Master of Science Thesis

Computational modeling of protein diffusion in the cell nucleus using the lattice-Boltzmann method.

Kateryna Filippovych

June 2012

University of Jyväskylä

Nanoscience center

Department of physics

Table of contents

Abbreviations

DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
NE	Nuclear envelope
PML	Promyelocytic leukaemia
NPCs	Nuclear pore complexes
EGFP	Enhanced green fluorescent protein
EYFP	Enhanced yellow fluorescent protein
ECFP	Enhanced cyan fluorescent protein
HeLa	Henrietta Lacks
FRAP	Fluorescence recovery after photobleaching
LSCM	Laser scanning confocal microscopy
FLIP	Fluorescence loss in photobleaching
FRET	The fluorescence resonance energy transfer
LBM	Lattice-Boltzmann method

1. Introduction

The number of fluorescence applications and methods of analysis is growing in the biological and related sciences. This work has been motivated by a previous work [Kühn 2011], where protein diffusion in the cytoplasm was studied. Here we study protein diffusion in the cell nucleus. Aiming to accurately determine the diffusion coefficient, diffusion in the cytoplasm and nucleo-cytoplasmic translocation should be considered as well. Fluorescence recovery after photobleaching (FRAP) experiments were used in order to study protein diffusion. To this end cells with fluorescent proteins were imaged with confocal microscopy.

2. Biological background

2.1 The Cell

The cell is the simplest fundamental unit of life. From the unicellular bacteria to multicellular animals, the cell is one of the basic organizational principles of biology. It can function as an individual and independent living system or as part of a larger organism. The cell has its own metabolism capable of independent existence, reproduction and development. All the cells are made of the same basic material: water, fats, salts, nucleic acids, carbohydrates and proteins. Like the internal organs in human body, the cell has organelles which play a unique biological role. To protect the organelles, the cell is surrounded by a membrane. This membrane plays a crucial role in biology. It serves as a delimiter of cells and has a double role. On the one hand the membrane separates compartments and functional units from each other to guarantee proper functioning of these units. On the other hand, it must be permeable for components necessary for the life and function of the cell. It consists of proteins and lipids (fats). The role of proteins is to regulate the chemical climate of the cell and lipids make the membrane flexible. [Lodish 2000]

Inside all cells there is cytoplasm, a jelly-like material that is 80% water. It covers everything inside the cell between the plasma membrane and the nucleus. Cytoplasm includes the cytosol, meaning cell substance, which is the area of the cytoplasm outside of the individual organelles. The cytosol covers the largest area in the cell. It composes 54% of the total cell volume. It facilitates the movement of materials around the cell. The cytoplasm, as seen through an electron microscope, appears as a three-dimensional lattice of thin protein-rich strands. This lattice serves to interconnect and support the other solid structures in the cytoplasm. The cytoplasm also contains a skeletal structure, called the cytoskeleton. This structure gives the cell its shape and allows it to organize many of the chemical reactions that occur in the cytoplasm. Additionally, the cytoskeleton is involved

in the movement of the cell and intracellular transport of vesicles and proteins. [Alberts 2002]

2.2 Membrane diffusion

Cellular and organelle membranes act as protectors and permit selective transport of molecules needed by the cell. Transport through a membrane happens either by diffusion through pores or actively across the phospholipids bilayer with embedded proteins in it as shown in Fig 1.

Figure 1. Transport mechanisms through membrane. ***A****: Direct diffusion, works for small uncharged molecules.* ***B****: Transport through a pore, also unspecific diffusion for larger molecules.* ***C/D****: Active transport through channels requires external energy, only for specific ions.* ***E****: Facilitated diffusion, transport by carrier proteins.*

The transport mechanism underlying routes A and B of Fig.1 is plain diffusion that follows Fick's law [Fick 1855]:

$$\boldsymbol{j} = D \cdot \nabla_x c(\boldsymbol{x}, t) \quad , \qquad (1)$$

Fick's law means that diffusive flux ***(j)*** is proportional to concentration gradient ($c(\boldsymbol{x}, t)$). The proportionality factor is a tensor, called diffusion coefficient *D*.

Direct diffusion of small unpolar molecules through the plasma membrane (**A)** can happen unspecifically everywhere. In facilitated diffusion (**E**) molecules associate with particular carrier proteins and are transported across the membrane from high to low concentration in order to equilibrate the concentration [Wilson 2000]. These carrier proteins change their conformation after binding to the ion they are transporting. Transport through the membrane is performed without additional external energy, but is highly selective.

Active transport through the channels (**C/D**) is induced by certain membrane proteins against concentration gradients, thus requiring external energy. This mechanism is highly specific for particular molecules and quite different from diffusion.

2.3 Eukaryotic cell

There are two main types of cells: prokaryotic cells and eukaryotic cells. Prokaryotic cells are simple, one-cell organisms, such as bacteria. They are far simpler in structure and much smaller than the eukaryotic cells.

Eukaryotic cells of plants and animals are highly structured. These cells tend to be larger than, for example, the cells of bacteria, and have developed specialized packaging and transport mechanisms that may be necessary to support their larger size. Eukaryotic cells contain two important things that prokaryotic cells do not: a nucleus and organelles with membranes around them, where specific metabolic activities occur [Alberts 2002]. These organelles serve specific functions within the eukaryotes, such as energy production,

photosynthesis, and membrane construction. Of all eukaryotic organelles, the nucleus is perhaps the most critical. The main feature of eukaryotic cells is the separation of the deoxyribonucleic acid (DNA) replication and ribonucleic acid (RNA) synthesis in the nucleus from the cytoplasmic machinery for protein synthesis.

2.4 Nucleus

Nucleus is a membrane bound organelle that contains the genetic material and regulates cell activity. It is the prime organelle of the eukaryotic cell and is the cell's control center responsible for DNA replication, transcription and RNA processing. DNA is the main information carrier molecule in the cell. Linear strands of DNA are entwined with histone and other proteins to form chromosomes. These structures can be identified and counted by staining them with dyes,.

The mammalian cell nucleus is usually oval-shaped with on axis radius of 10-15 μm. The nucleus is surrounded by two membranes, together known as the nuclear envelope (NE). NE consists of a double-bilayer of lipids with a diverse array of proteins embedded in it [Hetzer 2005]. It functions as a physical barrier, separating the nucleus from the cytoplasm [D'Angelo 2006]. The inner surface of NE has a protein lining called the nuclear lamina which binds to chromatin and other nuclear components [Bridger 2007]. Communication between the nucleus and the surrounding cytosol happens via numerous nuclear pores in NE. Besides functioning in the molecular trafficking, NE provides an important regulatory level in the eukaryotic cell by separating transcription and translation in-between the cytoplasm and the nucleus.

The prominent structure in the nucleus is the nucleolus. Nucleolus is a membraneless organelle that produces ribosomes which move out of the nucleus to positions on the rough endoplasmic reticulum, where they play an essential role in protein synthesis [Olson 2000]. Through a microscope, the nucleolus looks like a large dark spot in the nucleus.

In addition, the nucleus contains several substructures: Cajal bodies, promyelocytic leukaemia nuclear bodies (PML bodies), nuclear speckles, interchromatin compartments and many others. Cajal bodies are involved in the biogenesis of different nuclear RNA molecules [Gall 2011]. PML bodies are small dynamic intranuclear structures that are involved in various cellular functions like transcriptional regulation, apoptosis and antiviral defense [Everett 2007].

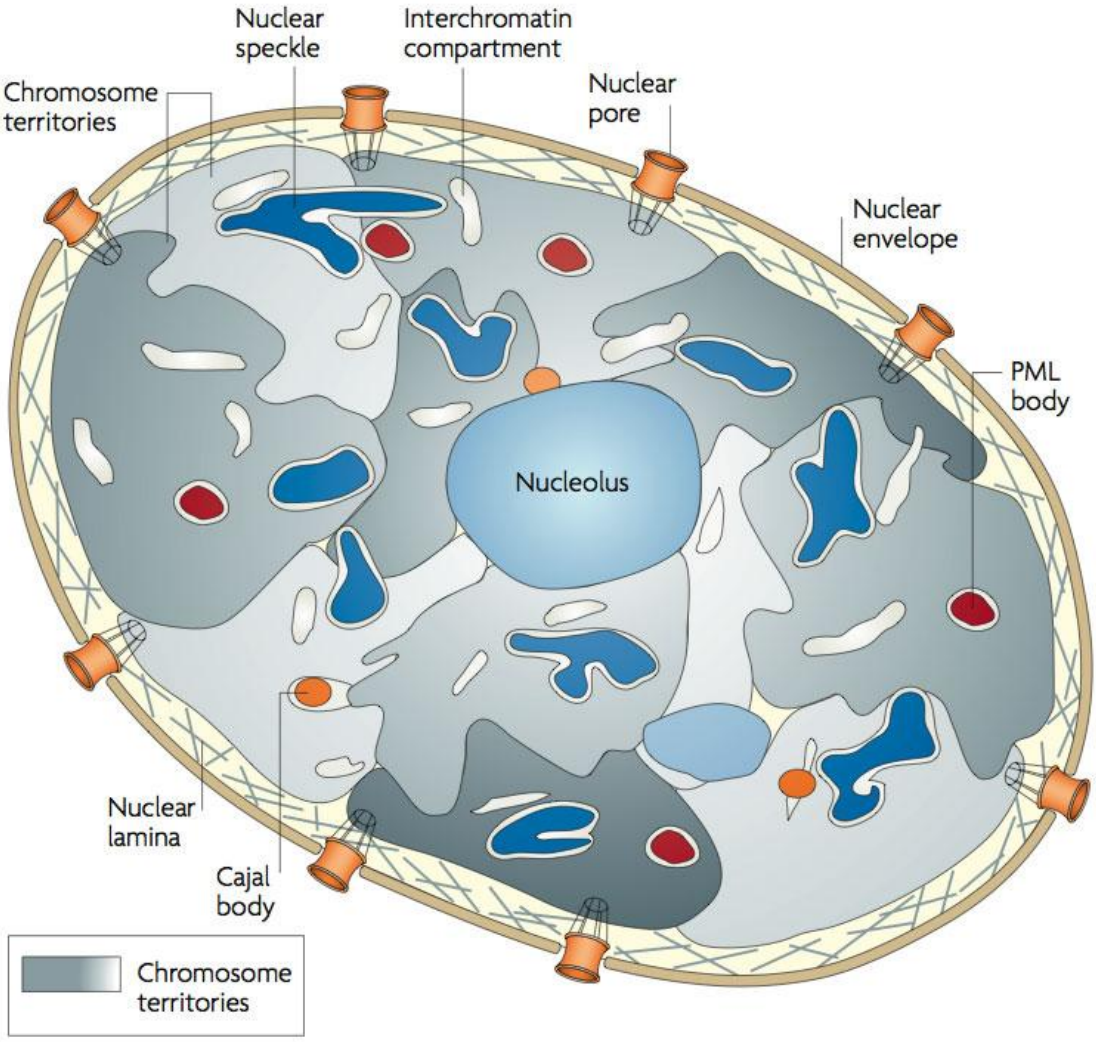


Figure 2. ***Nuclear organelles****. A typical mammalian cell nucleus, sliced open to reveal crosssections of organelles [Lanctôt 2007].*

2.5 Nuclear pore complex

Nuclear pore complexes (NPCs) are large structures consisting of approximately 30 proteins, termed nucleoporins (Nups) [Cronshaw 2002]. NPCs regulate bidirectional transport of molecules, including proteins and mRNA, between the nucleus and cytoplasm, while at the same time generating a diffusion barrier to separate the cytoplasm from the nuclear compartments.

NPCs are involved in numerous cellular functions, such as chromatin organization and regulation of gene expression. They are complex cylindrical structure with strong octagonal symmetry, and are formed by the two membranes of NE, inner and outer, shown in Figure 3.

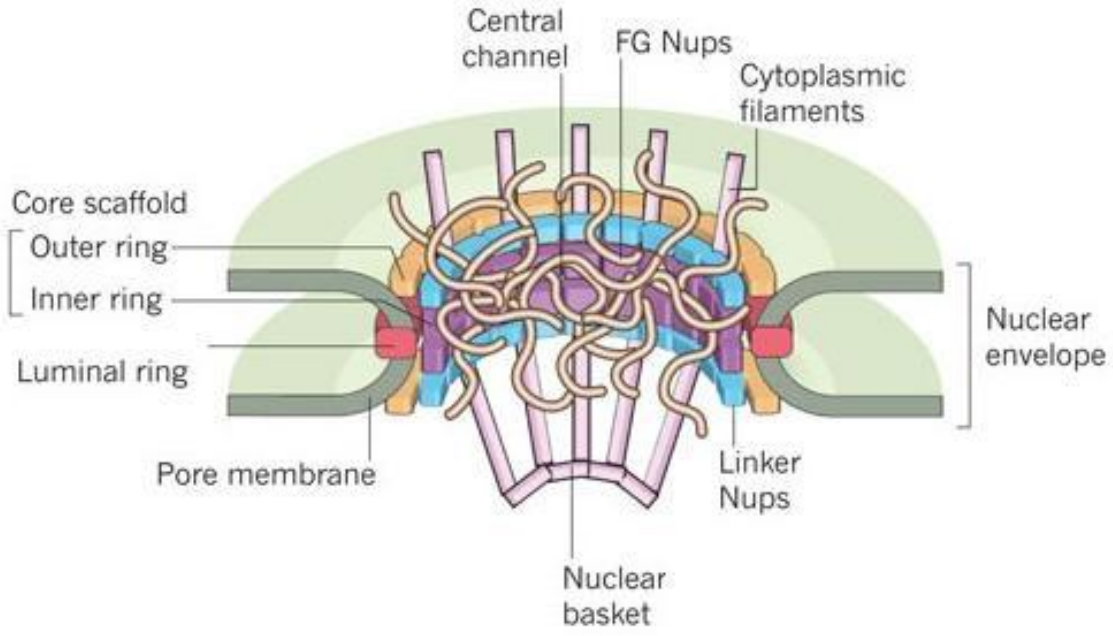


Figure 3. Major structural parts of a nuclear pore complex [Strambio-De-Castillia 2010].

The structure of NPC consists of a central transporter region or central pore, a core scaffold that supports the central channel, transmembrane regions, a nuclear basket, and cytoplasmic filaments [Alber 2007]. The central channel is filled with and surrounded by Nups that have numerous large domains of phenylalanine-glycine, termed FG Nups. The FG Nups mediate selective receptor-mediated transport [Jovanovic-Talisman 2009]. The nuclear basket consists of eight filaments that reach into the nucleoplasm, attached to each other by a ring at the end. The eight cytoplasmic filaments form highly mobile molecular rods projecting into the cytoplasm. The core scaffold is connected to a set of membrane proteins, which form a luminal ring [Alber 2007].

The interactions provided by NPC binding sites wholly define the function of NPC. All these interactions reflect the complex design of NPC and its important role in cellular processes.

2.6 Nuclear transport

One of the requirements for macromolecules to enter and leave the nucleus is their need to be recognized by sequence signal and the NPC's need to indentify it respectively. Transport through NPC is regulated by cargo-carrying factors that interact with FG Nups [Rout 2001]. Transport depends on diffusion, affinity of transport factors to NPC, and differences in these affinities between symmetric and asymmetric binding sites [Rout 2000].

NPC functions as a gateway to all nucleocytoplasmic transport and it provides a pathway that allows free diffusion of ions and small molecules, less than ~9 nm (40 kDa) in diameter [Paine 1975]. The diameter of diffusion channel is ~9–10 nm, which means that larger cargoes with a diameter of up to ~39 nm (25MDa) require active translocation by transport receptors [Wente 2010].

3. Fluorescence

The phenomenon of fluorescence was named by George Stokes in 1852. For biology purposes it started to be used in the 1930s. Nowadays fluorescence is widely used to study the structure and conformations of DNA and proteins.

Fluorescence is a short-lived type of luminescence created by electromagnetic excitation (light). Light excites fluorescent molecules to an energetically higher electronic state. The life-time of the excited state is about 1-10 nanoseconds. Fluorescent molecules lose part of their potential energy in the excited state due to vibrational relaxation. The relaxation of the electrons back to the ground state from a lower-energy excited state is accompanied by emission of light (fluorescence).The main steps of this process are thus excitation, non-radiative relaxation and relaxation by radiative (light) emission. The process of fluorescence is illustrated in Figure 4.

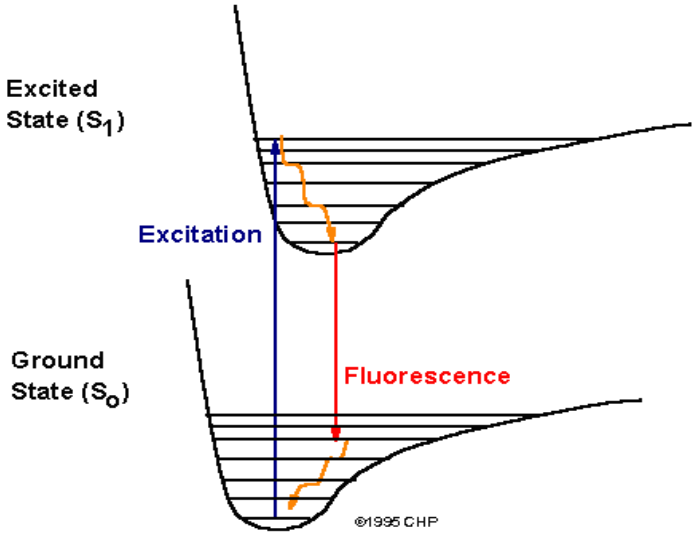


Figure 4. Mechanism of fluorescence showing excitation (blue arrow),non-radiative relaxation (orange arrow) and light emission (red arrow). The ground state S_0 represents the energy of the molecules in the ground state. Absorption of photons with high enough energy excites molecules from a ground state S_0 to a higher state S_1 (singlet state), followed by non-radiative and radiative transitions which bring it back to the ground state.

Therefore, a fluorescent molecule has two characteristic spectra: excitation and emission spectra. In fluorescence the emitted light has a longer wavelength than the exciting light. The difference in wavelength or energy between the maxima of those two spectra is known as the Stokes shift, shown in Figure 5. The Stokes shift is also a distinct characteristic of each fluorophore.

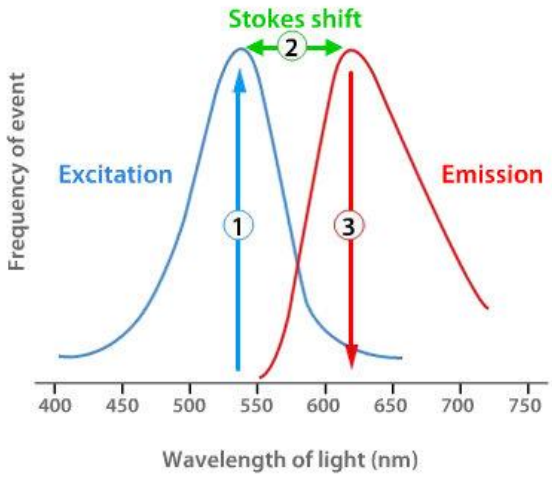


Figure 5. Stokes shift is the difference between excitation and emission peaks [Spectra are from the www.olympusmicro.com web site].

The Stokes shift is especially critical in multiple fluorescence applications, because the emission and excitation spectra of the fluorophore may overlap and therefore excitation of one fluorophore can lead to emission of another.

3.1 Fluorescent Proteins

With the discovery of fluorescence microscopy, fluorophores were used to study dye binding in fixed and living cells and have afterwards become an important part of cell biology. They are used to mark proteins, tissues and other cellular parts with a fluorescent label. Fluorophores differ by their absorption and fluorescence properties, including the

wavelengths of maximum absorbance and emission, and the fluorescence intensity of the emitted light. A fluorophore works by absorbing energy of a specific wavelength, causing excitation and re-emitting that energy at the emission wavelength, related spectra are shown in Figure 6.

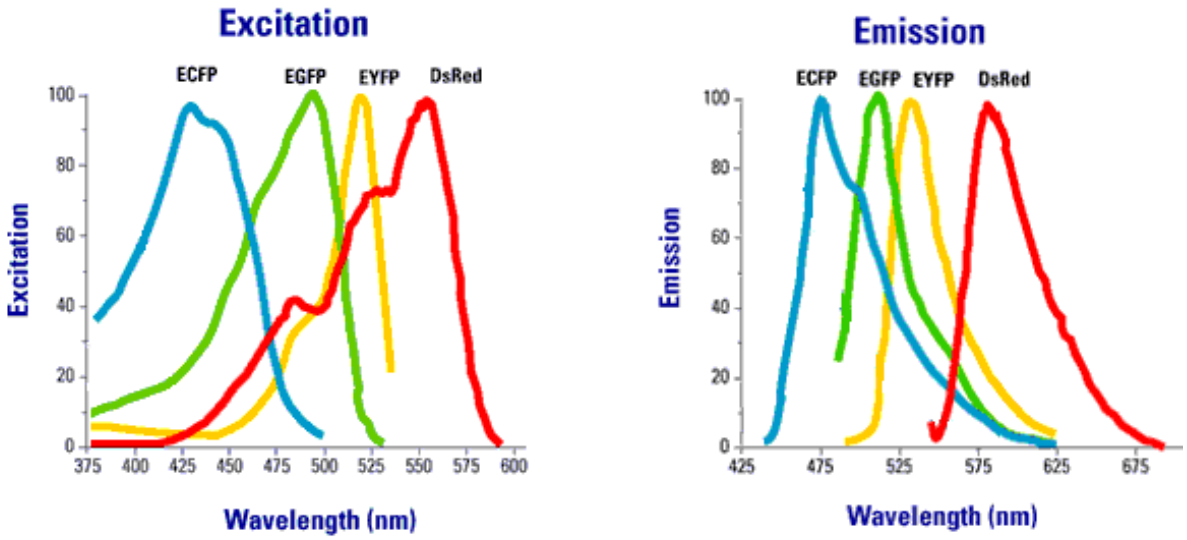


Figure 6. Excitation and emission spectra of fluorescent proteins that illustrate the relative shape and position of each fluorophore in the peak region of its excitation and emission [Spectra are from the Clontech web site].

The advantage of fluorescent proteins is the ability to specifically target subcellular compartments by fluorescent probes. These biological macromolecules offer a new frontier in live-cell imaging. Examples of fluorophores are fluorescent proteins, quantum dots and dyes.

The first use of a biological fluorophore in a research application took place in the 1990s, when the green fluorescent protein (GFP) was cloned from the jellyfish Aequorea Victoria and used as a gene expression reporter [Chalfie 1994].

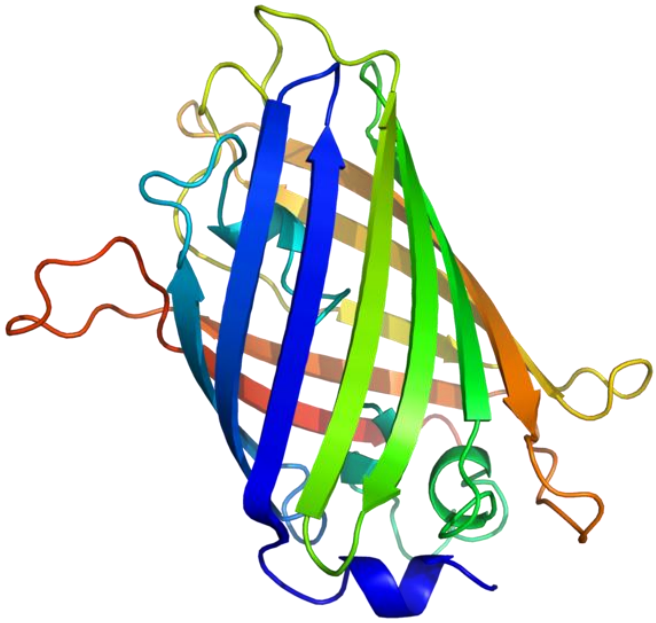


Figure 7. A vector drawing of the green fluorescent protein [derived from PDB: 1EMA].

Next the enhanced cyan fluorescent protein (ECFP) mutants were found by simple amino acid substitutions. Another popular fluorescent protein is the enhanced yellow fluorescent protein (EYFP). It was designed on the basis of crystalline structural analysis of GFP. EYFP is optimally excited by the 515 nm spectral line of the argon-ion laser, and provides a more intense emission than GFP. EYFP is widely used in photobleaching techniques. It can be used to monitor nuclear transport processes and intarnuclear dynamics. The approximate size of EYFP is ~2 nm (27kDa), which means that it can freely diffuses through NPC's [Erickson 2009].

Applications of fluorophores that can be utilized in confocal microscopy are rapidly growing.

3.2 Confocal microscopy

Marvin Minsky introduced the concept of confocal microscopy in the late 1950s. The technology was further enhanced by many scientists and at the end of the 1980s when commercial versions of confocal microscopes were also presented. The development of the high intensity light sources and computers, powerful enough to process large data sets, were important for confocal microscopy.

Confocal microscopy, or laser scanning confocal microscopy (LSCM), can provided much better resolution and magnification than a traditional bright-field microscopy or conventional fluorescence microscopy, and has become a general instrument in cell biology. A schematic layout of a confocal microscope is shown in Figure 6.

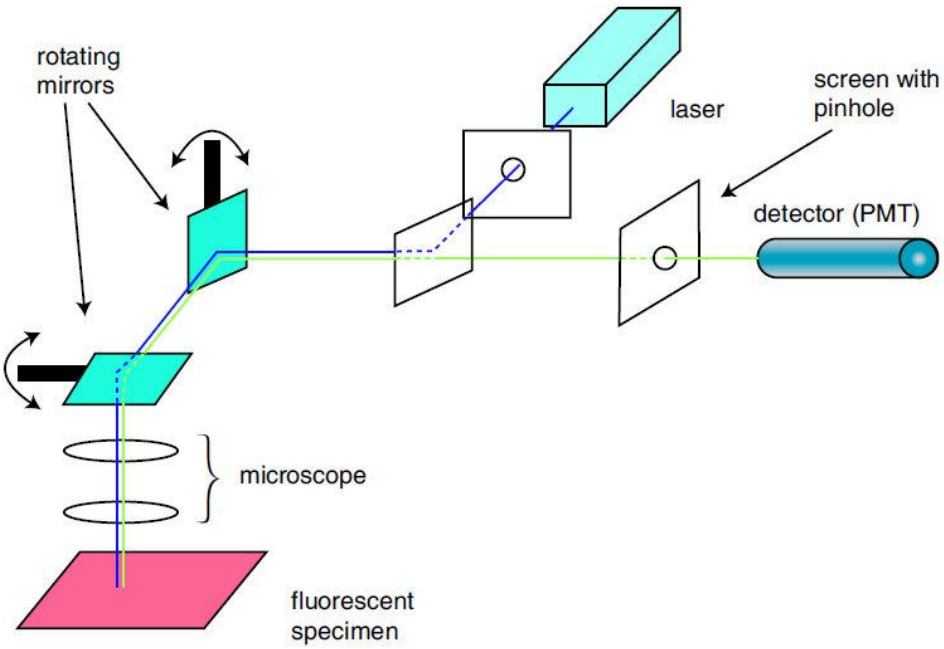


Figure 6. Layout of a confocal microscope. The laser scans across the sample by help of two scanning mirrors. Here the laser light is blue and the emitted light is green. The light emitted from the focal plane is focused onto the pinhole. That part of the emitted light passes through the pinhole, and is measured by a detector [Semwogerere 2005].

With confocal microscopy, it is possible to scan fixed cells or living cells. The samples can also be thicker than in traditional fluorescence microscopy because the laser scans the sample point by point at every cross section separately. The final 3D image can then be reconstructed. The advantage of fluorescence for microscopy is that one is able to attach fluorescent dye molecules to specific parts of the sample. It allows to indicate the precise location of the intracellular components labeled with the fluorescent dye. It is also possible to use more than one type of dye at the same time.

3.3 Photobleaching

Most confocal microscopes also provide the ability to photobleach fluorescent molecules. Photobleaching is caused by irreversible destruction of fluorescence due to a prolonged exposure by the excitation source to high-intensity light. Photobleaching can be minimized or avoided by reducing the level of light intensity. Another option to reduce photobleaching is to use a high numerical aperture lens so as to better collect fluorescent light. The exact mechanism of photobleaching is not known, but it is assumed to be linked to a transition from an excited singlet state to a triplet state. However, photobleaching is not always undesirable. A technique that takes advantage of it is fluorescence recovery after photobleaching (FRAP).

3.4 FRAP

Using fluorescence recovery after photobleaching (FRAP), it is possible to study the dynamics, diffusion and diffusion speed of specific protein (Sparague 2004).

The use FRAP in macromolecular kinetics has increased in recent years due to the development of fluorescent proteins. Fluorescent proteins make possible the assessment of molecular dynamics in vivo, for example, in the cytoplasm and nucleus. The method

of FRAP utilizes the phenomenon of photobleaching of fluorescent proteins followed by observation of the recovery of fluorescence caused by diffusion of fluorophores from other part of the cell to the bleached region, as shown in Figure 7.

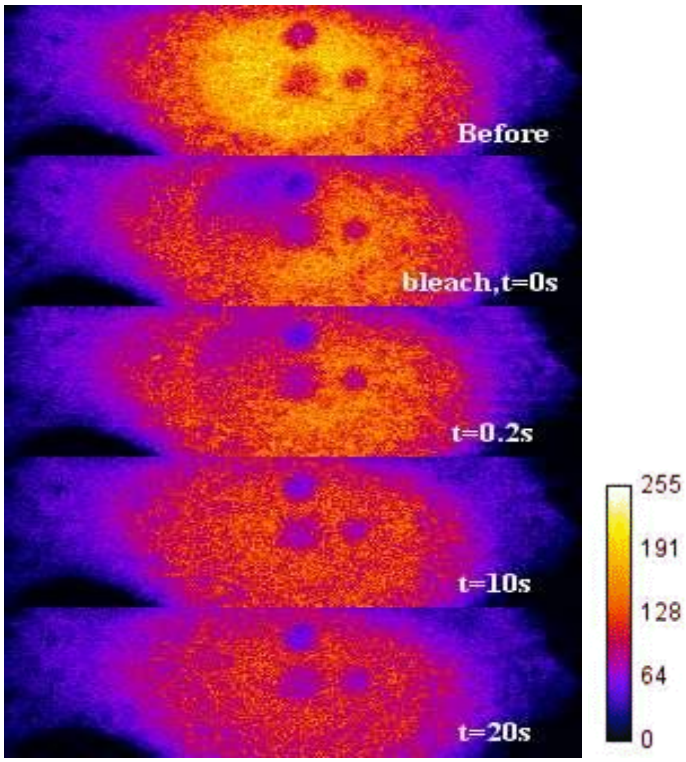


Figure 7. Visual representation of the FRAP process. On top there is an image of the cell before bleaching, then right after the bleach, and then at different times after the bleach. The last image (t=20s) shows the fluorophore distribution after a full recovery of the cell. The images are normalized with an arbitrary scale.

Recovery of fluorescence tells how fast molecules are exchanged between the bleached region and its surroundings, i.e. how fast they diffuse. A typical fluorescence recovery curve is shown in Figure 8.

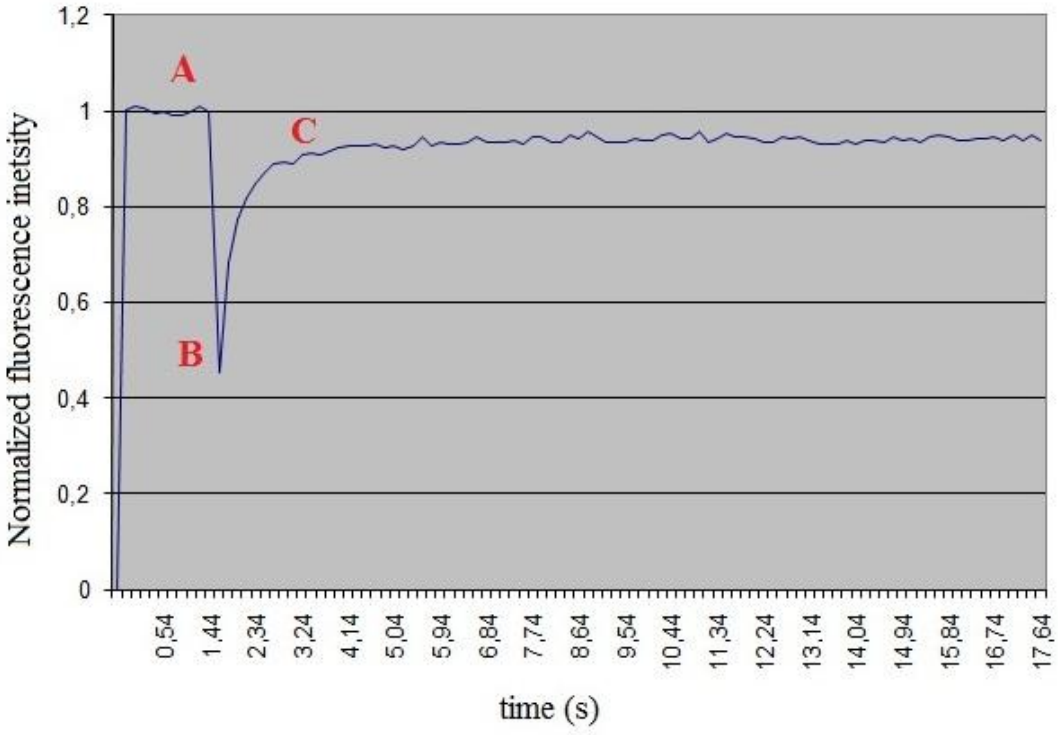


Figure 8. A typical fluorescence recovery curve. The phases indicated are prebleaching (A), bleaching (B) and recovery (C).

For FRAP experiments it is important to choose a protein which is bleached minimally at low illumination power so as to prevent photobleaching during image acquisition. Nowadays FRAP experiments are usually performed on a laser scanning confocal microscope capable of rapid bleaching of a small area using high intensity laser light. There after the dynamics of fluorescence recovery is recorded by sampling images at regular time intervals.

It has only recently become clear from FRAP studies that small molecules can rapidly diffuse through the complicated system of the cell and bind reversibly to dynamic scaffolds. Such studies have shown that molecules up to 200 kDa diffuse through the cytoplasm with lower values than those found in water [Swaminathan 1997].

FRAP experiments are very popular when analyzing intracellular and intranuclear protein diffusion, when gathering information about the protein's role in cellular functions, and how this role changes under stressed conditions like viral infection or cell damage.

3.5 Conventional FRAP analysis and its problems

The conventional way to analyze FRAP data is by the method of Axelrod/Soumpasis [Axelrod 1976,Soumpasis 1983], by which the FRAP recovery curve is fitted by an explicit mathematical expression derived from the diffusion equation.

Here fluorescence recovery was analyzed using the ImageJ and Excel software. Recovery was determined from a circular region of 1.4 μm radius in HeLa cells. The size of the region was measured by bleaching fixed cells stably expressing EYFP. ImageJ was used to construct an average shape and profile of the bleached region. Next, the data were exported to Excel where their normalization was performed. The first normalization was that of Phair and Misteli [Phair 2000]:

$$I_{PM}(t) = \frac{ROI(t)}{Cell(t)} * \frac{\langle Cell(t<0)\rangle}{\langle ROI(t<0)\rangle} \quad , \qquad (2)$$

where ROI(t) is the total fluorescence intensity in the bleached region at time t, ROI(t<0) is the time average of the local fluorescent intensity of the whole bleached region before the bleach pulse, and Cell(t) and Cell(t<0) are the respective quantities for the entire cell.

The time of full recovery can be difficult to determine and requires a long imaging phase. As a second normalization we used a modified normalization of Axelrod *et al.*, which took into account partial recovery [Axelrod 1976]:

$$I_{mA}(t) = \frac{ROI(t) - ROI(t=0)}{\frac{ROI(t=end)}{p} - ROI(t=0)} \quad , \qquad (3)$$

where RIO (t=end) is the fluorescence intensity at the end of the experiment and p is the recovery ratio at that time. With the Phair and Misteli normalization, fluorescence intensities vary between 0 and 1. They can be used to determine the recovery ratio p needed in the modified Axelrod normalization. The data where then fitted by the result for a free diffusion model by Soumpasis [Soumpasis 1983]:

$$I(t) = \exp\left(-\frac{\tau_D}{2t}\right)\left[I_0\left(\frac{\tau_D}{2t}\right) + I_1\left(\frac{\tau_D}{2t}\right)\right] \quad , \qquad (4)$$

where $\tau_D = r^2/D_f$, I_0 and I_1 are modified Bessel functions, r is the radius of the bleached region and D_f is the diffusion coefficient.

A problems with the Axelrod/Soumpasis method is that their assumptions (2D, infinite medium and infinite bleach time) do not correlate with the real process happening in living cells during FRAP experiments. Their model is too simplified and obviously the results are not exact.

3.6 Other fluorescence-based methods to study protein dynamics

Throughout the past decade, to overcome the invasive nature of immobilized protein-protein interaction assays, a new array of technologies has been developed. These new techniques are based on genetic labeling with fluorescent proteins.
These new approaches are devoted to the characterization and visualization of protein interactions , and have favored the possibility of carrying out experiments in vivo as well as in real time, thus allowing one to identify where and when protein interactions occur in the cell.

Inverse FRAP (iFRAP) is performed as a normal FRAP experiment with the difference that the molecules outside the region of interest are photobleached, and loss of fluorescence from the non-photobleached region is monitored over time. For example, iFRAP was used to monitor the dissociation kinetics of GFP tagged RNA polymerase components from sites of rRNA transcription [Dundr 2002].

In a similar manner as in FRAP, fluorescence loss from the area surrounding a repeatedly photobleached region can be measured. This technique is called fluorescence loss in photobleaching (FLIP). It allows to measure signal decay rather than fluorescence recovery, and is useful when analyzing protein mobility as well as protein shuttling between cellular compartments [Dundr 2003].

The fluorescence resonance energy transfer (FRET) methodology has become a classic approach in the study of protein-protein interactions. A fluorophore donor molecule is excited with matching monochromatic light, and if an acceptor fluorophore molecule is in the proximity, an energy transfer may occur between the molecules.

Photoactivation is photo-induced activation of an inert molecule to an active state. Before photoactivation, cells expressing photoactivatable proteins display only little fluorescence in the spectral region that is used for detecting enhanced fluorescence. After photoactivation of a selected region, an increase of fluorescence is observed. By directly highlighting specific populations of molecules, such as the nuclear pool of the fluorophore, the movement from this region of the cell can be monitored.

Alternatively, the entire cell can be photoactivated and the fate of fluorescence followed over time. The ability to ‘switch on’ the fluorescence of photoactivatable proteins makes them excellent tools for exploring protein behaviour in living cells [Elowitz 1997].

Photobleaching techniques provide a powerful method for highlighting intracellular transport and for analyzing the dynamics of protein-trafficking machinery.

4. Modeling and simulation

Modeling and simulation is essential in computational science. Modeling creates a mathematical abstraction of the problem. A mathematical object (formula, system of equations, algorithm, etc.) is called a model which is solved for the quantities of interest. The aim of modeling and simulation is to understand reality by quantification [Griebel 1998]. A model should be as simple as possible and as complex as necessary. It is always a simplification of reality. A model must be reliable, i.e. it must be derived from basic laws whose validity is beyond doubt, the so-called first principles. A typical issue in modeling and simulation is to compute the development of a system in time based on the known description of its initial state. For example, biological processes happen in physical space and time, and a model of this process is expected to describe the state of a system at later time.

Due to problems with the conventional FRAP analysis, explained in chapter 3.6, a new method of FRAP analysis has been developed [Kuhn 2011]. The new method is 'based' on generation of a digital model of the cell for the simulation environment, and the conditions of real experiment are fulfilled as closely as possible. The data needed for simulation of a FRAP experiment included a 3D image of the cell and nucleus. Result were then inferred by comparing simulations and experiments images.

4.1 Digital model cell

A numerical code has been constructed to simulate the spatial and temporal evolution of fluorescence intensity in digital realizations of the cells actually measured in FRAP experiments, [Kuhn 2011].

In this method the model cell was generated using a 3D LSCM scan of the fluorescence intensity distribution in the cell. For each FRAP experiment three sets of data were obtained here: a 3D stack of images representing the intensity profile of EYFP in the cell

before the bleach, a 3D stack of images representing distribution of H2B-ECFP in the cell nucleus and a stack of 2D images of the nucleus during the FRAP measurements, 10 frames before the bleach and the rest of the frames from the recovery phase. The distribution of fluorescence intensity measured right after the bleach was taken as the initial condition in the simulations, such that they could very accurately reproduce the experimentally observed fluorescence recovery.

After de-noising the 3D stack, the threshold function of the ImageJ software was used to segment the cytoplasm and nucleus in the cell using the EYFP and H2B-ECFP stacks. The nuclear envelope was represented as a two pixel wide layer. The spatial resolution of LSCM ~ 200 nm, did not allow segmentation of the more detailed structure. That is why the cytoplasm and nucleus were considered as effective porous media, assumed to be immobile for the duration of the FRAP measurement. Porosity of the medium showed up as a heterogeneous equilibrium distribution of fluorophores, low fluorescence intensity meaning high solids contents. Porosity was assumed to be given by the equilibrium fluorescence intensity $C_0(\boldsymbol{r})$ when normalized to one, $\varepsilon(\boldsymbol{r}) \equiv C_0(\boldsymbol{r})/\max\{C_0(\boldsymbol{r})\}$, and corresponded to the local ratio of the cytosol/nucleosol (liquid-phase) content to the solid-phase content. In the cytosol/nucleosol the equilibrium fluorophore density was assumed to be homogeneous [Kuhn 2011].

4.2 The lattice-Bolztmann method

The lattice-Boltzmann method (LBM) is an effective and accurate simulation tool for analysis of several different problems [Succi 2000, Chen 1998]. LBM is simple, explicit in time and local in space. It has been applied to a large variety of purposes. LBM is also well-suited to simulate non-ideal gases, multicomponent fluids, fluids with suspensions or in porous media, chemical-reactive flows and anisotropic fluids. Finally, numerous studies in turbulence modeling have been carried out with lattice Boltzmann models.

The previous constructed [Kuhn 2011] implementation of LBM was used here to simulate diffusive protein motion in the digital cell model discretized in a cubic lattice. The diffusive transport of particles with density ρ($\boldsymbol{r}.t$) in an inhomogeneous environment can be described by the continuity equation:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{J} = 0 \quad , \qquad (5)$$

where $\boldsymbol{J}$ is the total flux, $\boldsymbol{J} = \boldsymbol{J}_{\mathrm{D}} + \boldsymbol{J}_{\varepsilon}$, where $\boldsymbol{J}_{\mathrm{D}}$ is a diffusive contribution and $\boldsymbol{J}_{\varepsilon}$ is an extra flux term that has been added to take care of the regions not available to particle motion. As $\boldsymbol{J}_{\mathrm{D}} = -D\boldsymbol{\nabla}\rho$,the above equation can be written in the form:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{J}_{\varepsilon} = \boldsymbol{\nabla} \cdot (\mathrm{D}\boldsymbol{\nabla}\rho) \quad , \qquad (6)$$

which is a kind of advection-diffusion equation.

LB particles can move from a lattice site to one of its nearest neighbours or stay in rest, which means that they have seven possible velocities [Succi 2001]. The distribution function of these particles obeys a discrete version of the Boltzman equation. In the single relaxation time (τ) approximation this LB equation is given by

$$f_i(\boldsymbol{r} + \boldsymbol{\vartheta}_i\, \delta t, t + \delta t) - f_i(\boldsymbol{r}, t) = \frac{\delta t}{\tau}\left[f_i^{eq}(\boldsymbol{r}, t) - f_i(\boldsymbol{r}, t)\right] . \qquad (7)$$

The distribution function $f_i(\boldsymbol{r}, t)$ is that at lattice site $\boldsymbol{r}$ and time t of particles moving with velocity $\boldsymbol{v}_i$ in the i direction. The left hand side of the equation represents streaming

of particles during a time step δt, and the right hand side represents relaxation, owing to collisions at lattice site $\boldsymbol{r}$, of the particle density towards the local equilibrium f_i^{eq}. The total concentration of particles, $\rho(r,t) = \sum f_i\ (r,t)$, can be shown [Succi 2001], in the continuum limit, to satisfy the diffusion equation with the coefficient

$$D = c_s^2 \left(\frac{\tau}{\delta t} - \frac{1}{2}\right)\delta t \ , \qquad (8)$$

where $c_s^2 = \frac{2}{7}\frac{(\delta x)^2}{(\delta t)^2}$ is a free numerical parameter in units of velocity with δx the lattice spacing. Here the particle density $\rho(\boldsymbol{r},t)$ of the LB method was interpreted as the fluorescence intensity $C(\boldsymbol{r},t)$. Relevant structural features of the cytoplasm were taken care by porosity $\varepsilon(\boldsymbol{r})$ determined from experimental LSCM data. The volume excluded from protein motion was implemented by introducing an effective force field that prevented them from entering the regions occupied by the solid phase $(1-\varepsilon(\boldsymbol{r}))$. This field caused an additional flux ($\boldsymbol{J}_\varepsilon$) of particles, which opposed the diffusive fluxes that would have otherwise arisen from concentration gradients in the measured initial profile, $\boldsymbol{J}_\varepsilon{}^0 - D\nabla C_0 = 0$. Adding this flux, the local equilibrium distribution function of fluorophores in equation (7) was given by

$$f_i^{eq}(r,t) = w_i \left(C(r,t) + \frac{v_i \cdot J_\varepsilon(r,t)}{c_s^2}\right) \ , \qquad (9)$$

where w_i are LB weight factors, i = 0, …, 6. For correct diffusion dynamics, we had to include also additional features. From diffusion in porous media we know that one must distinguish diffusion in the liquid phase (cytosol/nucleosol with D_{csol}/D_{nsol}) from that in the medium with constrained motion (cytoplasm/nucleoplasm with D_{cp}/D_{np}) so that $D_{cp} = \varepsilon D_{csol}$, and similarly for the nucleus [Kaviany 1995]. The porosity varies locally,

$\varepsilon = \varepsilon(\boldsymbol{r})$, which leads to a spatially varying effective diffusion coefficient for the cytoplasm/nucleoplasm (D_{cp}/D_{np}). The diffusion coefficient of the nuclear envelope (D_{ne}) was assumed to represent that of a very thin permeable layer.

In the diffusion simulations the initial fluorophore distribution was the one found by LSCM imaging right after the bleach phase, bleaching was not modeled. The present LSCM equipment was quite slow, thus only a 2D cross section of the cell was analyzed at the post-bleach imaging phase, and the measured initial bleach profile was extrapolated vertically in the cell assuming that the relative amount of bleached fluorophores does not vary in the direction of the laser beam [Braga 2004]. At every time step during FRAP recovery, the fluorophore distribution was simulated in the whole digital model cell, while it was only recorded in the same cross section of the cell as actually imaged in the measurement. In the simulations, δx was fixed by the voxel size of the LSCM data and the relaxation time τ for the cytoplasm was fixed by numerical convenience. Comparison of the measured and simulated distributions was then made using a cross-correlation algorithm. The simulation frame that gave the global maximum in the correlation was plotted as a function of simulation time step n, and the best linear behaviour in this plot was sought by varying the relaxation times of the nucleosol and nuclear envelope. The slope of the most linear plot determined the simulation time step. Once δt and three relaxation times were thus determined, we could calculate the values for D_{nuc} , D_{env} and D_{cyt} from equation (8).

4.3 Simulation

To extrapolate the bleach profile vertically into the whole 3D digital cell, bleaching information was extracted from the experimental FRAP data. The first post bleach image was divided by an average of ten pre bleach images. This procedure removed the cell background and gave a map of relative fluorophore reduction ($p(x,y)$), which could only have a value between one and zero. It may happen that the reduction value was more than one, due to the noisiness of experimental data. That is why the reduction value was set to

one. It had not influence on the result of the analysis, because the cross-correlation analysis was invariant under change of intensity by a constant. The 3D bleach profile was then obtained by multiplying each cross section of the 3D stack distribution with $p(x,y)$.

4.4 Data analysis

During the simulation the time dependence of the fluorescence intensity distribution was recorded in the same cross section of the nucleus as in the FRAP experiment. The experimental and simulated images were compared by their cross correlation coefficient,

$$c_{k,l} = \frac{1}{N\sigma_k\sigma_l}\sum_{x.y}(v_k(x,y) - \langle v_k\rangle)\,(v_l(x,y) - \langle v_l\rangle), \qquad (10)$$

where $v_k(x,y)$ is the pixel intensity of the image, N is the number of pixels, v_k is their average intensity, σ_k is their standard deviation, subscripts k and l refer to the two series of images.

Cross correlation results were improved by removing the background from all the images, and a mask was used to restrict the region analyzed. These manipulations reduced perturbing effects caused by motion and deformation of the cell.

The different liquid phases of the cell were described by three relaxation times, one for the cytosol τ_{cyt} , one for the nucleosol τ_{nsol} and one for the effective substance of the nuclear envelope τ_{ne} . Simulation time step δt was a fitting parameter. For each experimental image k there was a global maximum $l_{max}(k)$ in the cross-correlation coefficient, as shown in Figure 9.

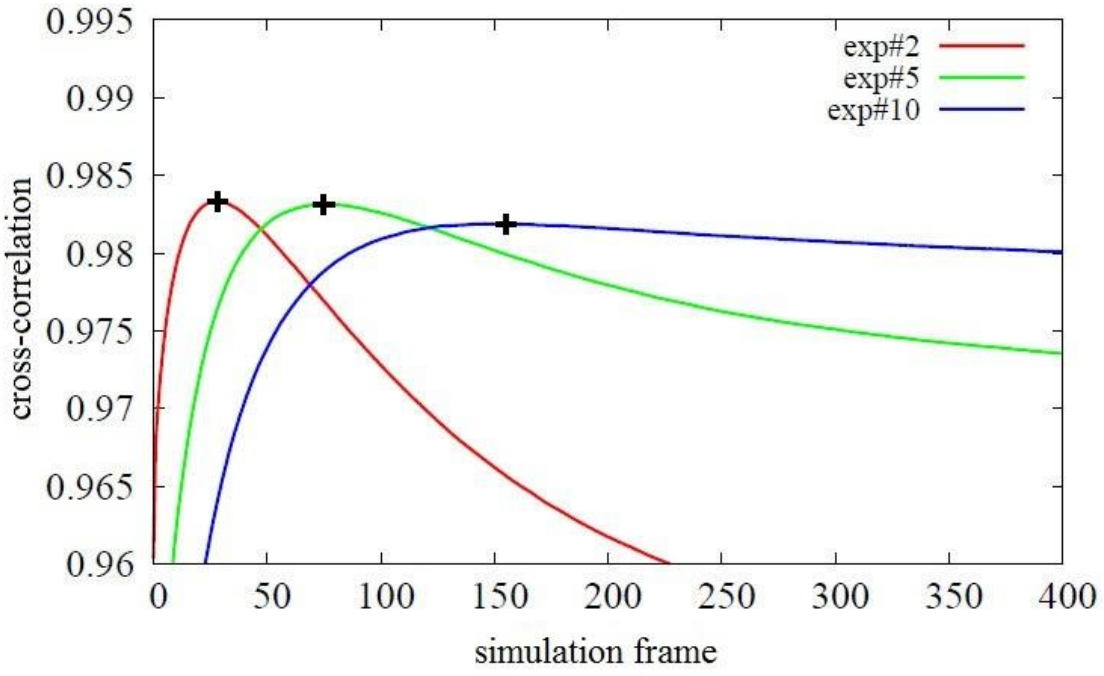


Figure 9. Cross-correlation coefficient $c_{k,l}$ (red, blue and green lines) of the 2^{d}, 5^{th} and 10^{th} frames of measured and simulated FRAP date. The cross-correlation crosses denote their global maxima.

When the global maximum was a linear function of k, the real and digital cells were assume to correspond to each other. The values of τ_{cyt} and τ_{ne} were varied so as to maximize the linearity of $l_{max}(k)$, which gave the simulation time as a function of real (experimental) time.

5. Results

5.1 The Axelrod/Soumpasis method

FRAP experiments were done on EYFP and H2B-ECFP-expressing HeLa cells. In these experiments 10 cells were measured and first analyzed by the method of Axelrod/Soumpasis. According to this analysis, the average diffusion coefficient of the measured cells was $3.1 \pm 1.1\ \mu m^2/s$. As can be seen from Figure 10, the model used did not fit the data well. The reason for this discrepancy is that the assumptions of the Axelrod/Soumpasis method were not really satisfied in the experiment, which caused the low value of the diffusion coefficient in the nucleus in comparison with more accurate values reported in [Kuhn 2011]. A curve fitted to a set of recovery data by the free diffusion model of Soumpasis is shown in Figure 10.

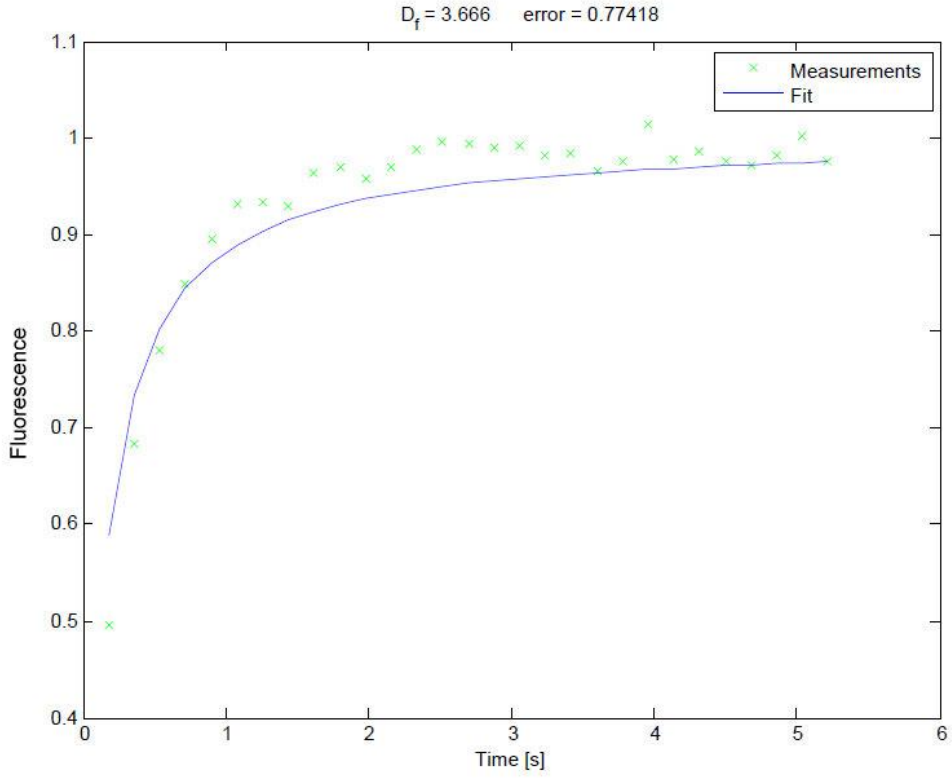


Figure 10. A set of measured recovery data with Axelrod normalization and a fit of that set by the free diffusion model of Soumpasis.

5.2 Results of FRAP experiments and simulations

With the new methods which assumed that transport was that through a porous medium allowed us to determine the diffusion coefficients in the nucleosol, cytosol and nuclear envelope shown in Table 1.

Table 1. The diffusion coefficients as found by LBM for the nucleosol, cytosol and nuclear envelope of HeLa cells, their average values and standard deviations (STD).

HeLa	D_{nuc} [µm^2/s]	D_{env} [µm^2/s]	D_{cyt} [µm^2/s]
1	29.8	0.11	44.6
2	29.9	0.05	43.1
3	24.9	0.01	19.9
4	37.2	0.59	37.2
5	29.1	0.11	44.5
6	21.4	0.06	32.2
7	25.3	0.2	12.7
8	29.1	0.4	29.2
9	36.3	0.4	36.3
10	32.1	0.4	16.1
11	28.6	0.1	28.6
12	27.2	0.1	27.3
Average	29.2	0.2	30.9
STD	4.5	0.2	10.8

The diffusion coefficients of EYFP in HeLa cells were found by bleaching the nucleus and comparing the resulting fluorescence distribution with that of the corresponding numerical simulation. The new approach gave excellent linear correlation between the

frames of the experiments and the related simulations, an example of which is shown in Figure 11. The resulting nucleosol diffusion coefficient, D_{nuc}, was 29.2 ±4.5 μm^2/s. The method also produced diffusion coefficients for the cytosol and nuclear envelope, although the main organelle explored was the nucleus. The former values were not determined accurately. Nevertheless, the cytosol diffusion coefficient, D_{cyt}, was found to be 30.9 ±10.8 μm^2/s and that of the nuclear envelope, D_{env}, was 0.2 ±0.2 μm^2/s.

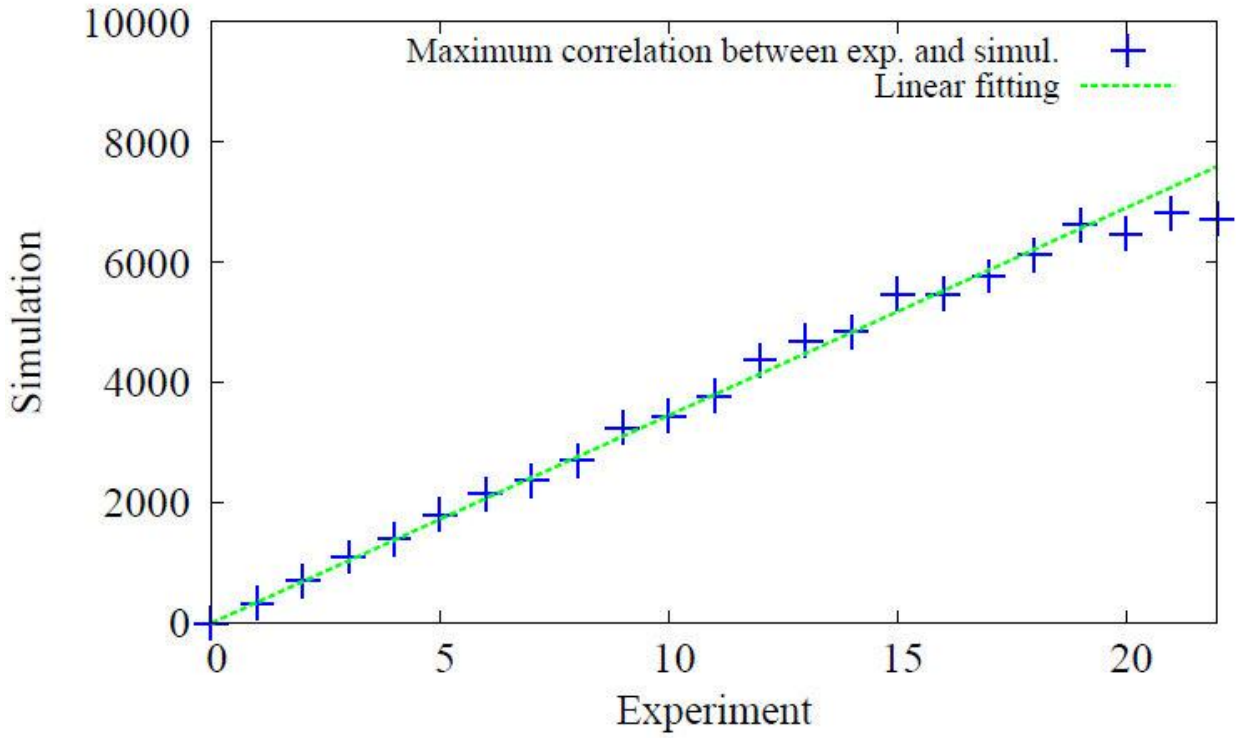


Figure 11. Correspondence between highest cross correlation values of an experiment and the corresponding simulation.

6. Discussion

Modeling involves developing a physical, conceptual and computer-based representation of the system considered. In this work a model which enabled analysis of nucleocytoplasmic diffusion of proteins in living HeLa cells was represented. Experiments were made in live cells with fluorescent proteins, confocal microscopy was used to acquire 3D data and ImageJ software was used to perform image analysis.

The results obtained in this way by FRAP showed that EYFP is freely diffusing inside the nucleus, which was demonstrated by the rapid recovery rate of free EYFP (Figure 8). On the other hand a conventional FRAP analysis produced very low diffusion coefficients, which means that the conditions in the measurements did not correspond to those assumed in the analysis.

A fully numerical modeling approach applied to diffusion was the LB method. The heterogeneous fluorescence intensity in the nucleoplasm was interpreted as a homogeneous distribution in the nucleosol, the liquid phase of the nucleoplasm. The plasma membrane was represented as an impermeable boundary and the nuclear envelope was considered as a permeable layer with a diffusion coefficient of its own. The method was not fine-tuned however so as to be able to determine τ_{cyt} and τ_{ne} very accurately, and that is why their values varied quite much.

The single colour fluorescence correlation spectroscopy (scFCS) technique in living cells has been used to show that the diffusion coefficient (D) of the fast fraction of EGFP molecules is 23.0 ± 1.0 $\mu m^2/s$ (SEM) and 25.1 ± 1.1 $\mu m^2/s$ in the nucleus and in the cytoplasm respectively [Maertens 2005]. These results are well comparable with the diffusion coefficients found here by the LB method.

There are a few possible sources of error in the present method. It is important that the nuclear envelope and the nucleus are reliably indentified. For that purpose the histone fusion protein was used to label the chromatin. Cells had a tendency to move during measurements, which had to be taken into account in the correlation analysis. Also, the

cross section analyzed at the imaging phase had to be indentified properly. Non-specific binding of the protein was not taken into account.

For clarity a freely diffusive and non-binding protein was selected. Binding reactions specific for the nucleus will obviously affect the protein mobility. Within the present methodology, binding/dissociation reactions with protein receptors can as well be taken into account.

7. Conclusions

Analysis of fluorescence recovery after photobleaching can be used to determine the dynamic parameters of proteins, including their diffusion coefficients, mobile fractions, transport rates and binding/dissociation rates. Here we focused on their diffusion coefficients.

A numerical model for FRAP was used to determine the diffusion properties of proteins by including the effect of the plasma membrane, the nuclear envelope, the cell nucleus, the fibrous structures of the cytoplasm and the chromatin, which reduced protein mobility. As this method simulated the fluorescence distribution in the entire cell, there was no need to make additional assumptions about the bleach process, such as e.g. the shape of the laser profile or its duration. The present method removed the difficulties of the conventional analysis, could produce, interesting results for protein diffusion in the cell as demonstrated above and can be applied in the future to define protein interactions beyond pure diffusion.

Appendix 1

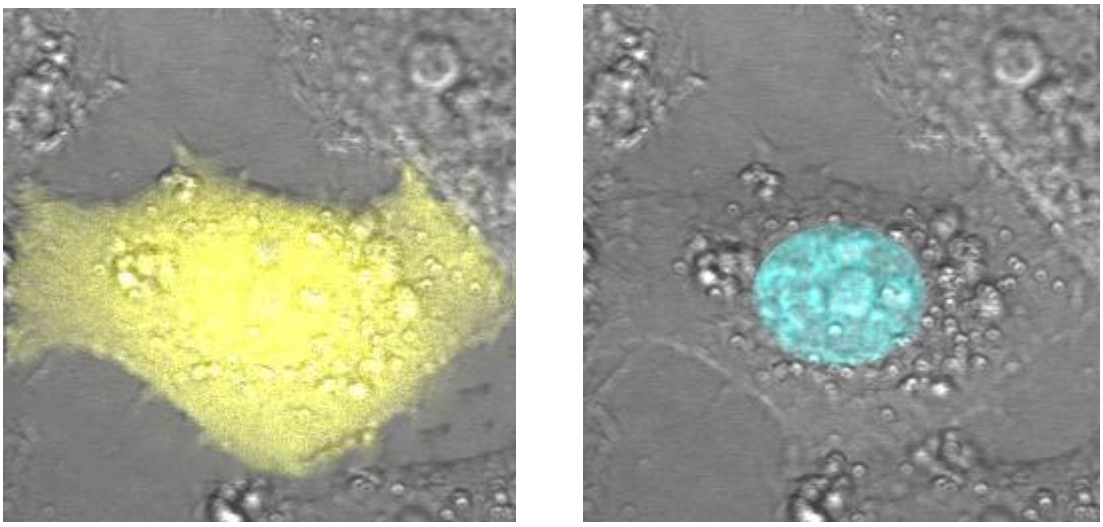
Materials and methods

1. Cell culture

HeLa cells were used in this research. Cells were maintained in Dulbecco's Modified Eagle Medium (DMEM, Gibco Introgen, Paisley, UK). They were grown as monolayer in 75 cm^2 culture flasks (Sarstedt Inc., Newton, USA) maintained in a 5% CO_2 incubator at +37 ºC. Cells were passaged twice a week.

2. Transfection

EYFP and H2B-ECFP constructs were transfected to HeLa cells with the TransIT Transfection Reagent (Mirus). Distributions of proteins are shown in Figure 12. Cells were growing on 50 mm culture dishes for live imaging. The protocol of cell transfection on 50mm culture dishes is described below. 15 µl of TransIT reagent was added to 650 µl of serum-free DMEM and incubated for 15 minutes. 4 µl of plasmid DNA was added to the transfection solution, and incubation at RT was continued for half an hour. The medium on the culture dish was replaced with fresh DMEM, and the transfection solution was added into the dish, followed by incubation overnight at +37 ºC.



A B

Figure 12. Images of HeLa cells showing the distributions of fluorescent proteins. ***A.*** *Distribution of EYFP which is a small, noninteracting protein that diffuses freely in the entire cell. EYFP is very good for photobleaching experiments.* ***B.*** *H2B ECFP shows the location and relative concentration of chromatin inside the nucleus and allows thus to deduce the position of the nucleus quite accurately.*

3. Live cell imaging with confocal microscope

Measurements were done with an Olympus confocal microscope. Object was UPLSAPO 60x (numerical aperture: 1.20). A specific region was bleached and then the sample was scanned to find how fast (in frames) the bleached region was recovered. A selected cell was bleached with the "tornado" method in a region of circular cross section of 10x10 pixels with 515 nm laser wavelength and 100 % laser power. Every sample was bleached only once. Bleaching was started after 10 frames were taken of the cell, and after the bleach additional 100 frames were recorded. Scanning was done with the maximum scanning speed. 12 nuclei were subjected to a FRAP experiment, and the data obtained were used to determine the diffusion coefficient of pEYFP-N3 in the imaged cells.

The bleached regions were chosen so that they were not near any cell organelle or in the nucleus near nucleolus because then the recovery would have taken longer and the results would have been less reliable.

4. Image analysis

The data obtained from the confocal microscope were analyzed with an image analysis program, in this case the ImageJ software. This kind of image analysis program makes the analysis of microscopy images straightforward and easy to perform.

References:

Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, Suprapto A, Karni-Schmidt O, Williams R, Chait BT, Sali A, Rout MP (2007) The molecular architecture of the nuclear pore complex. Nature 450, 695–701

Alberts B, Johnson A, Lewis J, Raff M, Roberts K and Walter P (2002) Molecular Biology of the Cell, 5th ed. Garland Publishing, New York

Angus I, Lamond E and William C (1998) Structure and function in the nucleus. Science 280, 547-553

Axelrod D, Koppel DE, Schlessinger J, Elson E (1976) Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. Biophys J 16, 1055–1069

Braga J, Desterro J and Carmo-Fonseca M (2004) Intracellular macromolecular mobility measured by fluorescence recovery after photobleaching with confocal laser scanning microscopes. Mol. Biol. Cell 15, 4749-4760

Bridger J, Foeger N, Kill I and Herrmann H (2007) The nuclear lamina both a structural framework and a platform for genome organization. FEBS Journal 274, 1354–1361

Carmo-Fonseca M, Platani M, Swedlow JR (2002) Macromolecular mobility inside the cell nucleus. Trends Cell Biol 12, 491–495

Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC (1994) Green fluorescent protein as a marker for gene expression. Science 263, 802-805

Cronshaw J, Krutchinsky A, Zhang W, Chait B, and Matunis M (2002) Proteomic analysis of the mammalian nuclear pore complex. JCB 158, 915-927

D'Angelo M, Anderson D, Richard E, and Hetzer M (2006) Nuclear Pores Form de Novo from Both Sides of the Nuclear Envelope. Science 312, 440-443

Dundr M, Hoffmann-Rohrer U, Hu Q, Grummt I, Rothblum LI, Phair RD, Misteli T (2002) A kinetic framework for a mammalian RNA polymerase in vivo. Science 298, 1623–1626

Elowitz MB, Surette MG, Wolf P-E., Stock, J. and Leibler S (1997) Photoactivation turns green fluorescent protein red. Curr. Biol. 7, 809–812

Everett R and Chelbi-Alix M (2007) PML and PML nuclear bodies: Implications in antiviral defence. Biochimie 89, 819-830

Gall J.G (2001) A role of Cajal bodies in assembly of the nuclear transcription machinery. FEBS let. 498, 164-167

Griebel M, Dornseifer T and Neunhoeffer T (1998) Numerical simulation in fluid dynamics: a practical introduction. 1-5

Harold P. Erickson (2009) Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. Biological Procedures Online 11, 32-51

Janicki SM, Spector DL (2003) Nuclear choreography: interpretations from living cells. Curr Opin Cell Biol 15, 149–157

Jovanovic-Talisman T, Tetenbaum-Novatt J, McKenney A, Zilman A, Peters R, Rout M (2009) Artificial nanopores that mimic the transport selectivity of the nuclear pore complex. Nature 457, 1023–1027

Kaviany M (1995) Principles of Heat Transfer in Porous Media. Springer-verlag, New York.

Kühn T, Ihalainen TO, Hyväluoma J, Dross N, Willman SF, Langowski J, Vihinen-Ranta M, and Timonen J (2011) Protein Diffusion in Mammalian Cell Cytoplasm. PLoS ONE 6, 1-12

Lanctôt C, Cheutin T, Cremer M, Cavalli G, Cremer T (2007) Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nat Rev Genet. 8, 104-15

Martin W. Hetzer, Tobias C. Walther and Iain W. Mattaj (2005) PUSHING THE ENVELOPE: Structure, Function, and Dynamics of the Nuclear Periphery. Annual Review of Cell and Developmental Biology 21, 347-380

Olson M, Dundr M, Szebeni A (2000) The nucleolus: an old factory with unexpected capabilities. Trends Cell Biol 10, 189–196

PDB: 1EMA Ormo M., Cubitt A., Kallio K., Gross L., Tsien R., Remington S. (1996) Crystal structure of the Aequorea victoria green fluorescent protein. Science 273, 1392-1395

Phair RD and Misteli T (2000) High mobility of the proteins in the mammalian cell nucleus. Nature 404, 604-609

Roix J, Misteli T (2002) Genomes, proteomes, and dynamic networks in the cell nucleus. Histochem Cell Biol 118, 105–116

Rout MP and Aitchison J D (2001) The Nuclear Pore Complex as a Transport Machine. The Journal of Biological Chemistry 276, 16593-16596

Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao Y and Chait B (2000) The Yeast Nuclear Pore Complex Composition, Architecture, and Transport Mechanism. JCB 148, 635-652

Soumpasis DM (1983) Theoretical analysis of fluorescence photobleaching recovery experiments. Biophys J. 41, 95-97

Sparague BL, Pego RL, Stavrev DA and McNally JG (2004) Analysis of binding reactions by Fluorescence Recovery after photobleaching. Biophys J. 86, 3473-3495

Strambio-De-Castillia C, Niepel M and Rout M (2010) The nuclear pore complex: bridging nuclear transport and gene regulation. Nature Reviews Molecular Cell Biology 11, 490-501

Succi S (2001) The Lattice Boltzmann Equation for Fluid Dynamics and Beyond. Oxford University Press, Oxford, UK

Swaminathan R, Hoang CP and Verkman A S (1997) Photobleaching recovery and anisotropy decay of green fluorescent protein GFP-S65T in solution and cells: cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion. Biophys. J. 72, 1900–1907

Wente S and Rout M (2010) The Nuclear Pore Complex and Nuclear Transport. Cold Spring Harb Perspect Biol 10, 1-20

Wilson K and Walker J (2000) Principles and thechniques of practical biochemistry, 5th ed. Cambridge university press, Cambridge