in white. The scale bar is (a-b) 10 μm and (c-d) 2 μm.

**Single particle tracking with tris-NTA dyes: Unspecificity**

Imidazole was used to reduce the unspecificity of the used tris-NTA dyes. The binding affinity of tris-NTA molecules to unspecific or specific cellular binding sites might be reduced by the simultaneous application of imidazole (Lata et al., 2006). The adaption of the MscL protocol made it possible to include an additional delivery step for the delivery of imidazole. Therefore, cells were stained with 10 nM tris-NTA-SiR/Cy3B (2.5 min MTSET), followed by an additional delivery step with 100 mM Imidazole (2.5 min MTSET).

The binding time analysis of control cells stained with tris-NTA-SiR and Cy3B showed that the addition of imidazole had no reducing effect on the unspecificity of the tris-NTA dyes (Figure 6.12 (a)). Furthermore, the addition of imidazole did not result in a clear separation of the binding time histograms of specific binding events, binding of labeled TBP and TFIIB, compared to the unspecific events (Figure 6.12 (b)). The experiments also showed that the unspecificity is not the sole effect of either the SiR or the Cy3B tris-NTA ligand. However, time-lapse microscopy experiments revealed that the unspecificity of tris-NTA dyes contributes solely to a fraction of binding events ranging in the milliseconds and the lower second range (Figure 6.13 (a)).
6 Fluorescent labeling of proteins in living cells

Figure 6.12: Comparison of specific and unspecific binding of tris-NTA. (a) Binding probability histogram of tris-NTA dyes in control cells (10 nM) with and without imidazole. (b) Binding probability histogram of labeled 10His-Halo-TBP and TFIIIB labeled with tris-NTA-Cy3B, with additional imidazole, in comparison to control cells. Data includes: #molecules >10,000, #cells > 10.

Single particle tracking with tris-NTA dyes: Specificity

To confirm that there is indeed specific labeling on the single molecule level, 10His-Halo-TBP constructs were labeled with Halo-TMR ligand and tris-NTA-SiR. Cells were stained with Halo-TMR as described in chapter 5.3, and tris-NTA-SiR labeling was done with 1 μM, as described in chapter 6.2.2. One cell nucleus was initially imaged in the SiR channel, followed by imaging of the TMR signal. Two different time-lapse conditions were used. First, continuous illumination (50 ms) to compare the fast kinetics, and second, time-lapse illumination microscopy with a dark time of 1 s, to compare the slow kinetics (= long-lasting binding events). In general, an increase of the dye concentration from 10 nM to 1 μM did not drastically increase the rate of the unspecific binding events (compare Figure 6.12 (a) and Figure 6.13 (a)). However, it increased the labeling frequency of tris-NTA labeled cells, expressing the 10His-Halo constructs, detected by costaining of the Halo-TMR ligand. As already mentioned in the previous section, the unspecific binding events showed a probability distribution in the milliseconds regime that was nearly identical to the probability distribution for 10-His-Halo-TBP stained by either TMR-Halo or tris-NTA-SiR ligand (Figure 6.13 (a)). Additionally, the comparison of tris-NTA labeled TBP molecules to Halo-TMR labeled ones, shows that the probability of detecting long
binding events decreased significantly faster for tris-NTA labeled molecules. This effect most likely arises from an additional kinetic off-rate introduced by the noncovalent nature of the tris-NTA labeling (Figure 6.13 (b)). This behaviour also excludes the use of tris-NTA dyes for the visualization of long bound molecules.

Figure 6.13: Comparison of specific and unspecific binding of tris-NTA. (a) Binding probability histogram of TBP labeled with Halo-TMR and tris-NTA-SiR, compared to tris-NTA dyes in control cells (1 μM tris-NTA-SiR, continuous illumination). (b) Binding probability histogram of TBP labeled with Halo-TMR and tris-NTA-SiR, compared to tris-NTA dyes in control cells (1 μM tris-NTA-SiR, 1 s time-lapse). Inset shows a zoom-in of (b). Data includes: #molecules >10,000, #cells > 10.

Quenching of the SiR dye by tris-NTA
Metal ions in close proximity to the electron system of organic dye molecules have a well known negative effect on the photophysics of the attached fluorophores (Lata et al., 2006). Intensity analysis of 10His-Halo-TBP labeled constructs labeled either with the Halo-SiR ligand, or the tris-NTA-SiR ligand showed that the tris-NTA moiety has indeed a strong negative quenching effect on the dye molecule (Figure 6.14). The observed quenching effect might be reduced by increasing the linker length between the dye and the tris-NTA moiety or by direct coupling of photostabilizers like triplet state quenchers to the fluorophore (Glembockyte et al., 2018).

Summary
Initial experiments on purified 6His-TBP and TFIIB showed the effective labeling of TFIIB with tris-NTA dyes and confirmed the protein’s intact biological functionality after the labeling process. Further experiments demonstrated the ability of the MscL
channel system to successfully deliver tris-NTA dyes across cellular membranes of living cells, and the fluorescent labeling experiments confirmed the specific labeling ability of tris-NTA dyes in the cellular environment, which was previously described by Kollmannsperger et al., 2016. However, single molecule experiments in the nucleus revealed that the tris-NTA dyes exhibit strong unspecific binding events in the millisecond's range. In addition to that, the long term tracking of tris-NTA labeled proteins is drastically limited by either an additional off rate caused by the noncovalent binding of the dye molecule to the tag, or by an a fast bleaching rate induced by the NTA moiety, or both. These combined characteristics highly restrict the application of tris-NTA dyes for single molecule tracking experiments.
6.2 Protein and organic dye delivery across cell membranes

6.2.4 TMP-Tag labeling

The TMP-Tag is based on the eDHFR (*E.coli*-Dihydrofolatreduktase), and labeling is facilitated by the ligand TMP (2,4-diamino-5-(3,4,5-trimethoxybenzyl) pyrimidine). Similar to the Halo-Tag, the TMP-Tag can be expressed as a fusion protein in living cells. Compared to GFP (27 kDa) and the Halo-Tag (33 kDa), the TMP-Tag has 18 kDa, a relatively small size, making the fusion protein less prone to any sterical hindrances in the biological environment. Another advantage is the high affinity of the TMP-ligand (1 nM) to eDHFR over the mammalian version of the DHFR (4 μM), which ensures specificity in human samples (Miller et al., 2005).

Initial experiments with a non-covalent version of the TMP-Tag already highlighted its use for live-cell imaging (Miller et al., 2005). The covalent version of the TMP, generated by the engineering of the proximity induced reactivity of the eDHFR, led to a cysteine mutant (L28C) with a slow *in vitro* reaction half-life of 50 min (Gallagher et al., 2009). The tag properties were further improved by optimizing the linker, making it more electrophilic and increasing the reactivity with the nucleophilic cysteine located just outside the TMP binding pocket. This engineering process cumulated in the publication of the "second generation covalent TMP-Tag" (Z. Chen et al., 2012). Live-cell super-resolution microscopy using dSTORM was used to successfully label and image the histone H2B with a resolution of 20 nm (TMP-ATTO655) (Wombacher et al., 2010), and the tag was recently also used for *in vitro* studies (Q. Yu et al., 2018).

Unfortunately, till now, no fluorescent variant of the TMP-Tag is commercially available, and as a non-chemical lab, it is not feasible to follow the published synthesis protocol (Jing and Cornish, 2013). However, in cooperation with Kai Johnsson’s lab, we received a cell-permeable SiR version of the first generation TMP-Tag as used in Q. Yu et al., 2018.

**Procedure**

The genetic DNA sequence coding for the eDHFR(G26C) was synthesized by GeneART and cloned N-terminally of the protein TBP into the LVTO backbone using EcoRI and XbaI restriction sites creating the plasmid LVTO-TMP-TBP. Additional plasmids containing only the TMP tag and an NLS version, using the SV40 NLS sequence, were created as well (LVTO-TMP and LVTO-NLS-TMP).
HeLa cells were transiently transfected with the TMP-TBP fusion construct using jetPrime. On the following day, the transfected cells were labeled with the TMP-SiR-ligand. The cells were incubated with 1 $\mu$M TMP-SiR-ligand for 1h. Afterward, the cells were subjected to multiple wash steps with PBS, followed by incubation of 15 min and another 45 min before imaging in OptiMEM. The labeled transfected and untransfected control cells were imaged using spinning disc microscopy (Figure 6.15).

Figure 6.15: Fluorescent imaging of HeLa cells expressing a TMP-TBP fusion construct labeled with TMP-SiR-ligand. (a) Cell overexpressing TMP-TBP, showing accumulation of cytosolic TBP (LUT:100-700). (b-c) Solely nuclear localization of TMP-TBP with varying expression levels (LUT: (b) 100-400, and (c) 100-200). (d) Untransfected labeled control cell showing spotty cytosolic stress particles that may contain unbound TMP-SiR-ligand (LUT:100-200). The scale bar is 10 $\mu$m.
6.2 Protein and organic dye delivery across cell membranes

Results
Initial experiments with TBP confirmed the reported live-cell imaging properties of the TMP-Tag. The fluorescent signal was mostly confined to the nucleus (Figure 6.15 (b-c)). A nonlocalized cytosolic signal was only observed at higher protein expression levels (Figure 6.15 (a)). A distributed spotty cytosolic signal was observed (Figure 6.15 (a-d)), that may arise due to transfection stress or it may represent confined particles including unbound TMP-SiR ligand, that could not be removed by the intense washing procedure.

Single particle tracking with TMP-Tag
The next step was the adaptation of the TMP-Tag for nuclear single molecule tracking in living cells, as already suggested in Gallagher et al., 2009. Therefore, HeLa cells were transiently transfected with TMP-TBP fusion constructs using jetPrime. The following day, the cells were incubated with 100 nM of TMP-SiR ligand for 10 min in DMEM. Previous tests showed that cells labeled with 1 nM and 10 nM TMP-ligand did not exhibit the required labeling efficiency. However, the shortening of the incubation time from 1h down to 10 min did not drastically decrease the labeling efficiency. Afterward, the cells were subjected to two washing steps with PBS separated by a 1 h incubation with DMEM to wash out unbound dye molecules. In the end, the cells were imaged in OptiMEM.

Results
As already shown in Figure 6.15, the labeled TBP molecules were confined and did localize to the nucleus (Figure 6.16 (a)). Experiments on labeled control cells, expressing the non-fused TMP tag (TMP-control) resulted in a small amount of molecules localizing and binding inside the nucleus (Figure 6.16 (b)). The quantification of detected molecules of multiple nuclei, confirms the visual conclusion drawn from figure 6.16 (a-b), that only a small fraction of TMP-control molecules is binding inside the nucleus (Figure 6.16 (c)).

The relevant question arising from the unspecifically bound molecules is still: How long do they bind and does the binding interfere with the data analysis? Track length analysis showed that most unspecific binding events (TMP-control) are smaller than a few hundred milliseconds, and only a minor fraction of binding events last around 1s. In comparison, TMP-TBP molecules show binding events that last a
Figure 6.16: Comparison of TMP-TBP and TMP-Tag (TMP-control). HeLa cells transiently expressing TMP-TBP (a) and TMP-control (b). Standard deviation of 300 frames of a continuous movies (50 ms) in HILO illumination mode. The nucleus is indicated in white with dashed lines and the scale bar is 2 μm. (c) Quantitative analysis of total number of detected molecules: TMP-TBP = 471 ± 260 (mean ± s.d.), TMP-control = 25 ± 18 (mean ± s.d.). Data include: # movies: TMP-TBP = 11, TMP-control = 9, # cells: TMP-TBP = 6, TMP-control = 8).

couple of seconds (Figure 6.17.). So far, the TMP-Tag seems similar to the HALO-Tag (Presman et al., 2017) to fulfill all relevant properties necessary for usage as a single molecule tag that can be used for tracking experiments inside the nucleus.

A distinct drawback of the used SiR-ligand was already detectable, during image acquisition of the TMP tagged TBP. SiR labeled fusion proteins did bleach faster than comparable SiR labeled Halo constructs. Detailed analysis and comparison of TMP and Halo tagged TBP showed that both constructs could be detected for several seconds. However, the TMP variant could be tracked significantly less long (Figure 6.18 (a)).

To distinguish bleaching from transition into long-lived transient dark states (e.g., triplet state), SiR labeled TMP-TBP expressing cells were imaged for 30 s, followed by a long dark time. The comparison of SiR labeled TMP and Halo tagged TBP
Figure 6.17: Probability distribution histogram of detected TMP-control and TMP-TBP molecules. (inset) Probability distribution with a cut off at p = 0.02 (1 frame = 50 ms).

molecules already shows that the TMP construct possesses faster bleaching kinetics than the Halo construct (Figure 6.18 (b)). Additionally, one can observe a recovery of the TMP version’s fluorescence signal to an almost equal height compared to the starting value. This dynamic behavior could not be observed for the Halo version (personal experience, no data acquired).

**Summary**

For obvious reasons, genetically encoded tags make experimental life much easier. The TMP tag combines the advantages of the Halo-Tag with its smaller size. The TMP-Tag was used as a fusion protein construct, and TMP-TBP could be successfully labeled with the first generation cell-permeable TMP-SiR ligand. Excess dye molecules were removed, similar to Halo-Tag ligands, through extensive washing steps, after the labeling process. Live-cell fluorescence single molecule tracking experiments showed that the TMP-Tag similar to the Halo-Tag showed minimal unspecific binding inside the nucleus. However, so far, the used TMP-SiR ligand showed
Figure 6.18: Tracklength and bleaching analysis of SiR labeled molecules. (a) Histogram of the tracklength of TMP-TBP molecules compared to Halo-TBP molecules. The data were recorded with an integration time of 50 ms and include # molecules: TMP-TBP = 15857, Halo-TBP = 11034, # movies: TMP-TBP = 11, Halo-TBP = 14, # cells: TMP-TBP = 6, Halo-TBP = 8). (b) Mean intensity analysis of two nuclei (100 $\mu m^2$) expressing either TMP-TBP or Halo-TBP.

fast and extensive transitions into a long-lived dark state. In contrast to the Halo-SiR ligand, the weak photophysical property of the TMP-ligand, limits the usability of the tag, in single molecule tracking experiments, dramatically. The observed phenomena are probably caused by the chemical characteristics of the eDHRF, and a change in the linker (e.g., lengthening) or a change of the attached dye molecule might lead to a more photostable ligand. The addition of a triplet state quencher might also improve the dye characteristics (Glembockyte et al., 2018). However, the described characteristics of the TMP-SiR ligand could make this system a candidate for STORM experiments.
6.2 Protein and organic dye delivery across cell membranes

6.2.5 Discussion and outlook

Two orthogonal live-cell labeling methods are required to perform dual-color single molecule tracking experiments. Apart from fluorescent proteins (Gebhardt et al., 2013), so far, only the Halo-Tag (Los et al., 2008) is a commonly accepted labeling method for tracking experiments. (Presman et al., 2017). Many challenges have to be faced in order to use a live-cell fluorescent labeling system. The system must possess minimal disturbance, high labeling efficiency with bright and photo-stable dyes, cell permeability, and a defined ratio of dye per protein (X. Chen and Wu, 2016). However, the main challenge of using a labeling technique for single molecule tracking remains the tag’s unspecificity in the cytosol or the nuclear compartment. Here, in addition to the Halo-Tag, several labeling approaches were tested for their usability for nuclear live cell single molecule tracking experiments.

Selective chemical labeling of purified proteins and protein delivery

First of all, direct chemical labeling of purified proteins was done by labeling free, accessible cysteine residues. Due to its small size, this quantitative labeling approach promised a minimal disturbance of the protein function and a final versatile dye exchangeability (Vinogradova et al., 2015). However, the main challenge was to deliver the labeled and functional proteins across the cellular membrane into living cells. As shown in chapter 6.1, cysteine labeling of purified TFIIB did not disturb the protein’s functionality. TFIIB-Alexa-647 did complex with TBP and a DNA promoter mimic. Protein delivery is highly protein-specific and depends on the size of the protein and its surface charge. Therefore, many different delivery methods ranging from electroporation to various lipid-based systems were tested, to deliver active fluorescently labeled TFIIB across the cellular membrane. Apart from the conclusion that no active protein could be delivered, the different methods showed individual drawbacks that could be addressed in future experiments.
6 Fluorescent labeling of proteins in living cells

**Cell viability**
The relatively simple bead loading approach was the most harmful method and resulted in very low cell viability. Next, in the context of cell viability, ranked electroporation, with the additional drawback that the cells had to reattach after the delivery procedure. The different chemical-based approaches also affected cell viability, but they were the least harmful ones. Cell viability could be improved by the addition of 20% FBS after the delivery procedure or special coating of the cell culture or imaging dishes (Hung and Young, 2006).

**Endosomal escape**
A significant challenge of protein delivery of lipid-based delivery approaches remains the cargo’s effective endosomal release into the cytosol (Shete et al., 2014). Fluorescent inclusions could be detected after lipid-based protein delivery, hinting at a defective endosomal release. Many different strategies could help to improve the endosomal escape of bioactive reagents or proteins. For example, receptor-mediated endocytosis can be impaired by decreasing the temperature during the delivery process (S. L. Schmid and Carter, 1990). The use of cell-penetrating peptides, originating from viruses (Y. Li, 2015), or other synthetic reagents like chloroquine, that raises the endosomal pH and thereby favors vesicle rupture (Lee and Tannock, 2006), could also significantly improve endosomal escape.

The delivery of bioactive molecules is a rapidly growing field that is particularly interesting for the pharmaceutical industry. Therefore various approaches that promise protein uptake are currently developed. Apart from the classical single cell microinjection method, that is highly effective, however time-consuming (Kovanen and Kallio, 2004), other mechanical delivery methods including cell squeezing (Kollmannsperger et al., 2016), or growing cells on protein-coated nanowires (Shalek et al., 2010) are currently developed. However, mechanical delivery methods mostly impose cellular stress. Therefore, biological delivery methods using bacterially derived membrane channels (Doerner et al., 2012), or protein toxins as the delivery vehicle (Beitzinger et al., 2012; Fahrer et al., 2013) might be more attractive tools.
His-Tag Labeling

After abandoning the direct chemical labeling approach, due to unresolved issues in protein delivery, other live-cell compatible tagging methods were investigated for nuclear single molecule tracking experiments. There are several metal chelation based peptide tags. Most commonly, this technique is used for affinity chromatography in protein purification. Based on this approach, apart from the tetra aspartate tags (Ojida et al., 2006), a His-Tag is the most commonly used tag. Similar to direct chemical labeling, a His-Tag is very small and yields high labeling efficiencies, if complexed with tris-NTA dyes (Lata et al., 2006). The first experiments with purified 6His-TFIIB showed that the labeled protein retains its biological activity, the minimal complex formation with TBP during PIC assembly. Unfortunately, the initially published cell squeezing approach, to deliver non-cell-permeable tris-NTA dye conjugates across cellular membranes was unavailable. However, the successful establishment of the MscL membrane channel system solved this experimental challenge (Doerner et al., 2012).

The first promising experimental results proved the general labeling ability of this approach. However, single molecule experiments revealed a prominent fraction of free dye molecules that show second-long nuclear binding events that would interfere with binding-time analysis and make short binding events in the range of up to 1-2 seconds indistinguishable from unspecific ones. On the other hand, tris-NTA dyes form noncovalent, high-affinity bonds with the His-Tag. The characterization of long bound TBP molecules labeled with tris-NTA-dyes showed that the bond’s noncovalent nature introduces an additional kinetic off rate, that impedes the use of the His-tag for the visualization of long binding events. One improvement to solve the present issue and improve its usability for single molecule tracking experiments might present the use of affinity maturated hexa-NTA dyes in combination with 12 histidine repeats (Gatterdam et al., 2018). The additional attachment of triplet state quenchers to the ligand (Glembockyte et al., 2018), could generate high-affinity ligands with bright, photostable dyes. In general, tris-NTA-dyes represents an interesting technique for STORM measurements, since the noncovalent, high-affinity nature of the bond imitates the necessary blinking events. It could also be used for general fluorescent protein labeling experiments. Especially for long term experiments, since the stochastic noncovalent nature of the binding events would compensate the photobleaching process.
6 Fluorescent labeling of proteins in living cells

**TMP-Tag**

The TMP-Tag combines the Halo-Tag advantages, like genetic encoding, specific, rapid, and covalent labeling in live cells under no-wash conditions with cell-permeable organic dyes and minimal biological disturbance, by occupying only half the space than the Halo-Tag. Initial experiments confirmed the reported live cell compatibility and the adequate labeling of nuclear proteins (Miller et al., 2005). The reported slow reaction half time of the first generation TMP-ligand (Miller et al., 2005), did, in the here performed, not present any major limitations during the live cell labeling procedure. In the here established procedure, the slow kinetics were potentially compensated by the excess of the applied dye molecules. However, this made extensive washing steps a prerequisite for single molecule tracking experiments. Further experimental improvements might be the implementation of the second generation TMP-ligands, that showed superior reaction half times (Gallagher et al., 2009).

Single molecule localization experiments showed that the TMP-Tag shows minimal amounts of unspecific binding events in the nucleus, and in addition to that, the detected molecules exhibited binding events that only ranged in the lower seconds regime. Both characteristics are important prerequisites for the use in single molecule tracking experiments.

So far, the TMP-Tag, combined with the Halo-Tag, exhibits an optimal choice for orthogonal fluorescence labeling in single molecule tracking experiments. However, the used TMP-SiR ligand showed surprisingly bad photo characteristics. The pursued experiments showed that in the specific context of the TMP-Tag, the SiR dye exhibits fast photobleaching kinetics and a high probability of transitions into long-lived dark states. These observed kinetics preclude its use in the selected context. Further TMP-ligand engineering as described in Glembockyte et al., 2018 is necessary to improve the dye properties.
Other live cell compatible fluorescent labeling approaches
As mentioned above, there is a high need for live-cell compatible orthogonal labeling techniques and many groups have recently investigated new labeling approaches.

FlAsH - ReAsH
Like the described and tested His-Tag, an orthogonal fluorescent labeling system has been described, which offers dual-color fluorescent microscopy with minimal biological perturbation. The system is based on the direct labeling of tetracysteine motives (CCXXXCC) and uses its high affinity to biarsenic probes (Martin et al., 2005). Two different cell-permeable versions in the green (FlAsH) and the red (ReAsH) channel are available. However, so far this method suffers from unspecific binding to other thiol rich proteins and the arsenic-induced toxicity (Hoffmann et al., 2010).

Unnatural amino acids
Genetic code expansion can be used to introduce noncanonical (unnatural) amino acid (nnAAs), with an additional modification into the host system. This nnAAc can than be used for site-specific fluorescent labeling (Bachmann et al., 2015). Specific tRNA synthases and tRNAs that recognize an unused or low abundant codon, in the host system, and that carry the unnatural amino acid are additionally introduced into the host cells. Most commonly, the Amber stop codon (UAG), with an abundance of 20% in human cells, is used to translate nnAAs. However, in principle, stop codon suppression can happen for every endogenous mRNA that terminates with this stop codon, leading to a high and uncontrollable unspecificity (Nööling et al., 2019). Endogenous stop codon suppression is a non critical process for in vitro studies, where the modified protein of interest is purified. However, the ill controllable stop codon suppression of endogenous proteins can heavily influence live cell studies. A new approach that uses designer membrane-less organelles tackles this issue by introducing a fully orthogonal translation system. This system harbors all necessary biological tools for genetic code expansion in a synthetic compartment inside the host system. This organelle is generated by engineered phase separation and generates a compartment for highly specific and selective codon suppression and the introduction of nnAA into the protein of interest. This approach was also used for orthogonal dual-color fluorescent labeling (Reinkemeier et al., 2019).
One major drawback of using fluorescent organic dyes is cell permeability. Most dyes are non-cell-permeable, which limits the number of usable dyes in live cell experiments. Another disadvantage is that cell-permeable dyes are most often used in excess to achieve high labeling efficiencies. Extensive washing steps are required to remove free unbound dye molecules. A recent approach uses genetic code expansion to encode cellular Cy dyes for single molecule labeling in frog oocytes (Leisle et al., 2016). This system combines the full set of advances for protein labeling, like being small and possible encodable at any site within the protein, being live-cell compatible, and it comes with bright and photostable dyes. This system’s remaining unspecificity might be addressed by combining this approach with the above-introduced system of membrane-less organelles.

**Self splicing inteins**

Intein mediated protein trans-splicing is one of the most promising and versatile approaches to study proteins in living cells. An intein is a protein self-splicing enzyme. Split inteins are a unique subgroup, which upon reassembly facilitates their splicing activity. This process results in the cleavage of the intein and the formation of the protein of interest (T. Y. Wang et al., 2014). The naturally occurring split intein Npu DNaE is the most commonly used intein for selective protein labeling (X. Chen and Wu, 2016). The protein of interest is tagged with the C-terminal half of the intein and expressed as a genetic fusion construct. The recombinantly expressed and labeled N-terminal half of the intein is transfected into the cells and facilitates the highly specific labeling reaction (Schütz and Mootz, 2014). Labeled, FRET quenched inteins constitute an additional improvement for the use of inteins in live cell experiments (Borra et al., 2012). In general, this technology offers a highly modular and biocompatible labeling approach with minimal biological disturbance and high specificity.

Most often, dual-color single molecule fluorescence experiments aim for the investigation of protein colocalization. Self complementing split fluorescent proteins like eGFP, mCherry, or mEos3.2, are widely used to study protein interactions in living cells (Feng et al., 2017). However, due to their fast photobleaching kinetics, these FPs are of limited use in single particle tracking experiments. As laid out here, many new approaches are currently the topic of scientific research, and in the future, further improvement or a combination of different fluorescent labeling techniques will increase the toolbox for live cell single molecule tracking experiments.
7 Results

7.1 Nuclear myosin VI

If not stated otherwise, the content of this section is based on previously published material (Grosse-Berkenbusch et al., 2020). New findings are highlighted in italic letters and modifications of figures are indicated.

Myosin VI has been shown to localize to the nucleus (Vreugde et al., 2006) and is there involved in transcription and hormone-dependent gene expression while also directly interacting with DNA (Cook et al., 2018; Fili et al., 2017; Fili and Toseland, 2020).

First, the proper localization of the used Halo-MVI construct was assessed by fluorescence microscopy. Stable HeLa and MCF7 cell lines did show cytosolic and nuclear localization of MVI and as previously described MVI was excluded from the nucleoli (Figure 7.1 (a,b)) (Fili et al., 2017). As expected for a polyclonal cell population, the expression level of Halo-MVI varied from cell to cell. Additionally, the fraction of cytosolic to nuclear MVI varied as well. The latter cellular difference might arise from different cell cycle stages or local stimuli.

Since the detailed behavior of nuclear MVI was of interest, single molecule fluorescence microscopy (HILO) was used for further investigations. The pool of nuclear MVI molecules could be divided into different populations: diffusive, static bound, and moving molecules. Diffusive molecules constituted the main fraction of the nuclear MVI pool. The static bound molecules could be further subdivided into four populations with distinctive binding times. Additionally one fraction of the intermediate bound molecules showed micrometer long directed motion events (Figure 7.2 (a-d)). The detailed analysis of single molecule fluorescence data will be described in the following chapters.
7 Results

Figure 7.1: Spinning disc fluorescence image of stable HaloTag-MVI in HeLa (a) and MCF7 (b) cell line. The scale bar is 10 µm. LUT in (a) 1000-4000 and in (b) 100-1000. (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0)

Figure 7.2: HILO imaging of nuclear myosin VI showed diffusive, static chromatin binding and micrometer-long moving MVI fractions. (a) Single molecule imaging of Halo-Tag-MVI. (b) HaloTag-MVI localizations (light blue) and tracked molecules lasting for > 150 ms (dark blue) or > 500 ms (magenta). (c) Kymographs of HaloTag-MVI molecules (upper and lower box in b). (d) Selected frames visualizing directed MVI motion (middle box in b). Scale bar is 10 µm in (a) and (b) and 2 µm in (d) (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0).
7.1 Nuclear myosin VI

7.1.1 ITM

ITM imaging (see chapter 5.5.1) was used to divide the whole pool into diffusive and bound molecules. The analysis showed that most of the molecules were diffusive, and only a population of 13% were statically bound molecules (Figure 7.3). This result already hints at a more dynamic biological function compared to statically bound nuclear molecules like histones.

![ITM measurements of Halo-Tag-MVI. Fraction of bound MVI molecules (mean 0.13 ± 0.05 s.d.). The data includes 30991 molecules from 22 cells (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0).](image)

7.1.2 Time lapse microscopy

Initial experiments showed that MVI molecules exhibited binding times in the range of several seconds. Time-lapse microscopy (see chapter 5.5.2) was used to resolve these long bound fractions without being limited by photobleaching. Time-lapse data sets with the following dark times were acquired: continuous movies (50ms), 100 ms, 200 ms, 400 ms, 800 ms, 1.6 s, 3.2 s, 6.4 s, and 12.8 s. Fitting of survival
Table 7.1: Data obtained by GRID analysis of MVI time lapse measurements (Grosse-Berkenbusch et al., 2020).

<table>
<thead>
<tr>
<th>dissociation rate interval (1/s)</th>
<th>0.01 – 0.14</th>
<th>0.14 – 0.7</th>
<th>0.7 – 7</th>
<th>7 – 70</th>
</tr>
</thead>
<tbody>
<tr>
<td>binding time</td>
<td>19s</td>
<td>4s</td>
<td>500ms</td>
<td>100ms</td>
</tr>
<tr>
<td>event spectrum weight</td>
<td>21 ± 2.6%</td>
<td>19 ± 2.1%</td>
<td>24 ± 0.9%</td>
<td>35 ± 0.7%</td>
</tr>
<tr>
<td>state spectrum weight</td>
<td>0.25 ± 0.06%</td>
<td>1.1 ± 0.1%</td>
<td>10 ± 0.6%</td>
<td>88 ± 0.7%</td>
</tr>
</tbody>
</table>

time distributions to all acquired data sets generated conflicts in the fitting routine. The small differences in dark times between the different time-lapse conditions resulted in highly overlapping data sets and each individual data set with few data points introduces additional noise into the fitting routine. The sum of the individual discrepancies can finally lead to a compromise of the fitting routine, which does not represent the data properly. Therefore, data sets with low amounts of data points were excluded from the analysis leaving five different data sets for the final analysis (50 ms, 200 ms, 400 ms, 1.6 ms, and 6.4 ms). Finally, four distinctive rate clusters referring to different binding times of MVI molecules were identified by using GRID (see chapter 5.5.3) (Table 7.1 and Figure 7.4).

MVI shows, apart from the diffusive population, two fast binding fractions, with binding times of 100 ms and 500 ms. Compared to the other fractions, the high amplitude for this fraction obtained by the event spectrum highlights the presence of significantly more nuclear MVI binding sites. In contrast, binding sites occupied by molecules with long binding times of 4 s and 19 s are less frequent. The state spectrum represents an alternative way of presenting the data. Here, the probability of detecting a molecule in a certain state in one snapshot is shown. For MVI, the occupation frequency in all identified rate clusters is comparable. The identified clusters propose four different biological mechanisms with an equal biological significance.
Figure 7.4: Time-lapse analysis of bound MVI molecules by GRID. (a) Survival time distributions of HaloTag-MVI molecules were measured at the time-lapse conditions indicated above. Dashed lines indicate theoretical survival time functions. Data includes 6752 molecules from 32 cells. (b) State spectrum of bound HaloTag-MVI molecules (magenta). (c) Event spectrum of bound HaloTag-MVI (magenta). Error estimation obtained by resampling of 80% of the data (blue) (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0)).

7.1.3 Motion analysis

Among the bound MVI molecules, a small fraction of directed moving molecules could be identified. Those molecules were identified by using SPLIT (see Chapter 5.5.5). The nature of the MVI runs was relatively diverse. Straight, curved, and angled runs were detected over the whole two-dimensional nuclear space. Individual runs were separated by pauses ranging from a few hundred milliseconds up to several seconds. Pauses most often occurred at the end of a run or at an angled turning point (Figure 7.5 (b)). Pausing molecules showed locally restricted diffusion. The run length, being either biologically limited or limited due to molecules moving out of focus and photobleaching, ranged from short runs of 500 nm up to several micrometers. The velocity with a mean value of 2.2 ± 0.1 \( \mu m/s \) ranged between a value of 1-4 \( \mu m/s \).

Nuclear Myosin V and I have been shown to be involved in the relocalization of heterochromatin breaks. This movement is directed towards the nuclear lamina (Caridi et al., 2018). Here, the direction of MVI runs in relation to the center of the nucleus was assessed by the direction of the end-to-end vector of the individual runs. However, the data did not support any preferred direction of nuclear MVI
Figure 7.5: (a) Three Examples of angled, straight, or curved MVI runs (purple) with interspersed pauses (blue). The motion of MVI was identified and separated into runs and pauses by a recurrence algorithm (SPLIT). The scale bar is 2 µm. (b) The directionality of MVI runs (direction of end-to-end vector relative to the center of the nucleus) ((b) As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0).

As already mentioned above, the frequency of MVI runs was relatively low compared to all MVI events, questioning their biological significance. However, since the labeling was restricted to exogenous MVI, the results lacked the endogenous fraction, and in order to achieve single molecule localization, the labeling density was adjusted to a very low level. To correct the measured MVI run frequency, Western Blot analysis and titration of Halo-Tag ligand was performed.

Western Blot analysis showed that Halo-Tag-MVI was overexpressed by a factor of $1.25 \pm 0.11$, compared to the endogenous protein level (Figure 7.6 (a)). Additionally, the minimal and maximal fluorescence signals of labeled Halo-MVI cells was measured by TMR titration. The differently labeled cells were measured by FACS to classify and compare the value used in single molecule experiments. Since there was a lack of SiR-ligand, the assumption was made that both dyes (TMR and SiR) show comparable labeling behaviours. The analysis revealed that roughly 1/4 of
7.1 Nuclear myosin VI

Figure 7.6: (a) Western Blot of endogenous MVI and HMVI expression in HeLa cells (lane: 1, 3, and 5), and a cell line stably expressing HaloTag-myosin VI (lane: 2, 4, and 6). γ-tubulin was used as a loading control. Lanes 1/2, 3/4, and 5/6 show biological replicates. The overexpression level of HaloTag-myosin VI was $1.25 \pm 0.11$ compared to endogenous myosin VI. (b) Titration experiment of TMR-HaloTag ligand. Stable cells expressing HaloTag-myosin VI were stained with TMR-HaloTag ligand with various concentrations (red dots). The mean TMR fluorescence intensity of 50,000 cells for every concentration was specified by flow cytometry. For single molecule tracking experiments, 31 pM of SiR-HaloTag ligand was used (blue square). By assuming similar labeling efficiencies for TMR-HaloTag ligand and SiR-HaloTag ligand, this results in a labeling degree 1/4 of exogenous expressed HaloTag-MVI (Grosse-Berkenbusch et al., 2020, reprinted under license CC BY-NC-ND 4.0).

Exogenous expressed Halo-MVI molecules were labeled in single molecule experiments (Figure 7.6 (b)). The total MVI motion frequency was determined by dividing the number of total runs identified by the recurrence analysis by the total observation time. For Halo-MVI, 92 runs were identified in a total of 179 movies, whereas each movie spanned an observation time of 15 s (= 2685 s). This gave a run frequency of 0.034 runs/s, which equals a rate of roughly one run every 30 seconds. Considering the overexpression level of 1.25 and the labeling level of 1/4, the real run frequency can be calculated, which gives a value of 1 run every 9 seconds (= 0.11 runs/s).

The classification of MVI runs by SPLIT resulted in a total of 92 runs (Figure 7.7 (a)). Most specified kinetic parameters showed a broad distribution, and the lack of an underlying model made the assessment of the individual parameters and their
interpretation difficult. The motion events could be separated into the characterization of the runs, with its run length and velocity, and the pause, with its dwell time and diffusion coefficient. The distribution of the run length with a median of 1.4 ± 0.6 µm (± s.d.) showed a peak at small values, whereas a few prominent long runs were identified as well (Figure 7.7 (b)). As already described, the velocity was the most defined parameter and showed a Gaussian distribution with a mean velocity of 2.2 ± 0.1 µm/s (± s.d.) (Figure 7.7 (c)). The pauses, either at the end of a run or separating two runs, were defined by a broadly distributed dwell time with a characteristic pause length of 1.3 ± 0.1 s (exponential fit ± confidence interval), and a slow mean diffusion coefficient of 0.01 ± 0.001 µm²/s.

Molecular motors like MVI are fueled by the hydrolysis of ATP to ADP + Pi, as described in chapter 3.1.1. The ATPase activity represents a fundamental function that might affect the identified motion events. The described MVI deafness mutant D179Y disrupts its ATPase activity (Hertzano et al., 2008) and presents a starting point for further experiments. Unfortunately, this mutant could not be used for
nuclear investigations since this MVI mutant did not localize to the nucleus (data not shown). TIP (2,4,6-triodophenol) specifically inhibits the ATPase activity of MVI (Heissler et al., 2012) and does impede MVI dependent transcription of ER target genes upon hormone stimulation (Cook et al., 2018). Previous live cell experiments showed that the use of 25 µM TIP did completely impede MVI dependent vesicular transport (Heissler et al., 2012). Based on those findings, TIP was used for further investigations. Nuclear single molecule experiments showed that blocking of the ATPase activity with 25 µM TIP did result in a significant decrease of MVI run frequency to 60%, compared to MVIwt motion (Figure 7.8 (a)), while not disturbing the general cellular morphology (Figure A.1).

![MVI motion frequency distribution](image)

Figure 7.8: (a) The MVI motion frequency distribution normalized to myosin VI wild type (wt) in the presence of the ATPase inhibitor TIP or the transcription inhibitor TPL. Frequencies were normalized to the number of detection events (n > 250,000). Detailed statistics are summarized in table 7.2. Error bars denote sqrt(N), P-values are from a Chi² test. (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0)

Another open question had been, if other molecular processes indirectly drive MVI molecules. MVI has been found to associate with RNA-PolII, being involved in the transcription process (Fili et al., 2017; Vreugde et al., 2006). The transcription inhibitor TPL (Triptolide) was used to exclude hitchhiking of MVI by RNA polymerase...
II during the transcription process. TPL binds covalently to the transcription factor TFIIH. TFIIH is part of the transcription initiation complex and is responsible for opening the DNA, thereby releasing built-up strain. Blocking of this fundamental process by TPL leads to inhibition of transcription initiation without directly interfering in the formation of the transcription initiation complex or other preceding processes (Titov et al., 2011). Investigation of nuclear MVI did show that the treatment of TPL did not significantly decrease the frequency of MVI motion (Figure 7.8 (a)).

The MVI inhibition with TIP did show that active ATPase cycling MVI molecules are involved in the identified motion processes. However, the treatment with the transcription inhibitor TPL showed that MVI motion is not correlated with any direct movements involved in RNA-PolIII related transcription processes. Surprisingly none of the kinetic parameters assessed to define the MVI movements did change either under the influence of TIP or the influence of TPL. Similar to non treated MVI (wt), broad distributions of the assessed parameters were found (Figure 7.9). The results for TPL matched with the expectations since TPL does not directly interfere with the motor function of MVI. The small reduction in the motion frequency might be due to a general reduction of gene transcription after long term TPL treatment. However, TIP acting as an ATPase inhibitor could have changed the kinetic parameters of MVI runs.
7.1 Nuclear myosin VI

(1) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.5 ± 0.6 µm, median ± s.d.) and (c) velocity (2.2 ± 0.1 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 1.4 ± 0.1 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.001 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses.

(2) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.6 ± 1.1 µm, median ± s.d.) and (c) velocity (2.2 ± 0.1 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 1.5 ± 0.1 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.001 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses.

Figure 7.9: Analysis of motion in presence of 25 µM TIP (1) and in presence of 125 nM TPL (2). (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0)
7 Results

7.2 Investigation of MVI motion

If not stated otherwise, the content of this section is based on previously published material (Grosse-Berkenbusch et al., 2020). New findings are highlighted in italic letters and modifications of figures are indicated.

MVI is composed of several subunits, the N-terminal motor domain, the neck domain, and the C-terminal tail domain with its single alpha helix (SAH) and the cargo binding domain (chapter 3.1). Further investigations were made by using different MVI constructs and fragments to reveal more details about the molecular mechanism of the MVI motion.

Four different cell lines stably expressing Halo tagged constructs of, (1) a monomeric version, unable to form MVI dimers via the SAH domain (SAH), (2,3) two truncated versions either only containing the cargo binding domain (CBD), or the motor domain (motor), and (4) an artificial dimeric mutant, by addition of a leucine zipper (Zipper), were created. All mutants where expressed within a range of 1.6x of the full-length HaloTag-MVI (Figure 7.10 (a)). The motion frequency and the kinetic parameters of all MVI mutants were compared to the MVI wildtype (wt).

Single molecule tracking experiments of the SAH mutant revealed that the potential monomeric mutant has, compared to the wt, an increased motion frequency by factor 1.5 with similar motion characteristics (Figure 7.10 (b) and 7.11 (1)(a)). Interestingly, all other mutants (CBD, Zipper, motor) showed a significant reduction in their motion frequency (>50%) (Figure 7.10 (b)). However, the assessment of the individual kinetic parameter distributions did not reveal any significant differences among the different mutants. However, neither the CBD nor the motor domain or the forced zipper showed single runs with a run-length above 3 µm (Figure 7.11 (2)(a) and Figure 7.12 (1)(a)).

The leucine zipper mutant has previously been used in vitro for the characterization of MVI movement (H. Park et al., 2006; Rock et al., 2001). The results obtained from the zipper mutant strengthens the argument that the full-length protein is required for the motion process since the forced zipper showed also no runs with significant run length (Figure 7.12 (2)(a)). These combined results suggest that the motion process is facilitated by MVI full length motors.
7.2 Investigation of MVI motion

Figure 7.10: (a) Number of single molecule detections in individual cells of wild type MVI (wt), a monomeric MVI mutant (SAH), the MVI cargo binding domain (CBD), a dimeric MVI mutant (Zipper) and the MVI motor domain (motor), each fused to HaloTag, under equal imaging conditions. (b) Frequency of MVI motion normalized to wild type (wt) of a monomeric MVI mutant (SAH), the MVI cargo binding domain (CBD), a dimeric MVI mutant (Zipper) and the MVI motor domain (motor). Frequencies were normalized to the number of detection events (n > 250,000). Detailed statistics are shown and summarized in table 7.2. Error bars denote sqrt(N), P-values are from a Chi2 test. (c and d) Fluorescence intensity analysis of MVI tracks and Kymographs with center position running along the tracks. Color maps indicate the intensity level. The scale bar is 2 µm ((b-d) As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0).

Until now, the biological mechanism behind the processive nature of MVI is not fully understood (Fili et al., 2020; Iwaki et al., 2006; Sivaramakrishnan and Spudich, 2009). Fluorescence intensity analysis of individual MVI tracks revealed double bleaching steps of moving and bound molecules (Figure 7.10), indicating that MVI might form dimers in its bound state and during its motion process. The experiments with the SAH mutant suggest that potential dimerization is not facilitated via the single alpha-helix domain and might be enabled through indirect binding of the cargo binding domain to other proteins as proposed by Sivaramakrishnan and Spudich, 2009 and Fili et al., 2020.

In summary, the motion frequency analysis showed that the full-length motor is required for MVI motion. The comparison of the kinetic parameters of the wt with the MVI mutants revealed no significant differences. The run length and velocity distributions are dominated mostly by short runs and partly by long runs separated
7 Results

(1) "(a) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.5 ± 0.1 µm, median ± s.d.) and (c) velocity (2.1 ± 0.1 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 1.2 ± 0.1 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.001 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses"

(2) "(a) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.2 ± 0.4 µm, median ± s.d.) and (c) velocity (2.4 ± 0.1 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 0.9 ± 0.1 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.002 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses"

Figure 7.11: Analysis of motion of HaloTag-MVI-SAH mutant (1) and motion of HaloTag-MVI-CBD deletion constructs (2). (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0).
7.2 Investigation of MVI motion

(1) *(a) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.4 ± 0.3 µm, median ± s.d.) and (c) velocity (2.3 ± 0.1 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 0.9 ± 0.2 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.002 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses*

(2) *(a) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.4 ± 0.3 µm, median ± s.d.) and (c) velocity (2.4 ± 0.1 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 1.0 ± 0.1 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.001 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses*

Figure 7.12: Analysis of motion of HaloTag-MVI-Zipper (1) and of motion of HaloTag-MVI-motor deletion constructs (2). (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0).
Results

by pauses. The small number of long runs did not significantly affect the analyzed kinetic parameters. The small quantity of obtained data points from those runs also did not allow for an individual analysis.
7.3 Nuclear actin

The presence of nuclear actin filaments has been a long and controversial scientific topic. It has been technically challenging to visualize any defined nuclear actin structures until new staining methods became available. The visualization of actin filaments was previously performed by using artificial actin labels that are known to induce or stabilize filaments. These properties have been shown for GFP-actin and are, of course, true for imaging with the labeled S14C actin mutant. These tagging methods may, therefore, induce non-natural actin structures inside the nucleus (Sliogeryte et al., 2016). Recently a new way of visualizing nuclear actin structures has been reported. This method is based on chromobodies. Chromobodies consist of target-specific nanobodies (VHH - antigen binding domain of heavy chain IgG) fused to a fluorescent reporter that can be expressed in living cells (Hamers-Casterman et al., 1993). Compared to the expression of actin fusion constructs, the labeling via nanobodies is thought to interfere less strongly with the kinetics of filament formation and potentially visualizes the natural structure of nuclear actin structures (Plessner et al., 2015).

Initially, nuclear actin filaments were visualized by the expression of NLS-YFP-actin fusion constructs. The expressed actin constructs contained one of two point mutations, namely either S14C or R62D. Whereas R62D is a nonpolymerizable actin mutant that blocks actin filament formation, S14C favors filament formation (Posern et al., 2002). Both constructs were transiently expressed in HeLa cells, and the presence of actin filaments was visualized by spinning disc microscopy (Figure 7.13).

The expressed fusion constructs were present in the whole cellular space. The added NLS facilitated the recruitment of both constructs into the nucleus but did not limit the localization to this defined cellular compartment. Interestingly, similar to MVI (Fili et al., 2017), both constructs did not localize to the nucleoli. Overall the observed actin phenotype was very heterogeneous. Independent of the expression level, which varied due to the transient transfection method, the localization efficiency to the nucleus and the presence of defined actin structures inside the nucleus varied from cell to cell. The filament stabilizing mutant S14C was successfully incorporated into regular cytosolic actin filaments and with a very low frequency into short, thick, and bundled actin structures inside the nucleus. These nuclear
Filaments were long-lived and stable for several minutes. Longer observation times were not applied. As expected, R62D did not stain any defined cellular actin structures. The reported filament destabilizing properties were not quantified (Posern et al., 2002).

Different stable cell lines expressing nAc-GFP (GFP coupled nuclear Actin chromobody) were created (HeLa, MCF7 and NIH3T3). Initial experiments recapitulating the visualization of actin filaments by using FCS for induction of nuclear actin or the formation of filaments during cell spreading after trypsinization reported in Plessner...
et al., 2015 failed in all tested cellular backgrounds. It has been reported that only 40% of cells showed nuclear filaments during cell spreading (Plessner et al., 2015). However, the nearly complete lack of nuclear actin structures, as shown in figure 7.14 (a), was puzzling. The inhomogeneous but low labeling density of actin with nAc-GFP was further reduced by using nAc-SNAP. SNAP tagged nAc molecules could be labeled with limited amounts of SNAP tagged ligands, thereby reducing any strong background signals from nonpolymerized nuclear actin monomers. However, the conducted experiments did not solve the visualization issue (data not shown).

Finally, the original cell lines used in Plessner et al., 2015 and the created stable cell lines could be compared by imaging those cells under the reported conditions in Freiburg. The visualization of induced nuclear actin filaments by FCS has not failed due to any preparative differences, nor by environmental conditions, neither during cell growth nor during imaging (Cell splitting, DMEM, FCS, Temperature, CO2) or by other technical limitations. The difference has been the used laboriously selected monoclonal cell line and the change of the induction reagent. Induction of nuclear actin filaments by addition of 8 µM ionophore A23187 (Y. Wang et al., 2019) resulted with a low but detectable frequency in the formation of nuclear actin filaments (Figure 7.14 (b-d)). The nature of those filaments differed from cell to cell. Most cells showed short-lived, transient, thin and dynamic filaments (Figure 7.14 (b,c)). However, a few cells formed more stable and thicker actin filaments (Figure 7.14 (d)). Since the goal of this study has not been the description of nuclear actin filaments, further nuclear actin experiments were conducted in HeLa cells expressing the Halo-MVI fusion construct by induction with A23187.
Figure 7.15: Nuclear actin filaments in polyclonal HeLa cells stably expressing nAc-GFP and Halo-MVI. (a,b) Cell after induction with the ionophore A23187. LUT in (a) 100-500 and in (b) 300-700. The nucleus is indicated with dashed lines (white), and the scale bar is 2 μm.

A stable cell line expressing nAc-GFP was generated by lentiviral transduction of the stable cell line expressing Halo-MVI. Induced nuclear actin filaments by adding the ionophore A23187 were visualized using HILO microscopy (Figure 7.15). The use of the ionophore A23187 did alter the cellular morphology. Therefore, cells were only imaged up until 30 minutes after the application of A23187. For colocalization experiments of actin filaments and dynamic MVI molecules, cells forming stable actin filaments were imaged since the visualization of the described dynamic, transient and thin filaments was technically challenging.
7.4 MVI moves on nuclear actin filaments

If not stated otherwise, the content of this section is based on previously published material (Grosse-Berkenbusch et al., 2020). New findings are highlighted in italic letters and modifications of figures are indicated.

Nuclear actin filaments play an important role in myosin facilitated motion (Chapter 3.1.2). Further investigations aimed to elucidate the direct involvement of nuclear actin filaments in the process of the identified directed MVI runs.

Actin filaments can be manipulated in many different ways. However, since the presence of naturally occurring nuclear actin is still under debate, the actin exporter Exportin 6 (XPO6) was initially used to assess the presented issue. XPO6 facilitates the active nuclear export of G-actin, maintaining a defined equilibrium of the cytosolic and nuclear actin pool (Stüven et al., 2003). The overexpression of XPO6 lowers the concentration of G-actin, thereby impeding nuclear actin filament formation (Onuma et al., 2019). The stably expressing HaloTag-MVI cell lines were transiently transfected with a plasmid coding for XPO6. For single cell analysis, the cells were positively identified by the coexpressed fluorescent marker eGFP. The frequency of MVI motion dropped significantly by factor 0.5 in cells expressing XPO6 (Figure 7.16 (a)), suggesting a relevant role of nuclear actin for MVI motion.

Further experiments supporting this finding were performed using Latrunculin A (LatA). Latrunculins are natural toxins which are expressed by certain sponges including Latrunculia and Negombata. The two main forms are Latrunculin A and B. Similar to XPO6, LatA reduces the available pool of actin molecules by binding actin monomers from the available actin pool 1:1 with high affinities, thereby affecting the dynamics of actin filament formation. In contrast to XPO6, the effect is not limited to disturbing the nuclear actin pool (Andrin et al., 2012), which finally leads to destabilization of the cell structure (Coué et al., 1987). To minimize structural effects, HeLa cells were treated with a small amount of 0.1 µM LatA to destabilize nuclear actin filaments. The treated cells were imaged 5 min after application, for up to 20 minutes, to reduce the systematic effect of the disturbance of the cytosolic filaments that has been observed (Figure A.1). Likewise XPO6, LatA reduced the frequency of MVI motion by factor 0.55 (Figure 7.16 (a)).
Different actin mutants can alter the dynamics of actin filament formation. Of particular interest were two actin mutants, namely R62D and S14C (Posern et al., 2002). R62D represents a non polymerizable actin mutant, whereas S14C constitutes a filament stabilizing mutant. Both mutants were individually coexpressed by transient transfection as NLS-tagged YFP fusion proteins in cell lines stably expressing Halo-Tag-MVI. For single cell analysis, positively transfected cells were selected by the fluorescence reporter YFP.

Cells expressing S14C did show a similar MVI motion frequency than untreated cells (wt) (Figure 7.16 (a)). The minor reduction might arise from disturbing interference of MVI molecules with the attached YFP reporter. In contrast, cells expressing R62D did show a significantly reduced MVI motion frequency. Similar to XPO6 and LatA, R62D interferes with MVI motion dynamics.

In summary, the negative effects on nuclear actin structures by direct and indirect interference on filament formation lead to a global reduction of nuclear MVI runs, whereas the positive stabilization of actin filaments did show no effect. Taking to-
7.4 MVI moves on nuclear actin filaments

(1) "(a) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.5 ± 0.7 µm, median ± s.d.) and (c) velocity (2.3 ± 0.1 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 1.0 ± 0.1 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.001 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses"

(2) "(a) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.5 ± 0.7 µm, median ± s.d.) and (c) velocity (2.6 ± 0.1 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 1.2 ± 0.1 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.001 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses"

Figure 7.17: Analysis of HaloTag-MVI motion in presence of conditions altering nuclear actin (1) (XPO6) and (2) (LatA). (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0).
(1) "(a) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.5 ± 0.8 µm, median ± s.d.) and (c) velocity (2.3 ± 0.1 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 1.0 ± 0.1 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.001 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses"

(2) (a) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.6 ± 1.2 µm, median ± s.d.) and (c) velocity (2.1 ± 0.1 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 1.2 ± 0.1 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.001 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses"

Figure 7.18: Analysis of HaloTag-MVI motion in presence of conditions altering nuclear actin (1) (R62D) and (2) (S14C). (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0).
7.4 MVI moves on nuclear actin filaments

Together these findings strongly support the hypothesis that nuclear myosins move along actin filaments. The kinetic analysis and comparisons of the motion parameters under actin filament stabilizing or destabilizing conditions did not reveal any significant changes, compared to MVIwt values. As already described above, the distributions are dominated by small runs, whereas the few present outliers could not be used for statistical analysis.

Finally, direct visualization experiments strengthening the statement that nuclear MVI moves on actin structures were conducted. Visualization of nuclear actin filaments was facilitated by NLS-tagged GFP fusion of nanobodies against actin (nAc-
GFP) (Plessner et al., 2015). Colocalization experiments were performed in stable cell lines coexpressing Halo-Tag-MVI and nAc-GFP. Nuclear actin filaments were induced by the addition of 8 $\mu$M Ionophore A23187 (Y. Wang et al., 2019), and cells were imaged for up to 30 min after induction. SiR labeled MVI, and GFP labeled filaments were illuminated in an alternating pattern: 100x 638 nm - 2x 488 nm.

Colocalization experiments of MVI nuclear actin confirmed that MVI moves along actin filaments (Figure 7.19 (a-c)). All motion events were detected on bundled and more stable, rather than on thin and highly dynamic actin structures. The additional quantification of colocalization events of all tracked molecules, including also statically bound ones (>150 ms), using a user defined threshold for the segmentation of actin regions, revealed that MVI binds preferentially to nuclear actin (factor 2, n = 3 cells). Figure 7.19 (c) shows a colocalization event during a dynamic change of the underlying structure. At t = 5s, the MVI track stops at its current location. This stop matches with the change of the underlying actin filament. The final position of both the MVI track and the actin filament is shown in the next image at t = 6.11.

In general, the regular static binding events of MVI colocalized but were not limited to areas of actin structures. Overall, the chance of detecting a colocalizing motion event was very low. Limitations have been the overall low MVI motion frequency combined with a low chance of visualization of transiently induced nuclear actin filaments, and the negative effect of the Ionophore A23187 on the cell physiology limited the observation time.
7.5 MVI supports long range chromatin rearrangements

If not stated otherwise, the content of this section is based on previously published material (Grosse-Berkenbusch et al., 2020). New findings are highlighted in italic letters and modifications of figures are indicated.

Recently, the investigation of the genome’s 3D organization and its relevance for transcription made substantial progress (van Steensel and Furlong, 2019; Vermunt et al., 2019). Additionally, investigations on the involvement of MVI in transcription and hormone-stimulated gene expression have been made (Loikkanen et al., 2009; Vreugde et al., 2006). Estrogen receptor (ER) target gene expression has been shown to be both dependent on MVI and nuclear actin filaments. Gene expression of target genes was reduced upon inhibition of MVI by TIP (Fili et al., 2017) and by destabilizing actin filaments by treatment with LatA (Q. Hu et al., 2008).

MCF7 cells originated from breast cancer cell tissues serve as a model cell line due to its relevant ER expression profile, thereby being sensitive to estrogen stimulation (Brooks et al., 1973). Initial experiments conducted by the Toseland group showed that MVI is actively involved in the nuclear rearrangement of chromosome 2 and chromosome 21 after hormone stimulation. This process was further shown to be dependent on actin filaments.

The nuclear behavior of MVI in MCF7 cells was investigated by single molecule tracking experiments. However, the binding characteristics were not quantified. MVI was assumed to bind in similar ranges to the ones characterized in HeLa cells. A small fraction of MVI molecules also exhibited directed and µm-long runs, similar to the ones observed in HeLa cells (Figure 7.20). The runs were characterized by using the same analysis as for HeLa cells. The tracked molecules showed a mean velocity of \(2.0 \pm 0.08 \mu m/s\) with a median run length of \(1.5 \pm 0.6 \mu m\). Nuclear MVI molecules mostly paused at the beginning or end of a run with a characteristic pause time of \(1.3 \pm 0.1\) s and a slow diffusion with a diffusion coefficient of \(0.02 \pm 0.001\).

In order to investigate the direct role of MVI in chromosome rearrangements, stable cell lines expressing Halo-Tag-MVI were initially starved in hormone depleted
7 Results

Figure 7.20: "Analysis of HaloTag-MVI motion in MCF7 (wt). (a) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.5 ± 0.6 µm, median ± s.d.) and (c) velocity (2.0 ± 0.08 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 1.3 ± 0.1 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.001 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses" (Grosse-Berkenbusch et al., 2020, reprinted under license CC BY-NC-ND 4.0).

medium. Then, cells were stimulated 5 min before the measurement with 100 nM β-estradiol and stimulated cells were measured in HILO mode for up to 30 min (Figure 7.21). The obtained results showed the hormone dependency of MVI motion events by revealing changes in the MVI run frequency under stimulating conditions, which complemented the chromosome paint data. The MVI motion frequency increased significantly by a factor of 1.8 in the first 30 min of hormone stimulation, compared to starved cells (Figure 7.22 (a)).
7.5 MVI supports long range chromatin rearrangements

(1) "(a) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.4 ± 0.5 µm, median ± s.d.) and (c) velocity (2.2 ± 0.1 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 1.2 ± 0.1 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.001 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses."

(2) "(a) HaloTag-MVI tracks including runs (magenta) and pauses (blue) identified by recurrence analysis. The scale bar is 5 µm. Distributions of (b) runlength (1.5 ± 0.6 µm, median ± s.d.) and (c) velocity (2.2 ± 0.1 µm/s, mean ± s.d.) of MVI motion. (d) Distribution of dwell times (decay time: 1.1 ± 0.1 s, exponential fit ± confidence interval) and (e) step-size histogram (diffusion coefficient: 0.02 ± 0.001 µm²/s, diffusion coefficient ± sem, Methods) of MVI pauses."

Figure 7.21: Analysis of HaloTag-MVI motion in MCF7 cells (1) treated 48h with charcoal stripped FCS (cs) and (2) stimulation with β-estradiol (E2). (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0)
Figure 7.22: Nuclear myosin VI supports long-range chromatin rearrangements. (a) MVI motion frequency normalized to the wild type (wt) in MCF7 cells in hormone depleted medium (cs) or after stimulation with β-estradiol (csE2). Frequencies were normalized to the number of detection events \((n > 250,000)\). Detailed statistics are shown in table 7.2. Error bars denote \(\sqrt{N}\), P-values are from a Chi2 test. (b) Chromosome PAINT of chromosome 2 (green) and 21 (magenta) after hormone stimulation with β-estradiol in absence (upper panel) or presence (lower panel) of the MVI inhibitor TIP in MCF7 cells. The scale bar is 10 µm. (c) The fraction of chromosome 2 overlapping with chromosome 21 in MCF7 cells in absence (cs) or presence of β-estradiol (csE2), the MVI inhibitor TIP, Latrunculin B (LatB) or Triptolide (TPL). P-values are from an unpaired two tailed t-test \((n > 60\) cells) (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0).

Chromosome paint experiments of chromosome 2 and 21 harboring the hormone sensitive genes GREB1 and TFF1 revealed that both chromosomes showed hormone regulated, MVI and nuclear actin dependent gene reorganization upon β-estradiol (Figure 7.22 (b), (c) and A.2). Additionally it was shown that this process, like the MVI motion process, is independent of the transcription mechanism, suggesting that this process occurs upstream of gene transcription. In total this complements the data that were found in single molecule tracking experiments.
7.6 Summary

If not stated otherwise, the content of this section is based on previously published material (Grosse-Berkenbusch et al., 2020). New findings are highlighted in italic letters and modifications of figures are indicated.

Myosin VI localizes to the nucleus of living cells, omitting the nucleoli. Single molecule fluorescence analysis showed that the major fraction of MVI are diffusive molecules, and only a fraction of 13% are bound molecules. The bound fraction can be further divided into four different binding time clusters ranging from 100 ms to 19 s. A small fraction of the intermediate bound molecules could be further identified as directed moving molecules.

Table 7.2 shows the data compilation of all above described experiments investigating the nuclear motion of MVI. The detected motion frequency of a single cell is directly correlated by its Halo-Tag-MVI expression level and the labeling density. Since both values varied from cell to cell, and the expression level could not be assessed on a single cell level (Figure 7.6), the cumulated motion frequency was normalized to the labeling density. For each data set, a total amount of cells were imaged (# cells), and the labeling density was assessed via the total number of detection events (total # events). The motion frequency was normalized by all detection events (motion #events) arising from the individual runs (# motion), and the results from the different measurements were normalized to the wildtype (wt).

The investigation of the identified tracks revealed that the movement is not caused by motion of RNA-PolII complexes (TPL), and second, is dependent on MVI’s own ATPase activity (TIP). Further, MVI motion is dependent on the full length motor (CBD, Zipper, and motor), and potentially requires dimerization, which is not facilitated through the single alpha helical domain. This could be shown by two independent results. First, by showing double bleaching steps during wt motion events and second through detecting motion of the SAH mutant. Additionally, individual runs require nuclear actin filaments (S14C, XPO6, LatA, and R62D), which was confirmed in fluorescent colocalization experiments. Biologically, MVI motion is involved in hormone-stimulated gene expression (cs and csE2) and involved in transcription related three-dimensional chromatin reorganization (Chromosome PAINT).
Table 7.2: Summarized MVI motion frequency statistics of all measurement conditions (Grosse-Berkenbusch et al., 2020, reprinted under license CC BY-NC-ND 4.0).

<table>
<thead>
<tr>
<th></th>
<th>#cells</th>
<th>total #events</th>
<th>#motion</th>
<th>motion #events</th>
<th>normalized frequency</th>
<th>normalized sqrt(#motion)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HeLa</td>
<td>wt</td>
<td>81</td>
<td>484857</td>
<td>92</td>
<td>1762</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>TIP</td>
<td>104</td>
<td>795202</td>
<td>112</td>
<td>1769</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>TPL</td>
<td>55</td>
<td>382691</td>
<td>63</td>
<td>1100</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>SAH</td>
<td>56</td>
<td>471180</td>
<td>104</td>
<td>2510</td>
<td>1.47</td>
</tr>
<tr>
<td></td>
<td>CBD</td>
<td>39</td>
<td>268357</td>
<td>16</td>
<td>195</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>zip</td>
<td>45</td>
<td>325672</td>
<td>35</td>
<td>424</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>motor</td>
<td>40</td>
<td>376680</td>
<td>42</td>
<td>522</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>S14C</td>
<td>55</td>
<td>536111</td>
<td>76</td>
<td>1716</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>XPO6</td>
<td>67</td>
<td>365389</td>
<td>39</td>
<td>666</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>LatA</td>
<td>57</td>
<td>285972</td>
<td>41</td>
<td>577</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>R62D</td>
<td>84</td>
<td>558446</td>
<td>54</td>
<td>904</td>
<td>0.45</td>
</tr>
<tr>
<td>MCF7</td>
<td>wt</td>
<td>44</td>
<td>275739</td>
<td>45</td>
<td>871</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>cs</td>
<td>76</td>
<td>284164</td>
<td>40</td>
<td>566</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>csE2</td>
<td>77</td>
<td>452567</td>
<td>100</td>
<td>1584</td>
<td>1.11</td>
</tr>
</tbody>
</table>

The motion events from the different measurements were divided into runs and pauses by using a recurrence algorithm (SPLIT). Figure 7.24 shows (a) the summarized run lengths (median ± s.d) and (b) velocities (mean of Gaussian distribution ± confidence interval). The analysis revealed no prominent differences between the different measurement conditions. Figure 7.24 shows (c) the summarized dwell-times (characteristic decay time ± confidence interval) and (d) the diffusion coefficients, which were calculated by the mean displacement as 2D isotropic diffusion, with t defined by the camera integration time. Similar to the characteristics of the runs, the pause characteristics showed no significant changes between the different measurement conditions.
7.6 Summary

Figure 7.24: Comparison of HaloTag-MVI motion parameters at different conditions. (A) The median of the different run length distributions. Error bars denote the standard deviation. (B) The mean velocity is determined by a Gaussian fit of the velocity distributions. Error bars denote confidence interval of 95%. (C) The individual average dwell time is determined by an exponential fit of the dwell time distributions. Error bars denote a confidence interval of 95%. (D) Diffusion coefficients were obtained from the displacement distributions. Error bars denote SEM. (As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0).
8 Discussion

The molecular motor myosin VI facilitates a wide range of cellular processes. It is involved in cytosolic transport processes (Arden et al., 2007; Aschenbrenner et al., 2003; Buss et al., 2001; R. Roberts et al., 2004) and anchoring mechanisms (Cramer, 2000; Hari-Gupta, Fili, dos Santos, et al., 2020; Self et al., 1999; G. Spudich et al., 2007). Apart from that MVI plays a direct role in gene transcription (Cook et al., 2018; Fili et al., 2020; Fili et al., 2017; Loikkanen et al., 2009; Vreugde et al., 2006). Despite being studied for nearly 30 years, the exact biological working mechanisms of MVI are not yet known, and only recent studies have shed light on its diverse biological functions.

Localization and binding of nuclear MVI

Figure 7.1 shows that the NI splice variant of MVI localizes to the cytosol and the nucleus, thereby being excluded from the nucleoli. Nucleoli are the site of ribosome biogenesis, which involves RNA polymerase I, II, and III (Abraham et al., 2020) and the nuclear fraction of MV has been found to play a role in this process (Lindsay and McCaffrey, 2009). Although MVI has been found to associate with RNA polymerase II during gene transcription (Vreugde et al., 2006), the motor seems to play no direct role in ribosome formation.

Recently many nuclear proteins have been characterized by single molecule fluorescence microscopy, and it has been found that many transcription factors that play a role in dynamic and stimulatory processes show high fractions of diffusive molecules (Hipp et al., 2019; Reisser et al., 2018). In contrast, structural proteins like histones or CTCF show smaller diffusive fractions (Agarwal et al., 2017; J. Chen et al., 2014). ITM analysis of MVI in non-stimulated cells showed that only 13% of MVI molecules are in a bound state, whereas the biggest fraction is in a
diffusive state. Similar to transcription factors, potentially the biggest fraction of MVI is passively waiting for their biological activation and recruitment. Several experimental results in the literature strengthen such a hypothesis. Unbound nuclear MVI is present in a backfolded confirmation, which by hiding its C-terminal binding motif, impedes protein binding (Fili et al., 2017). NDP52 has been found to resolve this confirmation, thereby exposing its protein binding motive. Upstream stimulation of NDP52 might therefore increase the fraction of actively binding motor complexes. Other stimulatory mechanisms might involve Ca\(^{2+}\), which by binding to CaM, resolves the binding of CaM to the motor neck region. This process is, on the other hand, thought to play a role in cargo binding (Batters et al., 2016). Further, KCl has been found to be another stimulator of MVI. It has been found that KCl treatment increases the nuclear MVI fraction by 25% (Majewski et al., 2018). However, until now no direct biological mechanisms have been associated with this mechanism. The biological relevance of the different stimulatory effects is not yet fully understood, and further studies might reveal a more in depth biological understanding. Additionally, cell cycle-dependent effects, as shown for CTCF (Agarwal et al., 2017), or nuclear actin related stimulators, which are further described below, might also impact the bound fraction of MVI, which might be easily studied by further ITM studies.

Binding time analysis of the bound fraction revealed four distinct rate clusters with binding times from 100ms to 19s (Figure 7.4). The obtained state spectrum reveals an equal probability of finding a MVI molecule in one of the four states. This finding proposes an equal biological significance for the different binding states. Whereas the event spectrum shows that the total amount of binding sites differs drastically from cluster to cluster. This translates into the finding that the studied, unstimulated nuclear environment harbors a much higher amount of short binding sites compared to long binding sites. However, the significance of the different binding times is not yet clear and needs to be assessed in future experiments. The motor domain and the CBD present two potential interaction sites that might be relevant for the nuclear binding of MVI. Starting from the N-terminus, MVI has a highly conserved motor domain that contains an actin-binding motif. Recently, the presence of nuclear actin filaments was confirmed (Plessner et al., 2015), which makes nuclear actin a potential target for MVI. Moreover, in the present study, it has been revealed that a minor fraction of 3% of bound MVI molecules that bind
an average of two seconds bind to nuclear actin structures. Other potential actin-binding sites might be RNA polymerase II-associated actin molecules (Hofmann et al., 2004b). These and similar kinds of interactions might also be required for the recently identified MVI tethering function (Hari-Gupta, Fili, dos Santos, et al., 2020). Interestingly, previously identified RNA polymerase II clusters were found to have lifetimes in a comparable range than the ones identified for nuclear MVI (Cho et al., 2016; Cisse et al., 2013). Additionally, since nuclear actin structures are induced by different stimulators such as DMSO, heat, serum, or during the cell cycle (Kelpesch and Tootle, 2018), binding-time analysis under these conditions might reveal which bound fraction is predominately involved in such potential binding events. In summary, further functional and microscopic studies, including a more diverse set of nuclear actin structures, might reveal a more conclusive image of these types of interactions.

On the other hand, the cargo binding domain presents two additional binding sites. One binding loop includes the conserved WWY motif, which can directly bind DNA (Fili et al., 2017). The other structure contains the RRL motif, which can bind NDP52 (Fili et al., 2017). Taken together, binding-time analysis of different MVI domains, including the motor domain or the CBD combined with the full-length protein, containing distinct mutations in the respective binding motifs, might give insights into the significance and meaning of different binding time clusters. Another potential approach might be the MVI binding-time analysis in a DAB2 overexpression system. DAB2 is the cytosolic antagonist of NDP52 and acts as a tumor suppressor (He et al., 2001) by potentially blocking the DNA binding motive of MVI (Fili et al., 2020). Since this effect is directly related to its nuclear function, it might be interesting how this interaction might influence the binding behavior of MVI.

Characterization of MVI motion events

Apart from the identified and described locally defined binding events, MVI shows micrometer long directed nuclear motion events. The MVI runs were split into active (runs) and passive parts (pauses) characterized by their run length and velocity and their dwell time and displacement, respectively. The presented data propose that the identified runs require the full-length motor, are not passively driven by RNA polymerase II, and that they require an active motor function. It was shown that the
Discussion

runs occur along nuclear actin filaments and that they potentially drive long-range chromatin rearrangements after hormone stimulation.

The general appearance of MVI runs was relatively diverse. As depicted in figure 7.5, molecules moved in straight lines, changed their direction at once, or moved in a curved fashion. Compared to the observed nuclear actin filaments, the observed MVI runs resemble the underlying actin structure surprisingly well. As shown in figure 7.15, induced nuclear actin filaments show straight, angled, and curved structures. Initial links between nuclear actin structures and MVI motion were found by subjecting the cells to destabilizing nuclear actin conditions (R62D, XPO6, and LatA). However, these results brought up the question, whether there is a direct correlation between MVI movement and actin structures?

The leading assumption for MVI runs has been that the introduction of filament destabilizing conditions would result in considerably less and shorter actin filaments. Whereas a decreased MVI motion frequency could confirm the first assumption, the second assumption was not reflected in the results. One reason for the lack of a decreased run-length might be that completely polymerized actin filaments are required for successful MVI motion. The unique directionality of MVI strengthens this assumption. Actin filaments polymerize from the minus to the plus end (Petty and Cassidy, 2005), whereas MVI moves towards the minus end (Wells et al., 1999). However, if the filament formation is interrupted during the polymerization process prior reaching its target site, at which the MVI molecules bound to their cargo accumulated, it is not surprising that the shortening of actin structures will not be resembled in the MVI run length distribution. Another shortcoming of this work, concerning the track parameters, is the relatively small amount of data points. Overall the few values obtained from the long runs are not well resembled in the final fittings of the distributions. However, they also did not resemble a complete and distinct separate population, which made their analysis rather difficult.

Recently, new live-cell compatible nuclear actin labeling methods, including life-act and chromobodies, made it possible to learn more about the hardly detectable nuclear actin structures (Melak et al., 2017). While many different studies described transient actin filaments formed upon specific stimuli or during certain cell cycle phases (Plessner and Grosse, 2019), it is not yet clear whether other structures are
regularly present during the interphase. The found MVI run length distribution might not only resemble the traveled MVI distance solely, but it might also present a distribution of the variable lengths of the underlying nuclear actin structure during the interphase. Since the study of nuclear actin structures is a complex field, further experiments on different nuclear motors might complement other direct visualization studies.

The identified mean velocity of MVI of 2.2 µm/s does significantly differ from in vitro characterized values. Previous in vitro measurements determined a velocity of 0.3 µm/s, for artificially zipped dimers (Rock et al., 2001). However, such discrepancies have also been observed for myosin V (Pierobon et al., 2009) and might arise from different experimental conditions. First, the cell has complex environmental conditions, like areas with tightly controlled ionic strengths and pH values. Additionally, molecular crowding effects, cofactors like CaM or other binding partners, and the local geometry and topology of the actin cytoskeleton might influence the velocity of MVI and are difficult to reproduce in vitro. Another factor might be the different structural properties that arise from the artificially zipped dimer in contrast to the wt motor. Until now, the dimerization mechanism of MVI is highly under debate. An indirect, less studied dimerization mechanism via the cargo binding domain might change the currently proposed stepping models, mostly based on tail-based dimerization mechanisms, reconstituted by the zipped dimer (J. A. Spudich and Sivaramakrishnan, 2010). Indirect measurements in yeast that measure the velocity of MVI facilitated vesicular motion, also propose higher velocities in the range of 2-3 µm/s (Schott et al., 2002). Similar comparisons for MV showed that the measured in vivo velocity is closer to the vesicular velocity than the respective in vitro value (Pierobon et al., 2009). Another experimental approach to study the identified motion process could be to change the length of lever arm of MVI. A previous study revealed that the MVI facilitated vesicular velocity is directly dependent on the lever arm length (Schott et al., 2002).

MVI motion events could be separated into runs and pauses. The identified pauses with a pause time of 1.3 s could be characterized by a slow diffusion coefficient of 0.01 µm²/s. The biological significance of these pauses could not be resolved since it was not further experimentally assessed. However, the identified pause
time clearly differentiates from the described motor dwell time that is below 100 ms, which characterizes the waiting time in between steps (Mukherjea et al., 2014). This suggests that the determined pause time arises from another molecular mechanism. The pauses might reflect starting points, endpoints, or MVI unbinding events where the motor loses its hold to the underlying actin structures. This could happen either due to structural changes in the filaments, a filament crossing site, or another actin-binding protein blocking its track. In this scenario, the motor could be hold in its position by its bound cargo. This might be plausible since the measured diffusion coefficient reflects the diffusion coefficient of chromatin (Gasser, 2002), and MVI can directly and or indirectly bind DNA via transcription factors (Fili et al., 2017). The rebinding of the MVI complex might also be resembled by the regularly detected angled turns in the runs.

After analysis of the runs, the initial question had been whether the motion is directly executed by MVI or evoked by another process and what can be further learned about this biological mechanism. First, the treatment with TIP showed a clear effect on MVI motion and suggests an MVI related ATPase dependent motion mechanism. It inhibited single motion events by a factor of 0.6, which is comparable to a partial ATPase activity block that has been previously characterized in vitro (Heissler et al., 2012). Interestingly the same TIP concentration induced a complete inhibition of MVI triggered vesicular fusion events (Heissler et al., 2012). This was comparable to the observed block of chromosome reorganizations after hormone stimulation in the performed chromosome paint experiments. These results suggest a cooperative motion model in which several motion events facilitate larger structural changes. To further confirm and strengthen such a model, additional TIP titration studies could be performed.

Next, initially, MVI has been found to associate with RNA polymerase II (Vreugde et al., 2006), and it has also been shown that actin is part of the transcription preinitiation complex (Hofmann et al., 2004a). However, TPL treatment revealed that the MVI runs are independent of RNA polymerase II motion. The results propose that the found motion occurs upstream of the RNA transcription process. Furthermore, the motion requires the full-length motor since the artificially zipped mutant, that can solely move along actin filaments (Rock et al., 2001), did not show any significant in vivo movement, whereas the SAH mutant did. Also neither the
motor domain nor the cargo binding domain showed any relevant movements. Additionally, observed single bleaching steps suggest that more than one motor is involved in the motion process. Recent experiments showed that the binding of NDP52 or DAP2 is required for MVI dimerization via the CBD (Fili et al., 2020; Fili et al., 2017). Together with the finding that the full-length motor is required for motion, this indicates that at least MVI dimerization or even multimerization via the CBD is required for motion events. Further experiments analyzing different MVI mutants that cannot bind specific cofactors via their CBD, especially through the WWY and the RRL core motifs, will reveal more details about this process. Double point mutations in the CBD and the tail domain (SAH) might additionally give insights into the debated dimerization mechanism. Additionally, overexpression of DAB2, which blocks the DNA binding site (Fili et al., 2020), will reveal if the dimerization via the CBD is solely required for MVI motion or if it also requires chromatin binding. Overall, further live-cell experiments will reveal more about the highly discussed MVI dimerization mechanism (Fili and Toseland, 2020; J. A. Spudich and Sivaramakrishnan, 2010; Sweeney and Houdusse, 2010), that has also been found to be dependent on locally high motor concentrations (Sivaramakrishnan and Spudich, 2009).

MVI moves on nuclear actin filaments

In colocalization experiments, it was confirmed that nuclear MVI motion occurs along actin structures, and the quantification of binding events showed that MVI preferably binds to actin-rich regions. However, the adopted actin visualization technique possibly underestimates the actual density of nuclear actin structures since it only shows ionophore induced filaments. Further experiments colocalizing MVI with actin structures formed by other stimuli are required to draw a more detailed conclusion. Also, the current labeling and imaging approach might have missed potential thinner and more dynamic structures. Both experimental drawbacks might end up in an underestimation of the true MVI:actin colocalization value. Mechanistically it has been proposed that Ca\(^{2+}\) stimulation facilitates cargo binding, whereas bound CaM is required for stiffening the lever arm (Bahloul et al., 2004). This suggests that the application of the ionophore might induce MVI binding but concurrently impede MVI motion. It might be interesting to study both effects individually to gain a more
conclusive answer about this process. Previously it has been found that certain nuclear actin structures constitute and reside in interchromatin spaces (Baarlink et al., 2017). Future experiments, such as staining the chromatin with Hoechst 33342, might reveal more details about the nature of the underlying actin structures that are required for MVI motion.

MVI is required for transcription stimulated chromosome reorganization

Lastly, MVI motion events could be linked to hormone-stimulated transcription processes. Previously it was shown that MVI (Fili et al., 2017) and actin filaments (Q. Hu et al., 2008) are required for gene expression of ER target genes. Initially, it was discovered that MVI motion frequency increased upon hormone stimulation. Chromosome paint experiments complemented this finding by revealing that chromosomes 2 and 21, which carry the genes GREB1 and TFF1, respectively, and whose expression is MVI dependent (Fili et al., 2017), are reorganized upon β-estradiol stimulation. These findings suggest an active nuclear actin-related mechanism that is responsible for the establishment of long-range chromatin contacts. Further studies, for example, like labeling the chromatin by using histone markers (e.g., SNAPTag, Figure A.3), will reveal a more direct involvement of chromatin in this process. Such experiments could later be complemented by selective labeling of relevant individual chromatin loci (Ma et al., 2015). Finally, the performed microscopy experiments could be complemented by HiC studies to investigate hormone-dependent chromatin contacts under various MVI knockdown or nuclear actin disturbing conditions.

The found MVI motion events occurred with a frequency of 0.11/s, which transfer into 9600 events per cell cycle. This event rate suggests that MVI motion might also drive other biological relevant mechanisms. Apart from the found involvement in ER related transcription mechanism, it is interesting to speculate about other potential processes. One other experimental subject might be the directly related androgen stimulation process (Loikkanen et al., 2009). However, other nuclear processes that involve nuclear actin filaments like DNA repair (Belin et al., 2015; Caridi et al., 2018; Schrank et al., 2018b), chromatin reorganization during cell spreading (Plessner et al., 2015), or nuclear actin formation at mitotic exit (Baarlink et al., 2017) might be biological processes that be dependent on active MVI motion.
MVI has been found to work against forces of 3 pN (Rock et al., 2001). Since the required forces that is associated with cohesin and condensin chromatin looping are below 2 pN (Ganji et al., 2018), it is reasonable to speculate that MVI is able to reorganize chromatin structures. The direct labeling of chromatin structures with ferritin molecules might allow further studies of this force dependent process (Monzel et al., 2017). Direct labeling of MVI direct or indirect chromatin binding sites and local force application with a magnetic tweezer under hormone stimulating conditions might give insights on the number of MVI molecules in these reorganization processes.

Another field of study would be investigating nuclear actin nucleation processes that need to precede MVI motion, which, for example, involves transcription-related actin polymerization. RNA-PolII clusters formation is enhanced by N-WASP-Arp2/3 mediated nuclear actin filaments, which involves an actin polymerization process, which starts at the polymerase clusters (Wei et al., 2020b). Since MVI moves towards the minus end, which is the nucleation site of the filament, it is reasonable to speculate that MVI drives its chromatin cargo into the center of the transcription cluster, thereby enhancing the transcription rate of the target gene. In such a scenario, CHIPseq studies (Mardis, 2007) could be used to initially reveal MVI target sites, which can later be verified by single molecule colocalization studies. Further colocalization studies of MVI and nuclear ARP2/3 complexes might reveal more about the target site of the motion process and the tethering function of MVI that might be succeed the motion process (Hari-Gupta, Fili, dos Santos, et al., 2020).
A Appendix

Figure A.1: Influence of TIP and LatA on the cell morphology. (A) Widefield image of untreated HeLa cells and widefield images of cells treated with (B) 25µM of the MVI inhibitor TIP for 4h or (C) 100 nM of LatA for 15 min. The scale bar is 400 µm.
Figure A.2: "Chromosome PAINT of chromosome 2 (green) and 21 (magenta) in (A) untreated MCF7 cells (wt), (B) after hormone depletion by growing cells in charcoaled stripped FCS (cs) and (C) with subsequent stimulation of transcription with β-estradiol in presence of the transcription inhibitor TPL (csE2 + TPL) or (D) in presence of acting filament destabilizing drug Latrunculin B (csE2 + LatB). Scale bar is 10 µm." ((B,D) As author modified from Grosse-Berkenbusch et al., 2020. Original published under license CC BY-NC-ND 4.0)

Figure A.3: Colocalization of Halo-Tag-MVI (SiR, colored lines) and SNAP-Tag-H2B (JF-549, grey). Sequential imaging of both channels (2x-2x). The nucleus is indicated in dashed lines and the scale bar is 2 µm.
References


References


References

References


References


References


References


Sakamoto, T., Amitani, I., Yokota, E., & Ando, T. (2000). Direct observation of pro-
cessive movement by individual myosin V molecules. *Biochemical and Bio-
physical Research Communications, 272*(2), 586–590.

pressed in striated muscles moves into the myonuclei upon differentiation. *Journal of Molecular Biology, 326*(1), 137–149.


tion, 422*(April), 45–55.


cytosis in intact cells. *Journal of Cell Biology, 111*(6 I), 2307–2318.


velocity in living cells depends on the myosin-V lever arm length. *Journal of Cell Biology, 156*(1), 35–39.

Schrank, B. R., Aparicio, T., Li, Y., Chang, W., Chait, B. T., Gundersen, G. G., Gottes-
man, M. E., & Gautier, J. (2018a). Nuclear ARP2/3 drives DNA break cluster-

Schrank, B. R., Aparicio, T., Li, Y., Chang, W., Chait, B. T., Gundersen, G. G., Gottes-
man, M. E., & Gautier, J. (2018b). Nuclear ARP2/3 drives DNA break cluster-

Schubert, H. L., Wittmeyer, J., Kasten, M. M., Hinata, K., Rawling, D. C., Héroux, A., 


References

XPB, a subunit of TFIIH, is a target of the natural product triptolide. *Nature Chemical Biology*, 7(3), 182–188.


Acknowledgment

The Acknowledgment has been removed for reasons of confidentiality.
Acknowledgment

The Acknowledgment has been removed for reasons of confidentiality.
The CV has been removed for reasons of confidentiality.

Erklärung

Ich erkläre, dass ich die Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Tübingen, den .................................................................

Andreas Große-Berkenbusch