Dissertation

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

presented by

Diplom-Biologe Markus Ulrich
born in: Ludwigshafen / Rhein, Germany
Oral-examination: April 2008

Tropical –
Software for quantitative analysis of FRAP experiments

-Identifying the dynamics of linker histone H1°-

Referees: Prof. Dr. Roland Eils
Prof. Dr. Harald Herrmann-Lerdon

Acknowledgements

First I would like to thank Prof. Dr. Roland Eils for welcoming me in a wonderful, international and interdisciplinary group at the renowned German Cancer Research Center, DKFZ. Prof. Roland Eils steadily supported my work and ideas, particularly in times when it was most needed. I'd like to thank him for showing continuous trust into the success of my project, for helpful discussions and the many valuable contacts to scientists in this field of research.

Prof. Dr. Harald Herrmann-Lerdon deserves many thanks for support of this thesis, creative ideas and critical discussions concerning the experimental part of my work. I also thank him for the opportunity to use his cell culture lab.

Special thanks go to Dr. Joel Beaudouin, who instructed me and critically followed the development of my work towards this thesis. His knowledge and his eager interest in microscopy and physics helped me a lot to understand the broad impact of this field of research.

I also want to acknowledge Constantin Kappel, a Ph.D. student, whom I was working closely together with during my thesis. The many discussions about the quantitative interpretation and the thorough performance of FRAP experiments definitely contributed to my point of view on those topics. He also participated in the programming of Tropical.

Next I would like to thank Jochen Ulrich, who helped a lot with the programming of Tropical's graphical user interface during a two month internship at DKFZ and Dr. Stefan Hezel, who also contributed to the development and improvement of Tropical during his stay at DKFZ.

Many thanks also go to Sven Mesecke (Ph.D. student), Dr. Martin Bentele and Dr. Hauke Busch for fruitful discussions about (bio)physical and biochemical theory.

Sabine Aschenbrenner, Dr. Benedikt Brors, Stephanie Geiger, Christina Grosch, Karlheinz Groß, Rolf Kabbe, Petra Kühnle-Ried and Dr. Michaela Reichenzeller receive special thanks for largely contributing to a pleasant and enjoyable working atmosphere and constant help with the many little difficulties arising during a Ph.D. student's work.

I further like to thank my very good friends, Nicolas, Sven, Oliver, Adrian, Andreas, Harry and Moe who always put me back to more relaxing sides of life.

The most cordial thanks go to my parents and grandparents for their long-lasting support and financial sacrifices all along my studies and for supporting my decisions throughout my life.

Last but not least, this Ph.D. thesis would not have been possible without the support of my wife Kerstin. Her love and inspiration are my source of happiness and always picked me up during hard times. Without her vast support she offered although she is studying towards a Ph.D. degree herself, I would never have been able to finish writing in such a short time.

For my beloved father, Dr. med. Günter Ulrich, who died of lung cancer on February 21st 2008.

Zusammenfassung

Fluorescence recovery after photobleaching (FRAP) ist inzwischen eine Stardardmethode der Zellbiologie zur Visualisierung des dynamischen Verhaltens von fluoreszenzmarkierten Proteinen *in vivo*. Mittels FRAP wird die Fluoreszenzverteilung der markierten Proteine durch Photobleichen in einem bestimmen Bereich der Zelle gestört. Das Widerherstellen der Fluoreszenzverteilung wird mittels Laserscanning Mikroskopie aufgezeichnet. Jüngst wurden geeignete Methoden zur quantitativen Analyse der gewonnen Daten entwickelt, mittels welcher es möglich ist Diffusionskoeffizienten und Dissoziationskonstanten von beweglichen Molekülen, die dynamische Bindungen mit unbeweglichen oder beweglichen Bindungsstellen eingehen, abzuschätzen. Geeignete Software, zur Umsetzung dieser Methoden war jedoch bis heute nicht erhältlich.

Darum entwickelte ich Tropical, eine Sofware zur Simulation und Parameterschätzung basierend auf Reaktions-Diffusionsmodellen. Mittels räumlich-zeitlicher Mikroskopiebilder schätzt Tropical Reaktions- und Diffusionskonstanten von benutzerdefinierten Modellen. Auch inhomogene Molekülverteilungen können analysiert werden, was Tropical zu einem geeigneten Tool zur quantitativen Analyse von FRAP Experimenten macht.

Tropical wurde in dieser Arbeit verwendet, um das dynamische Verhalten von Linkerhiston H1° zu untersuchen. Dieses stabilisiert das Nukleosom, eine Struktur, die wichtig für die Komprimierung von DNA ist und daher eine bedeutende Aufgabe bei der dynamischen Organisation von Chromatin in eukaryotischen Zellen hat.

FRAP Experimente wurden mit Wildtyp Linkerhiston H1° und zwei Formen mit Mutationen an möglichen DNA Bindungsstellen durchgeführt. Mit Tropical wurden die Diffusionskoeffizienten der drei Formen geschätzt, basierend auf einem Diffusionsmodell unter Annahme sehr schneller Bindungsreaktionen. Das Modell zeigte eine sehr gute Übereinstimmung zu den experimentellen Daten. Die Stelle Lysin 52 hatte einen starken Einfluss auf das DNA-Bindungsverhalten von H1°. Das Resultat der Mutation dieser Stelle war ein dreifach erhöhter Diffusionskoeffizient. Die H1° Variante mit sechs verschiedenen Punktmutationen zeigte sogar einen 15-fach erhöhten Diffusionskoeffizienten, was auf einen noch größeren Einfluss auf das DNA-Bindungsverhalten deutet.

Die Verwendung von Tropical zur Untersuchung der Dynamik von H1° war, nach dem jüngst publizierten (Ulrich et al. 2006), ein weiterer Beweis für seine Eignung zur Analyse von FRAP. Die Vorteile von Tropical sind, dass (1) es direkt mit Mikroskopie Bildern arbeitet, (2) auch inhomogene Verteilungen berücksichtigt und (3) die gewonnen Ergebnisse direkt validiert werden können.

In der vorliegenden Arbeit wird zunächst der aktuelle Wissensstand der Chromatinorganisation in eukaryotischen (Ulrich et al. 2006) Zellen dargestellt und die Rolle von Histon H1 hierbei erklärt. Danach folgt eine Einführung in die verfügbaren Mikroskopietechniken zur Visualisierung von Proteindynamiken und in die mathematischen Methoden zu deren quantitativen Analyse. Im darauf folgenden Teil werde ich Tropical und die darin verwendeten Methoden detailliert darstellen und schließlich die Ergebnisse meiner Untersuchungen zur Dynamik von Linkerhiston H1° präsentieren und kritisch diskutieren. Auch die Eignung von Tropical zur Analyse von FRAP Experimenten sowie von FRAP zur Untersuchung von Proteindynamiken wird schließlich kritisch beleuchtet.

Abstract

Fluorescence recovery after photobleaching (FRAP) experiments using laser scanning microscopes to follow the *in vivo* dynamics of proteins tagged to fluorescent markers like the green fluorescent protein (GFP) has become a standard method in cell biology. FRAP perturbs the fluorescence distribution by photobleaching GFP-tagged proteins in a specific area of a cell and monitors the fluorescence redistribution. Adequate methods to quantify the results of FRAP experiments have recently became available. Those methods allow the extraction of diffusion coefficients and dissociation constants of proteins diffusing inside distinct cellular compartments and undergoing dynamic binding and dissociation with immobile or mobile binding sites. However, software incorporating such methods was not available until now.

Therefore I developed Tropical, a software for simulation and parameter estimation of reaction–diffusion models. Based on spatio-temporal microscopy images, Tropical estimates reaction and diffusion coefficients for user-defined models. Tropical allows the investigation of systems with an inhomogeneous distribution of molecules, making it well suited for quantitative analyses of microscopy experiments such as FRAP.

Tropical was used in this thesis to analyze the dynamic behavior of linker histone H1°, which plays a crucial role in the dynamic organization of chromatin by stabilizing the nucleosome, a structure involved in DNA packing in eucaryotic cells.

FRAP experiments were performed using three forms of linker histone H1°, the wild type, and two forms with mutated sites, that are likely to play a major role in DNA binding. Diffusion coefficients on the three forms were estimated with Tropical by fitting a pure diffusion model, assuming binding to happen instantaneously. The model showed a very good fit to the experimental data. It could be shown that lysine 52 significantly influences the DNA binding properties of H1° and its mutation resulted in a 3-fold enhanced diffusion coefficient. The H1° form containing six point mutations however showed an even higher diffusion coefficient, about 15 times higher than the one of the wild type histone, revealing a much larger contribution to DNA binding of these six mutated sites.

Using Tropical to estimate the diffusion coefficients of linker histone H1° was another proof for the power and functionality of Tropical, besides the recently published one (Ulrich et al. 2006). Tropicals' main advantages are (1) that it directly operates on microscopy images, (2) an inhomogeneous distribution of binding partners can be considered and (3) the obtained result can directly be verified.

This thesis will first line out the current knowledge of eucaryotic chromatin organization, to clarify the role of linker histone H1. I will then give an overview of microscopy techniques available to reveal protein dynamics and their quantitative analysis using mathematical models. The next part will explain Tropical and its implemented methods in detail. Finally I will present the results obtained on the dynamics of H1° and critically discuss the applicability of Tropical to analyze FRAP data and FRAP as a method to reveal protein dynamics.

1 Introduction

The nucleus is the organelle that defines eukaryotes. It confines the DNA in the compacted form of chromatin inside it. The organization of chromatin and the mechanism of DNA compaction have been the purpose of investigation for many years. Genome sequencing has given and still is giving new insight into our knowledge of chromatin primary structure and the role of its sequence to code for proteins.

However how the processes of DNA compaction and chromatin organization influence gene expression and are coordinated in time and in space inside the nucleus is still not understood. Although we know the proteins, namely histones, involved in these processes, the way they organize and interact with their target, DNA, *in vivo* is much less clear.

Nevertheless electron and light microscopy techniques have recently started clarifying these issues.

In this introduction, I will present the current knowledge of the organization of chromosomes and the role of histone proteins in this context. I will especially focus on the role of the linker histone H1, which has a central role in stabilizing chromatin structure. Focus of this thesis was the dynamic behavior of linker histone H1 observed by light microscopy. Therefore Light microscopy techniques to characterize these dynamic properties will also be introduced.

Methods for quantitatively interpreting the results obtained with such microscopy techniques are already available, but software implementing these methods is still missing. I developed a software called Tropical that incorporates the most recent methods to quantify the dynamic behavior of proteins on the basis of microscopy images. In this introduction I will therefore describe the quantification methods available for microscopy data and give a brief introduction into the mathematics of diffusion of proteins.

1.1 Eucaryotic chromosome structure

Eucaryotic organisms have elaborate ways of packaging DNA into chromosomes as the total length of eukaryotic cellular DNA is up to a hundred thousand times a cell's length. In a single human cell, for example, the DNA measures about two meters in total length but must be contained within cells with diameters of less than 10 μm, a compaction ratio of greater than 10^5 (Horn and Peterson 2002; Razin et al. 2007). Therefore the packing of DNA in a way that it is still functional is crucial to cell architecture.

During interphase, when cells are not dividing, the genetic material exists as a nucleoprotein complex called chromatin, which is dispersed through much of the nucleus. Further folding and compaction of chromatin during mitosis produces the visible metaphase chromosomes, whose morphology and staining characteristics were detailed by early cell biologists (Lodish et al. 2003). This compression is performed by proteins that successively coil and fold the DNA into higher and higher levels of organization.

1.1.1 Nucleosomes, Histones and the packing of DNA

The proteins that bind to DNA to form eucaryotic chromosomes are traditionally divided into two general classes: the histones and the non-histone chromosomal proteins. The complex of both classes of protein with the nuclear DNA of eucaryotic cells is known as chromatin (Alberts et al. 2002).

The general structure of chromatin has been found to be remarkably similar in the cells of all eukaryotes, including fungi, plants, and animals. The most abundant proteins associated with eukaryotic DNA are histones, a family of small, basic proteins present in all eukaryotic nuclei.

Histones are present in such enormous quantities in the cell (about 60 million molecules of each type per human cell) that their total mass in chromatin is about equal to that of the DNA. Histones are responsible for the first and most basic level of chromosome organization, the nucleosome.

The five major types of histone proteins - H1, H2A, H2B, H3, and H4 - are rich in positively charged basic amino acids, which interact with the negatively charged phosphate groups in DNA. The amino acid sequences of the so called core histones H2A, H2B, H3, and H4, building the core of the nucleosome, are remarkably similar among distantly related species. The amino acid sequence of the linker histone H1 which is connected to the linker DNA (free DNA between two nucleosomes) varies more from organism to organism (Lodish et al. 2003).

Association of eukaryotic nuclear DNA with histone proteins and chromatin formation

Figure 1.1 gives a schematic overview of the principle mechanism of DNA packing to form the chromosomes located in the cell nucleus: rolling of DNA onto nucleosomes, which are connected via linker DNA to "beads-on-a-string", compaction of nucleosomes with formation of the so-called 30-nm fiber (first discovered by Finch and

Klug (1976)), and folding of the latter into giant (50-200 kbp) loops by scaffolding proteins, fixed onto the protein skeleton, the nuclear matrix (Horn and Peterson 2002; Razin et al. 2007).

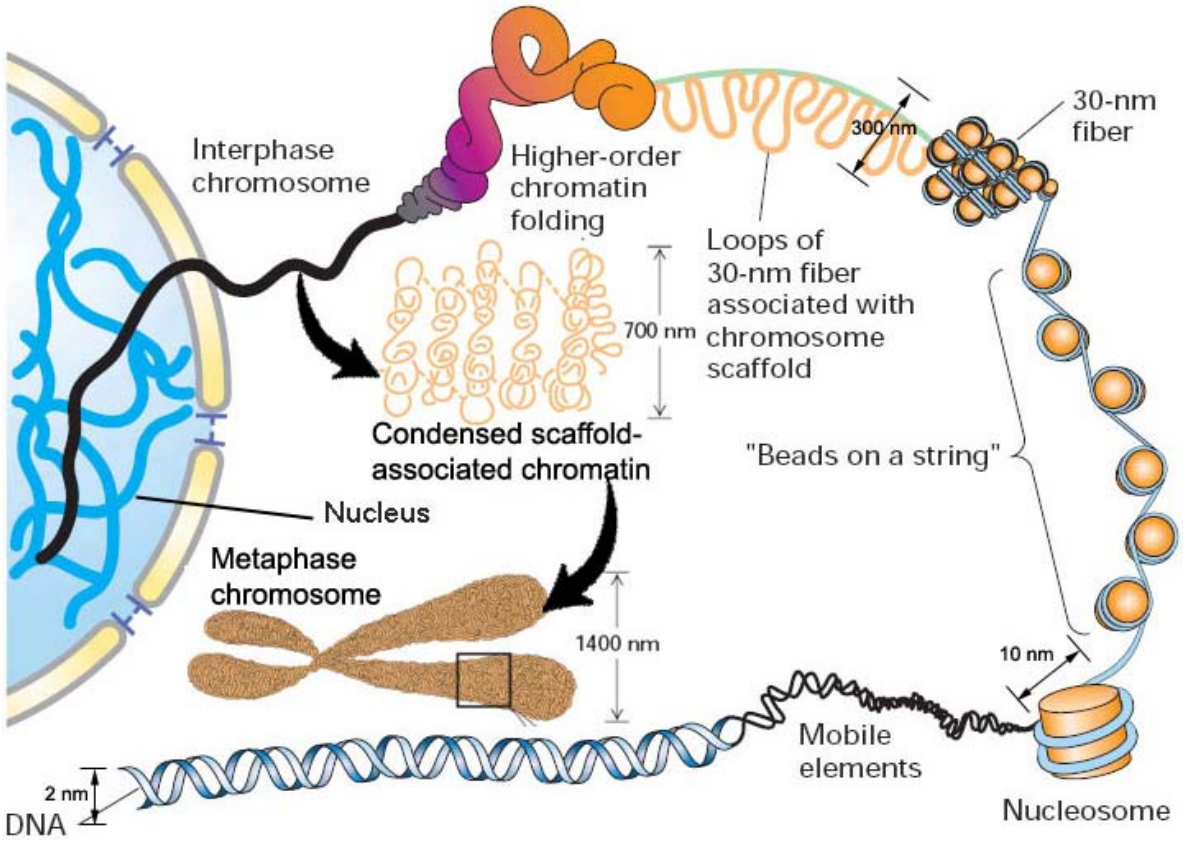


Figure 1.1 Overview of the DNA packing mechanism. The double-stranded DNA molecule is packed to form the nucleosome which is further compressed to a beads-on-a-string form. In a second step the beads-on-a-string are packed to form a 30-nm fiber, which is the basis for chromatin formation. Adapted from (Lodish et al. 2003) and (Horn and Peterson 2002).

Eucaryotic nucleosomes are about 10 nm in diameter and consist of a protein core with double-stranded DNA wound around it in ~ 1.65 turns, which equals 146-147 nucleotide pairs. The nucleosome core particle consists of a complex of eight histone proteins (octamer), two molecules each of the histones H2A, H2B, H3, and H4 (Figures 1.2 and 1.3) (Horn and Peterson 2002). The high-resolution structure of the nucleosome core particle, solved in 1997 by Luger et al. (1997), revealed a disc-shaped histone core with DNA tightly wrapped around it in a left-handed coil. In assembling a nucleosome, the histone folds first bind to each other to form H3-H4 and H2A-H2B dimers. The H3-H4 dimers further combine to form tetramers. An $(H3\text{-}H4)_2$ tetramer then unites with two H2A-H2B dimers to form the compact octamer core, around which the DNA is wound (Figures 1.2 and 1.3) (Razin et al. 2007).

The beads-on-a-string represent the first level of chromosomal DNA packing. *In vitro* observations by electron microscopy showed isolated chromatin as a series of "beads-on-a-string". The string is free DNA called linker DNA, and each bead is a nucleosome core particle.

The mechanism for forming the 30-nm fiber is not understood in detail. One model, the solenoid model (Finch and Klug 1976), suggests that nucleosomes are packed into an irregular spiral or solenoid arrangement, with approximately six nucleosomes

per turn (Robinson et al. 2006; Tremethick 2007). The linker histone H1 is thought to be bound to the DNA on the inside of the solenoid, with one H1 molecule associated with each nucleosome (Figure 1.2).

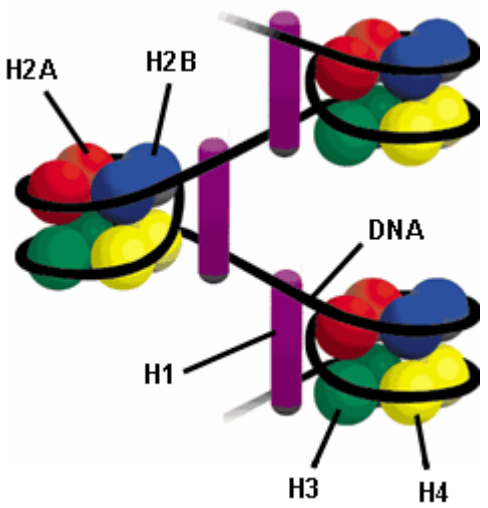


Figure 1.2 Section of the solenoid model of nucleosome formation. The histone octamers are formed of two copies of each histone H2A, H2B, H3 and H4, around which the DNA is wound. One histone H1 is bound to the linker DNA on the inside of each nucleosome. From (Pennington)

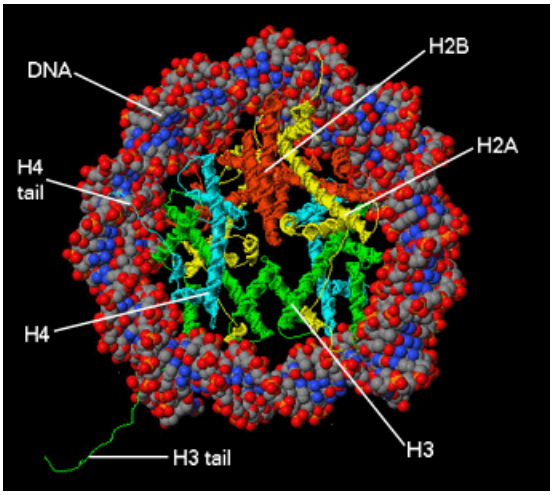


Figure 1.3 3D-structure of the nucleosome complex based on x-ray crystallography. The double stranded DNA is wound around the interlocking histone subunits (H2A, H2B, H3 and H4) giving the nucleosome core complex a disc like shape. (www.jmol.org 2007)

The current state of knowledge indicates that the globular domain and the C-terminal tail of histone H1 bind to the nucleosome surface, contacting both DNA and protein at the entry and exit points of the DNA (Figure 1.4A) (Hendzel et al. 2004; Robinson et al. 2006). Although the detailed mechanism, stoichiometry and role of H1 in the formation of the 30-nm fiber is not understood (Robinson et al. 2006; Woodcock et al. 2006), H1 is needed to stabilize rather than form the 30-nm fiber (Thomas 1999; Bustin et al. 2005; Woodcock et al. 2006).

A second mechanism for forming the 30-nm fiber is proposed to involve the tails of the core histones, which extend from the nucleosome (Figure 1.3). It is thought that these tails may help attach one nucleosome to another thereby allowing a string of them, with the aid of histone H1, to condense into the 30-nm fiber (Robinson et al. 2006). This second model, the zigzag two-start helix model of the 30-nm fiber (Woodcock et al. 1993) proposes that nucleosomes are arranged as a zigzag such that two rows of nucleosomes form, and the linker DNA, which is essentially straight, criss-crosses between each stack of nucleosomes (Figure 1.4B).

Loops of the 30-nm chromatin fiber, 50-200 kbp in length, have been found to associate with a flexible chromosome scaffold, yielding an extended form characteristic of chromosomes during interphase. Folding of the scaffold has been proposed to produce the highly condensed structure characteristic of metaphase chromosomes (Figure 1.1) (Lodish et al. 2003; Razin et al. 2007).

Numerous in vitro experiments have demonstrated that the interaction of H1 with nucleosome stabilizes the compact, higher-order chromatin structure and inhibits DNA-dependent activities such as transcription and replication. Histone H1 restricts nucleosome mobility, impedes the ability of regulatory factors to access their chromatin

targets and inhibits the action of chromatin-remodeling complexes (Zlatanova et al. 2000; Catez et al. 2006). The misregulation of the nucleosome core structure and function leads to a range of diseases, particularly cancer (Hoch et al. 2007).

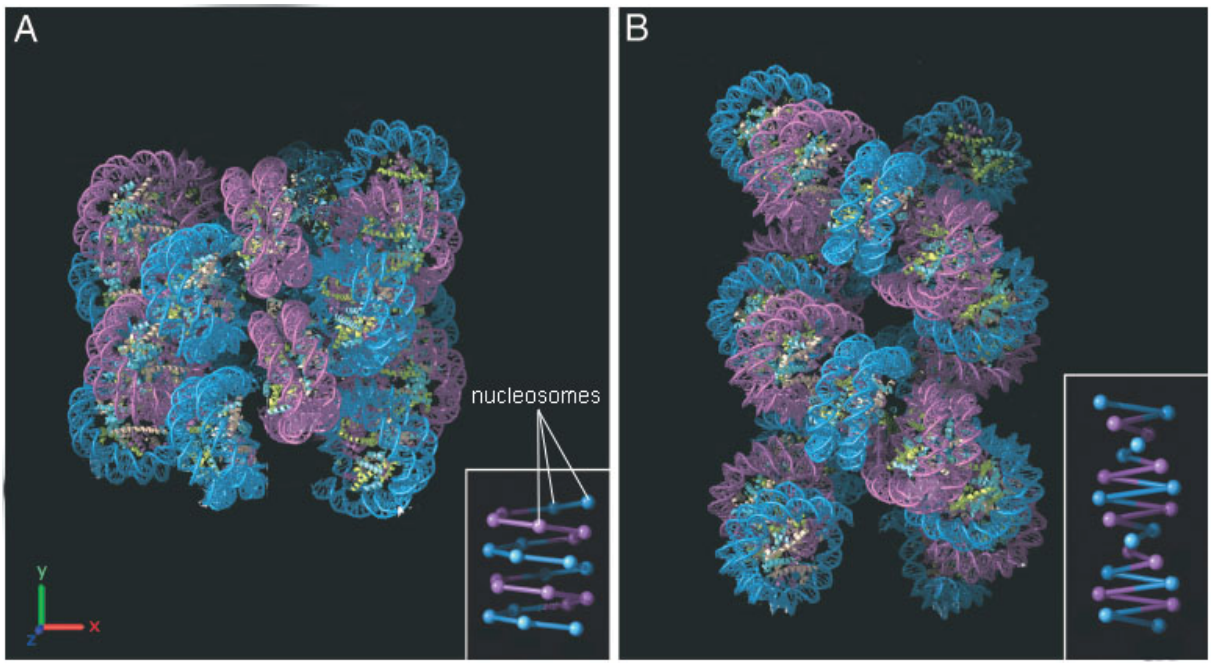


Figure 1.4 Two possible molecular models of the 30-nm fiber. Alternate nucleosome pairs are colored in blue and purple. **A.** The solenoid one helix model. Linker histone H1 is the driving force of packing 22 nucleosomes to a 33-nm structure. **B.** The zig-zag two helix model. Tails of the core histones together with linker histone H1 are packing the 22 nucleosomes together. This structure has a diameter of 28.4 nm. Adapted from (Robinson et al. 2006)

1.1.2 Histone dynamics

Based on results from *in vitro* observations the nucleosome was initially thought to be a static, highly regular packaging system. However, more recent *in vivo* studies revealed it to be a very dynamic structure (Phair and Misteli 2001).

The majority of nuclear proteins examined so far diffuses rapidly in the nucleoplasm and typically shows a fast exchange with their binding sites (Houtsmuller et al. 1999; Phair and Misteli 2000). This dynamic behavior is thought to play a major role in chromatin organization and plasticity (Misteli 2001; McBryant et al. 2006). Because the packaging of chromosomes must be accomplished in a way that allows rapid localized, on-demand access to the DNA, understanding the dynamics of the histones is of great importance.

Whereas core histone seem to exchange quite slowly and can still be considered rather static compared to most nuclear proteins (Verschure et al. 2003), linker histone H1 shows an extremely dynamic behavior. Since it has key role as a gate keeper to access the DNA of the nucleosome, understanding the diffusion and binding behavior of linker histone H1 is of particular interest for understanding the nucleosome.

In the following, I will give a short overview on the *in vivo* dynamics of core histones and explain the dynamic behavior and role of linker histone H1 in more detail.

Dynamics of core histones

The process of nucleosome core assembly from purified proteins is thought to be similar to that observed in living cells. The $(H3–H4)_2$ tetramer binds first and positions the central portion of the DNA wrapped in the nucleosome core. The H2A–H2B dimers bind to surfaces on each side of the tetramer, completing the two helical ramps that position the superhelical turns of DNA. This sequential order could be shown by Hoch et al. (2007) for the disassembly of the nucleosome core. The authors could also confirm the results found by Kimura and Cook (2001), indicating that the inner core of the nucleosome ($(H3\text{-}H4)_2$ tetramer) is much more stable. About 80% of its population is practically immobile, showing a half-time of recovery ~510 min. The surface histones (H2A-H2B) exchange more rapidly showing half-times of recovery of 6 to 130 min, with only 53% of the population being immobile. By perturbing the system with different salt concentrations and monitoring the distance between the H2A-H2B dimers and the $(H3\text{-}H4)_2$ tetramers *in vivo*, they found that under physiological conditions, H2A–H2B dimers are able to undergo rapid exchange between partially assembled nucleosome core particles and form unstable DNA/H2A–H2B complexes. A second important finding of their studies was that a salt-induced disassociation of the nucleosome was reversible, indicating a dynamic system.

The dynamic behavior of core histones were intensively investigated by Bhattacharya et al. (2006) using fluorescence correlation spectroscopy (FCS) and fluorescence recovery after photo bleaching (FRAP) techniques. These and related microscopy techniques will be explained in detail in chapter 1.2. The observed recovery times of core histones (H2A, H2B, H3 and H4) (half-times of about 130 minutes) suggest that core histones stay bound to chromatin for a much longer time than linker histones (half-times in the range of a few minutes, as we will see later) and even most other nuclear proteins (Phair and Misteli 2000; Brown 2003). However there is a difference in exchange rates between the H2A-H2B dimer and the H3-H4 tetramer (Kimura and Cook 2001). While most H3 and H4 stay bound to chromatin for a longer time ($t_{1/2}$ about 130 min), at least a small fraction (~ 3%) of H2B seems to exchange very rapidly ($t_{1/2}$ about 6 min).

Bruno et al. (2003) showed that the exchange of H2A-H2B dimers bound to chromatin could be ATP dependent. On contrary, linker histone mobility always occurs on a shorter timescale and is ATP independent (Bhattacharya et al. 2006).

Linker histone H1

A primary function of chromatin is compaction of DNA. This must be done such that the underlying DNA is potentially accessible to factor mediated regulatory responses. Linker histone H1 plays a key role in this regulatory process. Lever and colleagues (2000) have shown that histone H1 is dynamically associated with chromatin in living cells. Their results indicate that histone H1 exchange occurs through a pathway that involves dissociation, diffusion through the nucleoplasm, and reassociation with chromatin in a stop-and-go manner (Brown 2003).

Since the linker histone H1 is believed to play a key role in chromatin organization by stabilizing higher-order chromatin structure and several proofs exist that H1 is not statically bound to chromatin (Lever et al. 2000; Misteli et al. 2000; Hendzel et al. 2004; Beaudouin et al. 2006), determining the dynamic properties of histone H1 *in*

vivo is central to understanding chromatin functions. Furthermore the influence of histone H1 in cellular core processes via its chromatin binding properties is of great impact. It is, for example, directly involved in the regulation of specific genes (Strahl and Allis 2000; Bustin et al. 2005; Catez et al. 2006) as it prevents the access of transcription factors and chromatin remodeling complexes to DNA (Zlatanova et al. 2000).

Histone H1 is larger than the core histones and is considerably less well conserved. In fact, the cells of most eucaryotic organisms synthesize several histone H1 proteins of related but quite distinct amino acid sequences. To date, ten H1 homologous proteins have been described in mammals, which can be found in the NCBI protein database: H1° (H1.0 or H5 in birds), H1.1, H1.2, H1.3, H1.4, H1.5, H1t, H1Foo, H1x and HILS1. The members of the H1 histone family can be classified into three groups: the main class subtypes expressed in somatic cells, the developmental- and tissue-specific subtypes, and the replacement subtype H1°. Four of the five main class subtypes, H1.2, H1.3, H1.4 and H1.5, are present in all somatic cells, whereas the fifth, H1.1, is restricted to thymus, testis and spleen, and possibly lymphocytic and neuronal cells (Rasheed et al. 1989; Franke et al. 1998). The sixth subtype is the H1 replacement subtype H1°. Its expression is replication-independent and it is mainly restricted to cells that are arrested in proliferation or are terminally differentiating (Khochbin and Wolffe 1994; Zlatanova and Doenecke 1994). Three H1 subtypes are highly developmental and tissue specific. These are the male germ-cell specific H1t (Doenecke et al. 1997), H1Foo, which is expressed in the growing oocyte, the zygote and the very early embryo (Tanaka et al. 2001) and the spermatoid-specific H1-like protein HILS1 (Yan et al. 2003). The least knowledge exists about the tenth H1 subtype, H1x (Happel et al. 2005). At least two more H1 forms were described for *Caenorhabditis elegans* (Nematoda) named putative histone H1.6 and H1Q (Jedrusik and Schulze 2001).

All known H1 variants share a common domain structure: a short N-terminus, a central globular domain and a long C-terminal domain (Figure 1.5). The globular domain contains a helix-loop-helix motif. The C-terminus, in contrast, appears unstructured and is highly negatively charged (Allan et al. 1980; Hendzel et al. 2004).

Although the exact position of H1 on nucleosomes in vivo is still unknown (Becker et al. 2005), both the globular domain and the C-terminus appear to be essential for efficient binding and normal dynamic exchange (Goytisolo et al. 1996; Misteli et al. 2000; Khochbin 2001; Brown et al. 2006).

Lu and Hansen (2004) performed *in vitro* biochemical assays with H1° mutants having up to 24 amino acids of their C-terminal domain truncated. Their results indicate that the ability of H1° to alter linker DNA conformation and stabilize condensed chromatin structures is localized to specific C-terminal subdomains, rather than being equally distributed throughout the entire C-terminal domain. They propose that the functions of the linker histone C-terminal domain in chromatin are linked to the characteristic intrinsic disorder of this domain.

As shown by Hendzel and his coworkers (2004) this may also be true *in vivo.* They identified Tyr152 to significantly contribute to high-affinity binding of histone H1.1 to chromatin. They further found that phosphorylation can disrupt binding by affecting the secondary structure of the C-terminus. In their FRAP studies, when the GFP tag was placed on the C-terminus, the recovery rate was faster than when the tag was at the N terminus. The reduction of binding caused by the C-terminal fusion was almost

as great as that seen when the N-terminal domain was deleted. Deletion of the N terminus significantly reduced the binding affinity of the histone H1.1. When the C terminus was deleted, the GFP histone H1.1 deletion protein did not bind well and recovered at rates approaching diffusion. Analysis of recovery times of a Tyr152 point mutation identified this site as the most important of the mutated ones in H1.1 recovery.

Bhattacharya et al. (2006) examined histone H1.1 mobility using FCS and FRAP (see chapter 1.2 for details). They estimated a diffusion coefficient of ~0.3 $\mu m^2/s$ possibly arising due to H1.1-EGFP interaction with DNA. After deletion of the complete N- and C-terminal tails of H1.1, which is more than half of the complete molecule, the diffusion coefficient was estimated to be ~20 $\mu m^2/s$.

Several studies indicate that the linker histone H1 acts as a modulator of chromatin function. However, they fail to pinpoint a single distinct function for which a specific H1 variant is indispensable. For some cellular processes different H1 variants work in a redundant manner and the overall amount of H1 seems to be important rather then the amount of a specific variant (Bustin et al. 2005).

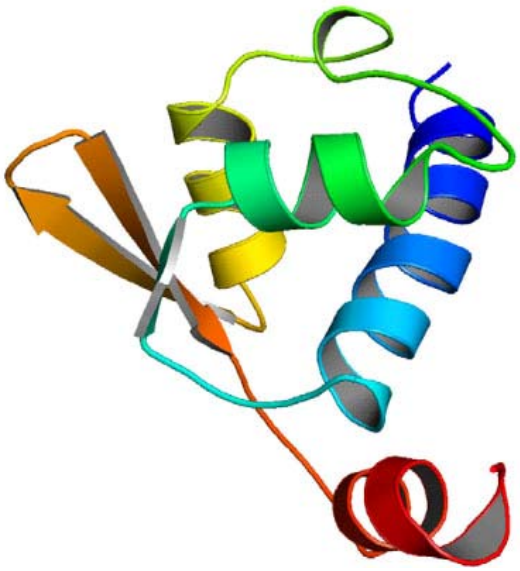


Figure 1.5 3D structure of histone H1. The N-terminus is shown in red, C-terminus in blue, globular helices in yellow and green.

The influence of different chromatin binding sites of histone H1° on its diffusion and binding behavior has been investigated for the C-terminal and the N-terminal tail as well as for the globular domain (Hendzel et al. 2004; Catez et al. 2006).

Investigations of the dynamic properties of different mutants of H1° by FRAP have been performed by Brown et al. (2006). There, the basic residues in the previously proposed primary (Lys69, Lys73 and Lys85) or secondary (Lys40, Arg42, Lys52 and Arg94) DNA binding site or both were simultaneously replaced with alanine, and the binding of the mutants was determined by FRAP. The multi-site mutants showed very rapid recovery after photobleaching, indicating severely compromised binding. The half-time for recovery (t_{50}) was used as an indicator of binding strength and was less than 1 s for all mutants but more than 50 s for the wild-type protein. For comparison, the half time for nonbinding GFP alone was on the order of 380 ms under their condi-

tions. To determine the contributions of individual residues to the recovery, stable cell lines expressing constructs containing the single point mutations of each residue were generated and analyzed by FRAP. Each mutant showed a recovery rate that was considerably faster than that of the wild-type protein. The strongest contributions came from Lys69 and Lys73 (t_{50} ~4 s and 8 s, respectively), with intermediate contributions from Arg42 and Arg94 (t_{50} ~16 s and 17 s, respectively) and relatively weak contributions from Lys40 and Lys52 (t_{50} ~39 s and 28 s, respectively). Nine mutants had dramatically faster recovery rates with t_{50} < 20 s. This group included the previously identified lysines at positions 69, 73 and 85 in the putative primary interaction site and arginines 42 and 94 in the putative secondary site, as well as His25. However, the data also identified three residues, Arg47, Arg74 and Lys97, that have not previously been implicated as contributing to nucleosomal binding. In all cases, replacement with glutamic acid further increased recovery speed. In contrast, replacement with another basic residue resulted in slower recovery than that observed after the neutral mutation, close to wild-type recovery. From their findings Brown and coworkers (2006) conclude that there are two distinct binding sites in H1°: The larger, site 1, comprises His25, Arg47, Lys69, Lys73, Arg74 and Lys85; the smaller, site 2, comprises Arg42, Arg94 and Lys97. All other sites, they claim, do not contribute to H1° binding. This led them to propose a structural model of H1° binding to DNA with the two proposed binding sites oriented towards the nucleosomal DNA and the remaining sites away from it.

Misteli et al. (2000) also looked at the recovery times of fluorescently labeled linker histone. In their FRAP experiments with H1-GFP constructs it was noted that, in both euchromatin and heterochromatin, the fluorescence signal was not completely recovered. This behavior indicates the presence of two distinct kinetic pools of H1 in the nucleus: a large mobile pool, which represents the continuously exchanging molecules, and a smaller, less mobile pool, which displays a residence time on the order or hours. Based on the recovery time they found that most labeled H1 molecules are bound to chromatin, but are continuously exchanged between chromatin regions. These observations led to the conclusion that most of the observed H1 molecules bind chromatin by a `stop-and-go' mechanism: a histone H1 molecule resides on chromatin for ~220 s. The molecule then dissociates and rapidly binds to another binding site (Lever et al. 2000; Misteli et al. 2000).

Although Misteli et al. (2000), Lever et al. (2000), Catez et al. (2006) and others propose that diffusion can be neglected when looking at H1° dynamics because the residence time of H1° being bound to chromatin is much longer than its free diffusional state, Beaudouin et al. (2006) as well as Sprague et al. (2006) could clearly show that relying on recovery times to ignore diffusion is not sufficient to characterize the dynamic behavior of nuclear proteins. Their investigations on histones H1.1, H1.2, H1.3, H1.4 and H1.5 clearly show that the dynamic properties of these linker histones are limited by the amount of free diffusive protein and not by the short residence time of the bound proteins. Individual proteins thus diffuse locally and bind to close binding sites, rather than diffusing globally before rebinding at random nuclear positions.

Concluding these findings it becomes obvious that the dynamic nature of diffusion and binding is an essential feature of linker histones in their functions as regulators of chromatin remodeling and chromatin structure in vivo. However, the kinetic properties in all studies were derived only from fluorescent recovery half-times and not based on a thorough analysis of diffusion and binding properties. Therefore quantita-

tively investigating the reaction limiting diffusion properties of H1° is crucial for understanding its complex dynamic behavior.

1.2 Fluorescence microscopy to reveal histone dynamics

Until the recent development of fluorescent proteins for use in cell biology, studies relating functions of nuclear proteins had been limited to *in vitro* reconstituted systems and electron or light microscopy in fixed cells. Investigating the dynamic behavior of proteins in living cells was therefore almost impossible.

Today large numbers of proteins are being followed in living cells by using a fluorescent marker called green fluorescent protein (GFP) or its variants. The availability of GFP as a fluorescent protein marker was a breakthrough in cell biology. GFP and other fluorescent dyes enabled the thorough investigation of protein dynamics *in vivo* for the first time.

The list of fluorescent proteins available nowadays is large and growing. Several dyes of different color for different purposes and methods have recently become available (Lippincott-Schwartz and Patterson 2003; Miyawaki et al. 2003).

Additionally, continuing development of microscopy techniques like wide-field, confocal and multi-photon microscopy have evolved a number of methods to visualize protein dynamics and interactions in living cells (Bastiaens and Squire 1999; Lippincott-Schwartz et al. 2001).

In the following part I will give a short overview about the use of fluorescent proteins and microscopy methods. In particular I will explain the method I used to analyze histone H1 diffusion dynamics, namely fluorescence recovery after photobleaching, short FRAP.

1.2.1 Fluorescent proteins

A good overview on available fluorescent dyes and their application in labeling of proteins, including typical problems occurring with specific dye-application combinations is given by Miyawaki et al. (2003).

Green fluorescent protein (GFP, Figure 1.6), a relatively small protein of 27 kDa was originally isolated from the light-emitting organ of the jellyfish *Aequorea victoria* (Shimomura et al. 1962) (reviewed in (Shimomura 2005)). Still more than 30 years passed before the complementary DNA encoding the protein was subsequently characterized by Chalfie et al. (1994). As *Aequorea* GFP is spontaneously fluorescent, chimeric GFP fusions offer the great advantage that they can be expressed *in situ* by gene transfer into cells. In addition, these GFP fusions can be localized to particular sites within the cell by appropriate targeting signals (Miyawaki et al. 2003). Numerous spectral variants with blue, cyan and yellowish-green emissions have been successfully generated from the *Aequorea* GFP. The discovery of novel GFP-like proteins from *Anthozoa* (coral animals) has significantly expanded the range of colours available for cell biological applications to approximately 30 significantly different members.

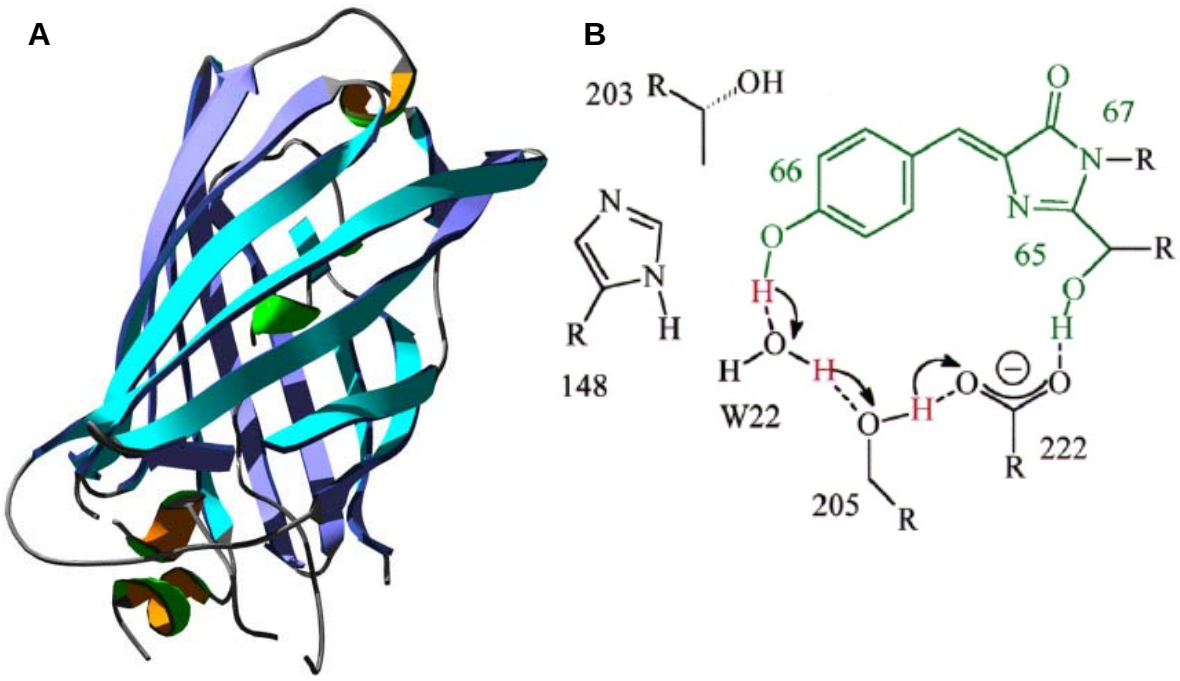


Figure 1.6 Structure of GFP. The relatively small protein of 27 kDa was originally isolated from the light-emitting organ of the jellyfish *Aequorea Victoria.* The 3D structure shows that GFP has a rather rigid globular domain. The tertiary structure has the chromophore (green) in its centre, shielded from the outer environment. In the secondary structure chromophore (p-hydroxy-benzylidene-imidazolinone, green line) is shown with some important amino acids. Secondary structure adapted from (Jung et al. 1998).

Depending on the purpose and application of the study a specific fluorescent dye will be chosen. N-terminal or C-terminal fusion may be impossible for some applications or very small tags may be needed. For such studies highly specialized fluorescent markers like circular or monomeric GFP, small organic dyes like fluorescein, FlAsH and ReAsH have been developed recently. For applications that require very long image acquisition times quantum dots, due to their high photostability and resistance to photodamage are the markers of choice, because they can be imaged over very long times. Other applications like FRET (fluorescent resonance energy transfer) require special pairs of dyes with overlapping spectra (e.g. CFP-YFP, GFP-DsRed) (Bastiaens and Squire 1999). New methods could be applied by the development of photoactivatable fluorescent markers like Kaede, DRONPA or PA-GFP, which can be "switched on" by a laser pulse of a certain wavelength (Patterson and Lippincott-Schwartz 2002; Ando et al. 2004; Chapman et al. 2005).

An early improvement of the original GFP was enhanced GFP (EGFP). It was developed to fulfill the needs of cell biologists working with mammalian cells: accelerated fluorophore maturation, correct folding at 37 °C and enhanced brightness (Lippincott-Schwartz and Patterson 2003). Features that make EGFP an ideal fluorophore for live cell imaging studies include its high quantum yield, its low tendency of photobleaching as well as its relative photostability.

Tagging proteins with GFP is as simple as attaching the gene for GFP to one end of the gene that encodes the protein of interest. Cells are transfected with the hybrid gene, which is then expressed in a cell so that the hybrid protein can be localized by its fluorescence. In many cases, the resulting GFP fusion protein behaves sufficiently

equal to the original protein, and its localization, movement and dynamic behavior can be monitored and visualized by following its fluorescence inside the living cell. Fusion of GFP to another protein is important to control. Because transfecting cells with fluorescently tagged proteins may never be possible without disturbing cells at all. The fusion may disturb the functionality of the protein and the overexpression may perturb the cell as the protein is present in much higher amount than normal. Checking the functionality of the tagged protein, its localization and the vitality of the transfected cells is therefore obligatory, whenever new fusion proteins are created.

Live cell imaging using GFP or its variants has become the standard way to determine the distribution and dynamics of any protein of interest in living cells.

1.2.2 Methods for fluorescence microscopy

Variant methods have recently been developed to apply fluorescence microscopy to reveal processes inside living cells. In this chapter, I will give a short overview of the most important ones to study protein dynamics, namely 4D microscopy, FCS and FRAP.

In **4D microscopy**, time-lapse observations of fluorescent molecules are collected as three dimensional data sets. This can be achieved with microscopes imaging different focal plains one after the other. In this way, information from an entire cell can be analyzed to provide spatial and temporal information about changes in a protein's distribution as it relates to complex cellular processes over time (Lippincott-Schwartz and Patterson 2003). However, scanning different focal plains is not a very fast process on most available microscopes and speed, besides accuracy of the *z* position change, is perhaps the most important consideration for 4D imaging. Thereby 4D microscopy rapidly generates enormous amounts of data, which can be interpreted with computer-based visualization programs that allow quantification and discrimination of fluorescence signals (Gerlich et al. 2001; Eils and Athale 2003). Time-lapse imaging can provide information about the distribution of a population of a protein over time but it does typically not reveal the kinetic properties of individual molecules, for example, whether the protein is immobilized or free to diffuse. This is due to the fact that when tagging a protein with a fluorescent dye and transfecting cell with it, what usually is recorded by the microscope represents the complete population of labeled proteins, not just a few or even single molecules. Even when after selection stable cell lines expressing the protein of interest are created, the amount of tagged protein is still so high, that it represents a protein population. Therefore what is seen on the microscope screen is the dynamics of a population, which is the steady state – the combined dynamics of many proteins observed at the same time and place.

In **FCS** (Fluorescence Correlation Spectroscopy), the fluctuations in the fluorescence intensity from fluorescently labeled molecules diffusing in and out of a small defined focal volume (~1 femtoliter) are measured over short periods of time. These fluctuations are recorded and analyzed by autocorrelation to reveal information about the protein's diffusion coefficient, binding constants, and concentrations (Bacia and Schwille 2003; Stephens and Allan 2003; Chen et al. 2006). Because the measured signal in an FCS experiment constitutes fluctuations from the time averaged fluorescence intensity, the lower the absolute number of molecules in the sample volume the greater the measured signal. This is accomplished not only by defining a very small observation volume but also by analyzing very dilute samples (~nM) such that

only very few molecules are present in the observation volume at any one time. The sample observation volume is in most cases defined by focusing a laser to a diffraction-limited spot in the sample plane. In general, FCS can measure diffusion coefficients over five orders of magnitude ranging from diffusion of small molecules in solution (~300 $\mu m^2/s$) to the lateral diffusion of membrane proteins within cell membranes (~0.01 $\mu m^2/s$) (Hess et al. 2002; Chen et al. 2006). However, FCS is best suited for measurements of particles with very low concentrations, diffusing rapidly (Bates et al. 2006). Whereas photobleaching does not play a major role in FCS when observing fast diffusing molecules or molecules moving even faster than diffusion, because only a small volume is illuminated and hence, the exposure to light is very short. This is totally different for slower movement, where single molecules might reside for a longer period exposed to light. In this case photobleaching has a huge impact on the diffusion coefficient (Bacia and Schwille 2003). FCS is best suited for observations of fast moving molecules over short timescales.

FRAP (Fluorescence Recovery After Photobleaching) was the method of choice for my studies on H1 dynamics. In contrast to FCS, where fluctuations of fluorescence are observed, in a FRAP experiment the distribution of fluorescent macromolecules is perturbed to observe the relaxation towards the steady state distribution. Technically, fluorescently labeled molecules are excited locally with high laser intensity to photochemically bleach the fluorophores. This generates two populations of molecules, bleached (invisible) and unbleached (visible), which are chemically equivalent but nevertheless distinguishable by fluorescence. The fluorescence recovery in the bleached region or the fluorescence loss in the unbleached region, driven by exchange of bleached molecules with unbleached molecules, can then be observed by timelapse fluorescence microscopy with low laser intensity. FRAP can be used with many fluorophores, but works particularly well with enhanced green fluorescence protein, EGFP (Rabut et al. 2004). EGFP allows the use of low laser intensity during post bleach image acquisition because of its high quantum yield. Its low tendency of photobleaching allows a long imaging period, however, essential for FRAP, the laser of a common laser scanning microscope is still powerful enough to destroy its fluorescence. FRAP is much better suited to study dynamics of proteins interacting with immobile scaffolds, situations for which the applicability of FCS is clearly limited. I will explain the principles of FRAP in more detail in chapter 1.2.4.

The following part will explain available microscopy techniques which are suitable for FRAP.

1.2.3 Microscope techniques

Regardless of the imaging technique to be used, it is crucial to consider the cells' health on the microscope stage. Cells are sensitive to photodamage, particularly in the presence of fluorophores which generate free radicals upon photobleaching, and have many ways of trying to limit light-induced damage. It is also vital to keep the cellular environment constant. There are a number of solutions to this problem, including the control of temperature, humidity, and CO_2. To maintain appropriate culture conditions for most cells it is necessary to have at least a heatable stage, ideally contained within a moisturized and CO_2-controlled box that fully encase a system (Stephens and Allan 2003; Rabut et al. 2004).

When selecting which system to use for imaging living cells, one should particularly consider two things: sensitivity of detection and speed of acquisition. Is the process to be observed fast or slow? Do you need to image for seconds, minutes, hours, or days? How bright is your signal? Because illumination of fluorophores causes photobleaching and over longer periods can even cause cell damage, everything possible should be done to limit the duration and intensity of illumination during recording. Finally, omitting phenol red and serum from the medium (if cells tolerate it) helps to reduce background fluorescence. The sensitivity of the photomultiplier tube, if using a confocal (or CCD camera, if using a widefield microscope) is vital, because the more sensitive the detector, the lower the illumination intensity needed (Stephens and Allan 2003).
A good overview on available microscope setups required for FRAP is given by Chen et al. (2006)

Confocal laser scanning microscopes (LSM)

Most confocal systems and epifluorescence microscopes are provided with a means of acquiring data series in four dimensions, since most cellular processes occur in three dimensions over time. However, in the context of FRAP, 4D imaging is rarely appropriate, because it takes time to switch from one z-layer to the next. The diffusion happening within this timeframe can not be addressed. Confocal microscopes suppress all structures out of focus by an arrangement of a detector pinhole which acts as a point light detector. Rays from out-of-focus are suppressed (Stephens and Allan 2003). Illumination is achieved by scanning one or more focused beams of light, usually from a laser, across the specimen. This point of illumination is brought to focus in the specimen by the objective lens, and laterally scanned using a scanning device under computer control. The sequences of points of light from the specimen are detected by a photomultiplier tube (PMT) and the output from the PMT is built into an image and displayed by the computer. An image is built up by focusing the laser beam as a diffraction-limited spot, pixel-by-pixel.

Many confocal laser scanning microscopes possess an AOTF (acousto-optic tunable filter), which permits the rapid (microsecond to millisecond) attenuation of the laser (Kao and Verkman 1996) and hence allows the rapid switching between a bleaching and sampling beam (Klonis et al. 2002). The AOTF is an electro-optical device that functions as an electronically tunable excitation filter, controlled by voltage, to simultaneously modulate the intensity and wavelength of multiple laser lines from one or more sources. The extremely rapid tuning speed of the AOTF is its primary virtue making a laser scanning microscope suitable for FRAP. The ability to perform extremely rapid adjustments in the intensity of the beam allows the rapid switching (in the range of µs) from the highest laser power, needed for photobleaching to a much lower laser power, needed for time-lapse imaging after the bleach. The other useful AOTF function allows the selection of small user-defined specimen areas (regions of interest; ROI) that can be illuminated with either greater or lesser intensity. It allows switching the laser power from 0% to 100% on a pixel-by-pixel basis, because the ATOF operates much faster than scanning speed. This feature enables bleaching of a strictly defined area. For a detailed explanation of the AOTF, see

http://www.olympusfluoview.com/theory/aotfintro.html.

The laser scanning device and the AOTF offer the possibility to use confocal laser scanning microscopes to perform FRAP experiments. Selecting single focal plains is

not necessary for most FRAP experiments, therefore the detection pinhole can be opened to record a brighter signal and reduce the laser power to minimize unwanted photobleaching.

The only limitation of an LSM is scanning speed. Because recording is done pixel-by-pixel, the speed of image acquisition, 10 images per second maximum, is much slower than when using a widefield microscope, which can easily record 100 images per second.

Widefield microscopes

Widefield microscopes do not exclude light from any plane out of focus; they collect it all. A key consideration, especially when observing fast dynamics like FRAP, is speed of data acquisition. Data acquisition rates of laser scanning confocal microscopes are fast enough for rapid imaging. Scanning systems, however, acquire data pixel by pixel, whereas CCD cameras, as used in Widefield microscopes acquire a whole field of view at once. This is much faster but implies blurry images showing not only focused light but also the emission from out-of-focus layers. For FRAP typical widefield microscopes cannot be used right away. Bleaching in a locally distinct area is needed to perform a FRAP experiment, which is not directly possible with such systems. However, customized widefield microscopes, equipped with an additional bleaching laser and automated shutting of the diaphragm can be used to perform FRAP.

Deltavision set up a microscope workstation based on a CCD camera recording that incorporates very sophisticated tools: real-time deconvolution, point-visiting (image many cells in one experiment) with integrated cell tracking, auto focus, background reduction, pattern photobleaching and others. Extremely rapid image acquisition is achieved by adding a laser beam into the back aperture of the microscope objective to provide a focused illumination spot in the center of the optical field. In a FRAP application the laser will bleach the region of interest while the CCD camera can immediately start recording, making it an ideal system for FRAP. However, this approach is limited to spot bleaching. Detailed information on this workstation and the available features can be found at

http://www.api.com/lifescience/DeltaVision.html.

Multiphoton microscopy

Multiphoton (or two-photon) confocal systems are now available from several companies. The two-photon effect excites a chromophore not by a single photon but from two photons being absorbed within a femtosecond time scale (Denk et al. 1990). This enables the use of longer wavelength excitation, which penetrates deeper into samples and reduces photobleaching (Coscoy et al. 2002; Stephens and Allan 2003). This feature makes it particularly suited for observing dynamics in z-dimension.

Total internal reflection microscopy (TIRFM)

Many cellular processes occur in specifically restricted areas of the cell, such as the plasma membrane. Total internal reflection fluorescence microscopy (TIRFM) pro-

vides a means of direct imaging of processes within very close proximity to the coverslip. Excitation at a critical angle generates an evanescent field of excitation light that decays rapidly with distance from the coverslip, limiting the depth of excitation to a distance of ~100 nm. TIRFM is therefore ideal for investigating protein dynamics inside cellular membranes or other small but distinct compartments, for which it has been successfully applied recently (Demuro and Parker 2006; Lavi et al. 2007; Nofal et al. 2007). TIRFM of live cells can also be combined with other techniques such as photobleaching (Sund and Axelrod 2000) or widefield imaging (Stephens and Allan 2003). In advantage to confocal FRAP techniques, TIRFM-FRAP offers greater resolution in the z-axis, ~100 nm vs. ~0.5 µm, and uses a limiting evanescent field to focus on near membrane signals rather than fluorescent reporters to define membrane boundaries (Pochynyuk et al. 2007). This is crucial when analyzing the dynamics of membrane proteins or channels. A good review on TIRF technique is given at

http://www.microscopyu.com/articles/fluorescence/tirf/tirfintro.html.

1.2.4 Fluorescent Recovery after Photobleaching (FRAP)

FRAP was the method of choice for my studies on H1 dynamics, because it is well suited to study dynamics of proteins interacting with immobile scaffolds, situations for which the applicability of FCS is clearly limited. Generally, to obtain information on the dynamics of a protein, a selected pool of fluorescent proteins must be distinguishable from the rest of the same protein in the cell and be monitored (Lippincott-Schwartz and Patterson 2003).

FRAP, and more recently photoactivation (PA) as well, have become methods of choice to visualize the dynamics of fluorescently tagged proteins in cells (Lippincott-Schwartz et al. 2001; Patterson and Lippincott-Schwartz 2002). These fluorescence perturbation methods can be readily accomplished on most confocal laser scanning microscopes and offer the possibility to quantitatively characterize diffusive processes and kinetics of interactions with binding sites in living cells.

Figure 1.7 gives an overview of the principle FRAP method.

With FRAP an area of the cell is photobleached with a high-intensity pulse and the movement of unbleached molecules from neighboring areas into the bleached area and vice versa is recorded by time-lapse microscopy. FRAP, much like FCS can provide insight into protein diffusion coefficient and dissociation constants. Photoactivation is an alternative and very similar method to photobleaching. By photoactivating locally, one can reproduce the situation of a FRAP, with two populations of activated and non-activated molecules whose equilibration can be followed over time.

Variant methods of FRAP using different geometries and bleach protocols have been developed. One of the most common variations is fluorescence loss in photobleaching (FLIP) which is performed by repeating local photobleaching to amplify the decrease of fluorescence and preventing recovery in regions that are connected to the bleach region. Over time, this leads to a loss of the fluorescent signal of certain cellular compartments and eventually throughout the cell, given that the fluorescent molecules are mobile and are able to enter the bleached region. The rate at which fluorescence is lost within the entire cell is monitored. The observation that molecules do not become bleached suggests that they are isolated (immobilized) in distinct cellular compartments. FLIP is typically used to reveal connections between different com-

partments, like nucleo-cytoplasmic shuttling of proteins (Birbach et al. 2004; Koster et al. 2005).

Classical interpretation of FRAP results

Although photobleaching experiments are easy to perform on most commercial confocal microscopes, their interpretation and especially their quantitative analysis to derive parameters describing molecular dynamics can be difficult (Saxton 2001; Klonis et al. 2002; Braeckmans et al. 2003; Braga et al. 2004; Beaudouin et al. 2006). Most studies restrict the analysis to a qualitative one and refer to interpreting the recovery curves only (Misteli et al. 2000; Brown et al. 2006; Lele et al. 2006). Two simple parameters are useful to qualitatively describe FRAP experiments: the mobile fraction and the half time of recovery (Figure 1.7) (Reits and Neefjes 2001; Lippincott-Schwartz et al. 2003). The mobile fraction measures to which extend recovery is complete over the time course of observation by comparing the intensity ratio of the bleached region at the end of the observation and before the bleach (Lippincott-Schwartz et al. 2001). As bleaching may not be complete, the mobile fraction M_f is typically calculated according to the intensity just after bleaching and is defined as

$$M_f = \frac{F_\infty - F_0}{F_i - F_0} \qquad (1.1)$$

where F_∞ is the intensity in the bleached region at the end of the recovery, corrected for photobleach artifacts during image acquisition and background-subtracted, when equilibrium is reached, F_0 the corrected intensity just after bleaching and F_i the corrected intensity before bleaching (Axelrod et al. 1976; Reits and Neefjes 2001; Chen et al. 2006). The mobile fraction depends on the ability of molecules to freely diffuse inside the cell or compartment. Long-lived interaction of proteins with other molecules or membranes will therefore change the mobile fraction. However, the calculation method for M_f assumes that no recovery occurs during the bleaching (from beginning of the bleach process until the end of the acquisition of the first post bleach image) and implies that recording after bleaching is infinitely fast. Practically, the precision of this method therefore depends on the image acquisition (scanning) speed and the speed of mobility of the observed particles. It is of great importance for calculation of M_f to wait until recovery is truly finished and equilibrium is reached. As a rule of thumb one should wait at least two times longer than observation time, when assuming recovery has finished.

The other parameter than can be directly accessed from the recovery curve is the half time ($t_{1/2}$) of recovery. This is the time it takes for the fluorescence to recover to 50% of the asymptote (plateau) intensity.

$$t_{1/2} = t\left(\frac{F_\infty - F_0}{2}\right) \qquad (1.2)$$

$T_{1/2}$ can be used to compare relative recovery rates, if it is not possible to estimate the diffusion constant (see chapter 1.3). However when the recovery is limited by diffusion, this parameter strongly depends on the size of the bleach region and is therefore not very informative to compare experiments performed with different bleach geometries (Lippincott-Schwartz et al. 2003; Beaudouin et al. 2006). The concept of the

half time is also very vague. For example, some FRAP experiments are performed using a fly-back mode, provided by some microscopes. This mode increases imaging speed by bleaching while the laser moves forward and imaging while moving backwards, instead of waiting until the laser has moved back and starting the imaging process on the next way forward. However, this does not increase imaging speed because it adds an additional recording time point already during the bleach (t_0) leading to a shift of the recovery curve towards a lower time. Without the fly-back mode t_0 is the first post-bleach image. When calculating the half time of recovery, the fly-back mode can lead to different results.

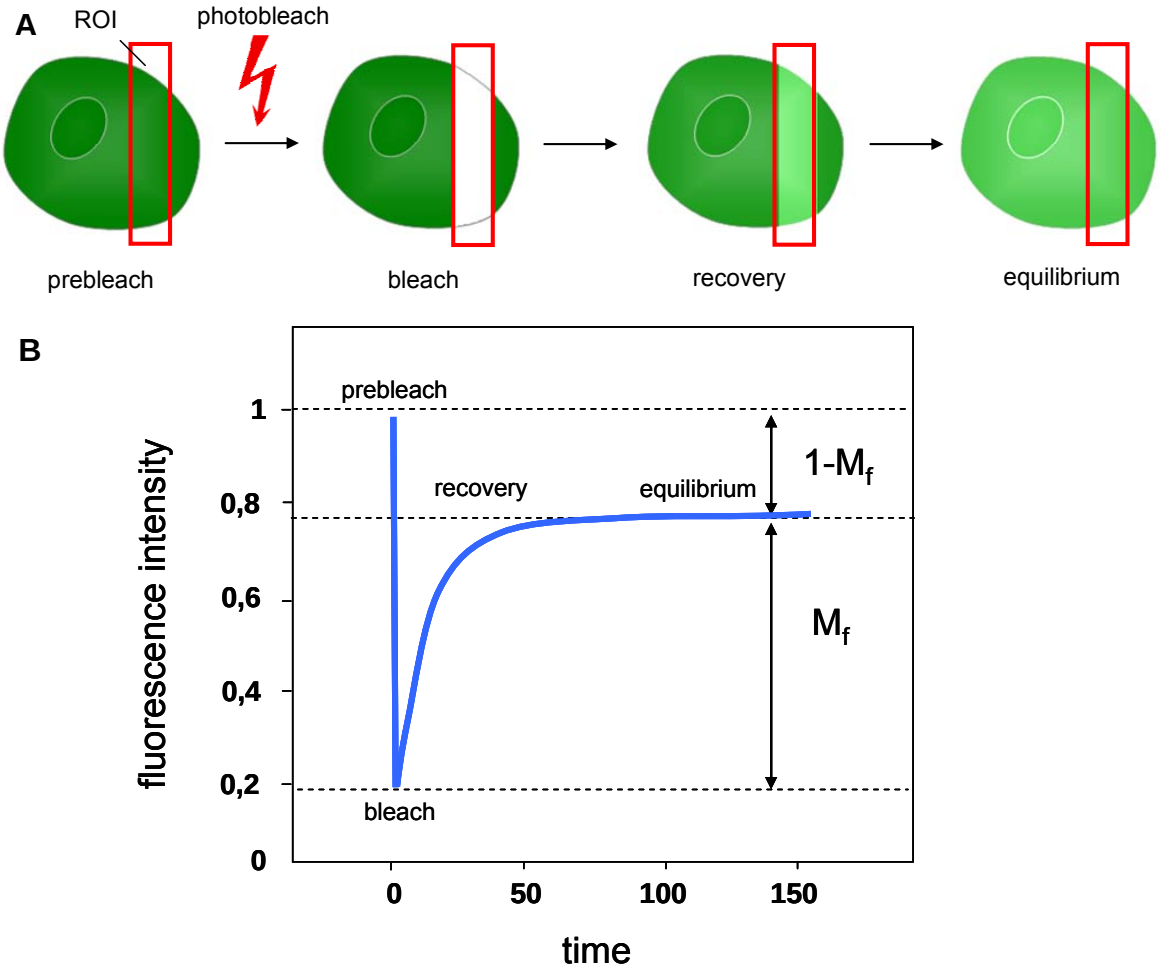


Figure 1.7 The principle of FRAP. A. The protein of interest, fused to GFP is bleached in a defined region of interest (ROI) by a laser pulse of very high intensity. This destroys the fluorescent properties of GFP but does not influence the protein of interest. The complete cell is than recorded in specific time intervals to follow the recovery of fluorescence due to diffusion of unbleached proteins from the neighboring regions. **B.** Idealized recovery curve plotting the relative fluorescence intensity over time. Before the bleaching event (prebleach), at t_0, the intensity is 1 and drops to a very low value at t_1 directly after the bleach. It does not drop to 0, since recovery may already occur during the bleach due to rapid movement of proteins. Further not all the fluorescence in the ROI may be destroyed. Recovery starts immediately until steady state equilibrium is reached which is typically not equal to the prebleach intensity. This difference is due to the immobile fraction ($1\text{-}M_f$) of proteins that may be bound to immobile scaffolds. The mobile fraction contributes to the recovery only.

1.3 Mathematical description of reaction and diffusion from FRAP data

Although FRAP experiments are relatively easy to implement on most confocal laser scanning and some widefield microscopes, analyzing the experiments quantitatively to extract physicochemical parameters of molecular mobility is nontrivial.

Quantitative interpretation of protein dynamics from FRAP data can be extracted in different ways.

In the previous chapter, I showed that most FRAP studies were interpreted on the basis of their recovery curves and the half-time of recovery, simplifying or even neglecting diffusion. However, extraction of diffusion coefficients or binding parameters from the recovery curve is basically impossible, because the shape of recovery curves does not reflect their underlying processes and deriving dynamic parameters off them is merely suspect (Sprague et al. 2006), although many scientists do still not consider this in their studies (Carrero et al. 2003).

The extracted half-time of recovery curves for example has a totally different meaning whether the process is diffusion limited or reflects only reaction. If diffusion does not contribute to the process at all, $t_{1/2}$ directly reflects the dissociation constant of the underlying unbinding process; if it represents free diffusion, however one can extract the diffusion coefficient from it.

The classical methods for analyzing FRAP described in chapter 1.2.4 may provide a first insight in the dynamics of the investigated proteins and in certain cases these methods will lead to a reasonable qualitative result. However, for most FRAP experiments only numerical simulations solving a spatio-temporal model taking the real geometry of the cell into account, will lead to truly quantitative results.

Detailed understanding of *in vivo* protein dynamics requires a model that first explains what the FRAP curve reflects, and then quantifies the underlying processes. Mainly two processes can be used to describe protein dynamics in living cells: reactions (binding and unbinding) and diffusion. Mathematically each of these processes can be analyzed analytically, limited to certain conditions in the case of diffusion. However in most cases the dynamics of proteins cannot be described by either one of them but only by a combination. The mathematical solution of such a reaction-diffusion system is nontrivial.

The most accurate approach towards a quantitative interpretation of FRAP data is to simulate a mathematical reaction-diffusion model and to derive the diffusion coefficient and binding or dissociation constants by parameter estimation, as shown by Beaudouin et al. (2006), Sprague et al. (2006) and Braga et al. (2007).

I will explain the principle of quantitative FRAP analysis in this chapter in three parts: modeling, simulation and parameter estimation of diffusion, reaction and reaction-diffusion. I implemented a procedure consisting of these three parts into Tropical, the software I developed for FRAP analysis.

1.3.1 Modeling diffusion

When motion due to active transport or unidirectional flow can be discounted, protein mobility in a cell is due to Brownian motion; expressed as the diffusion coefficient D. There are basically two different ways to spatiotemporally model diffusion: using a stochastic model describing Brownian motion in the sense of random walk or deterministically using the diffusion law formulated by Adolf Fick in 1855. Since FRAP experiments show diffusion of population of molecules and therefore reflect a concentration gradient, the deterministic approach is very well suited to model diffusion for FRAP. Fick's diffusion equation for the one dimensional case is written as

$$\frac{\partial c(x,t)}{\partial t} = D\frac{\partial^2 c}{\partial x^2} \qquad (1.3)$$

where c is the concentration of the particle of interest and D denominates the diffusion coefficient. For higher dimensional cases the diffusion equation can be written using the Nabla operator as

$$\frac{\partial c(\vec{x},t)}{\partial t} = \vec{\nabla}\cdot\left(D\vec{\nabla}c(\vec{x},t)\right) \qquad (1.4)$$

where $\vec{x}$ is a multi dimensional spatial vector. Considering isotropic diffusion, where D is constant rather then depending on space, equation (1.4) can be simplified to the Laplacian form

$$\frac{\partial c(\vec{x},t)}{\partial \mathrm{t}} = D\times\Delta c(\vec{x},t) = D\times\left(\frac{\partial^2 c}{\partial x^2}+\frac{\partial^2 c}{\partial y^2}+\frac{\partial^2 c}{\partial z^2}\right) \qquad (1.5)$$

representing diffusion in three dimensions. Equation (1.5) is a partial differential equation which is nontrivial to be solved, especially on a complex spatial domain. Therefore most studies analyzing FRAP in the past tried to assume a simplified geometry of the system to obtain an analytical solution.

For the *in vivo* situation, at least some of the published FRAP studies tried to derive a diffusion coefficient from the recovery information.

In experiments in which the protein of interest moves freely, the fluorescence will recover to the initial prebleach value. The diffusion coefficient of a protein from such data can be determined by any available analytical method (Axelrod et al. 1976; Kaufman and Jain 1990; Tsay and Jacobson 1991) and the mobile fraction M_f is 100%.

The first publication on assigning a diffusion coefficient to FRAP data was an analytical method presented by Axelrod et al. (1976), which has later been commonly used and modified (Soumpasis 1983; Phair and Misteli 2000; Braeckmans et al. 2003). In case the nature of transport is pure diffusion Axelrod and colleagues suggest using the simplified equation

$$D = 0.224 \frac{\omega^2}{t_{1/2}} \qquad (1.7)$$

to estimate the diffusion coefficient D, with ω being the radius of the bleached area and $t_{1/2}$ the half-time of fluorescence recovery. This equation assumes a Gaussian profile for the bleach and diffusion to occur only laterally, in 2D. For review see (Klonis et al. 2002; Carrero et al. 2003). Another method consists in comparing the half-time recovery with that obtained *in vitro*, if the size of the molecule is known (Seksek et al. 1997; Snapp et al. 2003). Many analytical methods presume the bleach region being a small spot and assuming the rest of the spatial domain to be infinitely large, thereby getting rid of the boundary conditions needed for equation (1.5). In many experiments these methods and their assumptions are not justified. In such cases the calculations should be considered with care.

To simulate diffusion without the compromises described above a spatial-temporal model taking the real geometry of the cell into account is required (Sbalzarini et al. 2005). Although such a model representing the most general description of diffusion cannot be solved easily there are numerical methods available to be applied.

Anisotropic diffusion due to highly organized cellular structures may play a role in certain cases (Tsay and Jacobson 1991; Volz et al. 2005; Travascio and Y. 2007), but nuclear proteins can usually be well described by isotropic diffusion. Applying the diffusion equation (1.5) to describe the distribution and movement of fluorescently labelled proteins in living cells tends towards a homogeneous distribution of fluorescence, which is true for free diffusive proteins (Seksek et al. 1997).

However, most protein distributions observed in living cells are not homogeneous, for example when they are observed in two dimensions in wide field conditions. Siggia et al. (2000) proposed an empirical method forcing the equation to tend towards the correct fluorescence distribution. In this method, diffusion is not driven directly by the differences of intensity between neighbors but by these differences divided by the intensities in steady state. This model was probably the first attempt incorporating the real geometry of a cell or nucleus.

Diffusion in living cells occurs in three dimensions; however reasonable experimental setups to image diffusion in 3D are not available. Experimental imaging systems therefore limit us to observe diffusion only in two dimensions. Simulations however could still be performed on a 3D domain. As each spatial dimension largely increases the dimension of the model which would result in the need of much longer computational time (curse of dimensionality), I used equation (1.5) in a two dimensional form to model and simulate diffusion in this study. Beaudouin et al. (2006) could show that a 2D approximation of nuclear diffusion occurring in 3D is valid. They found that although systematically smaller than the 3D values, the diffusion coefficient derived from 2D simulations was within 14% of the 3D value.

Typical diffusion coefficients of proteins in living cells are in the range of ~20-30 $\mu m^2/s$. Beaudouin et al (2006) measured a two-dimensional diffusion coefficient of PA-GFP in living cells of D = 40.6 $\mu m^2/s \pm 3.8 \mu m^2/s$, revealing the apparent viscosity of the nucleoplasm to be ~3.1 times higher than that of water, given that the GFP diffusion coefficient in water is 87 $\mu m^2/s$ at room temperature and that water viscosity drops from 1.00 to 0.69 mPa*s between 20° C and 37° C (Beaudouin et al. 2006).

1.3.2 Modeling binding

The second major mechanism influencing the dynamics of proteins *in vivo* is interaction with other proteins. Binding reactions to mobile or immobile obstacles are reversible in most cases and are described by a binding constant k_{on} and a dissociation constant k_{off}. The dynamics of binding and dissociation are represented by the chemical reaction

$$F + BS \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} B$$

and can mathematically be described by

$$\frac{\partial B(\vec{x},t)}{\partial t} = +k_{on} * F(\vec{x},t) * BS(\vec{x},t) - k_{off} * B(\vec{x},t) \qquad (1.8)$$

$$\frac{\partial F(\vec{x},t)}{\partial t} = -k_{on} * F(\vec{x},t) * BS(\vec{x},t) + k_{off} * B(\vec{x},t)$$

with F representing the unbound pool of the particle, BS denominates the free binding sites and B stands for the bound particles, x and t are the spatial and temporal dimensions, respectively.

Rabut et al. (2004) performed one of the first comprehensive investigations on binding and dissociation of FRAP studies using kinetic modelling. They analyzed several well characterized proteins of the nuclear pore complex (NPC) tagged to EGFP. They found that of the ten orders of magnitude (10^{-7} s^{-1} – 10^{3} s^{-1}), dissociation rates in living cells can cover, FRAP experiments can reveal those ranging from 10^{-6} s^{-1}–10^{-1} s^{-1}. At lower rates fluorescence would need days to fully recover and at higher rates it would be diffusion limited and therefore not possible to be fit with a pure reaction model. Using a binding-unbinding model with two distinct receptors as binding sites, Rabut and coworkers could cluster the NPC proteins into three clusters distinguishable by their k_{off}: Fast dissociating ones with $k_{off} > 10^{-3}$ s^{-1}, intermediate ones with $k_{off} \sim 10^{-5}$ s^{-1} and slowly dissociating ones with $k_{off} < 10^{-6}$ s^{-1}. When using a pure reaction model to extract dissociation constants from FRAP data, one has to be sure that the underlying process is not diffusion limited. If limited by diffusion, the extracted k_{off} rates always represent lower limits of dissociation rather than exact dissociation constants.

If recovery is only controlled by the release of bound molecules from immobile scaffolds and diffusion does not limit the recovery of fluorescence (the dynamics are limited by reaction), the recovery rate directly represents k_{off}.

1.3.3 Modeling reaction-diffusion systems

Many FRAP studies of nuclear proteins neglect diffusion justifying this by the shape of the recovery curve or a long halftime of recovery (Misteli et al. 2000; Phair and Misteli 2001). Most often, however, the dynamics of nuclear proteins are determined by diffusion, which is usually hindered by binding *in vivo*. In such cases, recovery kinetics of FRAP experiments are unlikely to be describable by free diffusion or reac-

tion only, but need to be described by complex reaction-diffusion systems (Beaudouin et al. 2006).

It is therefore crucial to check whether diffusion is the driving force of mobility before describing protein dynamics by it. A simple test for determining whether a fluorescent protein's mobility is limited by diffusion or binding is to vary the size of the bleached area A. The recovery rate will theoretically change proportional to A^2 for diffusion limited movement only.

Almost all biological studies assume that diffusion is not limiting the fluorescence redistribution of nuclear proteins (Dundr et al. 2004; Brown et al. 2006; Lele et al. 2006) therefore neglecting diffusion to simplify the analysis. They assume that if the time for fluorescence redistribution is long compared to the case of freely diffusing molecules, diffusion could be neglected in the analysis, which is not necessarily true (Beaudouin et al. 2006; Sprague et al. 2006; Braga et al. 2007). Moreover, in those cases where diffusion has been ignored, the consequences of this assumption have not been tested. Beaudouin and colleagues (2006) could clearly show that neglecting diffusion based on these assumptions is incorrect, because very transient interactions where diffusion is clearly limiting, can also lead to slow fluorescence redistributions. Sprague et al. (2006) have also shown how ignoring diffusion can lead to errors in the estimation of binding parameters up to two orders of magnitude with recovery curves still showing a reasonable fit. This was also confirmed by Braga et al. (2007). Further, Sprague and his colleagues analyzed the parameter domain of typical FRAP studies and found that diffusion is expected to contribute to the majority (~2/3) of FRAP recoveries, given typical values for effective diffusion and bleach spot size. The effective diffusion coefficient D_{eff} represents a mixed diffusion coefficient of he free and bound pool of a protein and can be calculated as $D_{eff} = D_{free}*Fraction_{free}$ (Sprague et al. 2006).

Some studies (Coscoy et al. 2002; Sprague et al. 2004; Bhattacharya et al. 2006) have so far considered reaction and diffusion occurring together in living cells, however by mostly theoretical investigations. Since both, diffusion and binding influence the dynamics of proteins, while diffusion is often limiting the binding process at least inside the nucleus, binding and dissociation must be considered when setting up a model.

Almost all methods analyzing reaction-diffusion systems so far have largely simplified or even ignored the cellular geometry, within which protein mobility occurs, and assumed a homogeneous distribution of binding sites (e.g. Lele et al. 2006).

Sprague et al. (2006) demonstrated that completely ignoring the localization of binding sites will introduce serious errors (about one order of magnitude) into the estimation of binding parameters and that the spatial localization of binding sites leads to characteristically different FRAP curves.

Including binding to an immobile binding site, the diffusion equation (1.5) will be extended by a binding term (1.8) representing the free pool of a protein, whereas the immobile bound pool is still represented by equation (1.8). The full reaction-diffusion model in its 2D Laplacian form will then be represented by

$$\frac{\partial F(\vec{x},t)}{\partial \mathrm{t}} = D \times \left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2} \right) - k_{on} * F(\vec{x},t) * BS(\vec{x},t) + k_{off} * B(\vec{x},t)$$
$$\frac{\partial B(\vec{x},t)}{\partial t} = +k_{on} * F(\vec{x},t) * BS(\vec{x},t) - k_{off} * B(\vec{x},t) \quad (1.9)$$

It is important to notice that BS (x,t), the spatial distribution of free binding sites, does effectively not depend on time, because photobleaching does not chemically modify the proteins and their affinity: if the distribution of molecules is in steady state before photobleaching, it will remain in steady state afterwards from a biochemical point of view. The only reason why one can still observe the dynamics is that only the fluorescent subset of the protein is visible, which does not affect the binding sites. This means that modeling can only extract the product of k_{on}*BS but not the two parameters separately.

Such a reaction-diffusion system requires the distribution of BS to be known, which is usually not the case. However Beaudouin et al. (2006) extracted a method to derive the binding sites distribution from the steady state of the fluorescent protein. The fluorescence of a FRAP image represents the visible sum of free diffusive and bound protein, resulting in an inhomogeneous fluorescence distribution.

Diffusion coupled with very fast reaction

Assuming now that binding and dissociation are very fast processes and diffusion is limiting the fluorescence recovery, the reaction-diffusion system (1.9) can be described by a pure diffusion equation (1.5) (Crank 1975). In such a case the diffusion coefficient represents effective diffusion. Beaudouin et al. (2006) could show that applying an extended diffusion equation, equal to the one proposed by Siggia et al. (2000) to describe a diffusing protein binding to immobile binding sites is valid, if binding and dissociation are much faster than diffusion. This could also be demonstrated for diffusion limited interactions of two mobile particles (Braga et al. 2007). This case, where reaction is instantaneous compared to diffusion was represented by a method published by Siggia et al. (2000). For this reason, I implemented this method within Tropical (see chapter 3.4); it is an easy way to address this special case.

Effective diffusion coefficients of the nuclear proteins have been shown to cover a range of some orders of magnitude, but typically being < 10 $\mu m^2/s$. Examples, to get an impression on derived values can be found in Beaudouin et al. (2006) for the hyperactive mutant of histone methyltransferase SUV39H1 (D_{eff} ~ 5.7 $\mu m^2/s$), Braga et al. (2007) for poly(A)-RNA (D_{eff} ~ 0.6 $\mu m^2/s$) and Schmiedeberg et al. (2004) for Heterochromatin protein 1 (HP1) (D_{eff} ~ 0.7 $\mu m^2/s$) and in many other studies.

On the basis of the recent findings on effective diffusion and the limiting behavior of diffusion one can distinguish different regimes of protein dynamic behavior in cells which can be fit with different models (Beaudouin et al. 2006; Braga et al. 2007):

- Pure diffusion with at most weak binding can be fit with a pure diffusion model
- Effective diffusion with very fast (instantaneous) binding can be fit with a pseudo-diffusive equation as just described
- Full reaction-diffusion which can only be fit by a reaction-diffusion model

By exploring extensively the reaction parameter space, Braga et al. (2007) studied the possible FRAP behaviors in the case of two interacting mobile species. They found that only in a small subset of points the full reaction-diffusion equations are required to fit FRAP data. For the majority of the cases, however, simple diffusion models yield good fits.

For linker histone H1°, which I analyzed for this thesis, the scenario of effective diffusion could very well fit the obtained FRAP images, as is shown in chapter 4. The typical approach for parameter estimation based on FRAP experiments is to first use a simple diffusion model, because it is easy to implement, and try to fit the observed data. If this fails one switches to a full reaction-diffusion model. Since different models have might be necessary when analyzing FRAP, depending on the observed processes, I implemented a method to include a user-defined model equation system within Tropical, allowing the user to be most flexible and applying any dynamic regime needed.

1.3.4 Simulation of partial differential equations

Simulation in the context of partial differential equations (PDE), such as the diffusion equation (1.5) or the reaction-diffusion equation (1.9) means integration of the equation over time and plotting the integrated curve or area of the solution per time step. Solving partial differential equations is nontrivial and usually requires numerical methods. Several standard methods for solving ordinary differential equations (ODE), in this context differential equations that depend on time but not on space, are available (Bock 1981; Deuflhard and Bornemann 2002). The typical way of solving a partial differential equation that depends on time and space is to discretize the equation in space to eliminate this dependence. This can be done using a finite differences approach.

Applying this method, one subdivides the spatial domain into a grid of equally distant grid points and defines an ODE, solely depending on time for each grid point. After discretization, the equation becomes a system of coupled ordinary differential equations, one equation for each point in space.

Partial differential equations, like equations (1.5) and (1.9) require spatial boundary conditions, which restrict diffusion to happen inside the spatial domain, preventing a flux out of the defined area. I used Neumann boundary conditions for the analysis for this thesis, which I also implemented into Tropical (compare chapter 3.4.5). Neumann boundary conditions allow diffusion only between nearest neighbors. At the border of the nucleus, diffusion is allowed only with the neighbors inside the mask, reflecting the fact that nuclear or cellular membranes are impermeable for the observed protein at the timescale of a FRAP experiment (compare chapter 3).

One major advantage of numerically simulating a PDE diffusion equation model to quantify FRAP data is the possibility to consider the real geometry of the cell or nucleus under observation.

1.3.5 Parameter estimation to reveal the diffusion coefficient D

The quantity characterizing a protein's movement in a cell is the diffusion coefficient D. Applying a model according to the diffusion equation (1.7) implies that D is known. Unfortunately this is usually not the case. On contrary the diffusion coefficient is the parameter which we want to derive.

Deriving an unknown parameter from an equation by simulation is a so called inverse problem because one knows the action and need to derive the cause from it. This is typical for a parameter estimation approach. Several methods for parameter estimation are described (Schloeder 1988) and have been applied to biological systems (Mendes and Kell 1998). The workflow of a parameter estimation process, starting from a mathematical model and experimental data and ending with an identified diffusion coefficient is shown in Figure 1.9. The methods implemented into Tropical are demonstrated in detail in chapter 3.

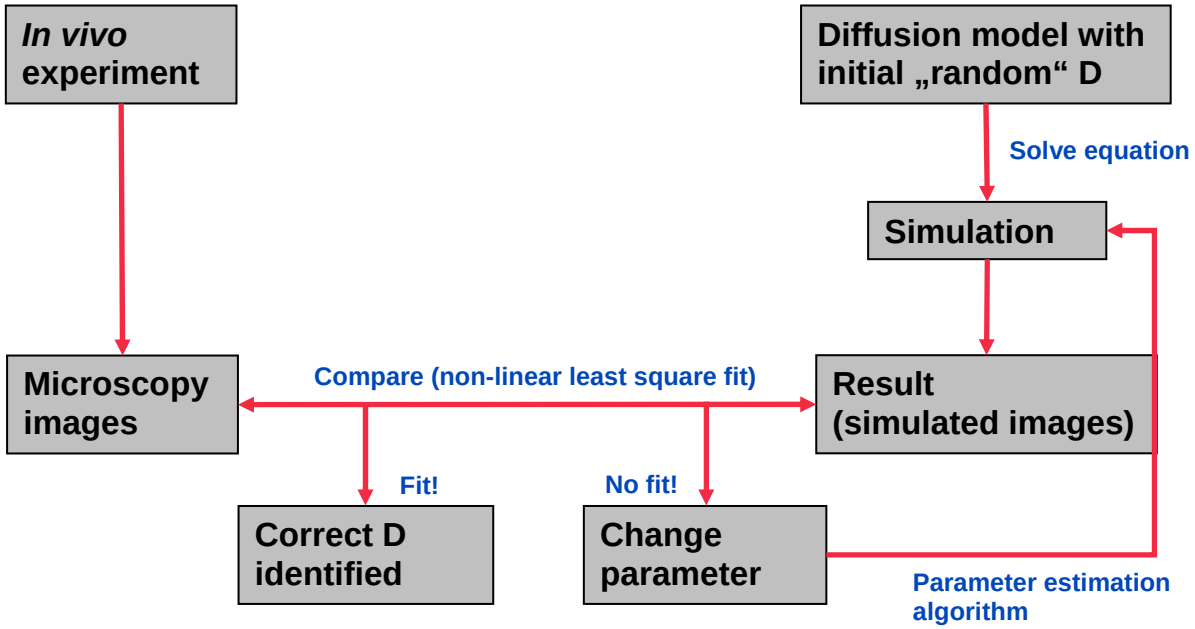


Figure. 1.9 Simplified workflow of a parameter estimation process. It starts from a diffusion model as described in equation (6) and the microscopy images obtained from an *in vivo* experiment. The first step is to assign a randomly chosen diffusion coefficient to the model. The closer the initial D is to the real D, the faster the parameter estimation might finish. The second step is to run a simulation with this initial random D that is solving the diffusion equation. The resulting simulated images will then be quantitatively compared to the original images by a non-linear least square fit (Bock 1981). If the difference between original and simulated images is large, D will be changed in a certain way and direction by the parameter estimation algorithm and simulation starts again. This will be iterated until the difference between original and simulated data are small enough to have estimated a diffusion coefficient which is close to the real D.

1.3.6 Software for modeling, simulation and parameter estimation

Several modeling software packages are available for building numerical models of any kind, which can also be used to interpret experimental data. Among the most frequent used are Copasi, Berkley Madonna, and Matlab. A huge variety of less known tools were designed for specific purposes of simulating the behavior of biological

molecules, estimating reaction coefficients, building models or performing *in silico* experiments like MCell (http://www.mcell.cnl.salk.edu/), Cell Designer, E-Cell, Jarnac, J-Designer or Biospice. A good overview of tools that support SBML, the systems biology markup language, is given on http://sbml.org/index.psp.

However, except Matlab and Berkley Madonna, which have been shown to be suited for FRAP analysis (Beaudouin et al. 2006; Sprague et al. 2006), applying other tools for this task is rather difficult and very limited. Specifically they are not able to simulate reaction-diffusion equation systems in spatio-temporal resolution or cannot deal with images as input for parameter estimation. This means most of them are not able to incorporate the real geometry of a cell and the exact description of the underlying processes. Although Matlab and Berkley Madonna are capable of these important demands, using them requires some knowledge about transforming images into numerical matrices and about matrix operations. In the case of Matlab, a very powerful tool for various applications even in engineering, it is necessary to know the program's specific programming language.

Hence, because all existing tools were either not suited to interpret FRAP images in a spatiotemporal reaction-diffusion model or rather limited or complicated, I developed Tropical, a tool which is able to simulate any reaction-diffusion differential equation system and estimate the relevant diffusion and binding or dissociation parameters based on the direct input of normalized FRAP time series images. Details of Tropical are presented in chapter 3.4.

2 Objectives

Tropical – software for FRAP analysis

FRAP has become a popular method to analyze protein dynamics inside single living cells. Adequate methods to quantitatively analyze FRAP results have recently been developed (Beaudouin et al. 2006; Sprague et al. 2006; Braga et al. 2007). However, suitable, user friendly software incorporating those methods is still missing. Available software for this purpose is either very rudimentary, simplifying FRAPs underlying processes to suit only very special situations or is very complicated to be used, requiring a lot o preprocessing of the FRAP data and even the knowledge of programming.

The primary objective of my work presented in this Ph.D. thesis was to develop a computer program, incorporating the latest and most accurate methods to analyze FRAP data. This program, named Tropical, should be as user friendly as possible, directly operate on microscopy images, and incorporate the true geometry of the observed cellular compartments. It should be able to simulate reaction-diffusion equation systems and extract the corresponding parameters of such equation systems, the diffusion coefficient and reaction parameters.

Dynamics of linker histone H1 in vivo

To apply the methods implemented into Tropical, I chose to analyze the dynamics of linker histone H1 *in vivo*. H1 has been the target of many FRAP studies, addressing its binding and diffusion properties to get insight into the dynamic organization of eucaryotic chromatin.

The second goal of this thesis therefore is to characterize the dynamic properties of linker histone H1° and two mutated forms of it by FRAP and in form of reaction an diffusion parameters. Further the goal is to identify the contribution of the mutated sites to DNA binding.

3 Materials and Methods

In this part of my thesis, I will present the fluorescently labeled proteins I observed using FRAP analysis, namely wild type H1° and two mutants of it: one single point mutation having Lys52 exchanged by alanine and one form including six point mutations in the primary and secondary DNA binding site, mutant “Six”.

In the second part of this chapter, I will give a short review on the microscope and its setup that I used for the FRAP experiments described in this thesis.

The latter part will deal with the main development which led to this thesis: Tropical, a computer program I developed to quantitatively analyze FRAP data. I will present the workflow of Tropical to give an overview of its principle concept and the idea behind it. I will also present some screenshots and describe its different parts in detail. The mathematical methods implemented for estimation of parameters and for the simulation of partial differential reaction diffusion systems will be explained and an overview of Tropicals’ output will be given. Furthermore, the image processing procedures that are relevant to obtain accurate and reliable results with Tropical are described and other prerequisites like the production of the time step file and the generation of compartment masks will be explained.

3.1 The three forms of H1°-GFP

Three forms of the nuclear protein linker histone H1°, C-terminally tagged with EGFP (Clontech), stably expressed in NIH 3T3 cells, were used to perform FRAP analysis. The two mutated forms were mutant “52”, containing a single point mutation at site Lys52, which was altered to alanine and mutant “Six”, containing six point mutations at the sites Lys40, Arg42, Lys52, Lys69, Arg74 and Lys85, all mutated to alanine (Figure 3.1). Mutations, GFP-fusion and transfection were done in the laboratory of Tom Misteli (National Institute of Health, NIH, Maryland, USA), who kindly provided the transfected cells. The same constructs were also used by Brown et al. (2006).

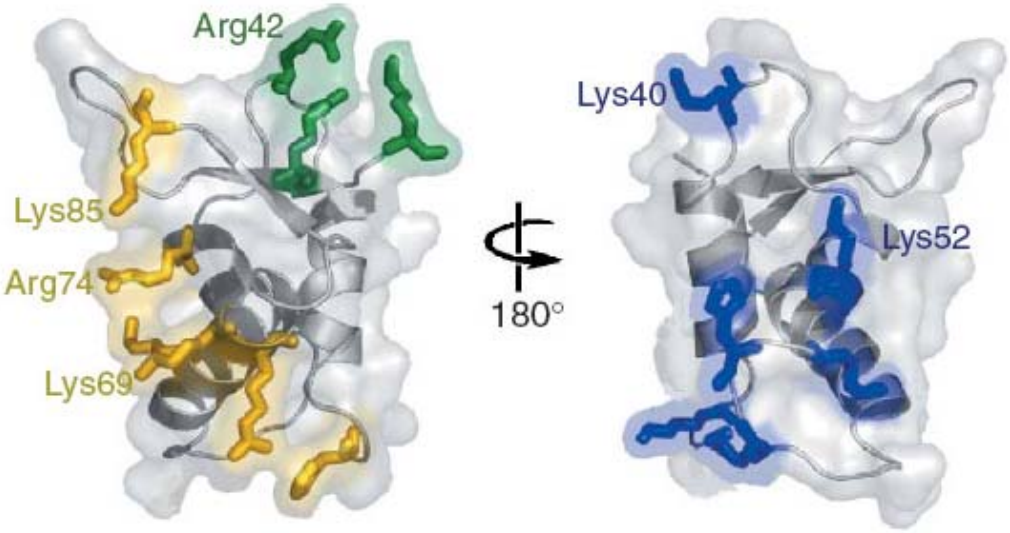


Figure 3.1 3D Structure of H1° showing the mutated sites. For mutant “52” Lys52 was changed to alanine, for mutant “Six” all shown sites were altered to alanine by (Brown et al. 2006).

The coding sequence for enhanced GFP was excised from pEGFP-C1 (Clontech) and was inserted after the last lysine residue of the H1° variants in pMTH1cneo or pMTH1°neo as published in Gunjan et al. (1999) and Misteli et al. (2000).

Plasmid MTH1°GFPneo has been described by Gunjan et al. (1999). In this plasmid, the coding sequence for enhanced GFP is fused to the C-terminus of the coding region for H1°, and expression is under control of the mouse metallothionein promoter. Point mutations were introduced with the QuikChange mutagenesis kit (Stratagene) or by introduction of annealed oligonucleotides between restriction sites (Brown et al. 2006).

On the basis of high performance liquid chromatography (HPLC) profiling, Misteli et al. (2000) estimate that the NIH 3T3 cell lines overexpress less than 5% H1-GFP in addition to endogenous H1. H1°-GFP proteins were released from nuclei upon micrococcal nuclease digestion with kinetics identical to that of the respective unmodified endogenous protein, suggesting proper positioning of the fusion proteins on DNA. The salt-extraction properties of the H1-GFP fusion proteins were indistinguishable from those of their endogenous counterparts. Like endogenous H1, the en-

tire pool of both H1-GFP variants was associated with chromosomes in mitotic cells, and expression of the fusion proteins had no effect on cell-cycle behavior or cell proliferation (Misteli et al. 2000).

Brown et al. (2006) confirmed that all H1 mutant proteins were structurally intact, in that all had circular dichroism spectra similar to that of the wild-type H1° globular domain. Becker et al. (2005) showed H1-GFP to be fully functional in somatic-cell nuclear transfer. Therefore, these mutations and fusion to GFP do not introduce major changes in the structure or function of the globular domain of H1.

All experiments were performed in NIH 3T3 cells cultured in complete DMEM (DMEM, 10%FCS, 2 mM glutamine, 100 U/ml streptomycin, 100 µg/ml penicillin).

For microscopy, cells were cultured in #1 LabTekII chambered coverglasses (Lab-Tek, Naperville, IL). For imaging, cells were maintained at 37°C in DMEM without phenol red complemented with 25 mM HEPES buffer, pH 7.3 and 20% FCS. For staining DNA *in vivo*, 0.2 µg/ml Hoechst 33342 was used.

3.2 FRAP laser scanning microscopy of H1°-GFP

FRAP experiments were performed on a Leica TCS SP2 confocal laser scanning microscope following a protocol established at our laboratory (Kappel and Eils 2004). An oil immersion objective with a 63x numerical aperture was used in all experiments. Imaging was done using the widefield mode with the detection pinhole opened to 0.000115 m (1 airy unit). Two dimensional images were recorded over time at a scan speed of 1400 Hz (app. 4 images per second). The laser was adjusted to excite EGFP at a wavelength of 488 nm. The recording spectrum was adjusted to the range of emission wavelength of EGFP, 498-600 nm. The general laser power was adjusted to ~75 % for the argon / krypton-argon laser and additionally, a laser exciting at 405 nm was used to guarantee effective bleaching.

All images were recorded at a resolution of 256x256 pixels with a voxelsize of approximately 0.095 µm, resulting in images of about 24x24 µm in size. A zoom factor of 10 was used for recording.

Laser power for bleaching and recording was adjusted using the AOTF. For bleaching at 488, 476, 458 and 405 nm a laser power of 100% was used; for recording the laser power of the 488 nm laser was turned down to ~10% via the AOTF, the other non-recording wavelengths were turned to 0%.

For imaging DNA labelled with Hoechst 33342 a DAPI laser was used, exciting at a wavelength of 405 nm. Emission of the DAPI light was recorded at a spectrum of 410-478 nm, with the AOTF adjusted to 10%.

The setup for all FRAP experiments was exactly the same, however two different bleaching ROI geometries were used with each H1° mutant, but in different individual cells. One bleached region showed a rectangular shape of ~15 x 4 μm^2, bleaching approximately half of the nucleus. The second bleach geometry used was a circular spot of 20 (~2 µm) pixels in diameter.

Five prebleach images were recorded with a time interval of 287 ms between each image, followed by one bleach image and 40 post bleach images using the same time interval. Then 40 additional images were recorded at time intervals of 1 s and finally a last series of 5 – 100 images, depending on the recovery time of the imaged H1° - mutant, was recorded at time intervals of 20 s.

3.3 Image processing

Before analyzing FRAP images, there are three necessary corrections to apply to the images produced by the microscope: Background subtraction, correction for photobleaching due to the acquisition of post-bleach images and finally image registration to compensate for cell movement during image acquisition. For this I used ImageJ 1.35s (http://rsb.info.nih.gov/ij/) including several plugins, which will be explained later. Two more computer tools were used for other image processing tasks: Adobe Photoshop 8.0 (www.adobe.com) for generation of the masks which represent the shape of the nuclei of the imaged cells and Matlab 6.5 (www.mathworks.com) to convert tiff images, which were written by ImageJ in a tiff byte order that is specific for Macintosh, to a tiff byte order which can be processed by any computer and software.

Figure 3.1 shows the workflow of the complete image processing.

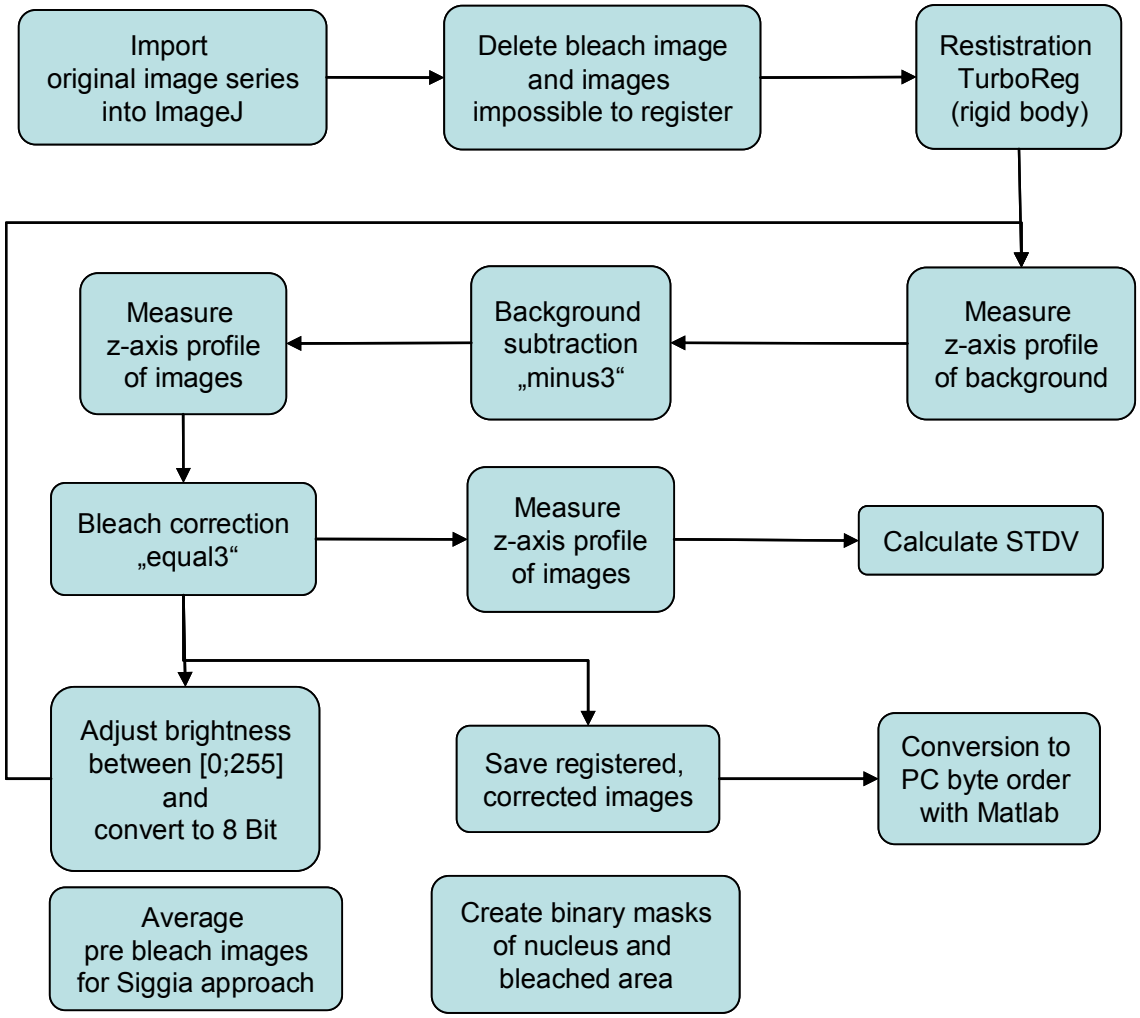


Figure 3.1 Image processing workflow. From recording images to using them for parameter estimation with Tropical the images need to be registered, background subtracted and corrected for unintended bleaching. Images were processed in 32 Bit mode and later converted back to 8 Bit for use with Tropical. Standard deviation of the image time series was calculated using Excel. Conversion from Macintosh byte order Tiffs as produced by ImageJ to PC byte order used by Tropical was done with Matlab. Prebleach images were averaged for Siggia correction (Siggia et al. 2000) and masks of geometries were produced using Photoshop.

Preparation of time series for image processing

After the import of the complete image sequence as it was produced by the Leica microscopy software into ImageJ, the first step in processing them to be used with Tropical was done using the stack function “delete slice”. The bleach image was deleted to avoid problems with this very bright image (Figure 3.2) during bleach correction and image registration. Furthermore those post bleach images were deleted which proofed impossible for registration because the cells moved too much (e.g. movement partly out of the image or out of focus). Since the “delete slice” function built in ImageJ can only remove one slice at a time, the plugin Slice Remover can be downloaded from http://rsb.info.nih.gov/ij/plugins/slice-remover.html.

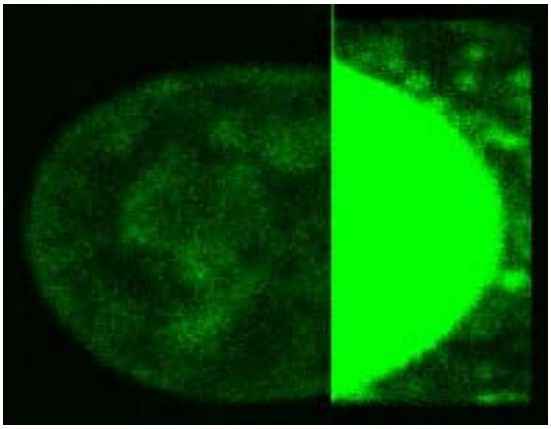


Figure 3.2 Nucleus transfected with H1° GFP during bleach. During the bleach process the microscope records an image where the bleach ROI appears as a very bright part. These images have to be deleted from the sequence before the images can be registered.

Image registration

After all images that make registration processes impossible are removed, the image series can be registered to compensate for cell movement or rotation. This was done using the ImageJ plugin TurboReg, which is explained in detail and can be downloaded at http://bigwww.epfl.ch/thevenaz/turboreg/. The first prebleach image was taken as reference (denominated “target” in TurboReg) image and the time series (denominated “source” in TurboReg) was then processed using rigid body registration and the option “accurate quality”. After the registration the time series was written to a new folder as a series of Tiff images.

Background subtraction

Using the registered images it was now possible to apply all other necessary corrections. For background subtraction an ImageJ plugin written by Dr. Joel Beaudouin called “minus3” was used (see chapter 6.2). For this plugin an area of the background was selected and a profile of this area was plotted over all images of the time series using the stack function “plot z-axis profile” of ImageJ. The results table was saved and subtracted from the intensities of the complete time series.

Bleach correction

After the background was subtracted the images were corrected for unintended photobleaching using the ImageJ plugin “equal3” written by Dr. Joel Beaudouin (see chapter 6.2). The intensity profile based on the complete images of the corrected time series was saved and used to calculate the standard deviation using Microsoft Excel. Since all ImageJ operations were done using 32-Bit images, whereas Tropical requires 8-Bit images as input, all images had to be converted to 8-Bit. As ImageJ keeps the intensity values displayed when converting from 32 to 8 Bit, the brightness of the image series had to be adjusted between 0 and 255 before conversion.

Averaging prebleach images for Siggia’s approach

For the Siggia’s approach to correct for inhomogeneous fluorescence distribution in steady-state in living cells (Siggia et al. 2000), the image series is divided by a prebleach image. To reduce noise occurring in a snapshot image of one particular time point, I used an average image produced by averaging all five prebleach images taken per time series with the averaging function of ImageJ’s image calculator.

Mask generation

To use the real geometry of the observed nuclei as spatial domain for the simulation, a binary mask of the nucleus is needed for Tropical. Further masks are needed to define the compartment or region which should be considered for parameter estimation and for plotting the recovery curves of the simulation.

For mask generation I used Adobe Photoshop 8.0. The averaged prebleach image was binarized by changing the image mode to “indexed color”. By increasing the contrast and finally deleting all pixels outside the nucleus and simply painting all pixels inside the nucleus black, the nucleus mask was finished. Masks were saved as 8-Bit raw image files for use with Tropical. For all other compartment masks the same process was used. Although this is a manual approach, it is very fast and accurate. However one could also use an automated batch process for this.

Conversion of Tiff images

Unfortunately, Tiff images written by ImageJ are arranged in a byte order unique to the Macintosh platform or software developed for Macintosh. Tropical on the other hand can only read Tiff images of PC byte order. Adobe Photoshop can read both and can safe images in either one, however only one at a time, which is very uncomfortable when processing hundreds of images. Therefore short Matlab scripts were used to first load the images into Matlab 6.5

[F, flag, Ffiltered] = tifLsmLoaderImageJ('*imagename*',[256 256],*nImages*,1);

where F is the matrix into which the images are written after the import to Matlab, *imagename* has to be replaced by the core name of the image series without the numbers (e.g. “cell1_” if the images are called “cell1_001”, “cell1_002”,…) and *nImages* is the number of images in the time series. The dimension of the images is identified in

brackets. *1* stands for not applying an additional filter. The function tifLsmLoarderImageJ() was written by Constantin Kappel. The images were then written back using

for i = 1:size(F) imwrite(uint8(F(:,:,i).*255), ['*new imagename*', checkindex(i,3), '.tif'], 'tiff', 'Compression', 'none'); end

with a new image name as Tiff images, where imwrite() is a standard Matlab function.

This had the effect of changing the byte order to PC conformity so that images could be imported by Tropical.

3.4 Tropical – Software for simulation and parameter estimation based on fluorescence microscopy images

For quantitative interpretation of microscopy images no software was available that could be used solitarily. Simulation of diffusion or reaction-diffusion equation systems on the spatial domain of the real geometry of a cell and parameter estimation on the basis of such a model and microscopy image data have so far not been possible. Therefore I developed Tropical (Ulrich et al. 2006) during the work leading to this thesis. Tropical is a software for simulation and parameter estimation of reaction–diffusion models. It consists of more than 19.000 lines of C++ program code. Based on spatio-temporal microscopy images, Tropical estimates reaction and diffusion coefficients for user-defined models. Tropical allows the investigation of systems with an inhomogeneous distribution of molecules, making it well suited for quantitative analyses of microscopy experiments such as fluorescence recovery after photobleaching (FRAP). The development of Tropical was the main part of my research. In this chapter, I will therefore explain the functionality of Tropical in detail.

3.4.1 Tropical workflow

Tropical consists of three main parts: input, data processing and output (Figure 3.8). The input is done via a graphical user interface (GUI) and is subdivided into five more parts.

All input data concerning the microscope image time series are collected via the first data tab called “input data” (Figure 3.3). These are the image time series itself, starting with the post bleach images, the initial images (usually the first post bleach image of each protein of interest) and all related specifications, like the number of time steps, the voxel size of the images and their standard deviation. Further the optional binning algorithm can be activated and its binning factor can be adjusted. This enables reduction of image size and therefore computational time, since the grid for the finite differences discretization is represented by the pixels of the images.

The second data tab called “model” (Figure 3.4) enables the input of the masks representing the geometry of the cell or nucleus which is the spatial domain for the simulation. Here, binary images in raw format can be loaded. The lower half of the screen enables the input of the reaction equations and the activation of the diffusion term. Once the button “add” is activated a screen appears that requires the input of the differential reaction equations and initial parameters as well as the initial diffusion coefficient for the molecule specified. Once a model is built, it can be saved in text format via the “Save model to file” button. Obviously, saved models can be loaded from such text files.

The third data tab named “simulation & parameter estimation” (Figure 3.5) contains the input for the Siggia normalization approach mention the context (Siggia et al. 2000), the input of the time steps of the experimental images via a file or vie equidistant time steps. Further, parameter estimation can be activated here and the compartments used for the spatial domain of parameter estimation and plotting of results can be imported here, again as binary masks in raw image format.

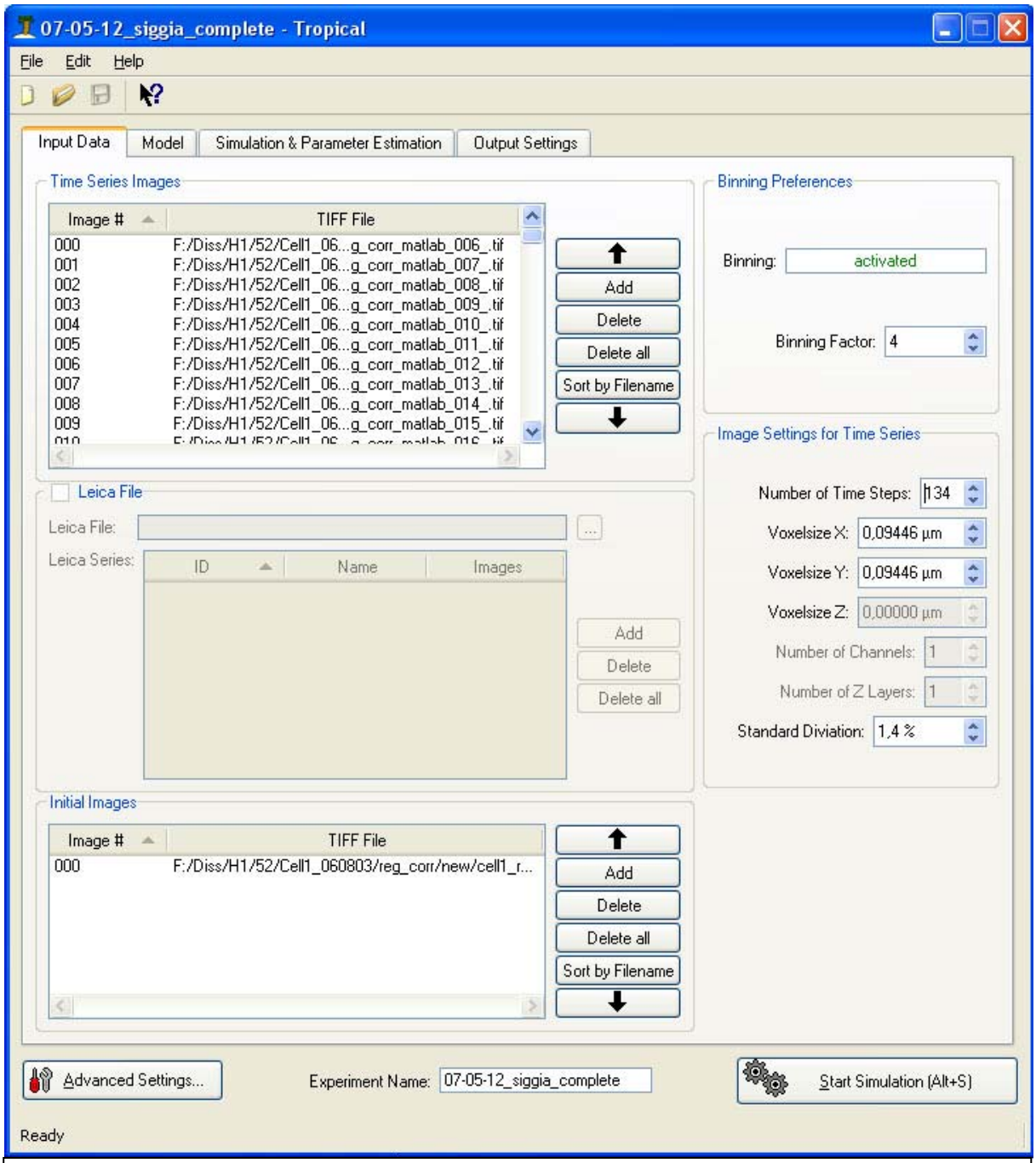


Figure 3.3. Input data tab of Tropical. Image time series can be loaded here and all settings concerning input images are set here.

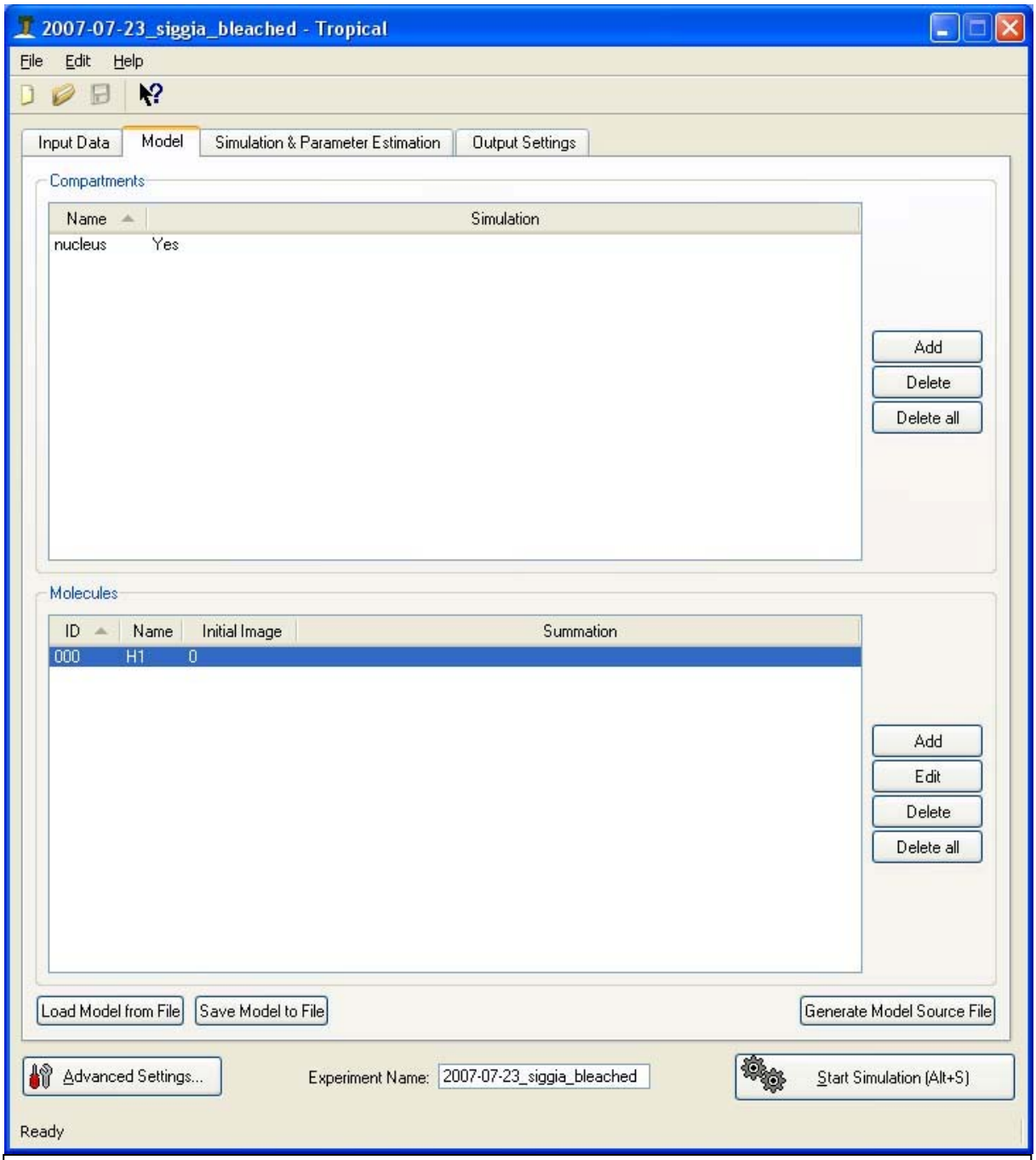


Figure 3.4 Model data tab of Tropical. The spatial domain for the model is loaded here as a black and white image of a cell or compartment and the molecules are specified with their corresponding mathematical reaction-diffusion equation.

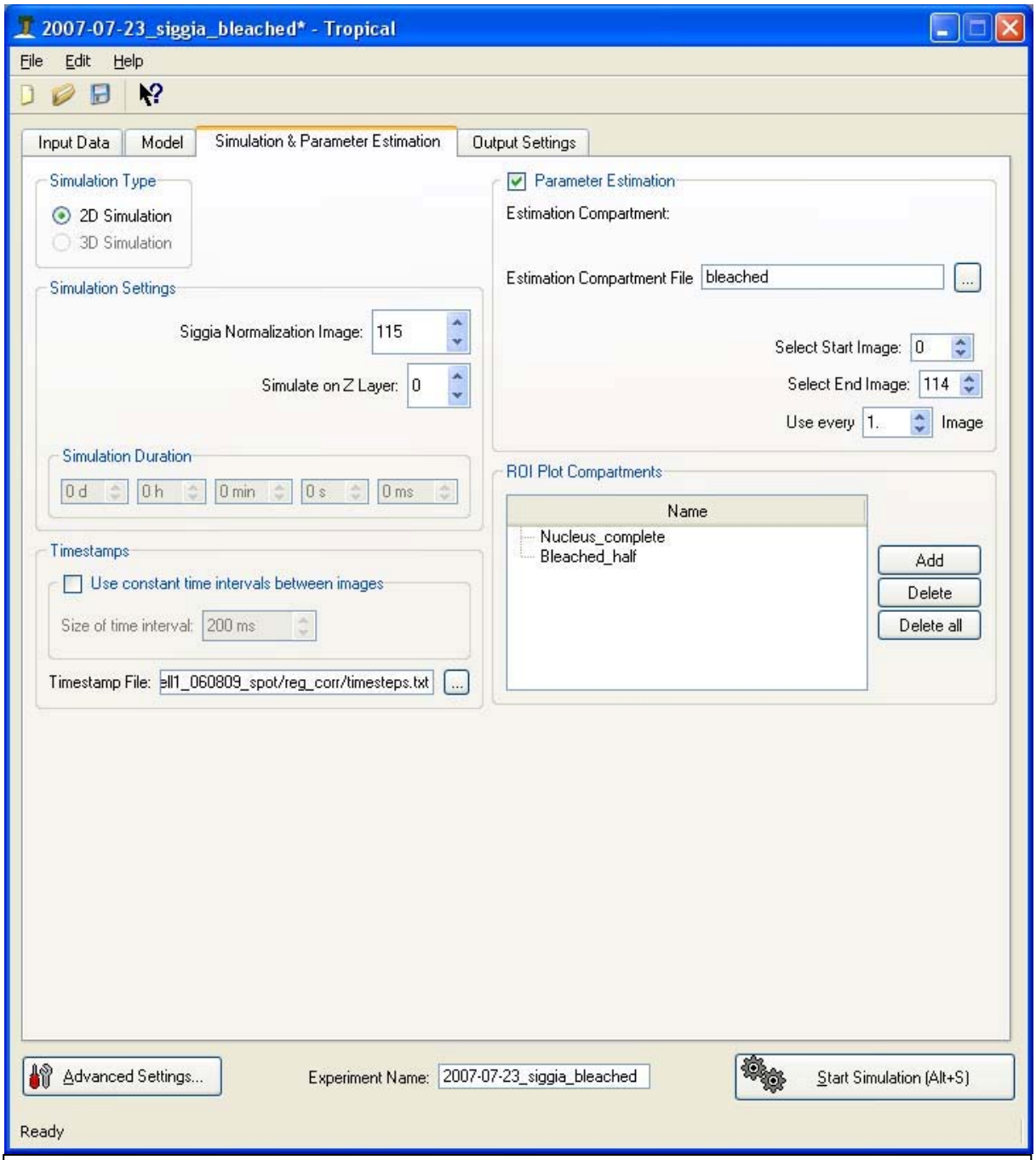


Figure 3.5 Simulation and parameter estimation data tab of Tropical. All settings related to simulation (time steps, estimation compartment, compartment for plotting recovery curves and the image used for the Siggia (2000) approach) have to be specified here.

The fourth data tab, called output settings lets one specify the path to where the output shall be written (Figure 3.6).

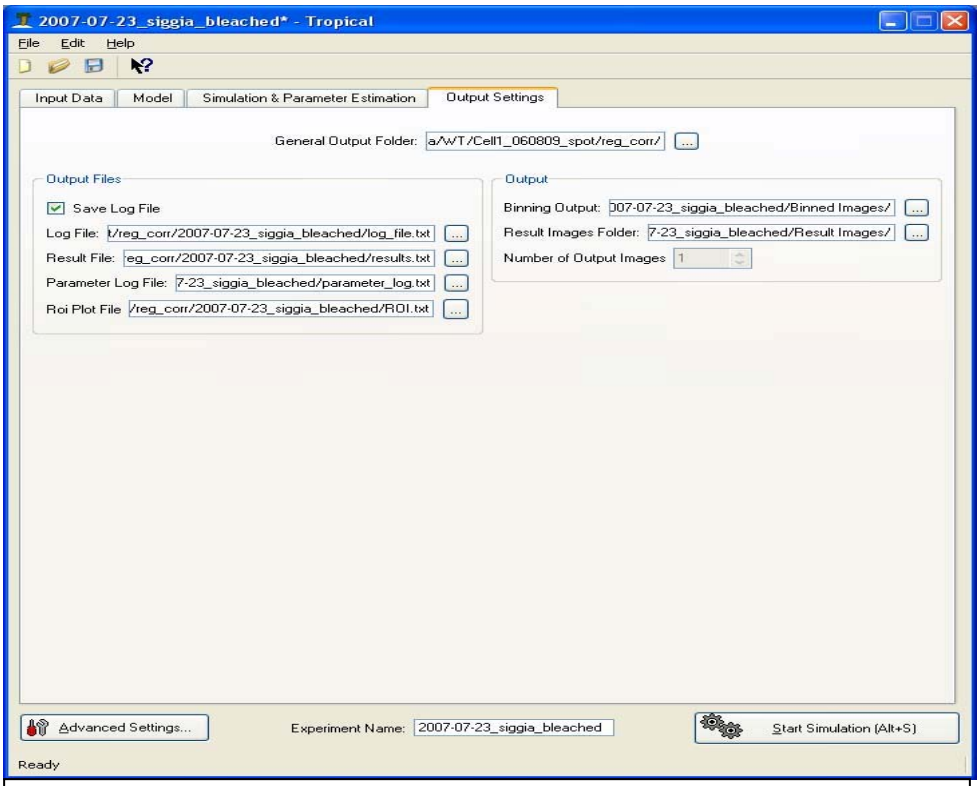


Figure 3.6 Output data tab of Tropical. Here the path to write all output data to has to be specified.

The fifth input part is the button “advanced settings” (Figure 3.7). Here some expert settings can be made to adjust the accuracy of the ODE solver and the abortion criterion of the parameter estimation algorithm as well as some parameters relevant for its convergence. These parameters are set to default values which are stored in a file called “preferences.opt” in the Tropical directory. Changing these parameters can result in significant changes in computational time but also accuracy of the result.

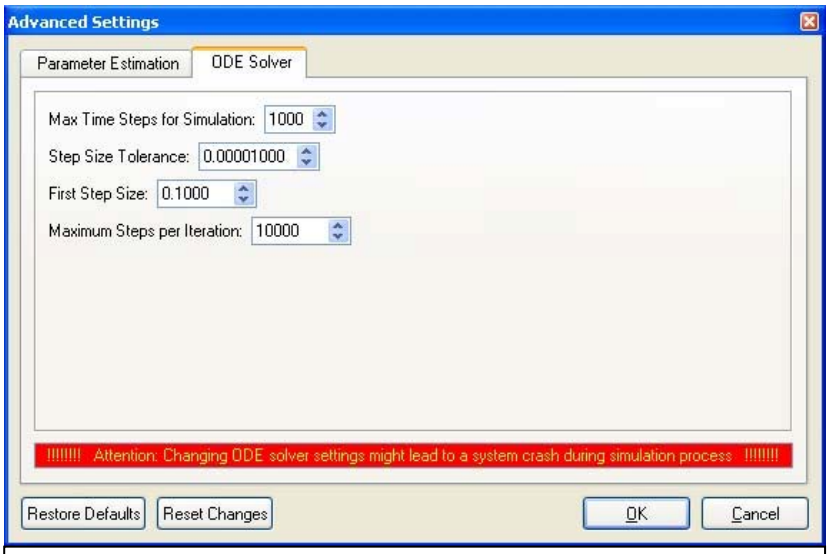


Figure 3.7 The advanced settings tab of Tropical. Expert setting to fine tune the parameter estimation and simulation algorithm can be made here.

Schemes of Tropical's structure and workflow is provided in figures 3.8 through 3.11.

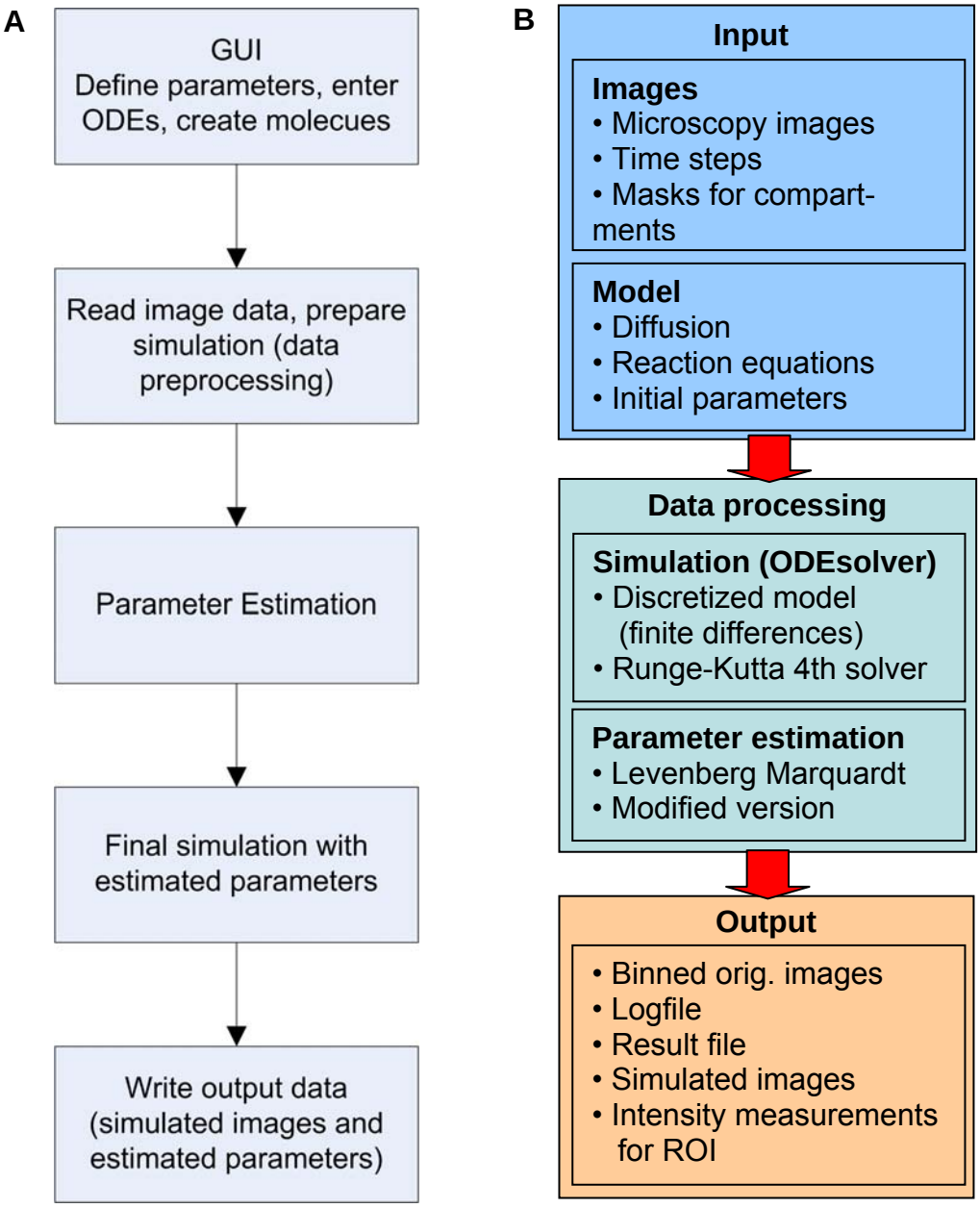


Figure 3.8 Main parts and workflow of Tropical. A shows the main workflow on which the program operates. Starting from the input via the GUI, then reading the images, masks and all settings, performing a parameter estimation in several iterations and calculating a final simulation with the best estimated parameters. Finally the result of this last simulation is written as output. **B** shows the three main parts of Tropical: The input, reading all information on the images, model and settings via the GUI, the data processing performing simulation and parameter estimation and the output writing the original and simulated images, intensity measurements and result files.

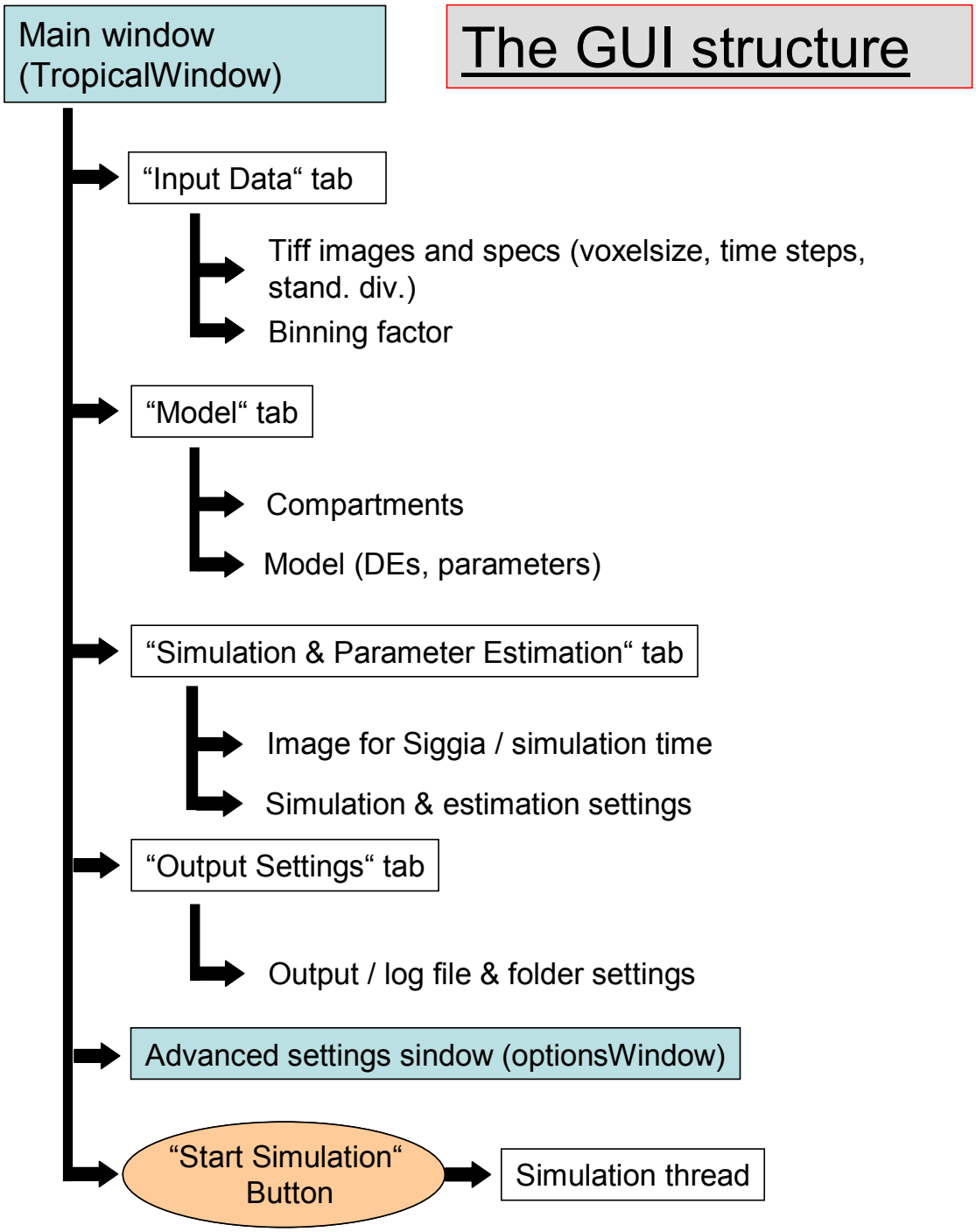


Figure 3.9 Structure of the GUI. The Tropical GUI consists of five input screens: The input data tab, the model tab, the simulation & parameter estimation tab, the output settings tab and the advanced settings window. Obligatory inputs or those that have predefined but changeable standard values are shown.

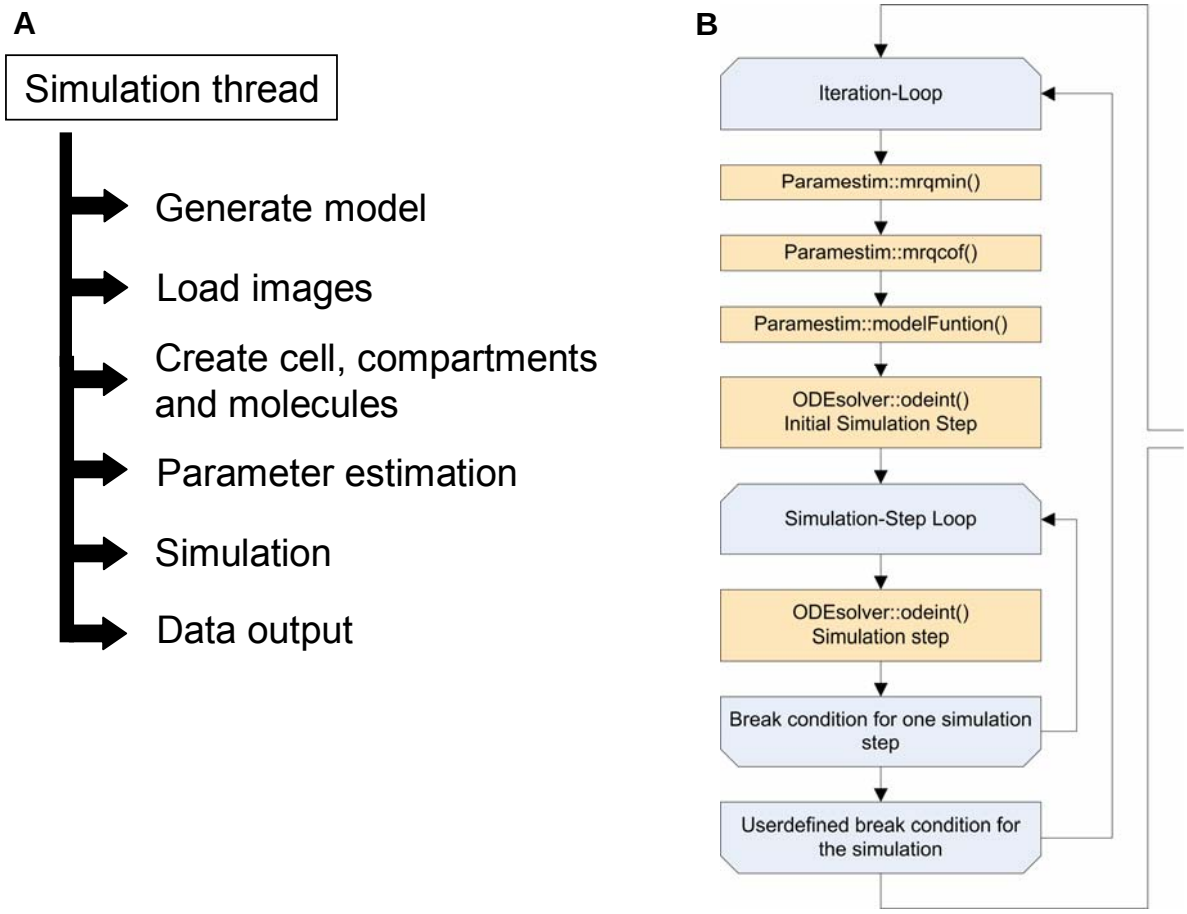
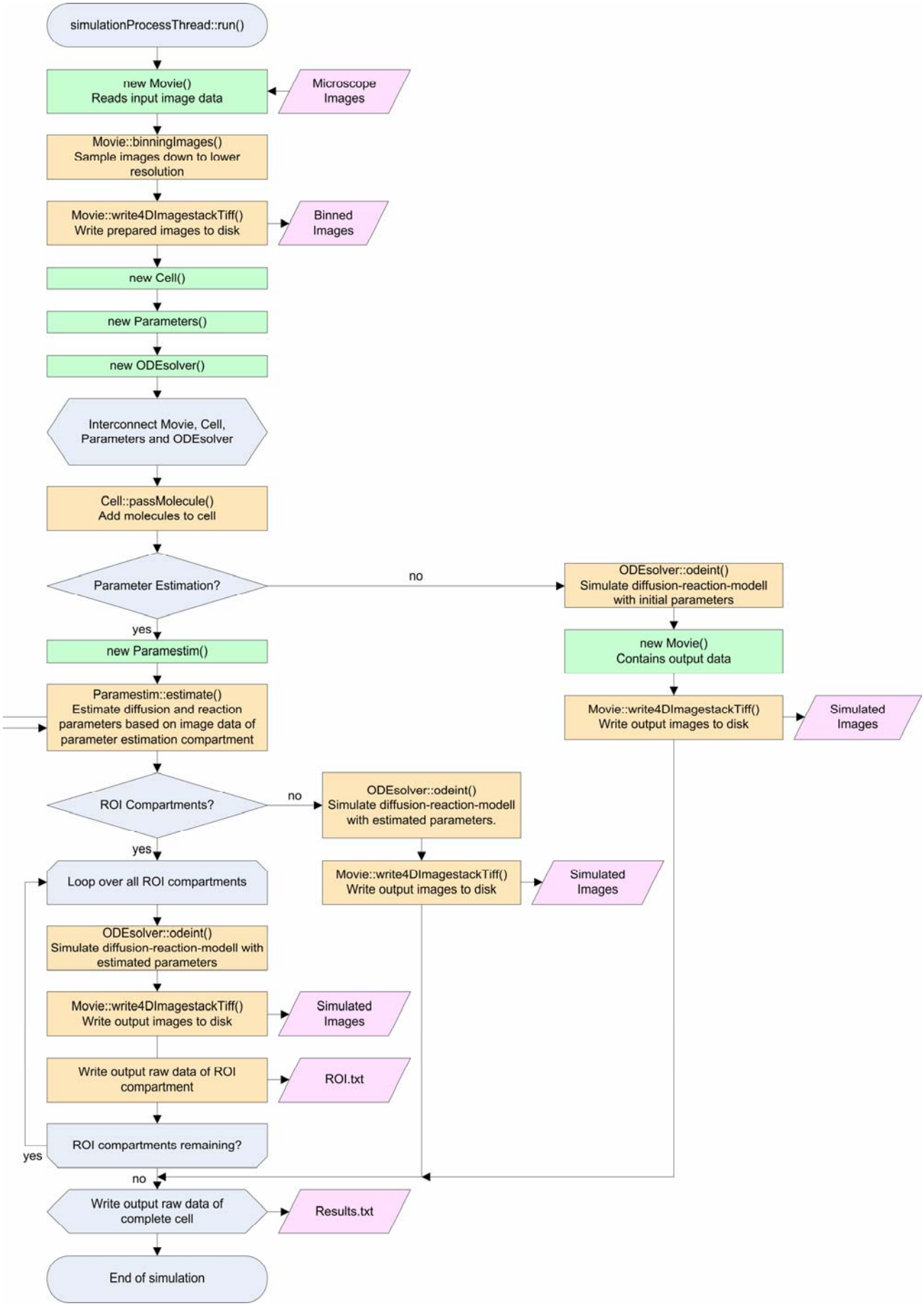


Figure 3.10 Workflow of the Tropical program code. After clicking "Start simulation" in Tropical, the simulationProcessThread is started. All next steps performed by Tropical are shown in **A** and together with the names of the methods that are called in **B (continued on next page)**. Green fields indicate that new memory is allocated, pink fields represent input or output of data and orange fields are functions of the program that execute a significant part of Tropical. Comments and explanations are plotted in grey fields. The main classes of which objects are constructed at the beginning are Movie, where all the image data of the time series are stored in, Cell, which represents the binary masks and hence the geometry of the cell and Parameters containing the initial parameters defined in the GUI. ODEsolver is the class that performs the complete simulation. Several functions belong to this class of which odeint() is the most important one, performing the integration of ODEs on each point of the spatial grid. The class that covers all functions relevant for parameter estimation is called Paramestim. Its functions are nested starting from estimate() which calls mrqmin(), calling mrqcof(), which itself calls the modelfunction containing the model. ModeFunction() then calls ODEsolver::odeint() which performs the simulation. With the best estimated parameters a final simulation is performed and the output files are written.


simulationProcessThread::run()
new Movie()
Reads input image data
Microscope Images
Movie::binningImages()
Sample images down to lower resolution
Movie::write4DImagestackTiff()
Write prepared images to disk
Binned Images
new Cell()
new Parameters()
new ODEsolver()
Interconnect Movie, Cell, Parameters and ODEsolver
Cell::passMolecule()
Add molecules to cell
Parameter Estimation?
no
yes
ODEsolver::odeint()
Simulate diffusion-reaction-modell with initial parameters
new Movie()
Contains output data
Movie::write4DImagestackTiff()
Write output images to disk
Simulated Images
new Paramestim()
Paramestim::estimate()
Estimate diffusion and reaction parameters based on image data of parameter estimation compartment
ROI Compartments?
no
yes
ODEsolver::odeint()
Simulate diffusion-reaction-modell with estimated parameters.
Movie::write4DImagestackTiff()
Write output images to disk
Simulated Images
Loop over all ROI compartments
ODEsolver::odeint()
Simulate diffusion-reaction-modell with estimated parameters
Movie::write4DImagestackTiff()
Write output images to disk
Simulated Images
Write output raw data of ROI compartment
ROI.txt
ROI compartments remaining?
yes
no
Write output raw data of complete cell
Results.txt
End of simulation

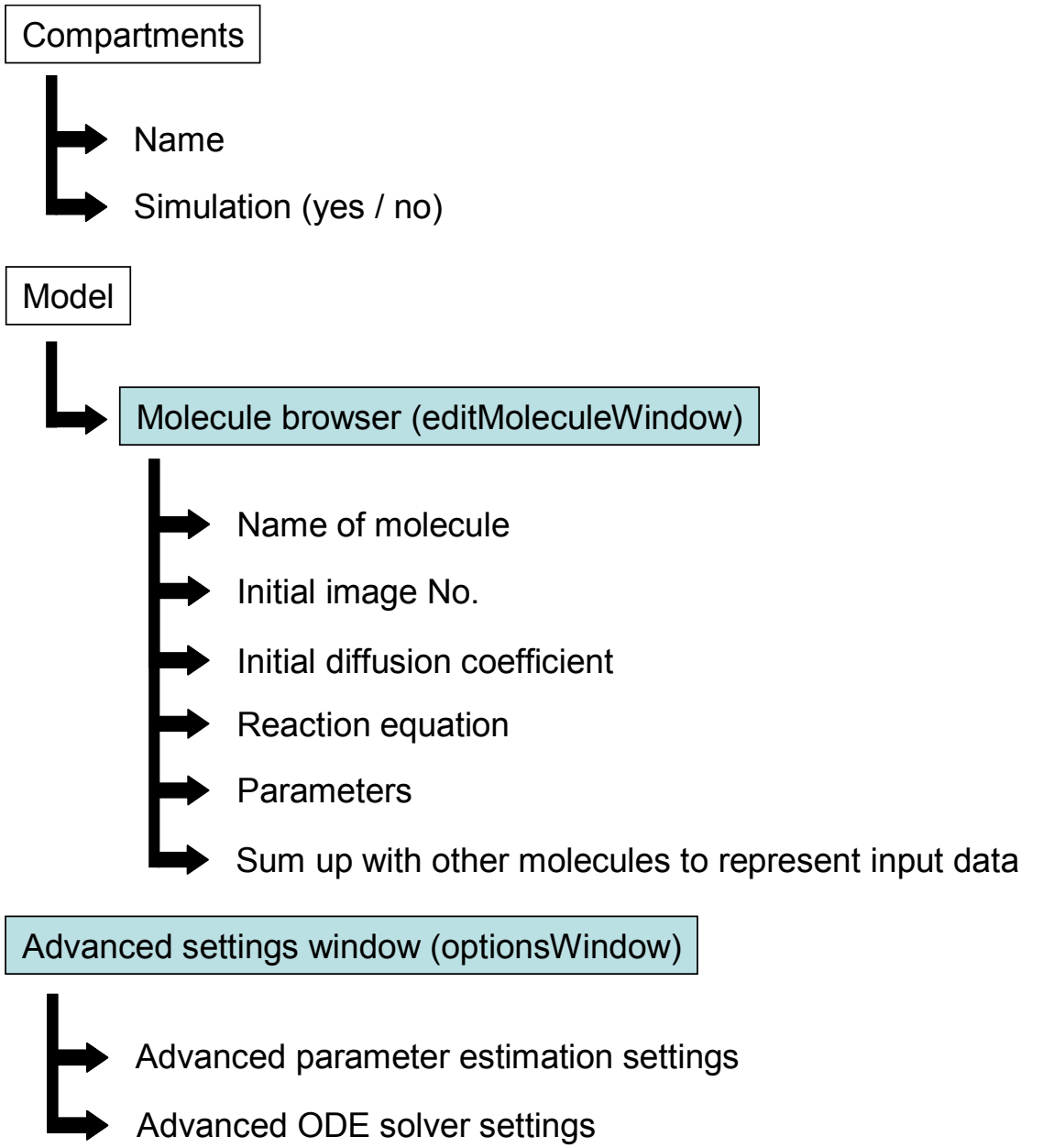


Figure 3.11 Structure of the three side windows. When clicking on "Add" in a compartment area of the GUI an additional window opens that allows loading a raw image file (mask) to be a compartment. The compartment has to have assigned a name and it has to be specified if simulation on the spatial domain of this compartment shall be performed. When adding a molecule on the model tab, an additional window awaits the input of the molecules name, the number of its representing initial image, its initial diffusion coefficient (which can be "0"), a reaction differential equation (default is dy/dt=0, no reaction) and the initial values for the reaction parameters (can be left blank if no reaction parameters are specified in the reaction equation). Finally the possibility is given to sum up the calculated intensities of some molecules, since it happens quite often that the measured fluorescence represents the intensities of one protein but in different states, which would be defined as to different molecules (state variables) in a mathematical model. The advanced settings window can be used to apply changes in the convergence parameters, abortion criterion and accuracy of the ODE solver and the parameter estimator. Changing these settings can result in significantly longer calculation time but also in large changes of the accuracy of the calculation.

3.4.2 Prerequisites: Preparation of time step files and image parameters

Binary compartment images serve as masks for the simulation grid. They define the area where parameters are estimated and specify regions for computing a recovery curve. The preparation of masks is described in chapter 3.3. Calculation of the standard deviation of the images was done by plotting the total intensity profile over all corrected images over time and calculating the standard deviation of the time series with Microsoft Excel's function for standard deviation. The standard deviation is a necessary input parameter for Tropical. Required by the Levenberg-Marquardt algorithm, it is used to include the uncertainty of the absolute grey values into the fit. This leads to a still reasonable fit if the images are truncated by noise.

Besides masks, Tropical needs to know the voxel size of the images. The correct voxel size is crucial, because it directly effects the scaling of the estimated diffusion coefficient. A wrong voxel size would therefore mock a diffusion coefficient to be correct, which is in fact scaled by a wrong voxel size. Especially when images are also binned, it will be a rather large effect. The voxel size of images taken with the Leica TCS SP2 software is saved in a text file attached to each experiment performed with the microscope and was taken from there.

Another need for performing parameter estimation is that the exact time of each acquired image (data point) is known. Typically the image sequence starts with the first post bleach image, representing time 0. For use with Tropical, the time of all following images has to be expressed in seconds after this first image. A time step text file for Tropical has to look like this (compare Annex 6.1, Tropical timestamp file document):

```
0
0.296
0.577
0.859
1.140
1.421
1.703
2.000
2.281
2.567
…
```

The Leica software writes the exact time into the text file of each experiment in the form "year:month:day,hours:minutes:seconds:milliseconds".

Therefore I extracted those data to Excel and transformed them into the correct format. If time steps are equidistant, this procedure is not necessary. Equidistant time steps can be declared in Tropical directly with a precision of 1 ms.

3.4.3 Input data

A number of input parameters are required for Tropical to thoroughly perform parameter estimation. In the following part I will explain the requirements for Tropical on the "Input Data" tab (Figure 3.12). In addition to creating a new Tropical experiment, an existing one can be saved with the file extension .xnp and later be loaded.

Time series images

The corrected and registered experimental image time series, which serves as the experimental data on which parameters are estimated, have to be loaded here. Clicking the “Add” button the complete series can be added. If the Siggia normalization approach is considered, the reference image has to be loaded as the last image of the series (separate) here. Images have to be in 8-Bit grey scale Tiff format.

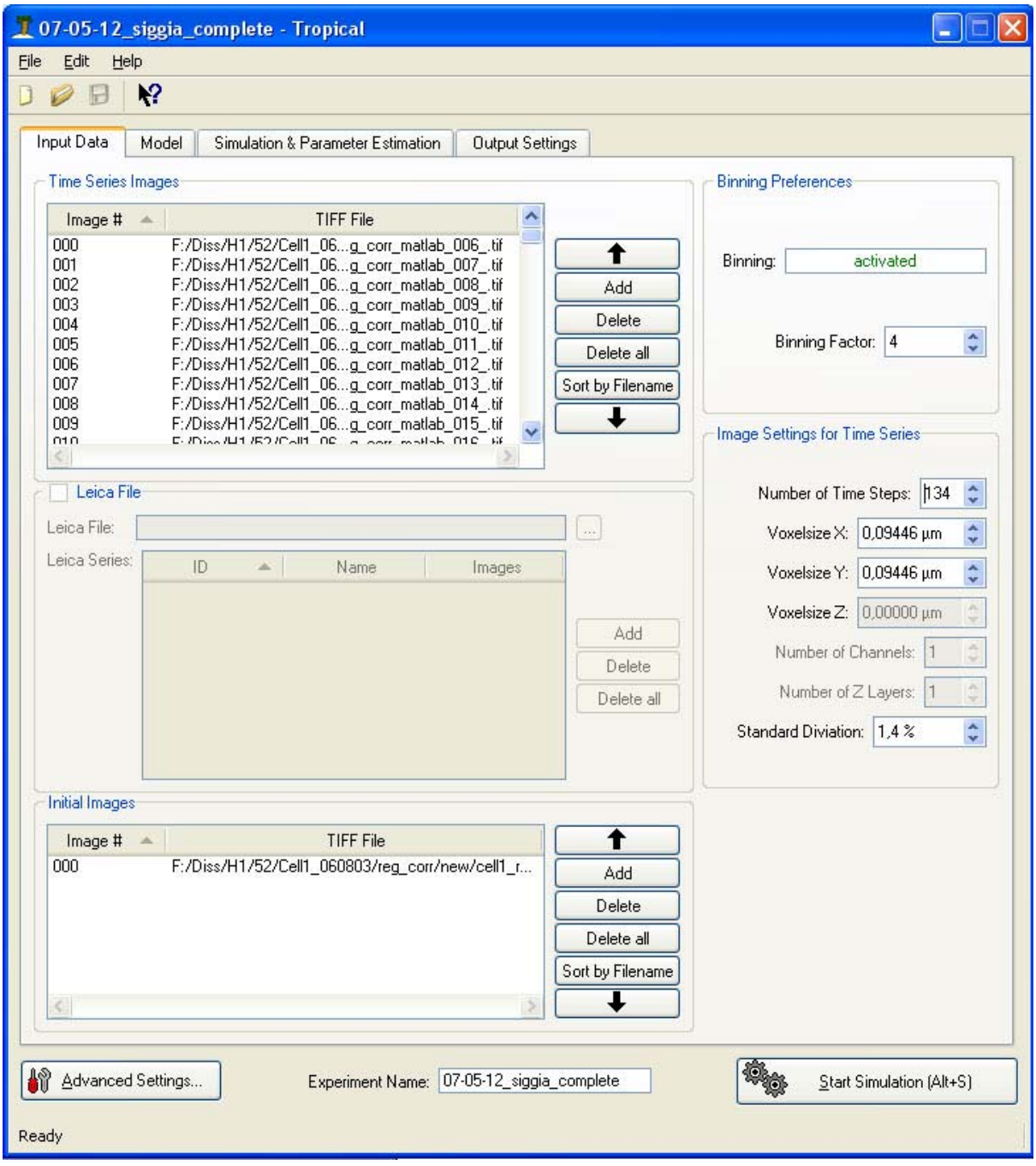


Fig 3.12 Tropical input data tab screen shot. Via the input data tab the images of the experimental time series and the initial image for each molecule of the model are loaded and the image settings (number of time steps, voxel sizes and standard deviation of the images) are applied. Furthermore binning can be activated.

Here the number of images in the time series has to be specified as well as the voxel size and the standard deviation of the images.

Initial images

Each initial image represents one state variable in the equation system, which is one molecule in a Tropical model. This means that for each molecule set in the model one initial image is needed.

Binning

An optional binning algorithm reduces the resolution and thus accelerates computation. By selecting the binning factor this function gets activated. Binning reduces the image size by combining a number of neighboring pixels equivalent to the binning factor in each direction to one new pixel. The new pixel is assigned the average intensity of the original ones. A binning factor of four would therefore reduce an image of 256x256 pixels to a resolution of 64x64 pixels. Binning also smoothes the images, thereby reducing noise and therefore improves the reliability of parameter estimation. For each pixel one differential equation has to be solved over time.

3.4.4 The model: compartments and molecules

The model tab requires the input of the compartment for the model, representing the spatial domain of the observed data, i.e. a mask of the nucleus. The lower part of the window, called molecules, is the input of the model. By clicking “Add” the input screen for a molecule, its initial diffusion coefficient, the option to estimate the diffusion coefficient and the input of a reaction equation and initial reaction parameters appears (Figure 3.13). For details on how to write reaction equations for Tropical see Annex 6.1: Tropical model file documentation.

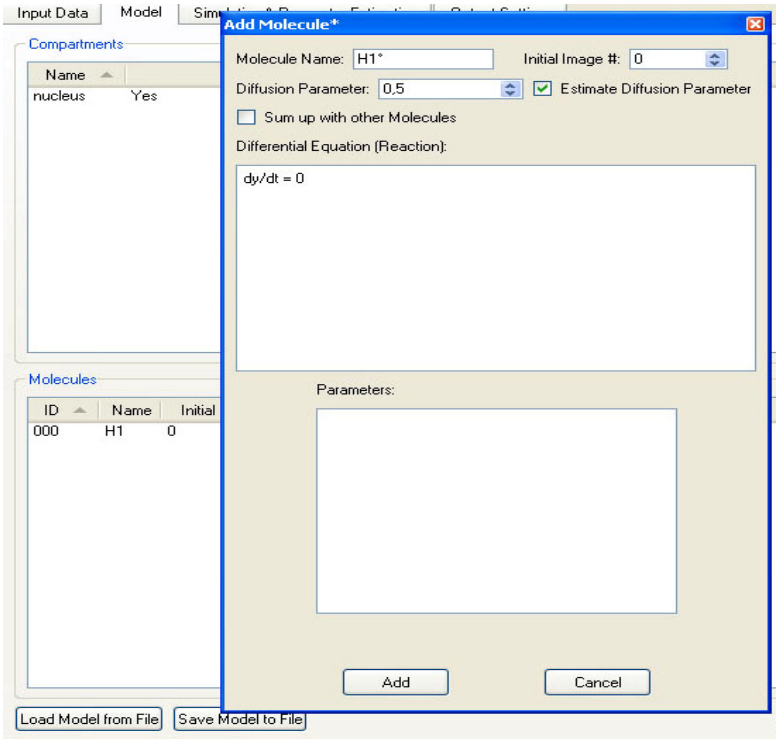


Figure 3.13 Screenshot of the „Add molecule“ window. For each molecule added this window lets the user specify the initial diffusion coefficient and the reaction equations.

3.4.5 Simulation of a diffusion model using finite differences and a Runge-Kutta 4th order algorithm with adaptive step size

Tropical estimates the specified parameters and runs a simulation with the best estimation result based on the input of (1) spatio-temporal microscopy images, (2) initial images for all state variables and (3) a user-defined model. The model is composed of one ordinary differential equation describing the reaction of each molecule. The diffusion term is added automatically. Diffusion was simulated using the geometry of the nucleus and the distribution of fluorescence just after photobleaching or photoactivation. As the diffusion equation (1.7) cannot be solved analytically in complex geometries like the nucleus, spatial discretization by a finite differences approach (Figure 3.14) with Neumann boundary conditions was used for approximation. The nucleus was modeled as a grid of square elements, with a grid step equivalent to the resolution of the images. The underlying grid for this discretization is always represented by the pixels of the input images when using Tropical. The grid distance is expressed by the voxel size of the images. If binning is applied, the voxelsize is automatically corrected by the binning factor. The advantage of this grid resolution is that the accuracy of the approximation is as good as the experimental information provided by the images. A finer grid would only lead to longer computational time but not to a more accurate result. A much coarser grid however would take much less time to compute but would result in the loss of information. However the grid can be downsized by binning the images and still lead to a reasonable result.

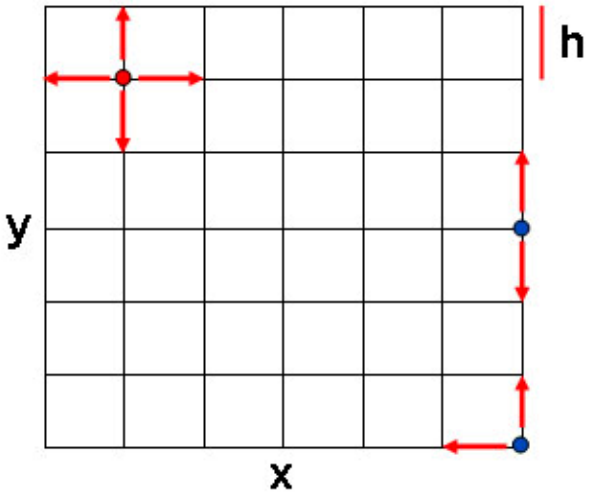


Figure 3.14 Finite differences discretization with Neumann boudary conditions on a 2D rectangular grid. Dots indicate diffusive particles, red arrows show allowed diffusion paths to the nearest neighbors in x and y direction. The grid distance h is equal in x and y direction.

Neumann boundary conditions allow diffusion only between nearest neighbors. At the border of the nucleus, diffusion is allowed only with the neighbors inside the mask, reflecting the fact that nuclear or cellular membranes are impermeable for the observed protein at the timescale of a FRAP experiment (Sprague et al. 2006). For a 2D grid, Neumann boundary conditions can be written as

$$\begin{aligned} c(x+1,y) &= c(x,y) \\ c(x,y+1) &= c(x,y) \\ c(x-1,y) &= c(x,y) \\ c(x,y-1) &= c(x,y) \end{aligned} \qquad (3.1)$$

where c is the concentration (or intensity) of the observed diffusive molecule and (x+1,y), (x,y+1), (x-1,y), and (x,y-1) describe positions outside of the grid and (x,y) are the corresponding neighbour pixels at the border of the grid.

Within Tropical, the diffusion equation (1.7) is therefore written as a system of coupled ordinary differential equations in 2D, one for each element of the grid in the form

$$\frac{\partial c}{\partial t}(x,y) = D/h^2 * (c_{x+1,y} + c_{x-1,y} + c_{x,y+1} + c_{x,y-1} - 4 * c_{x,y}) + f(x,y,t) \qquad (3.2)$$

where D is the diffusion coefficient, h is the gridpoint distance, c is the intensity of a gridpoint and f(x,y,t) represents a reaction term of a molecule.

Applying the correction method that was proposed for inhomogeneous fluorescence distribution during steady state in living cells (Siggia et al. 2000) instead of

$$c_{x+1,y} - c_{x,y} \qquad (3.3)$$

in equation (3.2) the term

$$\left(\frac{c^{\circ}_{x+1,y} - c^{\circ}_{x,y}}{2} \right) \left(\frac{c_{x+1,y}}{c^{\circ}_{x+1,y}} - \frac{c_{x,y}}{c^{\circ}_{x,y}} \right) \qquad (3.4)$$

is computed. An analogue replacement is done for $c_{x-1,y}$, $c_{x,y+1}$ and $c_{x,y-1}$

As seen in chapter 1.3.3 this algorithm is valid in the context of diffusion limited by very fast interaction.

Tropical solves the differential reaction-diffusion equations by a Runge-Kutta fourth order algorithm with adaptive step size adapted from (Press et al. 2002).

3.4.6 Parameter estimation using the Levenberg-Marquardt algorithm

For parameter estimation, the Levenberg–Marquardt algorithm, published by Press et al. (2002) and a modified version of it are implemented in Tropical. In this chapter I will give a very short introduction into the principle of parameter estimation and explain the algorithms implemented in Tropical.

Maximum likelihood estimation

In parameter estimation problems, a model M with m state variables $x_1,\ldots,x_m$ is given. Such a model contains a set of n unknown parameters $P_1,\ldots,P_n$. Parameter estimation is used to identify the parameter values, for which the model predictions are in the best agreement with the experimental data. The maximum likelihood estimation searches for a parameter set P, which maximizes the probability that model M produces the given experimental data (Gershenfeld 1999).

Consequently, the objective function

$$\chi^2(P) = \sum_t \frac{\left(y_{t,\exp} - y_{t,\mathrm{mod}}\right)^2}{\sigma^2} \qquad (3.5)$$

has to be minimized. Thus, the sum of squares of differences between experimentally measured ($y_{t,exp}$) and simulated data ($y_{t,mod}$), divided by the standard deviation of the experimental data (σ^2), has to be minimized in order to optimally fit the model with data from experiments.

Non linear least squares

If the χ^2-function depends on the parameters in a nonlinear way, a general method, which guarantees to find the global minimum of χ^2 does not exist. Instead, there is a variety of iterative methods starting from an initial parameter guess and proceeding iteratively. Most methods are based on the derivatives of the χ^2 function with respect to the parameters P. χ^2 and its derivatives are calculated locally and the parameters are changed accordingly to reach the next minimum. For this reason, computation of the gradient of χ^2 is required.

The steepest descent or gradient descent method seeks for the minimum by stepping into the direction of the gradient

$$P_{next} = P_{current} - \alpha \nabla \chi^2 \left(P_{current}\right) \qquad (3.6)$$

where α is a control parameter for the step size.

Newton's method, assumes that χ^2 can be expanded around the initial parameter set and calculated an iteration step according to

$$P_{next} = P_{current} H^{-1} \nabla \chi^2 (P_{current}) \qquad (3.7)$$

where H is the Hessian matrix, the matrix of second derivatives of χ^2 with respect to the parameters P

$$H_{jk} = \frac{\partial^2 \chi^2}{\partial P_j \partial P_k} \qquad (3.8)$$

This algorithm converges quickly if the initial parameters are chosen close to the minimum, whereas it is not appropriate far away from it. In the latter case, the steepest descent method is more adequate. A robust method combining both approaches is the Levenberg-Marquardt method (Levenberg 1944; Marquardt and SIAM J. Appl. Math. 11 1963), which I have also implemented in Tropical.

Levenberg-Marquardt algorithm

The Levenberg-Marquardt method can be considered an interpolation between the steepest descent and Newton's method. It works very well in practice (Bentele et al. 2004) and has become the standard of nonlinear least-squares routines (Press et al. 2002).

To interpolate between Newton's and steepest descent method, the dimensionless factor λ is introduced as a weighting factor for the contribution of both methods. For computation of the step size of the gradient descent, the diagonal elements of the Hessian are used since they can be regarded as a measure for the curvature of the χ^2-hyperplane. The steepest descent contribution to an iteration step is therefore given by

$$\delta P_j = -\frac{1}{\lambda H_{jj}} \frac{\partial \chi^2}{\partial P_j} \qquad (3.9)$$

Thus, high curvatures result in a low step width and vice versa. The contribution of Newton's method is given by the matrix α' with the elements

$$\alpha'_{jk} = \begin{cases} \alpha_{jk}(1+\lambda) & \text{for j = k} \\ \alpha_{jk} & \text{for j} \neq \text{k} \end{cases} \qquad (3.10)$$

where $\alpha = \frac{1}{2} H$ (compare equation (3.8)).

The iteration steps of the Levenberg-Marquardt method are

$$\delta P = -\alpha'^{-1} \nabla \chi^2 \qquad (3.11)$$

Obviously, for $\lambda = 0$ this is equivalent to Newton's method, whereas for high values of λ, the steepest descent contribution becomes dominant.

The computational workflow for the Levenberg-Marquardt algorithm is as follows (Press et al. 2002):

1. Compute χ^2 (P).
2. Pick a modest value for λ, say $\lambda = 0.001$.
3. Solve the linear equation (3.11) for δP and evaluate χ^2 (P +δP)

4. If χ^2 (P +δP) < χ^2 (P), increase λ by a factor of 10 (or any other substantial factor) and go back to 3.

5. If χ^2 (P +δP) > χ^2 (P), decrease λ by a factor of 10, update the trial solution P=P+δP, and go back to 3.

Also necessary is a condition for stopping. Iterating to convergence (to machine accuracy or to the round off limit) is generally wasteful and unnecessary since the minimum is at best only a statistical estimate of the parameters P. Furthermore, it is not uncommon to find the parameters wandering around near the minimum in a flat valley of complicated topography doing steepest descent in very un-steep degenerate valleys. These considerations suggest that, in practice, one might as well stop iterating on the first or second occasion that χ^2 decreases by a negligible fractional amount like 10^{-3} (Press et al. 2002). It might as well happen that such a flat valley is a local minimum rather than the global minimum in the parameter space. To circumvent a Levenberg-Marquardt algorithm from getting stuck in such a local minimum, a multiple random start routine is often applied. Here several runs of the algorithm are started with different randomly chosen initial parameters. This is a reasonable method but increases computational time a lot.

For Tropical a different approach was developed. By clicking on UH-Jump in the advanced settings tab, the algorithm counts how often λ is increased. After five increments in a row, λ is set to a negative value indicating an end of the current estimation process and forcing the algorithm to start over again with a Newton step, while taken the last estimated parameter as initial parameter. Five increments have proven to work well in practice, at least with all applications tried out during the development of Tropical. This method sometimes leads to a faster convergence.

3.4.7 Output of Tropical

Besides the various output files, that Tropical writes after a Tropical experiment is finished, an output window provides the intermediate results during runtime. In this window one can follow the intermediate results of the parameter estimation process. A log file of the complete parameter estimation process is written to the harddisc simultaneously. A result file with estimated parameters called parameter_log.txt contains information about the single iterations of the parameter estimation (Figure 3.15).

Iteration #	Parameters	Chi squared	Lamda	Sum of Data	Sum of Model	Best Chi squared
1	0.0764695	1.22862	0.01	3.0292e+006	2.6512e+006	1e+038
2	0.0927922	1.24094	0.1	3.0292e+006	2.6478e+006	1e+038
3	0.107779	1.25022	1	3.0292e+006	2.64614e+006	1e+038
4	0.116022	1.25439	-2	3.0292e+006	2.64555e+006	1e+038
5	0.0927211	1.2409	0.0001	3.0292e+006	2.64781e+006	1.2409
6	0.0850994	1.23537	1e-005	3.0292e+006	2.64912e+006	1.23537
7	0.0852901	1.23552	0.0001	3.0292e+006	2.64908e+006	1.23537
8	0.0854808	1.23566	0.001	3.0292e+006	2.64904e+006	1.23537
9	0.0856712	1.2358	0.01	3.0292e+006	2.64901e+006	1.23537
10	0.08586	1.23594	0.1	3.0292e+006	2.64897e+006	1.23537

Figure 3.15 Parameter log file of Tropical. For each iteration, the value of the parameters to be estimated, χ^2 of the iteration, λ, sum of intensities in the estimation compartment for the simulated and for the experimental data and the best χ^2 of all iterations are plotted.

Results of the simulation with the best estimated parameter set are also provided as Tiff formatted images and text images, containing the intensity floating point values. Finally a file called ROI.txt contains the total intensity values of the specified “ROI plot compartments” from the “simulation & parameter estimation” tab of Tropical. Those values can be used i.e. to plot recovery curves of the simulated images.

3.4.8 Source code and algorithms used

The complete source code is written in C++ and is available for Linux and Windows. The graphical user interface was developed using QT, version 4.1.2 (www.trolltech.com). Tiff image I/O is handled by the open source library libtiff (v.3.7.2) (www.remotesensing.org/libtiff/). The model description is compiled from a text file by the open watcom C++ compiler (www.openwatcom.org).

4 Results: Diffusion dynamics of H1°-GFP

Understanding the dynamic behavior of H1 is essential to understand chromatin organization. A number of studies exist, showing that the mobility and exchange of H1 at the nucleosome may be limited by binding and diffusion does not contribute significantly to these processes (Misteli et al. 2000; Catez et al. 2006; Lele et al. 2006). Other studies simply take this as a fact and neglect diffusion when analyzing the dynamics of H1 (Bustin et al. 2005). Nevertheless, Beaudouin et al. (2006) showed that diffusion largely contributes to the dynamics of five isoforms of H1. Therefore I performed FRAP experiments, computer simulation and parameter estimation with Tropical to reveal the diffusion of the linker histone H1°.

Three different forms of H1°, described in chapter 3.1, wild type, mutant “52” and mutant “Six” were tagged to GFP and observed with a confocal laser scanning microscope.

Although the development of the software Tropical was the main result and major work leading to this Ph.D. thesis, I decided to present Tropical in chapter 3 (Materials and Methods) rather than in this chapter and focus on presenting the biological results obtained with Tropical as a demonstration of its functionality and power.

Methods for quantitatively analyzing FRAP data and simulation of diffusion or reaction-diffusion equations have recently been developed (Siggia et al. 2000; Beaudouin et al. 2006), however user friendly computer tools to apply such methods are still missing. I developed Tropical to give scientists working with FRAP an opportunity to analyze their data with an easy to use computer tool that incorporates the most reliable methods. A proof of principle for solving reaction-diffusion equations with Tropical was given by Ulrich et al. (2006).

In this chapter, I show the results of the *in vivo* FRAP experiments and their quantitative interpretation. Furthermore, the results of the simulations and parameter estimations performed by Tropical are shown.

4.1 In vivo dynamics imaged with FRAP

To identify the residues within the globular domain of H1° that are involved in nucleosomal contacts *in vivo*, I used a comprehensive set of H1°–GFP fusion constructs containing individual point mutations kindly provided by Tom Misteli (US National Institutes of Health (NIH), Maryland). He introduced simultaneous point mutations at each of the putative DNA-contact sites within the globular domain (Brown et al. 2006) that have been suggested by in vitro analyses (Duggan and Thomas 2000). The basic residues in the primary site (Lys69, Arg74 and Lys85) and secondary site (Lys40, Arg42 and Lys52) were simultaneously replaced with alanine to form the mutant “Six”. The second mutation called mutant “52” consisted of a single point mutation in the secondary site, where Lys52 was replaced by alanine.

To evaluate the diffusive properties of those H1°-GFP forms, I performed *in vivo* FRAP experiments on a Leica TSC SP2 confocal laser scanning microscope. H1°-GFP and its two mutants were therefore stably expressed in mammalian cells. Furthermore, the images obtained with the microscope served as input data for the simulations calculated with Tropical.

The first images obtained by FRAP are the prebleach images, images of the steady state of fluorescence distribution of the GFP-tagged protein of interest, here H1°. The fluorescence of wild type (WT) H1°-GFP inside the nucleus in steady state shows an inhomogeneous distribution throughout the nucleus (Figure 4.1).

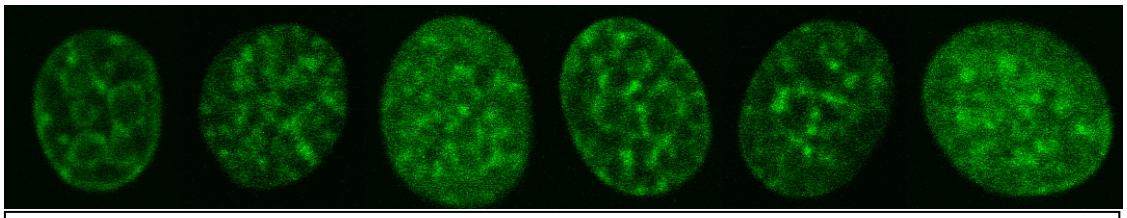


Figure 4.1 Fluorescence distribution of WT H1°-GFP in six different nuclei. Bright areas represent heterochromatin colocalization (het) whereas darker areas show euchromatin colocalization (eu) of stably expressed WT H1°-GFP in living NIH 3T3 cells. Differences in the appearance are a result of different focal planes for each cell imaged.

Brighter regions of WT H1°-GFP inside the nuclei correlate with heterochromatin distribution (Figure 4.2) whereas areas of lower fluorescence intensity indicate euchromatin colocalization. For the mutants “52” and “Six” a similar localization pattern was observed (Figures 4.2 B, C and 4.3).

FRAP experiments were performed by bleaching either a small spot of euchromatin or heterochromatin of approximately 20 pixels (~ 2 µm, depending on the voxelsize of the images) in diameter or a larger area of approximately half the nucleus (Figure 4.4 A-D).

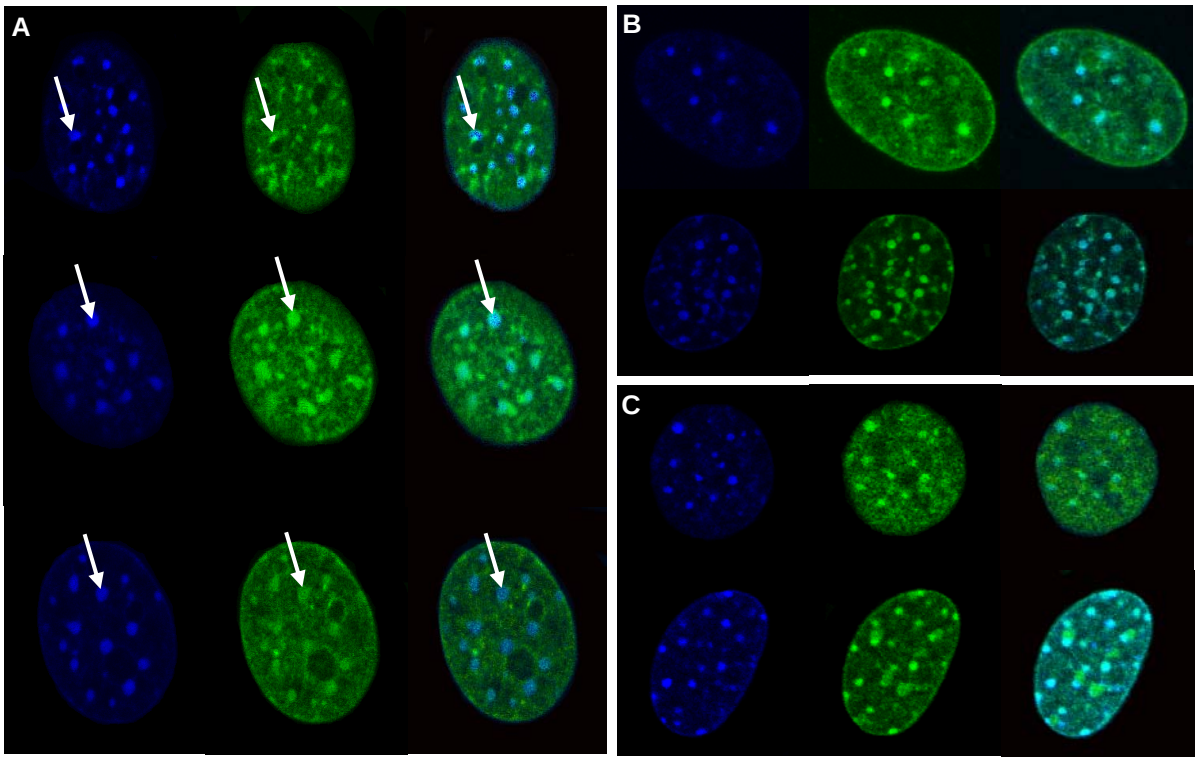


Figure 4.2 Localization of histone H1°-GFP. Stably expressed H1°-GFP colocalizes with the heterochromatin marker Hoechst 33342 (arrows) in different living NIH 3T3 cells. Images were contrast enhanced and background was deleted for a better view. **A** shows the wild type, **B** mutant "52" and **C** mutant "Six".

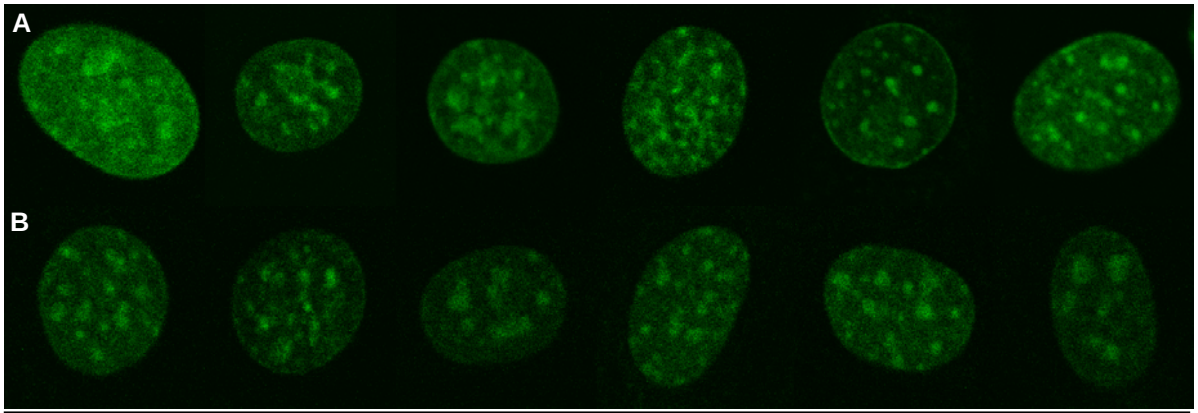


Figure 4.3 Fluorescence distribution of mutant "52" H1°-GFP (A) and mutant "Six" (B) in six different nuclei. Bright areas represent heterochromatin colocalization (het) whereas darker areas show euchromatin colocalization (eu) of stably expressed WT H1°-GFP in living NIH 3T3 cells. Differences in the appearance are a result of different focal planes for each cell imaged.

Normalized recovery curves from FRAP experiments show a significantly different shape that depends on the geometry of the bleached ROI (Figure 4.4 E-G), clearly demonstrating that the dynamic behavior of H1°-GFP is limited by diffusion. Recovery curves are based on fluorescence images after background subtraction and bleach correction. This was also reflected by the recovery half times which were approximately eight times faster with a bleached spot compared to bleaching half the cell for the WT, three to four times for mutant “52” and about 20 times for mutant “Six”. The recovery was much faster in the mutant “Six” with $t_{1/2}$ ~5 s for a bleached spot and ~100 s for a bleached half nucleus, compared to mutant “52” ($t_{1/2}$ ~ 15 s for spot bleaching and ~50 s for half the nucleus being bleached) and even more compared to the wild type form of H1°-GFP, which showed half times for recovery of ~ 160 s and ~1200 s for spot bleach and half the nucleus bleach, respectively. At this point it is crucial to mention that the half nucleus does not correspond to the same number of pixels within the different cells. Calculation of the mobile fractions M_f according to equation (1.1) generally showed a high amount of mobile proteins for all three forms. For the WT, M_f was typically 88 - 95 %, for mutant “52” 92 % - 98 % and for mutant “Six” the mobile fraction was > 98 %.

Quantitative results of the *in vivo* imaging clearly show that the dynamic behavior of H1°-GFP is mainly driven by diffusion. Since it is also known that H1° interacts with DNA, its dynamics are not solely influenced by pure diffusion. The most correct way to describe the dynamic behavior of H1° therefore is via a reaction-diffusion model. The difference in diffusion of the three mutations is most important and unexpected, as they have the same size and structure! Only the kinetics of interaction (related to the affinity) would be expected to change because of the different mutated sites.

This finding shows that estimation of the diffusion coefficient of the three mutants is obligatory to understand the role of the mutated sites in the diffusion and binding properties of H1° and hence contributes to understanding its role in stabilizing the nucleosome, a pseudo-diffusive model incorporating the Siggia (2000) equations, which is a subset of the reaction-diffusion model for the case of very fast reactions (compare chapter 1.3.3) was used first. The advantage of this algorithm is, that it computes much faster and it is already implemented into Tropical without the need to define further equations.

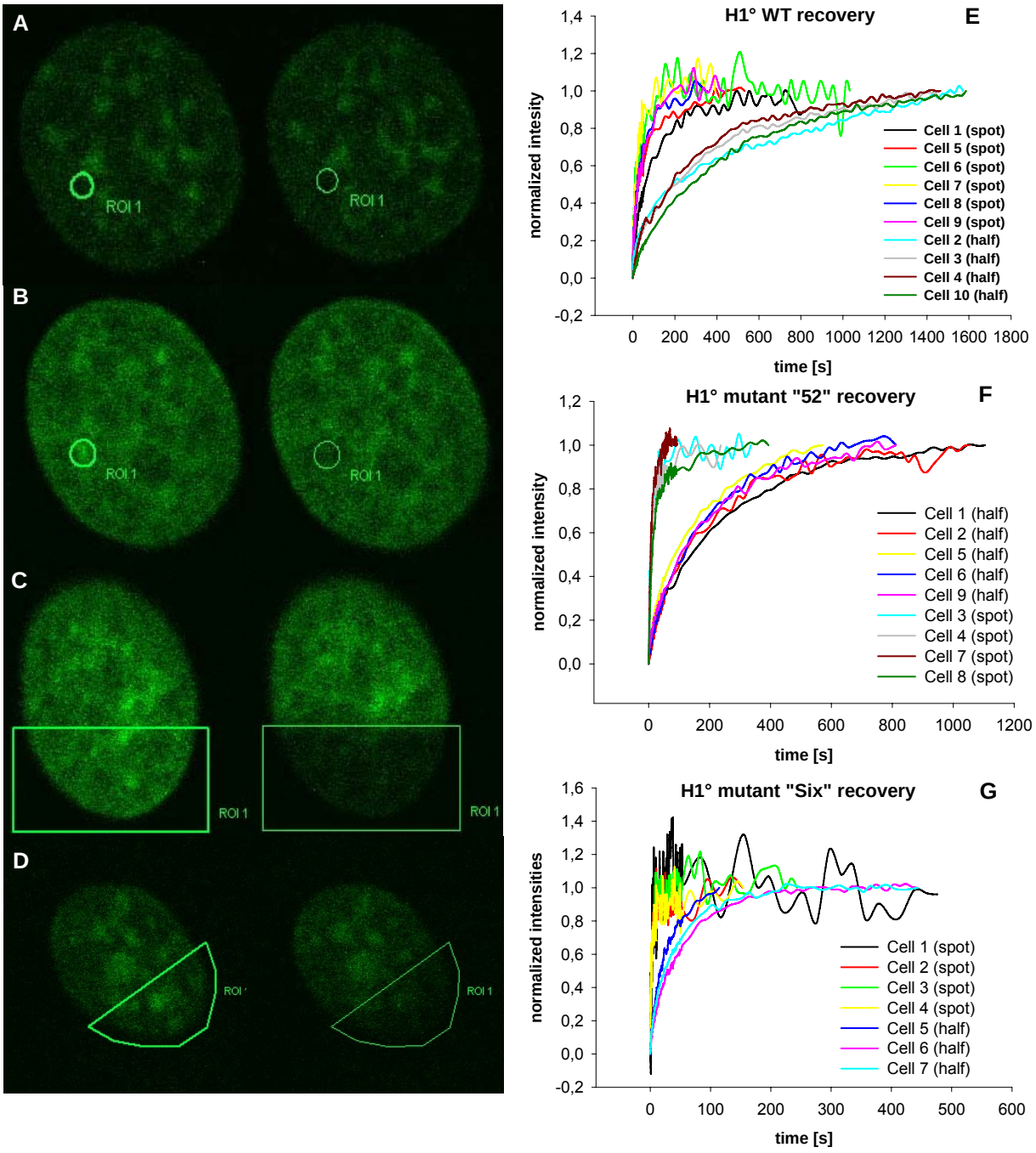


Figure 4.4 Different bleaching ROIs used in FRAP. For each mutation different bleach geometries were used in order to evaluate that recovery of H1°-GFP is diffusion limited. **A** shows a spot bleach of 20 pixels (~ 1.86 μm) in diameter in the euchromatin region. **B** shows the same but in a heterochromatin region, whereas in **C** a larger region of approximately one third to one half the nucleus was bleached. **E-G** show that the shape of the recovery curves depends on the bleached area. The steady state fluorescence distribution is reached much faster where a small spot is bleached compared to larger bleach ROIs. For the wild type H1°-GFP **(E)** and for the mutant "52" **(F)** this difference is very large whereas for the mutant "Six" **(G)** the difference is less distinct due to the rapid recovery of this mutant. Still, the shapes of the curves where half the cell was bleached indicates a slower recovery than the spot bleach curves.

4.2 In silico dynamics and parameter estimation with Tropical

In this thesis, I used Tropical to estimate the apparent diffusion coefficients of the wild type form of the linker histone H1° and two mutated forms of it. Beaudouin et al. (2006) could proof that a diffusion model is able to fit the behavior of different nuclear proteins. The *in vivo* FRAP results shown in chapter 4.1 reveal dynamics of H1° that are clearly limited by diffusion. Therefore, applying a pure diffusion model and estimating the apparent diffusion coefficients of the three H1° forms can elucidate their dynamic properties and lead to a better understanding of the role of H1 in chromatin stabilization.

4.2.1 Diffusion parameters of H1° estimated with Tropical

Tropical estimates diffusion coefficients on the basis of a spatially discretized partial differential equation system. The grid for spatial discretization is represented by the input images. A black and white raw image of the complete nucleus served as a mask for the parameter estimation. χ^2 was estimated comparing the simulated and original values inside this area. Diffusion parameters were estimated based on the *in vivo* microscopy images, which were preprocessed as described in chapter 3.3. Images were binned to a target size of 64x64 pixels from an original size of 256x256 pixels. Time series of the images used for parameter optimization contained between 86 and 155 images, representing a timescale between ~550 s to ~1600 s. Shortest time series were recorded for mutant “Six”, due to the faster recovery, whereas longest time series were recorded for wild type H1°-GFP, which showed the longest recovery times. The size of the recorded voxel was ~ 1 µm in all images. Standard deviations of the images were calculated to be between 1% and 5%, being 1% - 3% for most images. Initial diffusion coefficients were chosen between 0.01 $\mu m^2/s$ and 0.5 $\mu m^2/s$, firstly to check whether the initial parameter had a significant influence on the result of the parameter estimation and secondly, later on, to allow faster convergence by choosing an initial parameter close to the expected one. The fact that initial parameters do contribute to the final result to certain extend is due to the fact that criteria are assigned specifying the desired accuracy and number of iteration steps with a decreasing χ^2. Because step sizes in parameter space depend on χ^2, which depends on the parameter, starting with different initial parameters will result in different steps and hence in a slightly different result, if accuracy criteria are defined to ensure a reasonable calculation time. However, different initial parameters effected the final estimation only up to 1% of the final parameter, which has no influence on the recovery of any protein. Siggia’s correction method (Siggia et al. 2000) was applied to any estimation using a prebleach image, averaged over all recorded prebleach images. Several runs with different initial parameter estimation settings were performed per cell to evaluate the obtained result. Estimated effective diffusion coefficients of the three H1° forms confirm the observed differences in recovery time (Figure 4.11). For wild type H1°-GFP the lowest parameters were estimated, followed by those for mutant “52”, which were about three times higher. Mutant “Six” diffuses had the highest estimated diffusion coefficients and diffuses five times faster than mutant “52” and approximately 15 times faster than the wild type H1°-GFP.

To show that the estimated diffusion coefficients are independent of the reference compartment for parameter estimation, D was estimated for selected cells based on

two different compartments (Figure 4.12). It could be shown that the estimated parameters do not vary significantly using different estimation masks.

The effective diffusion coefficient D_{eff} represents a mixed diffusion coefficient of he free and bound pool of a protein and can be calculated as $D_{eff}=D_{free}*fraction_{free}$ (Sprague et al. 2006). The theoretical D of H1°-GFP WT calculated using the Stokes Einstein Equation, given a molecular mass of H1° of ~20 kDa and the one of EGFP being ~27 kDa, is ~25 µm^2/s in living cells. The observed D of 0.02 µm^2/s reflects a free pool of ~0.08%. For mutant "52" $D_{observed}$=0.07 µm^2/s, yielding a free pool of ~0.28% and mutant "Six" with $D_{observed}$=0.36 µm^2/s would give a free pool of protein of ~1.4%.

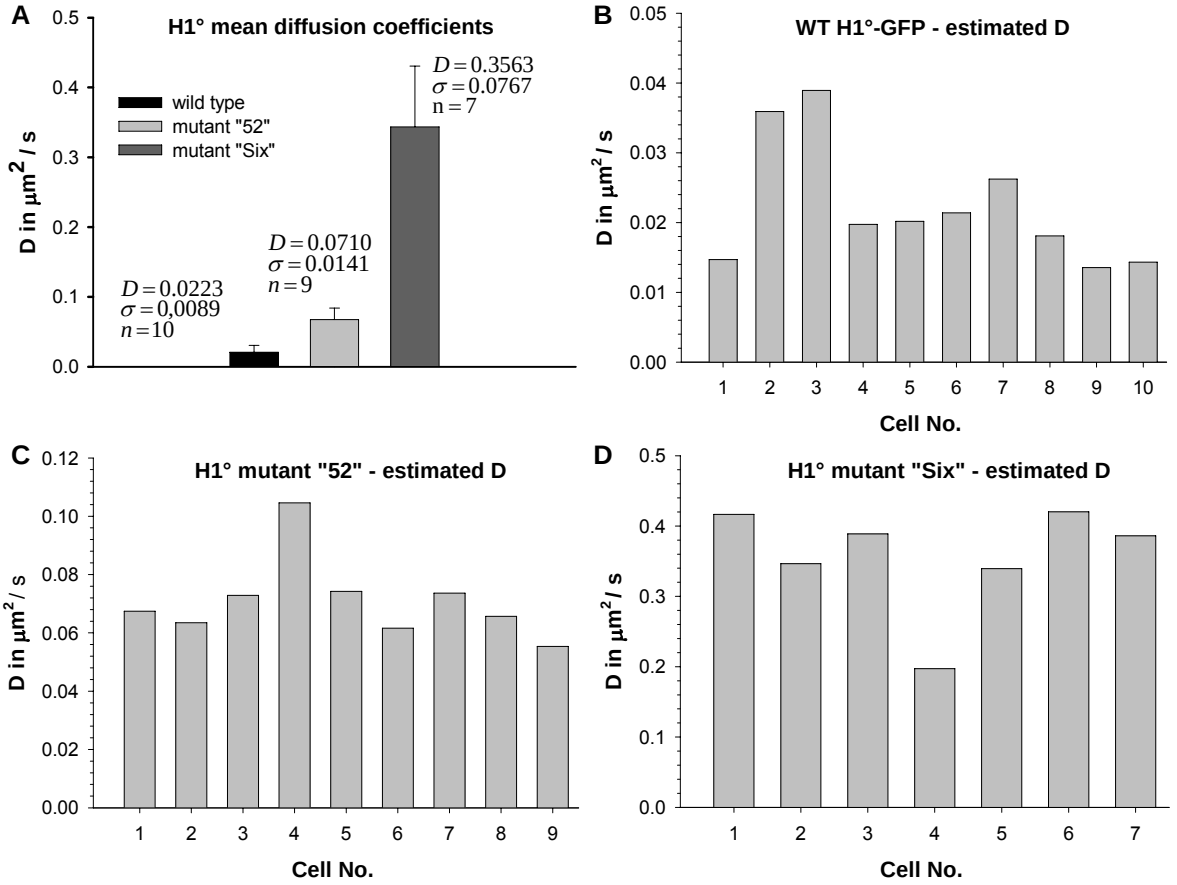


Figure 4.11 Estimated diffusion coefficients of H1°. The mean diffusion coefficients over all processed cells of each H1° form confirm the experimental findings of differences in recovery. WT H1°-GFP diffuses ~3 times slower than mutant "52" and ~15 times slower than mutant "Six" **(A)**. Most estimated diffusion coefficients were relatively homogenous within one form of H1°, however cells 2 and 3 of WT **(B)**, cell 4 of mutant "52" **(C)** and cell 4 of mutant "Six" **(D)** showed a slightly increased or decreased D. Nevertheless, the diffusion coefficients of all cells of one H1° form showed low standard deviations, reckoning living biological systems.

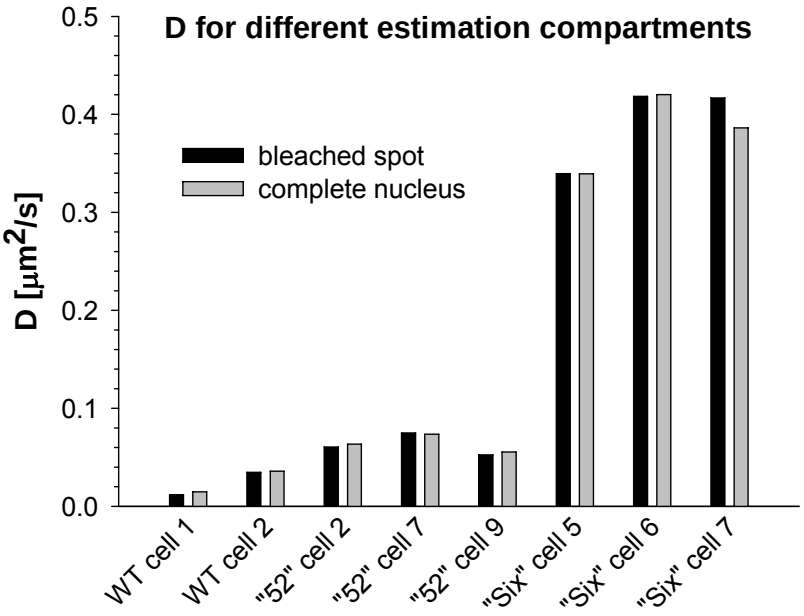


Figure 4.12 Comparison of different estimation compartments. Parameter estimation performed on different regions represented by different masks within Tropical showed that the estimated D is independent of the applied mask.

4.2.2 Simulated recovery of H1°

Tropical writes Tiff images as a visual result of a simulation based on the best estimated parameter set. These simulated images can directly be compared to the original binned images by eye or any image processing software. Furthermore, Tropical uses them to write the intensity values for the recovery curves, defined by a ROI that is loaded in the ROI plot compartments section in the simulation & parameter estimation data tab of Tropical. Observations of the images can give a first hint, if simulation and parameter estimation have performed completely wrong or not and show the order of magnitude of the complete recovery time.

Using Siggia normalization (Siggia et al. 2000) and a bleach-ROI that covered approximately half the nucleus, these images looked the same as the binned experimentally recorded images (Figures 4.5-4.7), indicating a good performance of the parameter estimation and simulation. Comparing the spot bleach images by eye is not possible, because of the small size of the spot and the relatively fast recovery.

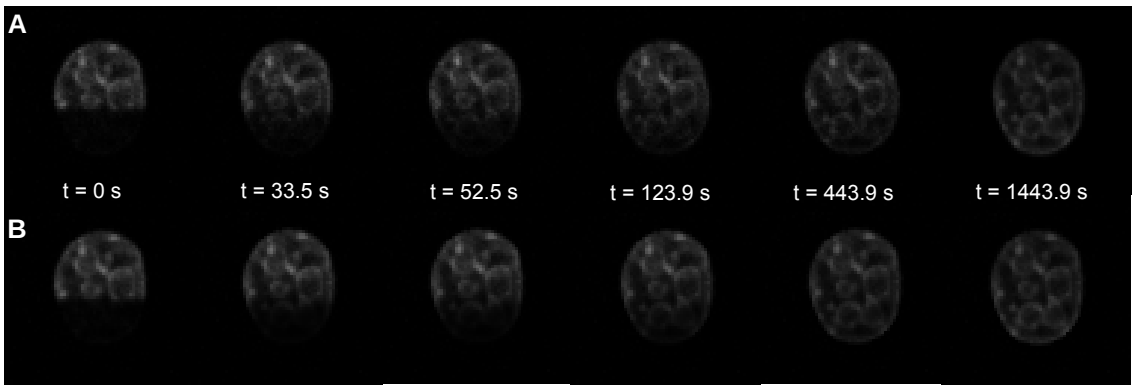


Figure 4.5 FRAP recovery of binned original (A) and simulated images (B) of cell 1 of WT H1°-GFP at selected time points. The simulated images are in good accordance to the experimentally recorded ones. Fluorescence recovery happens slowly. The plateau phase is not even reached after 1400 s.

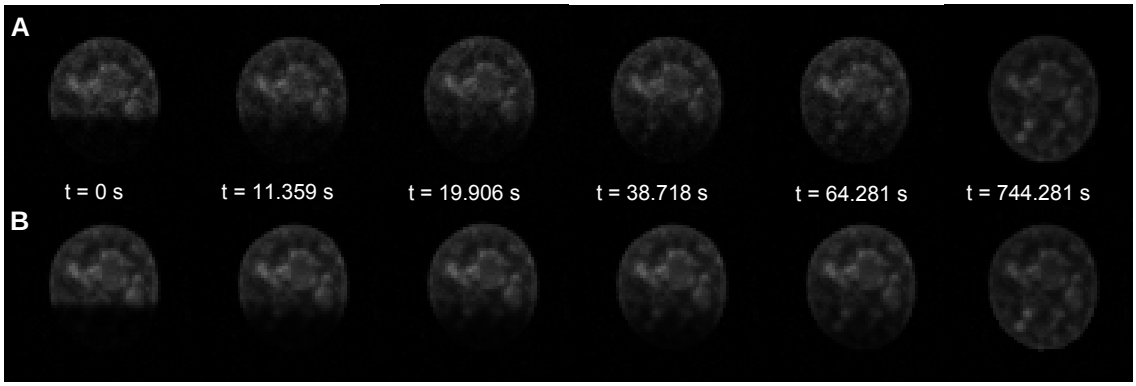


Figure 4.6 FRAP recovery of binned original (A) and simulated images (B) of cell 2 of mutant "52" H1°-GFP at selected time points. The simulated images are in good accordance to the experimentally recorded ones. Fluorescence recovery is much faster than for the WT. After ~744 s the plateau phase is reached.

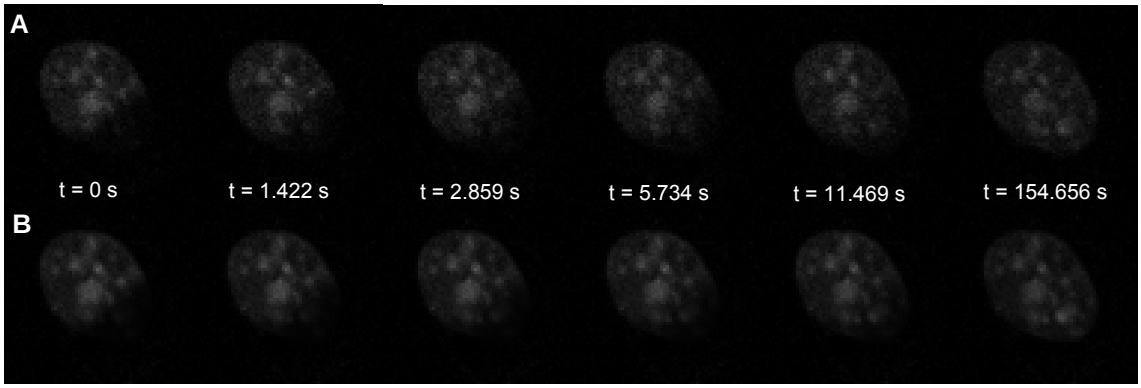


Figure 4.7 FRAP recovery of binned original (A) and simulated images (B) of cell 7 of mutant "Six" H1°-GFP at selected time points. The simulated images are in good accordance to the experimentally recorded ones. Fluorescence recovery reaches the plateau phase much faster (~ 154 s) than for the other two forms of H1°.

As a next step in detail, Tropical writes the file "ROI.txt", containing the floating point intensity values of the selected ROI enabling the plot of recovery curves. Simulated recovery curves can then be compared to original ones. χ^2 of the parameter estimation hereby indicates the goodness of the fit, because the difference of the original

and simulated intensities inside the user-specified parameter estimation ROI is the fundament of χ^2 calculation.

Simulated recovery curves were very similar to their experimentally obtained equivalent (Figure 4.8 – 4.10), showing a good performance of the parameter estimation and simulation with Tropical. This result also shows that the pseudo-diffusive Siggia (2000) approach already fit the experimental data very well. This is a very important finding demonstrating the validity of this model in the case of instantaneous reactions.

Wild type H1° showed a much slower recovery correlated with an estimated apparent diffusion coefficient in the range of 0.02-0.04 µm²/s, compared to mutant “52” with an estimated D ~0.06 µm²/s. Mutant “Six” however recovered most quickly with an estimated D ~0.4 µm²/s.

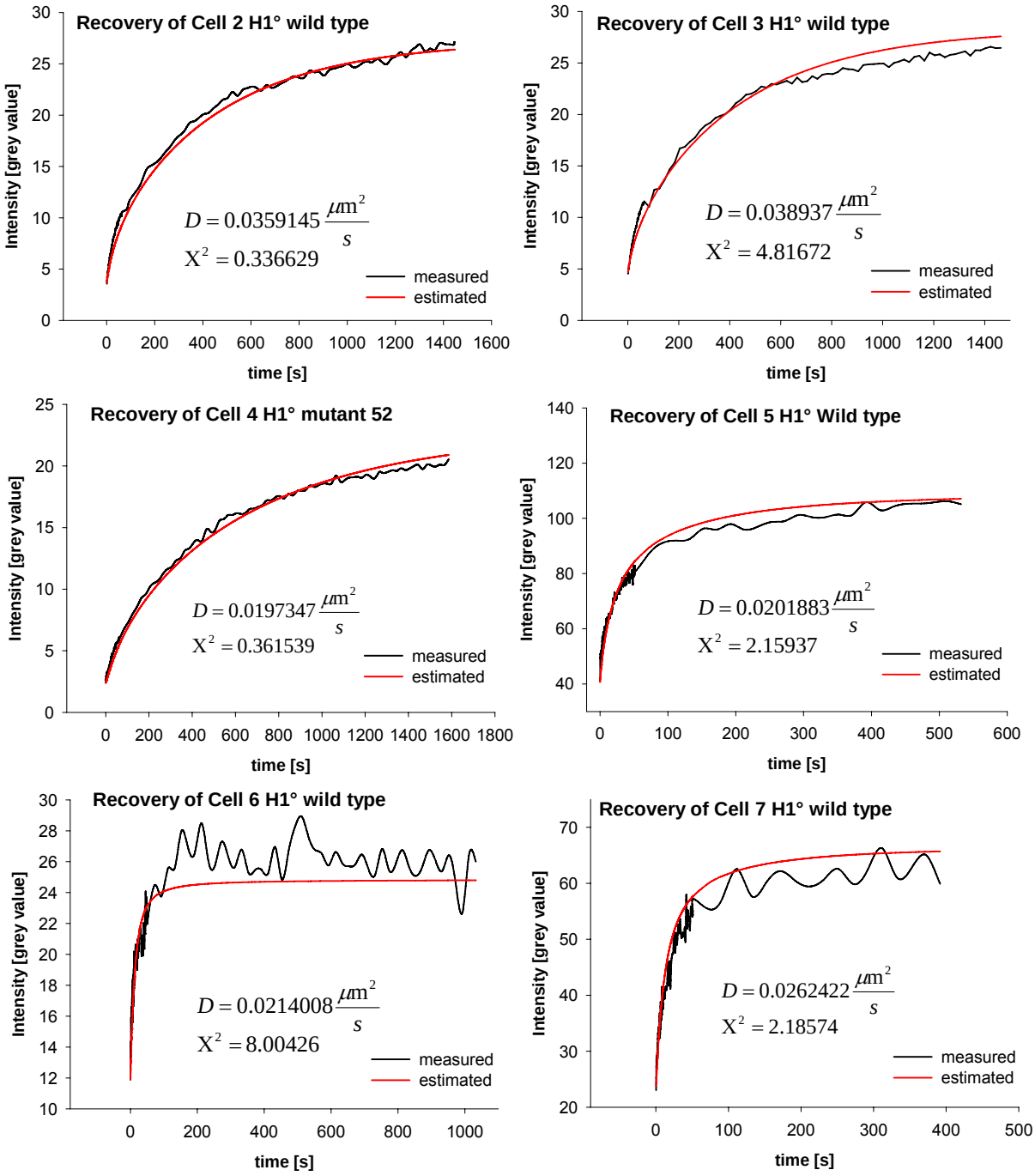


Figure 4.8 Examples of simulated and experimentally obtained recovery curves of wild type H1°. Most graphs show a very good fit, independent of reaching the plateau phase. The estimated effective diffusion coefficients were found to be in a narrow range, significantly lower than those of the mutants. In cells 2, 3 and 4 the bleached ROI was half the cell, whereas for cells 5, 6 and 7 a spot was bleached. The estimated diffusion coefficient was not significantly different for spot bleach or half-nucleus bleached ROI. Simulated recovery curves were calculated with Tropical and represent the mean intensity of specified ROI compartments. The graphs represent the fluorescence recovery in the bleached region of the nuclei.

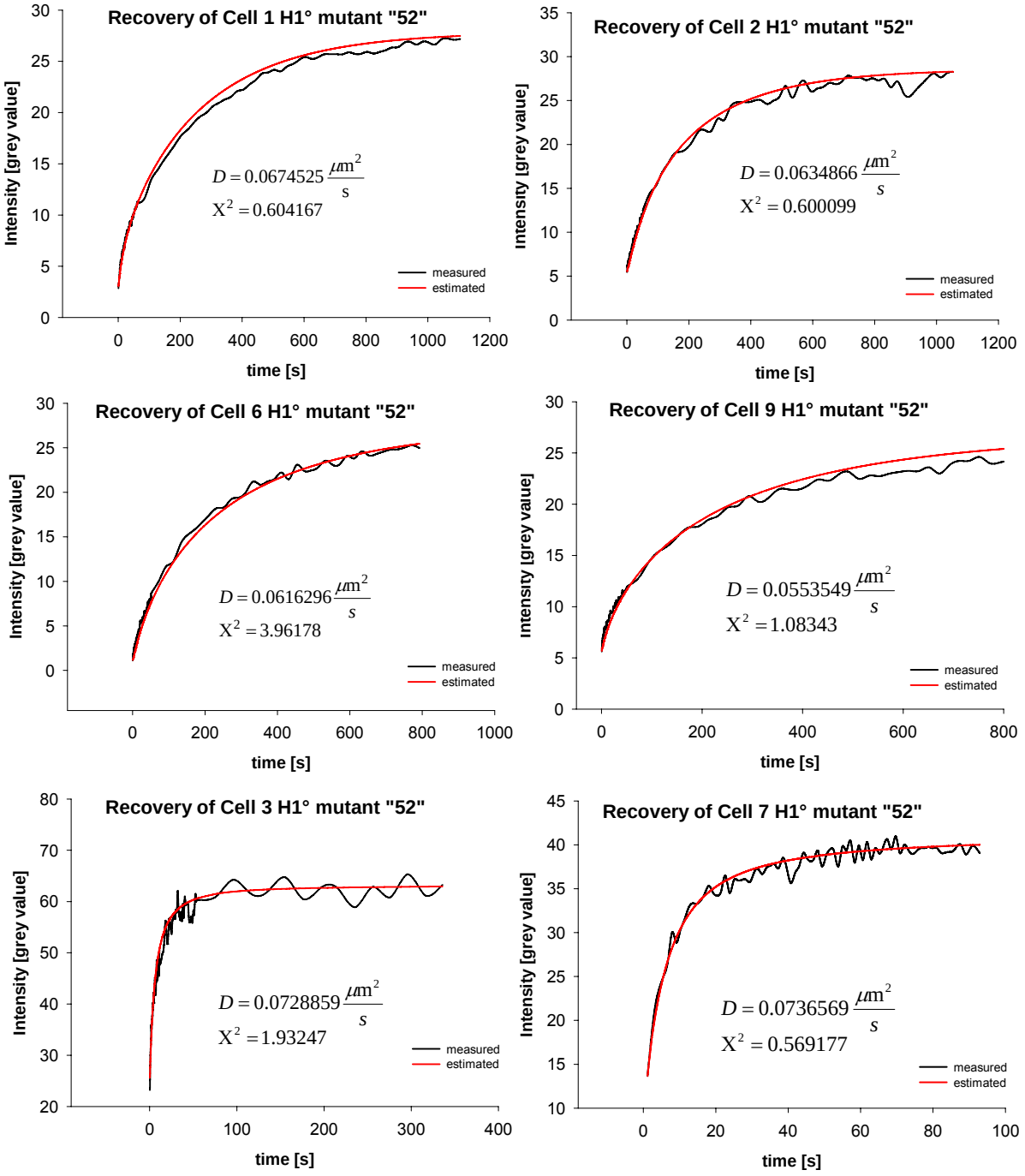


Figure 4.9 Examples of simulated and experimentally obtained recovery curves of H1° mutant „52“. Most graphs show a very good fit, independent of reaching the plateau phase. The estimated effective diffusion coefficients were found to be in a narrow range. In cells 1, 2, 6 and 9 the bleached ROI was half the cell, whereas for cells 3 and 7 a spot was bleached. Simulated recovery curves were calculated with Tropical and represent the mean intensity of specified ROI compartments. The graphs represent the fluorescence recovery in the bleached region of the nuclei.

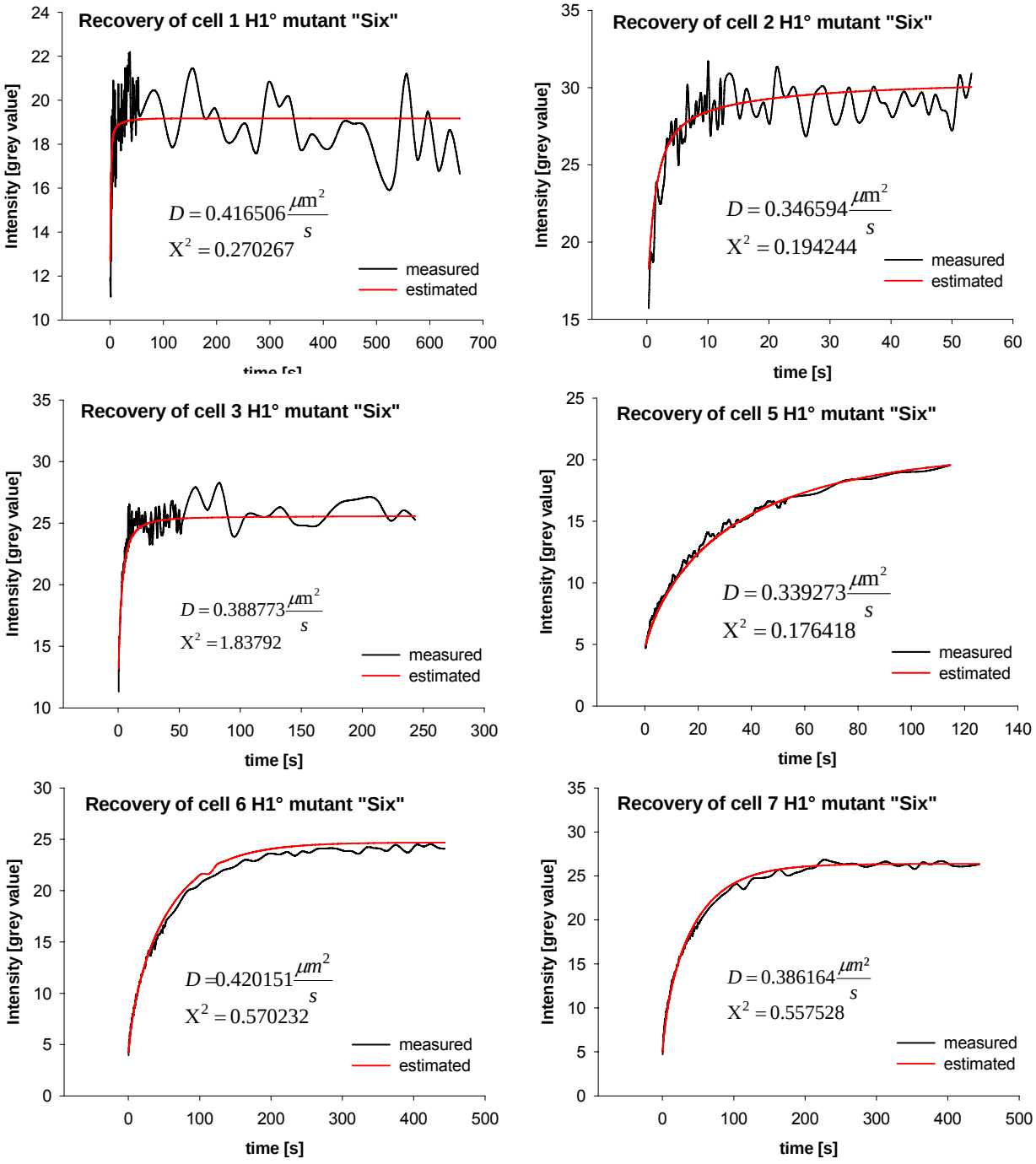


Figure 4.10 Examples of simulated and experimentally obtained recovery curves of H1° mutant „Six“. Most graphs show a very good fit, independent of reaching the plateau phase. However for this mutant, the noise within the experimental spot bleach data is very high due to the rapid diffusion. The estimated effective diffusion coefficients were found to be in a narrow range. In cells 5, 6 and 7 the bleached ROI was half the cell, whereas for cells 1, 2 and 3 a spot was bleached. Simulated recovery curves were calculated with Tropical and represent the mean intensity of specified ROI compartments. The graphs represent the fluorescence recovery in the bleached region of the nuclei.

5 Discussion and outlook

The primary objective of my work was to develop Tropical, a user friendly computer program, incorporating the latest and most accurate methods to analyze FRAP data. Tropical directly operates on microscopy images, incorporating the true geometry of the observed cellular compartments and estimates diffusion and reaction parameters based on a user defined reaction-diffusion model. Tropical delivers simulated FRAP images and FRAP recovery curves as well as the estimated parameters as a result. It is easy to use via a graphical user interface and requires only minor preprocessing of the FRAP data.

The second objective of this thesis was to characterize the diffusion properties of linker histone H1° and its two mutated forms by FRAP and to derive its diffusion coefficient using Tropical and to identify the contribution of the mutated sites to DNA binding.

The experimental procedure of FRAP to measure the dynamics of protein behavior and the interpretation of its results was presented in chapter 1-3.

In chapter 4, I presented the results obtained by FRAP experiments on the linker histone H1° and two mutated forms of it, one single point mutation having Lys52 exchanged by alanine and one form including six point mutations in the primary and secondary DNA binding site, mutant “Six”. I showed that for mutant “52” the effective diffusion coefficient was estimated to be two to three times higher than the one of the wild type H1°, the one of mutant “Six” was even ~15 times higher, indicating that the mutated sites significantly contribute to DNA binding.

In this chapter, I will discuss the role of the dynamics observed in the three H1° forms in living cells by FRAP experiments and the results from their quantitative analysis using Tropical. I will also examine critically the adequacy of FRAP experiments on a confocal laser scanning microscope to gain information on diffusion dynamics of nuclear proteins.

Further in this chapter, I will highlight the advantages of Tropical compared to classical methods of FRAP interpretation and give an outlook on possible applications for Tropical.

5.1 The role of the different mutations on the in vivo H1° dynamics

Several *in vitro* studies led to a rather static model of chromatin organization in the past. The role of the linker histone H1 was thought to be the final step in the formation of the 30-nm fiber by connecting the nucleosomes through static binding of one histone H1 to one nucleosome (Graziano et al. 1994; Thomas 1999).

Recent findings of different *in vivo* studies however found H1 to be very dynamic, frequently binding to and unbinding from DNA and diffusing throughout the nucleus (Misteli et al. 2000; Beaudouin et al. 2006; Woodcock et al. 2006). These studies demonstrated H1 to be necessary for stabilizing nucleosomes rather than forming them to a 30-nm fiber, which is observed only *in vitro*.

One of the surprising, and perhaps sobering conclusions to emerge from recent work on H1 is the extent to which the results from earlier *in vitro* experiments have been misleading. Subsequent *in vivo* work has required a reappraisal of the role of H1 as a chromatin architectural protein.

Hendzel et al. (2004) analyzed linker histone H1.1 dynamics in living SK-N-SH cells. They found that the C-terminal domain of histone H1 is responsible for high-affinity binding of histone H1.1 to chromatin *in vivo* and that high-affinity binding can be directly modulated by phosphorylation at Thr152 and Ser183. The single substitution of Thr152 with glutamic acid in their study reduced the affinity of histone H1.1 binding to a greater extent than did deletion of the C terminus at Lys182, a region that contains about one third of the positively charged amino acids found in the C-terminal domain. This indicates that the C-terminal domain of at least H1.1 may also play a significant role in DNA binding. However their conclusions were made on the basis of recovery times only and the differences shown by the recovery curves for the point mutations are much smaller than the differences between the recovery curves of H1° wild type and the mutated forms investigated in this thesis. This suggests that for binding of H1° to DNA the globular domain seems to have a higher impact than the C-terminal domain for binding of H1.1 to DNA.

Although the differences of the three forms of linker histone H1° analyzed in this thesis are mainly due to their different binding affinities to DNA, Beaudouin et al. (2006) could show that diffusion is the rate limiting step in the overall dynamics of this protein. A simple method to proof whether diffusion is limiting the fluorescence recovery of proteins is to vary the geometry of the bleached area. The resulting recovery will be slower, the larger the bleached area is only if it is limited by diffusion. If binding or dissociation is the limiting event, recovery will be independent of the bleached area.

Lele et al. (2006) like Beaudouin et al. (2006) performed this test, both for H1.1-GFP and got opposite results. I conducted the same test with applying the same correction methods for the three forms of H1°-GFP and was able to confirm the results from Beaudouin et al. (2006), as shown in figure 4.4 (E-G). This strongly indicates that H1 dynamics in living cells are clearly limited by diffusion. Since the three mutants still have the same structure and the same size, their diffusion coefficient should not be distinct. This leads to an interpretation of the estimated effective diffusion coefficients in terms of binding affinity. The estimated diffusion coefficients are therefore a direct readout of the binding affinity of H1° to DNA. Therefore I will talk of affinities in the following. Concerning the results found by Lele and coworkers (2006), this leads to

the conclusion that their findings are wrong, which may be due to mistakes in the data analysis. Following this they drew the wrong conclusions with neglecting diffusion as a prerequisite of their model.

Although the dynamic role of H1 and its isoforms is no matter of discussion anymore, the detailed mechanism of its binding to DNA still is (Catez et al. 2006; Woodcock et al. 2006). Whereas some studies indicate a key role of the globular domain of H1 for DNA binding (Brown et al. 2006), others observed the same for its C-terminus (Hendzel et al. 2004).

Within the globular domain, H1 is 97% homologous to the well characterized avian H5 variant (Pehrson and Cole 1981). The globular domain of H1 (H5 in birds) might bind to nucleosomal DNA at the exit and entry points of the DNA wound around the nucleosomal core. Previous biochemical studies (Draves et al. 1992; Thomas et al. 1992) provided evidence that the binding sites include two clusters of basic residues on opposite sides of the 3D structure of linker histone H1. These comprised Lys69, Arg73 and Lys85 at the exposed primary binding site on the globular domain of histone H5, and Lys40, Arg42, Lys52 and Arg94 at the putative secondary binding site (Ramakrishnan et al. 1993). Lys52 in this putative site and Lys69 in the proposed primary binding site were indeed protected from chemical modification by the association of H5 with chromatin, although to a lesser extent than Lys85 (Thomas and Wilson 1986). Basic residues at the seven positions occupied by arginine or lysine in the primary and second DNA-binding sites are highly conserved in the H1 family (Goytisolo et al. 1996). *In vivo* analysis of the dynamic behavior of these binding site candidates revealed their possible role in DNA binding of linker histone H1 (Brown et al. 2006). Still, the interpretation of these *in vivo* studies was based on rather qualitative methods, particularly the interpretation of FRAP recovery times. The extraction of binding parameters from recovery curves, neglecting diffusion as it has been done in this study, has been shown to lead to serious errors (Sprague et al. 2006).

In this thesis I applied more sophisticated methods (Siggia et al. 2000; Beaudouin et al. 2006) to model and simulate the diffusion kinetics of H1° and estimate the diffusion coefficients of the wild type and two mutants to address the influence of different sites on the DNA binding properties.

For the single point mutant “52”, having replaced Lys52 by alanine, a binding affinity (effective diffusion coefficient) of ~0.07 µm²/s was estimated, which is approximately 3.5 times higher than the affinity of the wild type H1°-GFP (~0.02 µm²/s), indicating that Lys52 contributes to DNA binding of H1°. This is in concordance with the results found by Brown et al. (2006), where this single point mutant was shown to have a ~2-fold increased recovery time (t_{50} and t_{80}) compared to the wild type H1°-GFP. Although my results show a slightly higher impact of this point mutation than those results do, they are in the quantitative range. The results I present in this thesis, however, are based not only on the measurement of recovery times in different cells, making them dependent on the bleach geometry; the results I present here rely on the estimation of the diffusion coefficient of several cells, using different bleach geometries and incorporating the real geometry of the cell. Interpretations of FRAP studies based on recovery times only is very much depending on the recovery time: In such cases it is crucial to image until full recovery, which is not always possible due to photodamaging cells. If recovery times are short, when looking at recovery times only it becomes very much important what the underlying mechanisms of recovery are. Recovery times can be biased, when comparing different proteins, if one protein recovers faster than another one, if the underlying mechanism is unknown.

Additionally, the recovery time is just a characteristic of the recovery curve. My findings however show that the recovery of H1° can be fitted with a reaction-diffusion model with very fast interactions. This provides insight into the mechanisms of H1° dynamics, which are crucial to understand before drawing conclusions on their effects. Brown and coworkers did not provide the scientific basis for their interpretations on H1° binding affinities because they only looked at the recovery times of H1°. In a complex reaction-diffusion mechanism where all parameters of a descriptive model are relevant the recovery time is a useless parameter for description.

This shows that interpretation of FRAP results based on recovery curves only, can sometimes lead to a reasonable qualitative result, however, since this depends on the intrinsic properties of a system, which are usually unknown at the time of performing a FRAP analysis the recovery time approach cannot be validated without further quantitative investigations. Therefore, parameter estimation on the basis of a numerical solution of diffusion or reaction diffusion equations is the method of choice.

Looking at the single point mutations of the two DNA binding sites located at the globular domain of H1°, which are mutated in mutant “Six”, Brown and coworkers (2006) identified alanine mutated Lys69 as the site with the highest impact on decreasing the recovery time (t_{80}) of H1° (~13-fold). The other mutated sites showed a decrease in recovery times ~2-5-fold for single point mutations. My results of the parameter estimation of mutant “Six”, where all six sites (Lys40, Arg42, Lys52, Lys69, Arg74, Lys85) are mutated at the same time, showed a binding affinity, approximately 15-fold the one of the wild type, which again is the same order of magnitude. Besides the more accurate result relying on numerical simulations and parameter estimation compared to recovery curves only, there might be a small synergistic effect of multiple sites being altered. This was also shown by Brown et al. (2006) for the mutation of different H1° site mutations.

Taking together, Brown and his colleagues suggest that Lys52 does not significantly contribute to the DNA binding of H1° and therefore must be oriented away from DNA, while the sites that bind DNA more efficiently, like Arg42, Lys69, Arg74 and Lys85, directly interact with the DNA at the nucleosomal dyad. However, if this was as absolute as proposed, mutation of this site would not alter the diffusion and binding properties of H1° at all; but in fact it does increase the effective diffusion 2-fold.

Summing up the role of the three different mutations on the in vivo H1° dynamics two main results became clear. Firstly, that diffusion is limiting the interaction with DNA for the wild type H1° as well as for mutant “52” and mutant “Six” and secondly, that Lys52 plays a role in the interaction of H1° with chromatin but together with the sites Lys40, Arg42, Lys69, Arg74 and Lys85 has a much greater influence on its dynamics, to which Lys69 may contribute significantly.

5.2 FRAP to analyze protein dynamics

Time-lapse imaging can provide information about the steady-state distribution of a protein over time but it does not reveal the kinetic properties of molecules within a complete population. FRAP however can and has become a standard method in cell biology to reveal protein dynamics *in vivo*. Its advantages and limitations will be discussed in this chapter.

FRAP vs. FCS

Two principles can be used to study dynamics of macromolecules when they cannot be observed directly by light microscopy, because single molecules are below its resolution limit. First, one can characterize fluctuations of fluorescently labeled macromolecules in a small volume, the method used in fluorescence correlation spectroscopy (FCS) (Hess et al. 2002). Second, one can perturb the distribution of fluorescent macromolecules to observe the relaxation towards the steady state distribution, by performing fluorescence recovery after photobleaching (FRAP) or photoactivation experiments.

Experimentally the two methods differ in their scale: while FCS works at a scale of around 200 nm in space and up to 1s in correlation time, FRAP is typically applied at the scale of complete cells or nuclei, several micrometers in space and on a much longer time range of up to hours. While both methods will give the same results if the medium has the same properties at the two different scales, this does not seem to be true in cells. For example FCS curves showed anomalous diffusion for GFP in the nucleus (Wachsmuth et al. 2000) whereas FRAP experiments could be fitted with a normal diffusion model (Beaudouin et al. 2006). This may reflect the fact that at the scale of FCS, one can observe the effect of local diffusion which is smoothed at the scale of FRAP. Although both methods should in principle lead towards the same conclusion, they are based on completely different physical concepts: fluctuations for FCS and perturbation for FRAP. However, their differences in scale make them complementary; FRAP characterizing global properties while FCS reveals local properties of dynamic molecules in living cells.

FCS is well suited for low protein concentrations, which may prohibit its use in high, heterologous expression systems. This is due to the fact, that FCS is based on the measurement of fluctuations in fluorescence intensity within the excitation volume, and with too high a concentration, the relative magnitude of the fluctuation is too low to be measured (Bates et al. 2006). Another limitation of FCS is that fluorophores must be relatively mobile; spending less than one second within the focal volume, otherwise significant photobleaching will falsify the result. Additionally, standard FCS is not able to detect immobile particles, because they do not cause fluctuations in time. This may also influence the accuracy of FCS for characterizing the diffusion of slowly moving proteins.

FRAP is a very good method for observing proteins and their behavior on the scale of the nucleus taking binding to DNA into account. However FCS might be better suited for observations on small scales, if diffusion is limiting, for example when observing very rapidly moving particles. For the application presented in my thesis FRAP was a good method. However, to reveal all information necessary to describe a full reaction-diffusion system with all its parameters, FCS combined with spatio-temporal

modeling and simulation would probably be a better method. Wachsmuth et al. (2003) presented a method combining photobleaching, confocal laser scanning microscopy and FCS which is well suited to quantify fast and slow processes. Using FCS one would possibly have been able to quantify the dynamics of the DNA binding of H1°, a process too fast to be monitored by FRAP. However, for future methods to quantify reaction-diffusion processes *in vivo* a numerical simulation approach like presented in my thesis in combination with FCS could be a major improvement.

GFP-fusion

Although in many cases, GFP fusion proteins behaves sufficiently equal to the original protein, and its localization, movement and dynamic behavior can be monitored and visualized by following its fluorescence inside the living cell, there are some important points to consider when using a functional fluorescent protein to investigate diffusion and binding. The first problem is size: Fusing a fluorescent protein to another protein increases the size of the protein of interest. This results in a slower diffusion of the target protein. If the protein of interest is small in comparison to the fluorescent tag, the decrease of the diffusion coefficient would be small and possibly negligible. EGFP has a molecular weight of ~27 kDa, being a bit larger than H1° (~20 kDa). This would have a non-negligible influence on the diffusion of H1°, which in the case of free pure diffusion would lead to errors in the measured parameters. A second problem is that the GFP fusion might influence the binding affinity of a protein due to conformational changes or steric hindrance. These are typical problems occurring when using fluorescent protein tags, which are extremely difficult to cope. Recently developed small fluorescent tags, like Flash and Reash, might be alternatives, which do at least not have such a high influence on the diffusion behavior of proteins, because they are app. 40 times smaller.

Further problems are: the fluorescent protein must fold correctly to fluoresce, the host protein also needs to fold correctly to be functional, and the integrity of the chimeric protein must be maintained (Miyawaki et al. 2003). But still, many factors have to be controlled before conducting the actual experiments with a GFP-tagged protein. In many cases, steric hindrance or folding interference can occur between the fluorescent protein and host protein; the decision of whether to fuse a fluorescent protein to the amino or carboxyl terminus of a protein depends on the properties of the protein. Poor folding of a fluorescent protein variant results in a non-fluorescent chimera. Accumulation of a large amount of such a protein inside cells will decrease the fluorescent signal, and potentially perturb cellular homeostasis if the labeled host protein retains its original function. It is therefore imperative to use fluorescent protein variants that mature efficiently. A further problem is the potential aggregation of fluorescent proteins, which impedes any cellular application and leads to cellular toxicity (Miyawaki et al. 2003). Those factors can be controlled by performing functional tests, checking if cells divide properly and the fluorescence shows the expected distribution. But still, introducing a fluorescently labeled protein into cells may never be possible without any effects at all. The illumination of fluorescent proteins produces toxic reactive oxygen species (ROS), which in the crowded environment within cells tend to react at sites very close to the site of generation (Remington 2006). For many applications, this photo toxicity is low enough to be ignored, which is not surprising given that fluorescent proteins evolved in organisms exposed to oxygen and plenty of sunlight. Dixit et al. (2006) present some guidelines to avoid these difficulties, and address issues associated with photobleaching and photoactivation.

Usually a cell is expressing the GFP-tagged protein additionally to the endogenous form of the protein. Even stably transfected cells slightly overexpress the protein of interest. Therefore the ongoing search for fluorescent markers with less effect on protein behavior is very important.

FRAP vs. PA

Photobleaching has been and still is a method of choice to perturb the steady state distribution of fluorescently tagged proteins and to analyze their molecular dynamics in living cells. However, FRAP inherently provides a large dose of excitation light sufficient to photobleach the fluorescent protein to be studied, enlarging the possibility of photo-oxidative damage (Dixit et al. 2006). The recent development of photoactivatable proteins (Patterson and Lippincott-Schwartz 2002; Ando et al. 2004; Chapman et al. 2005; Lukyanov et al. 2005) is therefore likely to facilitate protein dynamics studies.

In a photo activation experiment, one can mark a subpopulation of PAFP molecules with illumination at a specific wavelength, and then follow the signal of the photoactivated population with high signal to background ratio. Motion tracking of labeled species does not require continuous monitoring, lowering phototoxicity. This approach greatly reduces the problems associated with photobleaching and continuous turnover of FP markers within the cell (Dixit et al. 2006).

Photoactivation has some major advantages over photobleaching:

Newly synthesized GFP-fused proteins may contribute to the fluorescence recovery over long times in the case of FRAP, while newly synthesized PAGFP for example remains non-fluorescent and therefore unobserved.

Further, GFP photobleaching is partially reversible for the first few seconds after bleaching. This interferes with fluorescence recovery and can make FRAP difficult to analyze quantitatively over short time scales (Remington 2006). Sinnecker et al. (2005) have shown that experimental settings commonly used in microscopy experiments may induce reversible photobleaching of the fluorescent proteins ECFP, EYFP and EGFP, which is more pronounced at acidic pH. In their findings, fluorescence intensities recovered spontaneously with time constants of 25-58 s. Therefore they conclude that the observed steady-state fluorescence reflects a variable equilibrium between reversible photobleaching, spontaneous recovery, and light-induced recovery and that these processes can cause significant artifacts in commonly applied imaging techniques like FRAP. However Sinnecker and colleagues (2005) present considerations, which, taken into account in FRAP experiments, may avoid artifacts due to reversible photobleaching. On the other hand, photoactivation reversibility would only affect the molecule population that has been activated, the same way as photobleaching due to acquisition does, and is therefore easy to correct for by normalizing ROI intensities with total fluorescence.

The process of photoactivation is fast compared to photobleaching, about ten times faster in the case of PA-GFP, EGFP and fusions with nuclear proteins (Beaudouin 2003), which can prevent any significant movement during photoactivation. Importantly, this simplifies the definition of initial conditions for computer simulations when several populations of fluorescent molecules are present in the photoactivated region by considering that no significant diffusion occurred during photoactivation. However, for models like the one described in my thesis it is necessary to know the steady

state distribution of fluorescence, which can readily be measured in FRAP before bleaching. With PA, the steady state is dark and can therefore not be imaged.

Taken these recent findings into account, photoactivation for some applications might be a preferable method compared to photobleaching in the future.

5.3 Applicability of Tropical to analyze FRAP data

I developed Tropical to fulfill the needs of scientists performing FRAP measurements to quantitatively interpret the generated data, namely images, to gain knowledge on the diffusive properties of proteins but also on their reaction mechanisms. Tropical estimates any specified parameters, such as diffusion coefficients, binding or dissociation constants and runs a simulation with the best estimation result based on the input of (1) spatio-temporal microscopy images, (2) initial images for all state variables and (3) a user-defined model. The model is composed of one ordinary differential equation describing the reaction of each molecule. The diffusion term is added automatically. I applied a normal diffusion term which can be modified as described by Siggia et al. (2000). Binary compartment images serve as masks for the simulation grid. They define the area where parameters are estimated and specify regions for computing a recovery curve. Additionally, an output window provides the intermediate results during runtime. A log file of the complete parameter estimation process and a result file with estimated parameters and numerical control values are also written. Results are also provided in tiff formatted images and a text file.

Tropicals' main advantages are that (1) that it directly operates on microscopy images, (2) an inhomogeneous distribution of binding partners can be considered and (3) the obtained result can directly be verified.

The fact that Tropical estimates the parameter based on real microscopy images and a user defined partial differential equation model is new and has many advantages. Compared to classical approaches, which quantify FRAP results only on the basis of recovery curves it provides a much higher precision. Further, the validity of the used model is implicit in the result and hence, information on the underlying mechanism can be obtained, which is not possible when looking at recovery curves only. On the contrary, when deriving parameters from recovery curves one assumes a certain mechanisms being present without knowing anything about it. Recovery curves plot the mean or total intensity in a specified area of interest. They plot a single value for each time point (Lippincott-Schwartz et al. 2001). By fitting this curve to an exponential function, the classical approach estimates a diffusion coefficient dependent only on the recovery time based on this single value per time point (Weiss 2004). Using Tropical all parameters are estimated on the basis of intensities of thousands of pixels per time point representing the complete area of events under observation. Furthermore, the shape of a recovery curve depends on the size and geometry of the bleached area in diffusion limited systems (Lippincott-Schwartz et al. 2003). Fitting a recovery curve to estimate a diffusion coefficient in such a system needs thorough definition of the assumptions made to the model. It is valid when bleaching a small circular spot, assuming a Gaussian profile of the bleach. Recent theoretical work shows that recovery curves that seem to contain two recovery phases cannot necessarily be separated into two distinct processes occurring at different timescales (Braga et al. 2007). Hence, this qualitative approach only allows the comparison of experiments using the same bleach geometry, which might rarely be encountered when comparing experiments from different studies. Diffusion coefficients estimated with Tropical on the other hand are totally independent of the bleached area, because it considers all pixels of the observed object, as I have shown in chapter 4.

So far only a few studies have incorporated the inhomogeneous distribution of binding sites inside living cells (Beaudouin et al. 2006; Sprague et al. 2006). Even recent

models dealing with binding and dissociation events observed by FRAP usually neglect diffusion and assume a homogeneous distribution of binding sites (Lele et al. 2006). This is unrealistic in most cases and not sufficiently accurate for a quantitative interpretation of binding. For Tropical I implemented a method published by Siggia et al. (2000), which allows the correction of the diffusion equation to address inhomogeneous distribution of fluorescence. This way Tropical is able to simulate diffusion in a way that forces the resulting fluorescence distribution towards the observed one resulting in simulated images comparable to the experimentally recorded ones. But Tropical can also handle inhomogeneous distribution of immobile and even mobile binding sites, as long as their distribution is known and can be represented by an image. Such an image can either be obtained directly by microscopy through recording the fluorescently labeled binding partners or can be calculated by image processing techniques as this was done for the proof of Tropicals' concept in Ulrich et al. (2006). This way, Tropical can simulate complex reaction-diffusion systems and estimate the corresponding parameters, representing an entirely new concept of FRAP analysis software.

The third main advantage of Tropical is related to its output. Tropical writes simulated tiff images, values to plot recovery curves for any specified region plus a file containing the estimated parameter per iteration together with a number of numerical control parameters like χ^2 and λ. This way it is not only possible to directly judge the goodness of the estimated parameters but more important it is possible to directly compare any recovery curve of the simulated images with the corresponding one of the original images. Furthermore the simulated and the original images themselves can be compared by any image processing software. This allows a fast and accurate verification of the generated results. Compared to the classical approaches this is a great advantage, since estimated diffusion coefficients from recovery curves cannot be verified easily. For binding parameters this is even more difficult.

Taken together, Tropical incorporates advanced methods while still being user friendly. It has several advantages compared to existing methods: it is more precise, can directly handle microscopy images, addresses inhomogeneous fluorescence distribution, its results can directly be verified and it provides information about the underlying mechanisms of the observed process. Therefore Tropical is a new software well suited to quantitatively analyze many dynamic processes observed by fluorescence microscopy data.

Outlook

Tropical and its manuals (which are also attached to this thesis in the Annex) is available for download at

http://www.dkfz.de/tbi/projects/modellingAndSimulationOfCelluarSystems/tropical.jsp

So far we have used Tropical to investigate the protein dynamics of linker histone H1 and of the nuclear protein B23 (Ulrich et al. 2006).

Tropical can be used to analyze any protein dynamics, especially those limited by diffusion. It was developed to estimate parameters of user defined models. The main task towards the knowledge of protein dynamics is now the development of such models, which is in many cases nontrivial. However, having a tool that allows the direct processing of developed models is a major step forward. Since biology is di-

verse, models describing this biology will always be diverse and model description will stay assigned to scientists' brains.

The processing of massive upcoming biological data from systems biology and recent cell biology methods remains a field for software development. With Tropical, I provide a tool that fills one gap – the accurate quantitative analysis of fluorescence microscopy data.

6 Tropical Manuals

In the following the manuals for the use of Tropical are provided, which can also be downloaded in the latest version from

http://www.dkfz.de/tbi/projects/modellingAndSimulationOfCelluarSystems/tropical.jsp

6.1 Installation instructions for Microsoft Windows

• To install Tropical, simply unpack the file to any directory. To start Tropical, double click tropical.exe located in the folder Tropical1.0.

• To use the example files pure_diffusion.xnp and reaction_diffusion.xnp located in the folders pure_diff and react_diff, provided with the .zip archieve, it is necessary to copy the directory Tropical1.0 and all included subdirectories into the home directory C:\. Otherwise the paths to the example images are incorrect and Tropical will not find the images!

• If you don't want to use the provided .xnp examples, you can build your own experiment with the provided examples following the instructions of the manual (Tropical handbook, Timestamp file documentation and modelfile documentation).

• To open the example files, klick "open" and select one of the two .xnp files, located in the directories pure_diff or react_diff. To start the parameter estimation of the example file klick "Start Simulation" or press Alt+S.

• If something goes wrong, check if the molecule settings are correct: Go to the "Model" tab, click on one molecule and click "Edit". For the pure_diffusion.xnp, the checkbox for "Estimate diffusion coefficient" should be checked, the "sum up with other molecules" box must be unchecked. For the reaction_diffusion.xnp, the checkbox for "Estimate diffusion coefficient" should be checked for molecule "free" only, not for the others, the "sum up with other molecules" box should be checked for the molecules "free" and "bound", for "ratio" it remains unchecked.

For further instruction see "Tropical handbook", "Tropical timestamp file documentation" and "Tropical modelfile documentation".

6.2 Tropical Version 1.0 handbook and documentation

Contents of the manual

Preface

Software for modeling, simulation and parameter estimation supports mostly models of ordinary differential equations (ODEs) that do not include spatial information. Although it is possible, with some effort, to apply some of the available tools to spatio-temporal modeling, very few tools allow the usage of spatio-temporal data acquired from microscopy experiments for parameter estimation. *Tropical* allows simulation and parameter estimation of diffusionreaction models based directly on microscopy images.

Tropical consists of a model generator for the ODE system (reaction equations), an ODE solver for the spatially discretized reaction-diffusion equations and a parameter estimator to estimate kinetic parameters directly from microscopy images. Spatial discretization of the reaction-diffusion partial differential equation (PDE) is performed with finite differences on a grid represented by the pixels of the images. A normalization method proposed by Siggia (2000) is implemented in the software for optional use.

The spatially discretized reaction-diffusion system is solved by a Runge-Kutta 4th order algorithm using adaptive step size. Parameter estimation is performed by a Levenberg-Marquardt algorithm or an optionally available modified form of it: Instead of a widely used multiple random starts approach for the initial parameters *Tropical* allows the algorithm to continue with estimated parameters that produce a larger error until it reaches a certain number of worse steps (the numerical parameter λ is increased 5 times in a row). If this criterion is reached, the algorithm reinitializes with the last estimated parameter. This way in some cases the Levenberg-Marquardt algorithm is able to overcome local minima and can still find the minimum in a reasonable time. Observation showed that in some cases this modification leads to convergence of the parameter estimation in a shorter time. The algorithm uses less iterations until the optimal parameter set is found.

The software is well suited to estimate diffusion coefficients and reaction rates from FRAP (Fluorescence recovery after photobleaching) experiments. Input data and simulation parameters are:

• Tiff images as initial conditions for each molecule (for FRAP experiments post bleach images are used)

• The tiff image time series of the FRAP experiment beginning with the first post bleach image

• For the tiff images an optional binning is available to reduce the image size which results in shorter computational time

• Number of time points of the post bleach images

• Voxel size of the experiment

• Standard deviation of the images

• A binary raw file as a mask for the compartment one wants to simulate on (e.g. a cell or a nucleus)

• The model itself consists of the involved molecules which are assigned one of the initial images, a reaction equation in form of an ODE, initial parameters for the ODEs and a diffusion coefficient (the diffusion term is automatically added to the equations)

• If parameter estimation is selected it is necessary to select a compartment (binary raw file) on which the parameters will be estimated as well as a first and last image of the time series that are used for estimating parameters

• Also a constant time interval between two images of the input series or a text file that contains the time steps in detail must be selected

• If parameter estimation is not selected and only a simulation with fixed parameters will be performed, the simulation time has to be specified

• Optionally the Siggia normalization approach can be selected for simple diffusion equations (Siggia et al., 2000)

• Output settings and folders as well as a name for the simulation experiment should be specified

Most FRAP experiments make it necessary to sum up different state variables (molecules) of a model, since the experimental data show only one channel (e.g. GFP) and no distinction between e.g. a free and a bound molecule is possible.

Therefore *Tropical* offers the possibility to sum up different molecules to represent the experimental time series. Approaches published so far are limited to models with only one diffusing molecule.

(1) System Requirements

- Intel Pentium IV 2,0 GHz or higher / AMD Athlon 2200+ or higher recommended, however *Tropical* may run on slower machines, the faster the CPU the faster the calculation

- At least 512MB of RAM (2048MB required for complex simulations with high resolution images)
- 20MB of available hard-disk space (on top of the microscopy data)
- An imaging or painting software able to create binary RAW images

(2) Installation

To install Tropical just extract the zip file into one directory. The main folder can be copied to any desired path. However it is required to maintain the structure of all subdirectories and files contained in the main folder. To start Tropical for Windows, run tropical.exe.

(3) Using Tropical

To use *Tropical* for parameter estimation it is necessary to have a time series of tiff images, tiff images of initial conditions for each molecule, the specifications of those images including voxel size, time steps, standard deviation of the images, binary raw files as masks for simulation, parameter estimation and regions of interest (ROI) for writing out data and most important a model written as ordinary differential equations.

(3.1) A word on the Data

Sometimes the images need some pre-processing before they can be used with Tropical. Feel free to try any image pre-processing you would like to improve your results but make shure that the following requirements are fullfilled: Tropical reads only 8 bit TIFF images with an IBM / PC byte order. Therefore if your images are of a different format you have to convert them to TIFF before you can use them with Tropical. This can be done with programs like Irfan View, Adobe Photoshop or others. ImageJ, a widely used Image processing program can only write tiff images with Macintosh byte order. Tropical can not read those images until the byte order is changed to PC byte order! The images must contain only grey-values (0-255) (most image processing software has a feature to convert images to grey-value images). The images must be of quadratic size (e.g. 64x64, 512x512). The compartment masks must be binary RAW images. This means the RAW images must only contain two colors: black and white, whereas black means “in the Compartment” and white means “out of the Compartment”. Also we tried to include as many scenarios as possible in the development of Tropical, it is impossible to write software that fits to every possible experiment.

Therefore we strongly recommend looking critically at the results and checking them closely to be able to exclude a misinterpretation because of a wrong model or some wrong settings.

For a simulation Tropical needs only images of the starting conditions. That means one image for every molecule. For parameter estimation Tropical needs the images of the starting conditions and the images of at least one further time step. However a larger number of time steps are necessary to calculate a reasonably precise result. Working with images always needs a lot of memory and a fast processor. To have some influence on the resources Tropical needs for a simulation, you can adjust the

size of the images without changing the images themselves. This is done by binning. The binning reduces the size of your images by a certain factor and thereby reduces the memory and the processor time the simulation needs. Another positive side effect of the binning is noise reduction.

We recommend image sizes of 64x64, 128x128 or 256x256 for a simulation in an acceptable time. In principle you could reduce the image size by every factor you like but we strongly recommend reducing the image size to an integral size, otherwise interpolation errors can occur.

Example: Original image size: 640x640

Binning Factor: 5

Image size used in the simulation: 128x128

Since binning replaces a pixel value by the average of a number of neighbouring pixel values, it can happen that objects shown in an image become smaller if they have very low grey values. For example an image with an object of grey values 1 will certainly be reduced in relative size since pixels at the border of the object will be replaced by the integer average of (1+0+0+0) / 4 = 0 (with a binning factor of 4). If you have images with such low grey values just multiply them with an integer value and correct for that in your equations or results.

(3.2) Running parameter estimation

1. Start Tropical.
2. Enter a name for your experiment in the box at the bottom of the window.
3. **Go to the “Input data” tab.**
4. Click on “Add” in the “Time Series Images”-box.
5. Select the Tiff images to be used for parameter estimation (usually a time series) and press “Open”.
6. Enter a binning factor (see “3.1 A word on the data”).
7. Enter the image settings.
   - 7.1 Enter number of time steps (equals the number of images in the time series images box)
   - 7.2 Enter the voxel size of your images (this information is usually provided within your microscopy experiment)
   - 7.3 Enter the standard deviation of your time series images (this can be an averaged value over all images)
8. Click on “Add” in the “Initial Images”-box and select the images representing your initial conditions for the simulation. In a FRAP experiment these are the first post bleach images.
9. **Go to the “Model”-tab.**
10. Click “Add” in the “Compartment”-box.

11. Enter a name for the compartment you want to perform the simulation on.
12. Add a binary .raw file (mask) representing the simulation compartment by clicking on “Open Raw File” and selecting the .raw file. (Remember: in the .raw file “black” represents the compartment)
13. Click “Add” to accept the compartment settings. The compartment settings **can’t be changed afterwards**. If you want to change some settings you will have to delete the compartment and recreate it.
14. If you have a model file (see “Tropical_Model_File.pdf”) click “Load Model from File”, select the model file and click “Open”. If you have a model file and don’t want to add additional molecules go on with step 28.
15. Click “Add” in the “Molecule”-box if you want to add additional molecules or if you don’t have a model file.
16. Enter a name for the molecule
17. Select the number of the initial image (as stated in the “Initial Images” – box) representing the initial conditions for this molecule.
18. Enter a diffusion coefficient. If the molecule is non-diffusive set the value to 0.
19. Select whether you want to estimate the diffusion parameter (the value will be used as initial parameter for the parameter estimation) or not (the value will be treated as a fixed parameter).
20. Check the box “Sum up with other molecules” if you wish to get the sum of this and other molecules as output. This can be if your time series data represent a sum of several initial conditions, e.g. free plus bound pool of a GFP-tagged protein, but you have initial conditions for free and bound pool separately.
21. Enter the differential equation for the reaction (see “4.2 Reactions” for the syntax of the differntial equation).
22. If the molecule participates in a reaction, enter the initial values for the reaction parameters, used in the equation above, in the “Parameters” box, (see “4.3 Reaction Parameters” for the syntax of the parameters).
23. Click “Add” to accept the molecule settings. These settings **can** be changed afterwards by clicking “Edit” in the “Molecule”-box and using the molecule browse control on the bottom of the dialog to browse through the molecule.
24. Repeate the steps 15-23 until you have created all molecules you need for your model.
25. With the “Save Model to File” button you can now save your model to a seperated text file (model file, see “Tropical_Model_File.pdf”), which you can import later via the “Load Model from File” button.
26. **Go to the “Simulation & Parameter Estimation”-tab.**
27. Select which image you want to use as reference image for the Siggia normalization (Siggia et al., 2000, Biophys. J. 79) in the “Simulation Settings”-box or select “No Normalization” if you don’t want to use this approach for the simulation. This approach is only valid if you do not need to sum up initial images to represent the time series data.

28. Check the “Parameter Estimation”-box. If you don’t want to perform parameter estimation, see “3.3 Running a simulation without parameter estimation”.
29. Select the compartment on which the parameter estimation will be performed by selecting a binary .raw file representing the compartment. This can be e.g. a specific region of interest, the bleached part of your cell or your complete cell.
30. Select “Start-“ and “End Image” of the simulation and parameter estimation, representing the first and last image of your time series that will be used for the parameter estimation and select whether Tropical should skip images (for example: use only every 2. image).
31. Select whether you want to use constant time intervals between the images or use a timestamp file (see “Tropical_Timestamp_File.pdf”) by checking the box in the “Timestamp”-box.
32. If you want to use a timestamp file open it by clicking the “...”-button, selecting the file and then clicking “Open” or typing the full path to the file into the box by hand.
33. In the “ROI Plot Compartments” box you can select additional .raw files which will be used as compartments for which the grey values will be written into the ROI.txt file. Those values can later be used to plot e.g. recovery curves.
34. **Go to the “Output Settings”-tab.**
35. Select a “General Output Folder” by clicking the “...”-button, selecting the folder and clicking “OK” or entering it into the box by hand. The general output folder is the folder where every data output Tropical produces is written to (see “3.4 The Output”).
36. Select whether you want to save the log file or not in the “Output Files”- box.
37. You may now change the directory where a specific output is written to (see “3.4 The Output”).
38. Click “Start Simulation” or press Alt+S on your keyboard to start the simulation process.
39. Select “Yes” if you like to save the experiment or “No” to continue without saving the experiment.

(3.3) Running a simulation without parameter estimation

If you want to run a simulation without estimating any parameters of your model, everything works exactly as described in 3.2 with the following exceptions:

Input Data Tab

1. No time series images are needed.

2. The number of time steps and the standard deviation are irrelevant.

Simulation and Parameter Estimation Tab

1. Don’t check the parameter estimation check box

2. Specify simulation duration.

This denotes the time of simulated real time. Attention: The simulation duration is **not** the time Tropical will need to perform the simulation but the time the simulated process should have last if it had been a real process. That's it. Now Tropical simulates and estimates the parameters (if selected). You will always have a report of the current status with the log window and simulation process window. Sometimes it can happen that iterations take very long (depending on the stiffness of the model equations). This can give the impression that nothing is happening. In most cases this is not true. If you really want to be sure if Tropical is still working or not in these cases it is best to check if the process is still using CPU power (via the task manager / processes / tropical.exe for windows or via the "top" command for Linux). If you want to abort the simulation click "Abort" in the simulation window and click "Yes". Aborting the simulation might take a few minutes depending on what Tropical currently calculates. If the aborting takes too long click the "x"-button in the upper right corner of the main window and select "Yes". This will terminate the simulation immediately but will also close Tropical!

(3.4) The Output

The output Tropical produces consists of 4 files and 2 folders containing image files.

The files are:

- the log file "log_file.txt"
- the results file "results.txt"
- the parameter log file "parameter_log.txt"
- the ROI file "ROI.txt"

The folders are:

- the binned images folder
- the results folder

The log file contains all log messages Tropical produces during the parameter estimation process (much similar to the information printed in the log window). This file allows a detailed analysis of the parameter estimation based on important numerical numbers and parameter values of each iteration step of the estimation. The results file contains the "intensity" matrices of all molecules, which are the result of the simulation. The file contains a huge number of values. The values are ordered exactly like an image: lines represent the "intensity" of one molecule in x-direction; columns (e.g. lines 1-64, 65-128) represent the "intensity" in y-direction. The results file can be used to represent the simulation result with external programs, like Matlab.

The parameter log file contains a table with the developing of the parameters during the parameter estimation and numerical control values. It contains the values of the parameters which are chosen to be estimated, the X2 value representing the goodness of the fit by the sum of squared differences of the original – simulated data, divided by the variance.

$$\chi^2 = \frac{\sum (origData - simData)^2}{\sigma^2}$$

Further values are λ, which should be very small for the correct parameter set (e.g. 10e-6), the sum of grey values of the original data and the simulated data in the parameter estimation compartment and the best χ^2 of all previous iterations. The parameter set with the best fit is the one which produces the lowest χ^2. However only if λ is small at the same time, the parameter set is likely to represent a reliable result. The ROI file contains a mean grey value of the output molecules in the specified compartments for each time point. These values can be used to plot e.g. a recovery curve of a simulated FRAP experiment with an external program. The binned images folder contains the binned original images. The results folder contains the result images produced by the simulation.

(4) The Molecules

Tropicals' molecules are the state variables of the equations. Each molecule needs a corresponding initial image.

(4.1) Reactions

The syntax of the differential equations for the reactions is very simple (see also "Tropical model file").

The equations always start with one of the following left hand sites:

- dy/dt =
- dc/dt =
- d/dt =
- dc(Molecule)/dt =
- d/dt[Molecule] =
- d[Molecule]/dt =

The left hand sites are all equal. So it doesn't matter which style you use.

The right hand side looks like a "normal" term. Every function of the C Standard Library math is accepted (Beware: C is case-sensitive!). And there exist also a function for the concentration of a molecule:

c(Molecule) or C(Molecule) or [Molecule] indicate a concentration of a molecule.

Examples:

The molecule of the following examples will always be "H2O" and there exists another molecule called "c2h6".

- dy/dt = 0
- d/dt = c(c2h6) + [H2O]
- d[H2O]/dt = fabs(15-C(c2h6))*(-3)
- dy/dt = k1*[c2h6]-k2*[H2O]/[c2h6]

For further examples see the model.txt files of the examples for Tropical!

Comments are allowed and are specified in the typical C style: // denotes a line comment, slash and asterisk denote longer comments: /* start of comment….end of comment */.

(4.2) Reaction Parameters

The syntax of the reaction parameters is very simple, too (see also “Tropical model file”). The list of parameters looks like this:

Parameter1 = Value1, Parameter2 = Value2, ...

So the parameters are seperated by a “,”. The sign for comma for values must be a dot “.”. To tell Tropical to estimate a reaction parameter you just have to surround the name with “**” or “[]”.

Examples:

- k1 = 2.05, k2 = 1e-16, pi = 3.14
- *param1* = 5, [param2] = 1.3e-3
- *k1*=1.2e-2, *k2*=2.0

(5) Advanced Settings

Clicking the “Advanced settings” tab opens a window where numerical parameters for the Runge-Kutta ODE solver (simulation) and the Levenberg- Marquardt algorithm (parameter estimation) can be changed. While some of these settings need to be set for each experiment specifically, some others don’t. The following list helps to judge which parameters should be changed for which purpose.

Parameter estimation

• Criterion for a good step (Chi-square difference in %): after a number of “good steps”, the parameter estimation stops. A good step in this case is defined here as a step in which the estimated parameters do not lead to a significantly better result compared to the previous estimation step. In the standard settings this is set to 1%, meaning that the difference between the error of the previous step and the recent step is smaller than 1%. The estimation did not lead to a significant better result than the previous one whereas the significance interval is 1%. If you need higher accuracy decrease the percentage value, however this may lead to longer computational time.

• Number of good steps before stopping: defines the number of “good steps” in a row that must be achieved before the parameter estimation process ends. Standard set-

ting is 10 steps. For higher accuracy increase the number, however this may lead to longer computational time.

• Number of maximum steps for parameter estimation: This defines the maximum number of steps the parameter estimator is allowed to take, even if no “good step” is done. Standard value is 50. Increase the number if the parameter estimation has problems to calculate a good enough result.

• Use UH-Jump: If checked, the modified Levenberg-Marquardt is used. It may lead to faster convergence in some cases. Standard setting is “checked”.

ODE solver

• Maximum time steps for simulation: This is the maximum number of (time) steps the ODE solver is allowed to make for one simulation process. This is related to the adaptive step size control. Increasing the standard value of 1000 can lead to more accurate results but increases computational time.

• Step size tolerance: This is a parameter to determine the maximum allowed error the ODE solver is allowed to make. We recommend setting this value as small as possible. The standard value is 1.0e-8.

• First step size: This is an important parameter that should be set individually for each experiment. It is displayed in milliseconds and determines the first step size the ODE solver will try. It is recommended choosing a first step size at least ten times smaller than the time steps between first and second image of your microscopy experiment.

• Number of steps per iteration: This determines how many steps the ODE solver is allowed to make during one iteration. For time steps in the range of milliseconds up to a few seconds in the original microscopy experiment the standard value of 10000 is well suited. For longer times between two images a higher value is recommended. Increasing the value can result in much longer computational time, but may increase the accuracy significantly.

6.3 Tropical model file

General

Every model file must contain **one** model starting tag and **one** model end tag. These tags denote the start and the end of the model information. Model information outside of these tags will be ignored:

```
<Model>
…
</Model>
```

Comments (see 2. Comments) are always ignored because they are removed before anything else is done. Therefore a model file like the following would be correct:

```
<Model>
// <Model>
</Model>
# </Model>
```

Generally the model file is case-insensitive. Only literal strings (text marked with “”, e.g. “<Molecule “H2O”>”) are case-sensitive.

The molecules of the model can be created in the model file, but the model file can also access the molecules created inside the Tropical GUI.

The model file is processed line by line and therefore molecules are only known if they were created in the model information above the current line or if they were created in Tropical itself.

Comments

There exist three kinds of comments in a Tropical model file.

There are two styles for single-line comments:

```
# This is single-line comment
…
// This is another single-line comment
```

And there is one style for multiple-line comments:

```
/* This is a multiple-line
comment
*/
```

The Tags

There is only one rule about the order of the tags:

The first tag inside a model tag must always be a molecule tag.

Example:

The following model-file would cause an error because the first tag in the model is not a molecule tag:

```
<model>
   <reactions>
      “H2O” : “dydt = 0”;
   </reactions>
   <molecules>
      “H2O” : 30.53, 0;
   </molecules>
</model>
```

There exist two ways to write a model file:

The **“list-oriented” style** and the **“molecule-oriented” style**.

Both styles can be used in one model and even access the molecules written in the other style.

List-oriented Style

The general structure of the list-oriented style looks like this:

```
<Model>

   <Molecules>
   …
   </Molecules>

   <Parameters>
   …
   </Parameters>

   <Reactions>
   …
   </Reactions>

</Model>
```

<Molecules>-Tag

The syntax of the <Molecules>-Tag looks like this:

```
...
   <Molecules>
      “Molecule-Name” : Diffusioncoefficient, Initial Image of the Molecule;
   </Molecules>
...
```

Example:

```
<Model>
   <Molecules>
      “H2O“: 33.06, 0;
      “ch4“: 28.1, 1;
   </Molecules>
</Model>
```

This model file would create two molecules. The first one would be called “H2O”, would have a diffusion coefficient of 33.06 and would be associated with initial image 0.

The second molecule would be called “ch4”, would have a Diffusioncoefficient of 28.1 and would be associated with the initial image 1.

<Parameters>-Tag

With the <Parameters>-Tag the **reactions parameters** which can be used in the reaction equations are defined.

The syntax looks like this:

```
…
   <Parameters>
   “Molecule-Name“ : “Parameter1 = Value of Parameter1, Parameter2 =
                        Value of Parameter2, …;”
   </Parameters>
...
```

Example:

```
<Model>

   <Molecules>
      “C2H6“ : 26.2;
      “H2O” : 33.7;
   </Molecules>

   <Parameters>
      “C2H6“ : “k1 = 0.632, k2 = 2.0e-3, *k3* = 1.3“, [k4] = 40;
   </Parameters>
```

This model file contains a molecule named “C2H6” with a diffusion coefficient of 26.2 and four reaction parameters named “k1”, “k2”, “k3” and “k4” with the values 0.632, 2.0e-3, 1.3 and 40.

The reaction parameters exist for every molecule and have to be defined only once.

Therefore the reaction parameters exist also for the molecule “H2O” and don’t have to be defined again for the molecule “H2O” (see Example in chapter 4.4).

If a reaction parameter called “k1”, “k2”, “k3” or “k4” had been defined for the molecule “H2O”, too, then it must have the same value as the corresponding reaction parameter of the molecule “C2H6” or an error will occur.

The “**” and the “[]” around the reaction parameters “k3” and “k4” indicate that this parameters should be estimated. So, Tropical estimates the values for “k3” and “k4” during the simulation.

“**” means exactly the same as “[]” and is just provieded for convenience.

<Reactions>-Tag

With the <Reactions>-Tag the **chemical reaction** of every molecule is defined a differential equation.

The syntax looks like this:

```
…
   <Reactions>
      ”Molecule-Name“ : ”Differential Equation“;
   </Reactions>
...
```

Example:

```
<Model>

   <Molecules>
      “C2H6“ : 26.2;
      “H2O” : 33.7;
   </Molecules>

   <Parameters>
      “C2H6“ : “k1 = 0.632, k2 = 2.0e-3“;
      “H2O” : “k2 = 0.3e14”;
   </Parameters>

   <Reactions>
      “C2H6” : “dy/dt = k1 * 13 + 1”;
      “H2O” : “dy/dt = k2 * k3”;
   </Reactions>

</Model>
```

In this example the molecule “C2H6” reacts in a way described by the differential equation “dy/dt = k1*13+1” and the molecule “H2O” reacts in a way described by the differential equation “dy/dt = k2*k3” where k1 is 0.632, k2 is 2.0e-3 and k3 is 0.3e14.

There exist two styles to include the concentration of a molecule in the differential equation of the reaction. These two styles mean exactly the same and look like this:

```
“dy/dt = [H2O] *k1“;

“dy/dt = c(H2O) *k1“;
```

Molecule-oriented Style

The general structure of the molecule-oriented style looks like this:

```
<Model>

   <Molecule “Molecule-Name“>
      Channel : Initial image of the molecule;
      DiffParam : Diffusion coefficient;
      Parameters : “Reaction parameters”;
      Reaction : “Differential equation”;
   </Molecule “molecule-name”>

</Model>
```

The syntax of the single parts of the <Molecule>-Tag looks the same like the corresponding tags in the list-oriented style.

Example:

```
<Model>

   <Molecule “H2O”>
      Channel : 0;
      DiffParam : 33.6;
      Parameters : “k1 = 0.34”;
      Reaction : “dy/dt = k1*2.506”;
   </Molecule “H2O”>

</Model>
```

6.4 Tropical Timestamp File

General

The timestamp file contains the timepoints of all TIFF images loaded in Tropical. The timestamps are directly indentified with the images and therefore they must have the same order and if there are more than one channel every timestamp must be entered twice or 3-times depending on the number of channels.

Relative Timestamp Files

The first line must be “0”. Then the time in seconds passed between the first time-stamp and the current timestamp follow line by line.

Example:

```
0
0.213
0.420
0.517
0.680
0.800
1.036
…
```

This timestamp file contains the relative timestamps for 7 images.

The order of the images must be:

Timestamp 0, Channel 0 (= Image 0)

Timestamp 1, Channel 0 (= Image 1)

Timestamp 2, Channel 0 (= Image 2)

…

7 Bibliography

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2002). Molecular Biology of the Cell, Fourth Edition. New York, Garland.

Allan, J., Hartman, P. G., Crane-Robinson, C. and Aviles, F. X. (1980). "The structure of histone H1 and its location in chromatin." Nature **288**(5792): 675-679.

Ando, R., Mizuno, H. and Miyawaki, A. (2004). "Regulated Fast Nucleocytoplasmic Shuttling Observed by Reversible Protein Highlighting." Science **306**(5700): 1370-1373.

Axelrod, D., Koppel, D. E., Schlessinger, J., Elson, E. and Webb, W. W. (1976). "Mobility measurement by analysis of fluorescence photobleaching recovery kinetics." Biophys. J. **16**(9): 1055-1069.

Bacia, K. and Schwille, P. (2003). "A dynamic view of cellular processes by in vivo fluorescence auto- and cross-correlation spectroscopy." Methods **29**(1): 74-85.

Bastiaens, P. and Squire, A. (1999). "Fluorescence lifetime imaging microscopy: spatial resolution of biochemical processes in the cell." Trends Cell Biol. **9**: 48-52.

Bates, I. R., Wiseman, P. W. and Hanrahan, J. W. (2006). "Investigating membrane dynamics in living cells." Biochem. Cell Biol. **84**: 825-831.

Beaudouin, J. (2003). Structural and molecular dynamics of nuclear proteins revealed by fluorescence microscopy. UFR de physique, Université Paris. **Ph.D. Thesis:** 117.

Beaudouin, J., Mora-Bermudez, F., Klee, T., Daigle, N. and Ellenberg, J. (2006). "Dissecting the Contribution of Diffusion and Interactions to the Mobility of Nuclear Proteins." Biophys. J. **90**(6): 1878-1894.

Becker, M., Becker, A., Miyara, F., Han, Z., Kihara, M., Brown, D. T., Hager, G. L., Latham, K., Adashi, E. Y. and Misteli, T. (2005). "Differential In Vivo Binding Dynamics of Somatic and Oocyte-specific Linker Histones in Oocytes and During ES Cell Nuclear Transfer." Mol. Biol. Cell **16**(8): 3887-3895.

Bentele, M., Lavrik, I., Ulrich, M., Stosser, S., Heermann, D. W., Kalthoff, H., Krammer, P. H. and Eils, R. (2004). "Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis." J. Cell Biol. **166**(6): 839-851.

Bhattacharya, D., Mazumder, A., Miriam, S. A. and Shivashankar, G. V. (2006). "EGFP-Tagged Core and Linker Histones Diffuse via Distinct Mechanisms within Living Cells." Biophysical Journal **91**(6): 2326-2336.

Birbach, A., Bailey, S. T., Ghosh, S. and Schmid, J. A. (2004). "Cytosolic, nuclear and nucleolar localization signals determine subcellular distribution and activity of the NF-{kappa}B inducing kinase NIK." J Cell Sci **117**(16): 3615-3624.

Bock, H. (1981). Numerical treatment of inverse problems in chemical reaction kinetics. Modelling of Chemical Reaction Systems. Ebert, K. H., Deuflhard, P. and Jäger, W. Heidelberg, Springer. **18:** 102-125.

Braeckmans, K., Peeters, L., Sanders, N. N., De Smedt, S. C. and Demeester, J. (2003). "Three-Dimensional Fluorescence Recovery after Photobleaching with the Confocal Scanning Laser Microscope." Biophys. J. **85**(4): 2240-2252.

Braga, J., Desterro, J. M. P. and Carmo-Fonseca, M. (2004). "Intracellular Macromolecular Mobility Measured by Fluorescence Recovery after Photobleaching with Confocal Laser Scanning Microscopes." Molecular Biology of the Cell **15**: 4749–4760.

Braga, J., McNally, J. G. and Carmo-Fonseca, M. (2007). "A Reaction-Diffusion Model to Study RNA Motion by Quantitative Fluorescence Recovery after Photobleaching." Biophys. J. **92**(8): 2694-2703.

Brown, D. T. (2003). "Histone H1 and the dynamic regulation of chromatin function." Biochem. Cell Biol. **81**: 221-227.

Brown, D. T., Izard, T. and Misteli, T. (2006). "Mapping the interaction surface of linker histone H1° with the nucleosome of native chromatin in vivo." Nature Structural and Molecular Biology **13**: 250-255.

Bruno, M., Flaus, A., Stockdale, C., Rencurel, C., Ferreira, H. and Owen-Hughes, T. (2003). "Histone H2A/H2B Dimer Exchange by ATP-Dependent Chromatin Remodeling Activities. ." Molecular Cell **12**(6): 1599-1606.

Bustin, M., Catez, F. and Lim, J.-H. (2005). "The Dynamics of Histone H1 Function in Chromatin." Molecular Cell **17**(5): 617-620.

Carrero, G., McDonald, D., Crawford, E., de Vries, G. and Hendzel, M. J. (2003). "Using FRAP and mathematical modeling to determine the in vivo kinetics of nuclear proteins." Methods **29**: 14-28.

Catez, F., Ueda, T. and Bustin, M. (2006). "Determinants of histone H1 mobility and chromatin binding in living cells." Nature Structural and Molecular Biology **13**(4): 305-310.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. and Prasher, D. C. (1994). "Green fluorescent protein as a marker for gene expression." Science **263**(5148): 802-805.

Chapman, S., Oparka, K. J. and Roberts, A. G. (2005). "New tools for in vivo fluorescence tagging." Current Opinion in Plant Biology **8**(6): 565-573.

Chen, Y., Lagerholm, B. C., Yang, B. and Jacobson, K. (2006). "Methods to measure the lateral diffusion of membrane lipids and proteins." Methods **39**(2): 147-153.

Coscoy, S., Waharte, F., Gautreau, A., Martin, M., Louvard, D., Mangeat, P., Arpin, M. and Amblard, F. (2002). "Molecular analysis of microscopic ezrin dynamics by two-photon FRAP." PNAS **99**(20): 12813-12818.

Crank, J. (1975). The Mathematics of Diffusion. New York, Oxford University Press.

Demuro, A. and Parker, I. (2006). "Imaging single-channel calcium microdomains." Cell Calcium **40**(5-6): 413-422.

Denk, W., Strickler, J. H. and Webb, W. W. (1990). "Two-photon laser scanning fluorescence microscopy." Science **248**(4951): 73-76.

Deuflhard, P. and Bornemann, F. (2002). Numerische Mathematik II. Gewöhnliche Differentialgleichungen. Berlin, New York, de Gruyter.

Dixit, R., Cyr, R. and Gilroy, S. (2006). "Using intrinsically fluorescent proteins for plant cell imaging." The Plant Journal **45**(4): 599-615.

Doenecke, D., Albig, W., Bode, C., Drabent, B., Franke, K., Gavenis, K. and Witt, O. (1997). "Histones: genetic diversity and tissue-specific gene expression." Histochem. Cell. Biol. **107**: 1-10.

Draves, P. H., Lowary, P. T. and Widom, J. (1992). "Co-operative binding of the globular domain of histone H5 to DNA." Journal of Molecular Biology **225**(4): 1105-1121.

Duggan, M. M. and Thomas, J. O. (2000). "Two DNA-binding sites on the globular domain of histone H5 are required for binding to both bulk and 5 S reconstituted nucleosomes." Journal of Molecular Biology **304**(1): 21-33.

Dundr, M., Hebert, M. D., Karpova, T. S., Stanek, D., Xu, H., Shpargel, K. B., Meier, U. T., Neugebauer, K. M., Matera, A. G. and Misteli, T. (2004). "In vivo kinetics of Cajal body components." Journal of Cell Biology **164**(6): 831-842.

Eils, R. and Athale, C. (2003). "Computational imaging in cell biology." Journal of Cell Biology **161**(3): 477–481.

Finch, J. T. and Klug, A. (1976). "Solenoidal model for superstructure in chromatin." PNAS **73**(6): 1897-1901.

Franke, K., Drabent, B. and Doenecke, D. (1998). "Testicular expression of the mouse histone H1.1 gene." Histochem. Cell. Biol. **109**: 383–390.

Gerlich, D., Beaudouin, J., Gebhard, M., Ellenberg, J. and Eils, R. (2001). "Four-dimensional imaging and quantitative reconstruction to analyse complex spatiotemporal processes in live cells." Nature Cell Biol. **3**: 852-855.

Gershenfeld, N. (1999). The Nature of Mathematical Modeling Cambridge, UK, Cambridge University Press.

Goytisolo, F. A., Gerchman, S.-E., Yu, X., Rees, C., Graziano, V., Ramakrishnan, V. and Thomas, J. O. (1996). "Identification of two DNA-binding sites on the globular domain of histone H5." EMBO Journal **15**(13): 3421-3429,.

Graziano, V., Gerchman, S. E., Schneider, D. K. and Ramakrishnan, V. (1994). "Histone H1 is located in the interior of the chromatin 30-nm filament." Nature **368**(6469): 351-354.

Gunjan, A., Alexander, B. T., Sittman, D. B. and Brown, D. T. (1999). "Effects of H1 Histone Variant Overexpression on Chromatin Structure." J. Biol. Chem. **274**(53): 37950-37956.

Happel, N., Schulze, E. and D., D. (2005). "Characterisation of human histone H1x." Biological Chemistry **386**(6): 541-551.

Hendzel, M. J., Lever, M. A., Crawford, E. and Th'ng, J. P. H. (2004). "The C-terminal Domain Is the Primary Determinant of Histone H1 Binding to Chromatin in Vivo." Journal of Biological Chemistry **279**(19): 20028–20034.

Hess, S. T., Huang, S., Heikal, A. A. and Webb, W. W. (2002). "Biological and Chemical Applications of Fluorescence Correlation Spectroscopy: A Review." Biochemistry **41**(3): 697-705.

Hoch, D. A., Stratton, J. J. and Gloss, L. M. (2007). "Protein–Protein Förster Resonance Energy Transfer Analysis of Nucleosome Core Particles Containing H2A and H2A.Z." J. Mol. Biol.: doi:10.1016/j.jmb.2007.1005.1075.

Horn, P. J. and Peterson, C. L. (2002). "Chromatin Higher Order Folding: Wrapping up Transcription." Science **297**(5588): 1824-1827.

Houtsmuller, A. B., Rademakers, S., Nigg, A. L., Hoogstraten, D., Hoeijmakers, J. H. J. and Vermeulen, W. (1999). "Action of DNA repair endonuclease ERCC1/XPF in living cells." Science **284**: 958-961.

Jedrusik, M. A. and Schulze, E. (2001). "A single histone H1 isoform (H1.1) is essential for chromatin silencing and germline development in Caenorhabditis elegans." Development **128**(7): 1069-1080.

Jung, G., Wiehler, J., Göhde, W., Tittel, J., Basché, T., Steipe, B. and Bräuchle, C. (1998). "Confocal microscopy of single molecules of the green fluorescent protein." Bioimaging **6**: 54-61.

Kao, H. P. and Verkman, A. S. (1996). "Construction and performance of a photobleaching recovery apparatus with microsecond time resolution." Biophysical Chemistry **59**: 203-210.

Kappel, C. and Eils, R. (2004). Resolution: Fluorescence Recovery After Photobleaching with the Leica TCS SP2 Confocal Application Letter. Heidelberg, Leica Microsystems GmbH**:** 12.

Kaufman, E. N. and Jain, R. K. (1990). "Quantification of transport and binding parameters using fluorescence recovery after photobleaching. Potential for in vivo applications." Biophys. J. **58**(4): 873-885.

Khochbin, S. and Wolffe, A. P. (1994). "Developmentally regulated expression of linker-histone variants in vertebrates." European Journal of Biochemistry **225**: 501-510.

Khochbin, S. (2001). "Histone H1 diversity: bridging regulatory signals to linker histone function." Gene **271**: 1-12.

Kimura, H. and Cook, P. R. (2001). "Kinetics of Core Histones in Living Human Cells: Little Exchange of H3 and H4 and Some Rapid Exchange of H2B." J. Cell Biol. **153**(7): 1341-1354.

Klonis, N., Rug, M., Harper, I., Wickham, M., Cowman, A. and Tilley, L. (2002). "Fluorescence photobleaching analysis for the study of cellular dynamics." European Biophysical Journal **31**: 36-51.

Koster, M., Frahm, T. and Hauser, H. (2005). "Nucleocytoplasmic shuttling revealed by FRAP and FLIP technologies." Current Opinion in Biotechnology **16**(1): 28-34.

Lavi, Y., Edidin, M. and Gheber, L. A. (2007). "Dynamic patches of membrane proteins." Biophys. J.: biophysj.107.111567.

Lele, T., Wagner, S. R., Nickerson, J. A. and Ingber, D. E. (2006). "Methods for Measuring Rates of Protein Binding to Insoluble Scaffolds in Living Cells: Histone H1-Chromatin Interactions." Journal of Cellular Biochemistry **99**: 1334-1342.

Levenberg, K. (1944). "A method for the solution of certain problems in least squares." Quarterly of Applied Mathematics **2**(2): 164-168.

Lever, M. A., Th'ng, J. P. H., Sun, X. and Hendzel, M. J. (2000). "Rapid exchange of histone H1.1 on chromatin in living human cells." Nature **408**(6814): 873-876.

Lippincott-Schwartz, J., Snapp, E. and Kenworthy, A. (2001). "Studying Protein Dynamics in Living Cells." Nature Rev. Mol. Cell. Biol. **2**: 445-456.

Lippincott-Schwartz, J., Altan-Bonnet, N. and Patterson, G. (2003). "Photobleaching and photoactivation: following protein dynamics in living cells." Nature Cell Biol. **5**: S7-S14.

Lippincott-Schwartz, J. and Patterson, G. H. (2003). "Development and Use of Fluorescent Protein Markers in Living Cells." Science **300**: 87-91.

Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., Zipursky, L. and Darnell, J. (2003). Molecular Cell Biology, W. H. Freeman.

Lu, X. and Hansen, J. C. (2004). "Identification of Specific Functional Subdomains within the Linker Histone H1° C-terminal Domain." J. Biol. Chem. **279**(10): 8701-8707.

Luger, K., Maeder, A. W., Richmond, R. K., Sargent, D. F. and Richmond, T. J. (1997). "Crystal structure of the nucleosome core particle at 2.8A°resolution." Nature **389**(6648): 251-260.

Lukyanov, K. A., Chudakov, D. M., Lukyanov, S. and Verkhusha, V. V. (2005). "Photo-activatable fluorescent proteins." Nature Reviews Molecular Cell Biology **6**(11): 885-890.

Marquardt, D. and SIAM J. Appl. Math. 11 (1963). "An algorithm for least-squares estimation of nonlinear parameters." SIAM Journal of Applied Mathematics **11**(431-441).

McBryant, S., Adams, V. and Hansen, J. (2006). "Chromatin architectural proteins." Chromosome Research **14**(1): 39-51.

Mendes, P. and Kell, D. B. (1998). "Non-linear optimization of biochemical pathways: applications to metabolic engeneering and parameter estimation." Bioinformatics **14**(10): 869-883.

Misteli, T., Gunjan, A., Hock, R., Bustink, M. and Brown, D. T. (2000). "Dynamic binding of histone H1 to chromatin in living cells." Nature **408**: 877-881.

Misteli, T. (2001). "Protein Dynamics: Implications for Nuclear Architecture and Gene Expression." Science **291**: 843-847.

Miyawaki, A., Sawano, A. and Kogure, T. (2003). "Lighting up cells: labelling proteins with fluorophores." Nature Cell Biol.: S1-S7.

Nofal, S., Becherer, U., Hof, D., Matti, U. and Rettig, J. (2007). "Primed Vesicles Can Be Distinguished from Docked Vesicles by Analyzing Their Mobility." J. Neurosci. **27**(6): 1386-1395.

Patterson, G. H. and Lippincott-Schwartz, J. (2002). "A Photoactivatable GFP for Selective Photolabeling of Proteins and Cells." Science **297**(5588): 1873-1877.

Pehrson, J. R. and Cole, R. D. (1981). "Bovine H10 histone subfractions contain an invariant sequence which matches histone H5 rather than H1." Biochemistry **20**: 2298–2301.

Pennington, S. (09. Jul 2007). "11th Hour Introduction to Genetics." Book 5, from http://www.blackwellpublishing.com/11thhour/book5/index.html.

Phair, R. D. and Misteli, T. (2000). "High mobility of proteins in the mammalian cell nucleus." Nature **404**(6778): 604-609.

Phair, R. D. and Misteli, T. (2001). "Kinetic modelling approaches to in vivo imaging." Nature Rev. Mol. Cell. Biol. **2**: 898-907.

Pochynyuk, O., Staruschenko, A., Bugaj, V., LaGrange, L. and Stockand, J. D. (2007). "Quantifying RhoA Facilitated Trafficking of the Epithelial Na+ Channel toward the Plasma Membrane with Total Internal Reflection Fluorescence-Fluorescence Recovery after Photobleaching." J. Biol. Chem. **282**(19): 14576-14585.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (2002). Numerical Recipes in C++. The Art of Scentific Computing. Cambridge, UK, Cambridge University Press.

Rabut, G., Doye, V. and Ellenberg, J. (2004). "Mapping the dynamic organization of the nuclear pore complex inside single living cells." Nat Cell Biol **6**(11): 1114-1121.

Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L. and Sweet, R. M. (1993). "Crystal structure of globular domain of histone H5 and its implications for nucleosome binding." Nature **362**(6417): 219-223.

Rasheed, B. K., Whisenant, E. C., Ghai, R. D., Papaioannou, V. E. and Bhatnagar, Y. M. (1989). "Biochemical and immunocytochemical analysis of a histone H1 variant from the mouse testis." Journal of Cell Science **94**: 61-71.

Razin, S. V., Iarovaia, O. V., Sjakste, N., Sjakste, T., Bagdoniene, L., Rynditch, A. V., Eivazova, E. R., Lipinski, M. and Vassetzky, Y. S. (2007). "Chromatin Domains and Regulation of Transcription." Journal of Molecular Biology **369**(3): 597-607.

Reits, E. and Neefjes, J. (2001). "From fixed to FRAP: measuring protein mobility and activity in living cells." Nature Cell Biol. **3**: E145-E147.

Remington, S. J. (2006). "Fluorescent proteins: maturation, photochemistry and photophysics." Current Opinion in Structural Biology **16**(6): 714-721.

Robinson, P. J. J., Fairall, L., Huynh, V. A. T. and Rhodes, D. (2006). "EM measurements define the dimensions of the 30-nm'' chromatin fiber: Evidence for a compact, interdigitated structure." PNAS **103**(17): 6506-6511.

Saxton, M. J. (2001). "Anomalous Subdiffusion in Fluorescence Photobleaching Recovery: A Monte Carlo Study." Biophys. J. **81**(4): 2226-2240.

Sbalzarini, I. F., Mezzacasa, A., Helenius, A. and Koumoutsakos, P. (2005). "Effects of Organelle Shape on Fluorescence Recovery after Photobleaching." Biophysical Journal **89**: 1482-1492.

Schloeder, J. (1988). Numerische Methoden zur Behandlung hochdimensionaler Aufgaben der Parameteridentifizierung. Bonn, Universität Bonn.

Schmiedeberg, L., Weisshart, K., Diekmann, S., Meyer zu Hoerste, G. and Hemmerich, P. (2004). "High- and Low-mobility Populations of HP1 in Heterochromatin of Mammalian Cells." Mol. Biol. Cell **15**(6): 2819-2833.

Seksek, O., Biwersi, J. and Verkman, A. S. (1997). "Translational Diffusion of Macromolecule-sized Solutes in Cytoplasm and Nucleus." J. Cell Biol. **138**(1): 131-142.

Shimomura, O., Johnson, F. H. and Saiga, Y. (1962). "Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea." Journal of Cellular and Comparative Physiology **59**: 223-239.

Shimomura, O. (2005). "The discovery of aequorin and green fluorescent protein." Journal of Microscopy **217**(1): 3-15.

Siggia, E. D., Lippincott-Schwartz, J. and Bekiranov, S. (2000). "Diffusion in Inhomogeneous Media: Theory and Simulations Applied to Whole Cell Photobleach Recovery." Biophys. J. **79**: 1761–1770.

Sinnecker, D., Voigt, P., Hellwig, N. and Schaefer, M. (2005). "Reversible Photobleaching of Enhanced Green Fluorescent Proteins." Biochemistry **44**(18): 7085-7094.

Snapp, E. L., Altan, N. and Lippincott-Schwartz, J. (2003). "Measuring protein mobility by phtobleaching GFP chimeras in living cells." Curr. Prot. Cell. Biol. **21**: 1-24.

Soumpasis, D. M. (1983). "Theoretical analysis of fluorescence photobleaching recovery experiments." Biophys. J. **41**(1): 95-97.

Sprague, B. L., Pego, R. L., Stavreva, D. A. and McNally, J. G. (2004). "Analysis of Binding Reactions by Fluorescence Recovery after Photobleaching." Biophysical Journal **86**: 3473-3495.

Sprague, B. L., Müller, F., Pego, R. L., Bungay, P. M., Stavreva, D. A. and McNally, J. G. (2006). "Analysis of Binding at a Single Spatially Localized Cluster of Binding Sites by Fluorescence Recovery after Photobleaching." Biophys. J. **91**(4): 1169–1191.

Stephens, D. J. and Allan, V. J. (2003). "Light Microscopy Techniques for Live Cell Imaging." Science **300**: 82-86.

Strahl, B. D. and Allis, C. D. (2000). "The language of covalent histone modi®cations." Nature **403**: 41-45.

Sund, S. E. and Axelrod, D. (2000). "Actin Dynamics at the Living Cell Submembrane Imaged by Total Internal Reflection Fluorescence Photobleaching." Biophys. J. **79**(3): 1655-1669.

Tanaka, M., Hennebold, J. D., Macfarlane, J. and Adashi, E. Y. (2001). "A mammalian oocyte-specific linker histone gene H1oo: homology with the genes for the oocyte-specific cleavage stage histone (cs-H1) of sea urchin and the B4/H1M histone of the frog." Development **128**: 655-664.

Thomas, J. O. and Wilson, C. M. (1986). "Selective radiolabelling and identification of a strong nucleosome binding site on the globular domain of histone H5." EMBO Journal **5**(13): 3531–3537.

Thomas, J. O., Rees, C. and Finch, J. T. (1992). "Cooperative binding of the globular domains of histones H1 and H5 to DNA." Nucl. Acids Res. **20**(2): 187-194.

Thomas, J. O. (1999). "Histone H1: location and role." Current Opinion in Cell Biology **11**: 312-317.

Travascio, F. and Y., G. W. (2007). "Anisotropic Diffusive Transport in Annulus Fibrosus: Experimental Determination of the Diffusion Tensor by FRAP Technique." Annals of Biomedical Engineering: DOI: 10.1007/s10439-10007-19346-10432.

Tremethick, D. J. (2007). "Higher-Order Structures of Chromatin: The Elusive 30 nm Fiber." Cell **128**(4): 651-654.

Tsay, T. T. and Jacobson, K. A. (1991). "Spatial Fourier analysis of video photobleaching measurements. Principles and optimization." Biophys. J. **60**(2): 360-368.

Ulrich, M., Kappel, C., Beaudouin, J., Hezel, S., Ulrich, J. and Eils, R. (2006). "Tropical—parameter estimation and simulation of reaction–diffusion models based on spatio-temporal microscopy images." Bioinformatics **22**(21): 2709-2710.

Verschure, P. J., van der Kraan, I., Manders, E. M. M., Hoogstraten, D., Houtsmuller, A. B. and van Driel, R. (2003). "Condensed chromatin domains in the mammalian nucleus are accessible to large macromolecules." EMBO Reports **4**(9): 861–866.

Volz, D., Eigel, M., Athale, C., Bastian, P., Hermann, H., Kappel, C. and Eils, R. (2005). "Spatial Modelling and Simulation of Diffusion in Nuclei of Living Cells." Lecture Notes in Bioinformatics **3082**: 161-171.

Wachsmuth, M., Waldeck, W. and Langowski, J. (2000). "Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy." Journal of Molecular Biology **298**(4): 677-689.

Wachsmuth, M., Weidemann, T., Muller, G., Hoffmann-Rohrer, U. W., Knoch, T. A., Waldeck, W. and Langowski, J. (2003). "Analyzing Intracellular Binding and Diffusion with Continuous Fluorescence Photobleaching." Biophys. J. **84**(5): 3353-3363.

Weiss, M. (2004). "Challenges and Artifacts in Quantitative Photobleaching Experiments." Traffic **5**: 662-671.

Woodcock, C. L., Grigoryev, S. A., Horowitz, R. A. and Whitaker, N. (1993). "A chromatin folding model that incorporates linker variability generates fibers resembling the native structures." PNAS **90**(19): 9021-9025.

Woodcock, C. L., Skoultchi, A. I. and Fan, Y. (2006). "Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length." Chromosome Research **14**: 17-25.

www.jmol.org. (2007). "Jmol: an open-source Java viewer for chemical structures in 3D." Version 11. Retrieved 09. Jul 2007.

Yan, W., Ma, L., Burns, K. H. and Matzuk, M. M. (2003). "HILS1 is a spermatid-specific linker histone H1-like protein implicated in chromatin remodeling during mammalian spermiogenesis

" PNAS **100**: 10546-10551.

Zlatanova, J. and Doenecke, D. (1994). "Histone H1 zero: a major player in cell differentiation?" FASEB J. **8**: 1260-1268.

Zlatanova, J., Caiafa, P. and Van Holde, K. (2000). "Linker histone binding and displacement: versatile mechanism for transcriptional regulation." FASEB J. **14**(12): 1697-1704.