Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

MSci in Natural Sciences

Sofya Mikhaleva

born in: Moscow, Russia.

Oral examination: 06.11.2019

In situ high resolution studies of the FG-nucleoporins in the central channel of the nuclear pore complex

Referees: Dr. Martin Beck

Prof. Dr. Hans-Georg Kräusslich

Summary

The nuclear pore complex (NPC) is the largest protein complex in the nuclear envelope of eukaryotic cells. It provides a permeability barrier for rapid and selective transport of biomolecules in and out of the nucleus. While the scaffold structure and composition of the NPC are well understood, little is known about the permeability barrier, which is formed of multiple copies of about ten different intrinsically disordered nucleoporins. These proteins characteristically contain multiple phenyl-glycine repeats in their sequence and are referred to as FG-Nups. Due to technical limitations in the study of these highly flexible and dynamic proteins, the conformational dynamics and spatial arrangement of FG-Nups in the permeability barrier of the NPC remain elusive.

In this thesis I developed and established two complementary imaging techniques to visualise FG-Nups in the NPC in mammalian cells. To achieve this, I further developed a combination of genetic code expansion using unnatural amino acids and click-chemistry technologies suitable for high resolution fluorescence imaging. This method allows for efficient site-specific labelling of two sites in FG-Nups with small, photostable organic fluorophores with residue precision *in situ*. The first imaging approach used super-resolution localisation microscopy to precisely map the labelled site of the FG-Nups to a reference in the NPC. This resulted in images showing two-dimensional projection of FG-Nups distribution in NPCs, resolving distances down to 10nm. The second approach involved a complimentary technique based on confocal scanning microscopy: using fluorescence lifetime imaging microscopy (FLIM) to directly measure end-to-end distances between two sites in an FG-Nup less than 10nm apart using Förster resonance energy transfer (FRET).

The ultimate goal of my work was to use the approaches to determine the structural arrangement of FG-Nups in the permeability barrier of the NPC. I further applied polymer physics concepts of scaling laws to my results from FLIM-FRET studies. I was able to show that the two major players in the permeability barrier FG-Nups, Nup62 and Nup98, tend to resemble a state between a collapsed coil and an ideal chain. This finding demonstrates the first experimental evidence of the actual scaling of FG-Nups *in vivo*, which has implications for existing transport models. Aside from its contribution to understanding the NPC permeability barrier function, the developed approach lays the groundwork for a more detailed understanding of disordered protein dynamics, dimensions and functions inside the cell.

Zusammenfassung

Der Nuclear Pore Complex (NPC) ist der größte Proteinkomplex der Kernmembran eukaryontischer Zellen. Seine Permeabilitätsbarriere reguliert den Biomolekültransport zwischen Zellkern und Zytoplasma. Im Gegensatz zu Struktur und Zusammensetzung der Ringstruktur des NPC ist die Permeabilitätsbarriere, welche im Transportkanal lokalisiert ist, sehr schlecht erforscht, wobei bekannt ist, dass sie aus verschiedenen ungeordneten Nukleoporinen besteht. Typisch für diese ist, dass ihre Sequenz mehrere Phenyl-Glycin-Wiederholungen enthält, weswegen sie auch als FG-Nups bezeichnet werden. Aufgrund technischer Beschränkungen sind die Konformationsdynamik und die räumliche Anordnung dieser hochflexiblen und dynamischen Proteine in der Permeabilitätsbarriere des NPCs schwer erforschbar.

In dieser Arbeit habe ich zwei sich ergänzende Mikroskopieverfahren entwickelt, um FG-Nups in ihrem physiologischen Zustand im NPC zu abzubilden. Dazu habe ich eine Kombination aus genetischer Code-Erweiterung und Klick-Chemie-Technologien weiterentwickelt. Diese Technik erlaubt die aminosäurengenaue standortspezifische Kennzeichnung der FG-Nups mit zwei kleinen photostabilen organischen Fluorophore. Das erste Mikroskopieverfahren nutzte hochauflösende Lokalisationsmikroskopie, um eine markierte Stelle der FG-Nups einer Referenz im NPC genau zuzuordnen und dadurch Entfernungen oberhalb von 10nm zu messen. Das zweite Mikroskopieverfahren nutzte eine ergänzende Technik, die auf konfokaler Abtast-Mikroskopie basiert. Dabei nutzte ich Fluoreszenzlebensdauer-Mikroskopie (FLIM), um mit Hilfe des Förster-Resonanzenergietransfers (FRET) Entfernungen von weniger als 10nm zwischen zwei Stellen im FG-Nup zu messen.

Das Ziel meiner Arbeit war, mithilfe der hier entwickelten Verfahren die strukturelle Anordnung der FG-Nups in der Permeabilitätsbarriere des NPC zu bestimmen. Dazu wandte ich auch polymerphysikalische Konzepte der Skalengesetze auf meine Ergebnisse aus dem FLIM-FRET-Verfahren an. Hierdurch konnte ich zeigen, dass die beiden Hauptbestandteile der FG-Nups in der Permeabilitätsbarriere, Nup62 und Nup98, einer Anordnung zwischen einem 'collapsed coil' und einem 'ideal coil' nahekommen. Dieser Befund zeigt die ersten experimentellen Beweise zum tatsächlichen Zustand der FG-Nups in vivo, woraus sich Rückschlüsse auf die bestehenden Transportmodelle ziehen lassen. Darüber hinaus schafft dieses Verfahren die Grundlage für weitere Erkenntnisse über Dynamik, Dimensionen und Funktionen ungeordneter Proteine in Zellen.

Publication List

1. **Mikhaleva, S.**, Lemke, E. A. (2017). Beyond the Transport Function of Import Receptors: What's All the FUS about? *Cell*, Cell 173, 549–553.

2. Nikić, I.*, Girona, G.E.*,Kang, J.H.*,Paci, G., **Mikhaleva, S.**, Koehler, C., Shymanska, N., Ventura Santos, C., Spitz, D., Lemke, E. A. (2016). Debugging Eukaryotic Genetic Code Expansion for Site-Specific Click-PAINT Super-Resolution Microscopy. *Angewandte Chemie International Edition*, 55 (52), 16172–16176.

Contents

List of Figures

List of Tables

List of Abbreviations

aa: Amino acid

aaRS: Aminoacyl-tRNA synthetase

AFM: Atomic force microscopy

AIDS: Acquired immunodeficiency syndrome

APD: Autoproteolytic domain or avalanche photodiode detector (context dependent)

BCN: Bicyclo [6.1.0] nonyne-l-lysine (here *endo*)

BSA: Bovine Serum Albumin

CA: capsid

CRISPR: Clustered regularly interspaced short palindromic repeats

CTD: C-terminal domain

CypA: Cyclophilin A

DMSO: Dimethyl sulfoxide

DNA: Deoxyribonucleic acid

dSTORM: Direct stochastic optical reconstruction microscopy

EDTA: Ethylenediaminetetraacetic acid

EGFP: Enhanced green fluorescent protein

EMCCD: Electron multiplying charge-coupled device

FG: Phenylalanine-Glycine

FKBP: FK506-binding protein 12

FLIM: Fluorescence lifetime imaging microscopy

FRB: FKBP12-rapamycin binding domain

FRC: Fourier ring correlation

FRET: Förster resonance energy transfer

GDP: Guanosine-5′-diphosphate

GFP: Green fluorescent protein

GLEBS: Gle2-binding sequence

GLFG: Glycine-Leucine-Phenylalanine-Glycine

GLOX: Glucose oxidase with catalase

GSD: Ground state depletion

GTP: Guanosine-5′-triphosphate

GUI: Graphical user interface

HBV: Hepatitis B

HIV-1: Human immunodeficiency virus 1

H-tet: 1,2,4,5-tetrazine

IDP: Intrinsically disordered protein

IDR: Intrinsically disordered region

IPTG: Isopropyl β-D-1-thiogalactopyranoside

IRF: Instrument response function

Kap: Karyopherin

kDa: Kilo Dalton

LB: Lysogeny broth

MEA: Mercaptoethylamine

mRNA: Messenger ribonucleic acid

NC: Nucleocapsid

NES: Nuclear export signal

NLS: Nuclear localisation signal

NPC: Nuclear pore complex

NTD: N-terminal domain

NTR: Nuclear transport receptor

Nup: Nucleoporin

PA: Photoactivatable

PALM: Photoactivated localisation microscopy

PAM: Pulsed interleaved excitation (PIE) Analysis with MATLAB

PBS: Phosphate buffer saline

PCR: Polymerase chain reaction

PFA: Paraformaldehyde

PIC: Pre-integration complex

PKR: Protein kinase R

PMSF: Phenylmethane sulfonyl fluoride

PMT: Photomultiplier tube

POI: Protein of interest

PylRS: Pyrrolysyl-tRNA synthetase

RESOLFT: Reversible saturable optical fluorescence transitions

RNA: Ribonucleic acid

ROI: Region of interest

rRNA: Ribosomal ribonucleic acid

SAXS: Small-angle X-ray scattering

SCO: Cyclooctyne-l-lysine

SIM: Structured illumination microscopy

smFRET: Single molecule Förster resonance energy transfer

SMLM: Single molecule localisation microscopy

snRNA: Small nucleolar ribonucleic acid

SPAAC: Strain-promoted azide-alkyne cycloaddition

SPIEDAC: Strain-promoted inverse electron demand Diels-Alder click reaction

SSIM: Saturated structured illumination microscopy STED: Stimulated emission depletion

STORM: Stochastic optical reconstruction microscopy

TB: Transport buffer

TCEP: Tris(2-carboxyethyl)phosphine

TCHD: Trans–cyclohexane–1,2-diol

TCO*: Trans-cyclooct-2-en – 1 - lysine

TCSPC: Time-correlated single photon counting

TIRF: Total internal reflection fluorescence

tRNA: Transfer ribonucleic acid

WGA: Wheat germ agglutinin

UAA: Unnatural amino acid

1 Introduction

1.1 The Nuclear Pore Complex

The nuclear pore complex (NPC) is the largest macromolecular machine found in the nuclear envelope. These machines, besides being the prime transport gate between the nucleus and the cytoplasm, are also involved in chromatin organisation, regulation of gene expression and DNA repair mechanisms [1]. NPCs are comprised of multiple copies of 30 different proteins termed nucleoporins (Nups), which can be categorised into two types: scaffold Nups and phenylalanine-glycine (FG)-Nups. Scaffold Nups are folded proteins responsible for the solid frame of the NPC and the exterior elements such as the nuclear basket and the cytoplasmic filaments. FG-Nups, in contrast, are largely disordered and contain multiple FG-repeats. Scaffold Nups serve as the architectural elements of the NPC, to which the FG-Nups are grafted, filling the central channel.

Other important elements of the nucleocytoplasmic transport machinery are nuclear transport receptors (NTRs), also referred to as karyopherins. FG-Nups and NTRs form a permeability barrier with specific properties, allowing only for small molecules and ions to freely diffuse but preventing larger molecules (5nm or 40kDa) to pass. NTRs' prime function in this system is to offer shuttle large proteins if they present a special localisation signal, specifically recognised by some NTRs. The energy for the release of cargoes in this process comes from an asymmetric distribution of Ran GTPase regulators across the nuclear membrane [2, 3]. RanGEF and RCC1 guanine nucleotide exchange factors are predominantly in the nucleus, where they exchange RanGDP to RanGTP, while RanGAP is in the cytoplasm, hydrolising RanGTP to RanGDP form.

There exist about 30 different NTRs for different transport pathways [4]. Most NTRs are monodirectional and are referred to as importins or exportins, depending on the direction in or out of the nucleus respectively, but some are bidirectional [5, 6, 7]. Typical import complexes consist of a molecule presenting a nuclear localisation signal (NLS) and an import receptor, for example a complex of importin-α and importin-β [8]. RanGTP in the nucleus disassembles the import complexes. Export complexes involve a molecule presenting a nuclear export signal (NES) and an export receptor, for example exportin-1 and RanGTP. RanGAP then hydrolyses RanGTP into RanGDP, releasing the cargo from the export complex. NTRs are then

recycled back to bind the next cargo and RanGDP is imported back into the nucleus by the nuclear transport factor 2.

The FG-Nups are considered responsible for rapid and selective transport of the NTR-cargo complexes and are essential for NPC function [9, 10, 11]. Vertebrate NPCs have around 11 different FG-Nups that together provide more than 5000 functional FG motifs per pore [12].

1.1.1 Structural organisation and plasticity of the NPC

NPC architecture and function are rather well conserved among different species. However, their composition and size vary not only across the tree of life but also between different cell types, tissues and in development. Nuclear pore complex
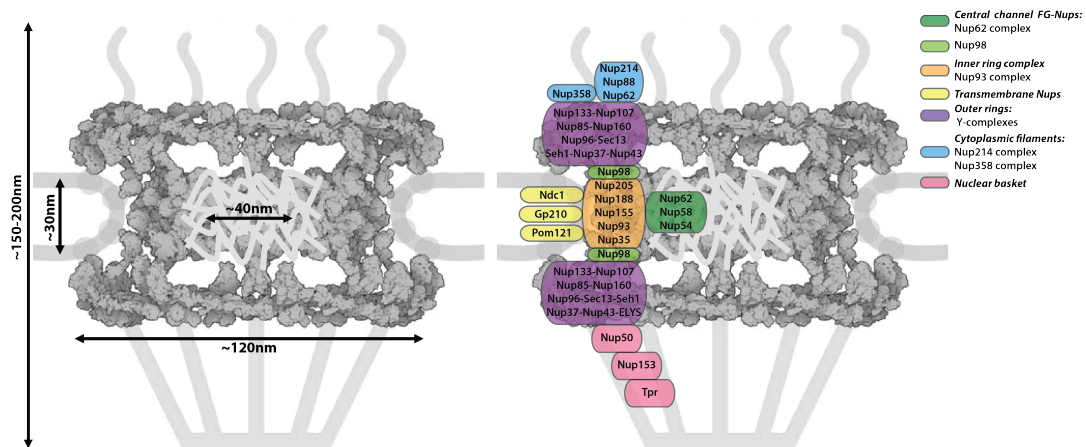
FIGURE 1.1: Schematic representation of the vertebrate nuclear pore complex dimensions and composition. The scheme is colour coded according to the Nup complexes. No Aladin, Rae, NDC are presented.

is assembled from several modularities or repetitive building blocks in a hierarchical way. Overall the pore consists of three connected rings: one inner ring in the centre and two outer rings on the cytoplasmic and the nucleoplasmic sites. Most of the structural components in the cytoplasmic and the nuclear rings are identical and symmetrical, although they vary between different species. Outer rings are mostly comprised of the Y-complex or Nup107-Nup160 complex. In human cells the Y-complexes dimerise in a head-to-tail arrangement in 16 Y-shaped assemblies on each side of the NPC, while in *Saccharomyces cerevisiae* yeast there are only 16 Y complexes total and in *Chlamydomonas reinhadatii* alga 16 Y copies are on the nuclear side and only 8 on the cytoplasmic [13]. Very similarly to the Y-complexes in human cells, the inner ring is formed by four horizontally stacked rings, each made of 8 core modules: 4 inner and 4 outer (also called spokes), that are also rotationally shifted with respect to each other and symmetric across the nuclear envelope plane. Each core module consists of 32 copies each of Nup205 or Nup188 , Nup93, Nup155 and the Nup62 complex; another 16 copies of Nup155 connect the inner ring to the outer

rings [14].

The inner ring is one of the most plastic structural assemblies in the NPC. There have been studies observing the diameter of the pore in intact cells suggesting more dilated [15] and constricted conformations [16]. Moreover, the inner ring of the NPC in *C.reinhardtii* alga has a significantly larger diameter (63nm) than that of a human (43nm), even though the composition and relative arrangement of subunits are identical [17]. There is therefore a strong belief that the central channel of the NPC is highly flexible and its dimensions may vary according to a physiological state.

Another important candidate to explain plasticity of the NPC is the permeability barrier itself. Especially, since FG-Nups provide multiple physical links between and within many structural complexes of the NPC [18]. The FG-Nups in the central channel have been previously observed reversibly collapsing by addition of amphipathic alcohols like trans–cyclohexane– 1,2-diol (TCHD) [19] and also deoxyglucose and chilling [20] and calcium [21]. The FG-Nups may also undergo multiple rearrangements contributing to dilation specifically for accommodating large transport cargoes like viral capsids or ribosomes. However, it should be noted that the actual rearrangement of collapsed FG-Nups in the NPC has never been visualised.

Pores not only differ in size between different human cells or even within the same cell [15] but also have cell type–specific composition and stoichiometry [12]. Furthermore, NPCs change their composition in mouse embryos [22] and likewise size and number during the development of *Xenopus laevis* embryos [23]. NPCs can also adapt to multiple cellular changes and alter their composition by varying the expression levels of its subunits [24]. Consequently, it is becoming increasingly clear that NPCs are in fact quite heterogeneous structures. One of the main questions remaining is how and whether this heterogeneity affects the function of the NPC. Some evidence suggests that, for example, HIV-1 virus infection and integration may take different pathways through the NPC in different cell types, which in turn have different Nup compositions [25].

1.1.2 Nup62 and Nup98

Nup62

Mammalian Nup62 (Nsp1 is a homologue in in *S. cerevisae*) is an FG-Nup that is found in the NPC in two complexes: Nup62 (Nup62, Nup58, and Nup54) and Nup214 (Nup214, Nup82 and Nup62). Nup62 complex is a stable 1:1:1 complex that heterotrimerizes from the coiled-coil domains [26], which might also be the case for Nup214 complex even though it has not yet been observed [13]. Nup62 can be O-GlcNAcylated and this modification may be important for its interaction with Nup88 to the Nup214 complex [27]. Nup62 complex associates to the NPC via the short linear motifs in the N-terminal of the Nup93 only after it is formed [28]. The
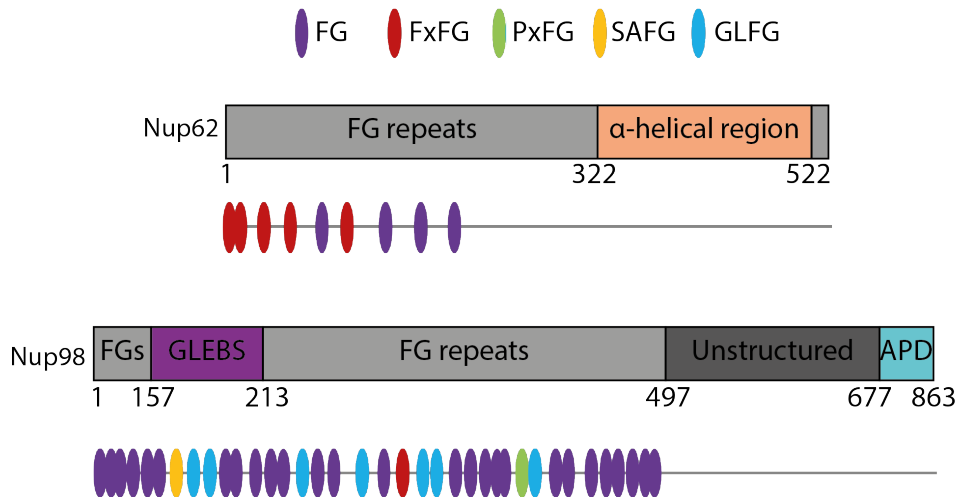
FIGURE 1.2: Schematic representation of FG-Nups Nup62 and Nup98 involved in this study. Coloured ellipses demonstrate different type of FG repeats in the polypeptide sequence; shaded rectangles represent functional domains of the Nups.

complex is present in 32 copies, forming approximately two 16-membered rings in the N-termini with FG-regions potentially projecting circumferentially into the center of the pore [29]. Nup62 consists of an N-terminal, predominantly FxFG region, a largely unstructured region and a coiled-coil pore-targeting C-terminal domain (Figure 1.2).Interestingly, in contrast to Nup98, depletion of Nup62 complexes does not completely abolish the permeability barrier in *X. laevis* egg extracts [10]. Nup62 also plays roles outside the NPC context. Nup62 was shown to be involved in mitotis regulation, spindle assembly and chrosomes alignment [30, 31]. Misregulation of posttransaltional control of Nup62 has also been shown to be a marker of primary biliary cirrhosis [32], diabetes [33] and was linked to chronic stress [34].

Nup98
Nup98 is one of the essential FG-Nups to maintain the permeability barrier of the NPC [10] and is encoded by an vital gene, deletion of which is lethal [35, 36]. Initially, mammalian Nup98 (Nup145N in *S. cerevisae*) is expressed as a fusion with the scaffold component Nup96 (Nup145C in *S. cerevisae*). Cleavage in the autoproteolytic domain (APD) in the middle of the polypeptide chain [37] relocates the C-terminal component (Nup96) to associate with Sec13 protein and integrates into Y-complexes [38, 39], while the N-terminal Nup98 associates with the inner ring and the cytoplasmic filaments [40]. Nup98 consists of three main regions [41]: (1) an N-terminal FG and GLFG-repeats rich region with a Gle2-binding sequence (GLEBS) domain, where Gle2 is a *S. cerevisiae* homologue of human mRNA export factor Rae1 [42]; (2) a largely disordered region that mediates interactions with inner ring nucleoporins Nup155, Nup205, and Nup188 (Nup170,157, Nup192, Nup188 in *S. cerevisae*) [40, 43]; (3) the C-terminal pore-targeting region containing an APD [44] (Figure 1.2). Nup98 has been estimated to occur in ≈48 copies per NPC [12].

The N-terminal FG-rich domain of Nup98 (1-497aa) is known to interact with multiple nuclear transport receptors like Crm1 [45] and also to contribute to multiple gene regulation mechanisms [46].

Nup98 is the only GLFG Nup in the mammalian NPC and was previously observed to associate into nuclear bodies termed GLFG bodies in HeLa cells [47]. Together with Nup98, some components of the Y-complex, Elys and Crm1 dynamically localise to the endogeneous GLFG bodies [48]. Nup98 is a highly mobile nucleoporin, potentially directly exporting mRNA, snRNA, and rRNA from the nucleus [47]. The FG-domain of Nup98 has also been to shown to spontaneously phase separate *in vitro* forming droplets that rapidly turn into a hydrogel [49] at critical concentrations ranging from 20 to 700nM. Phase separated particles exhibited permeability barrier properties in accordance to the selective phase model (described in the next section). Nup98 like Nup62 also plays multiple roles outside that of the NPC. One of the major roles it plays is in gene and transcription regulation [46].

1.1.3 Models of the NPC permeability barrier

The nuclear pore complex is truly a "Golden Apple of Discord" of biology. Many models attempt to describe the state of matter of the permeability barrier of the NPC. They look at it from thermodynamics, polymer physics and mechanistic perspectives, however, to this day the question remains unresolved. The following section briefly describes each of the existing models in chronological order and then gives a fairly recent polymer physics perspective on the problem.

Selective phase model
This model was proposed by the Görlich group by Ribbeck *et al.* [50] in 2001 and states that the FG-Nups in the central channel interact with each other to form a selective three-dimensional mesh that excludes large molecules. This meshwork restricts inert molecules but NTRs can locally outcompete the FGs attractions and get incorporated into the network.

Oily spaghetti model
This model was introduced in the same year as the previous by I.G. Macara [51] stating that FG-Nups form extended chains layer like "oily spaghetty" around the "central tube" (was believed to be around 10nm in diameter channel right in the middle of the NPC) about 7 nm thick. In this model NTRs loaded with cargoes can easily push FG-Nups aside and slide through via multiple interactions, similar to the virtual gate model described next.

Virtual gate model

Virtual gate model was a prototype for the polymer brush model and was first proposed in 2003 by Rout *et al.* [52]. This model doesn't take into account any cohesiveness of FG-Nups and simply considers them 'entropic bristles', whiping and writhing around their anchor points at the NPC. These bristles can repel or 'push' large macromolecules away from the central channel, thus creating an entropic barrier. NTRs carrying cargoes can overcome this barrier by binding to the FG domains, thus lowering the entropic barrier, allowing for transport.

Reduction of dimensionality model

This model was proposed in 2005 by R. Peters [53]. It categorises FG-Nups into two groups, one being coherent FG-Nups lining the inner walls of the scaffold and the other group of loose peptide chains forming a narrow flexible 'tube' 8–10nm in diameter at the very center. The narrow tube is thought to primarily be for passive diffusion, while NTRs carrying a cargo tend to make interactions with the FG-network on the sides making their way through via a two-dimensional random walk. Large macromolecules are excluded from being transported because they do not specifically interact with the FG-surface.

Polymer brush model

This model was proposed in 2006 by R. Lim *et al.* and suggests that interactions between FG-Nups are entropically dominated and associate less strongly with each other than suggested by the selective phase model [54]. In polymer theory the extension of the chain away from the grafted surface depends on the density of grafting: the more polymers are in vicinity of each other the more extended they become to avoid overlapping [55]. Lim *et al.* use this principle in atomic force microscopy (AFM) on FG-Nups tethered to gold nanodots to show that one of the FG-Nups (Nup153) acts as a 'brush' that can be collapsed by NTRs-cargo complexes during transport. The Lim group has continued to develop the model and has shown permeability barrier properties also in nanopores with grafted FG-Nups and proposed a transient highly dynamic interaction inside the central channel of the NPC [56, 57]. Overall, these results suggest that the selective gating mechanism is predominantly an interfacial effect arising from the entropic fluctuations of several unfolded FG domains within the peripheral near-field of individual NPCs instead of a bulk-like hydrophobic gel or meshwork.

Hydrogel model

This model was proposed by S. Frey and D. Görlich and is an extension to the selective phase model [58]. The hydrogel model is based on the ability of FG-domains of yeast Nups to form saturated hydrogels resembling permeability barrier properties. The properties of FG-Nups to phase separate into hydrogels were further confirmed

for multiple FG-Nups also of other species*in vitro* [59, 10, 60, 61, 49]. However, evidence of such state has not yet been found in cells.

Two-gate model

This model was proposed by Patel *et al.* proposing the existence of two different phases between FG-Nups anchored at the periphery or at the centre of the NPC [62]. Using a combination of *in vitro* assay and deletions in live yeast authors concluded that the cohesive Nups in the centre follow the selective phase model and form a selective meshwork, while others that are not cohesive do behave more like repulsive bristles in virtual gate model.

Forest model

This model was proposed in 2010 by Yamada *at al.* suggesting both 'hydrogel-like' and 'brush-like' behaviours coexist in the NPC in two different 'zones'. In their study the authors used FG-Nups purified from yeast and looked at their amino acid compositions to split them into groups according to their charge-to-hydrophobicity ratios (cohesiveness strength parameter). The class with low charge content adopts a collapsed coil configuration ('hydrogel'), while the class with high charge content form more extended coils repelling each other ('brush'). Smaller cargo molecules bound to NTRs are more likely to interact with the shrubs while bigger cargo molecules like ribosomal subunits will interact with trees and be transported through the centre of the transport conduit.

Kap-centric model

This model was first formulated in 2015 by R. Lim *et al.*, who suggested the polymer brush model [63] and by A. Lowe *et al.* [64]. The Kap-centric model is based on the observation that Kaps or NTRs are always present at fairly high concentrations in the FG-Nups phase and can modulate its behaviour [64, 63]. The model suggests two distinct pools of NTRs, one stably associated with the Nups, almost saturating the pore, and one transiently. The transiently associated pool of NTRs allows for fast transport. In essence, the model states that the permeability barrier properties of the FG-Nups arrangement in the NPC and the transport speed depend directly on the NTRs in a concentration dependent manner. Key experimental evidence for this model comes from the observations that there is always an endogenous pool of NTRs like importin-α and importin-β in nuclear envelopes in permeabilised cells and in isolated *X. laevis* envelopes [65, 66, 67].

New models

The increasing number of studies of intrinsically disordered proteins (IDPs) and phase separation in biology has emphasised the importance of the properties of the amino acid composition in determining structure and behaviour of many IDPs [68, 69, 70, 71]. IDPs are proteins that lack any tertiary structure, are very flexible and

play vital roles in many essential biological processes [72, 73]. The intrinsically dis-ordered regions (IDRs) of some proteins and entirely disordered IDPs can be charac-terised by the low sequence complexity (over-representation of a few residues) and a small number of hydrophobic amino acids. Due to low complexity composition of most of IDPs they often tend to transiently and non-covalently interact with each other resulting in a phase transition, forming supramolecular condensates. FG-Nups are also IDPs and were previously shown to undergo phase transitions into a dense hydrogel that resembles permeability barrier properties [49, 74].

In the context of FG-Nups Yamada *et al.* suggested different levels of organisation of intrinsically disordered domains of FG-Nups based on their sequence [75]. They proposed that from charge and hydrophobicity content of FG-Nups one can estimate their spatial configuration, such that globular structures are formed by FG-Nups with a low charge content, while those containing many charged residues tend to be in extended coil conformation. This charge-to-hydrophobicity ratio is often used in multiple simulations of FG-Nups as cohesiveness strength parameter, which can also be related to the permeability properties. In these models regions of high cohe-siveness would correlate with higher permeability [76, 77]. Other models emphasise the impact of the charges in forming an electrostatic potential in the NPC, which would inflict structural arrangement of FG-Nups and effect their phase separation [78, 79]. Many of these models also take into account enthalpic and entropic ef-fects of NTRs partitioning into the FG-Nups phase. Recently authors of one of the mesoscale computational simulations suggested that there is in fact a perfect bal-ance between repulsive and attractive interaction of FG-Nups in the NPC, such that FG-Nups would behave as an ideal chain [80]. Thus, there are multiple new com-putational models arising that well-describe the experimental results and take into account multiple parameters. What is still missing is good experimental data from *in situ* measurements to advance these developments.

1.2   Genetic code expansion technology for site-specific pro-tein labelling in mammalian cells

Labelling proteins for super-resolution microscopy and finding high quality probes is a huge field of research. Due to rapid developments in resolution enhancement there is a constant demand in improving the sample labelling quality and minimis-ing any perturbations to protein functionality. Many methods for small molecule based protein labelling usually involve either using a specific primary antibody in combination with a labelled secondary antibody or a smaller nanobody [81], usu-ally against a fluorescent protein [82], or using specific fluorescent probes against an enzymatic tag such as Halo-tag [83], SNAP-tag [84] or CLIP-tag [85]. These meth-ods are typically limited to targeting only the termini of the protein and present

relatively large label (> 5nm in case of antibodies [86]) that may localisation precision and achievable labelling density. Alternatively there are also methods that use a small peptide tag at the protein termini to be specifically recognised by a fluorophore [87] or a nanobody [88]. The most direct and versatile way to site-specifically target the POI is a combination of genetic code expansion and click-labelling.

Genetic code expansion and, in particular, Amber suppression are powerful techniques which allow for site-specific encoding of unnatural amino acids (UAAs) into any part of the polypeptide sequence of a protein of interest (POI). Amber suppression involves replacing an endogenous amino acid codon with a STOP codon TAG via site directed mutagenesis [89]. An orthogonal tRNA recognises this codon and an aminoacyl-tRNA synthetase (aaRS) acylates only this tRNA resulting in specific incorporation of a UAA exactly in the Amber codon position in the polypeptide sequence by the ribosome. Orthogonality in this context means aaRS and tRNA do not cross-react with any of the tRNA/aaRS components of the endogenous system. The TAG termination codon is typically chosen as it is least frequently used in translation termination in mammalians [90]. Pyrrolysine synthetase (PylRS) from the *Methanosarcinaceae family* (*M. mazei, M. barkeri*) is the most commonly used aaRS in mammalian cells [91, 92].

The UAA can be used to introduce a uniquely reactive group into the protein. To site-specifically label a protein, one of the most commonly used groups is one undergoing a catalyst-free click reaction such as a strain-promoted azide-alkyne cycloaddition (SPAAC) or the strain-promoted inverse electron demand Diels-Alder click reaction (SPIEDAC) reaction [93, 94]. SPIEDAC reactions with tetrazine-dyes are the fastest [95] and were used to show site-specific labelling with organic dyes of mammalian insulin receptors [96], cytoskeletal proteins [97] and a component of the NPC - Nup153 [98], using super-resolution microscopy. Enforcing the PylRS/tRNA system to be localised to the cytoplasm using the addition of the NES (NESPylRS/tRNA) significantly improved of the efficiency of the Amber suppression system [98].

1.3 Super-resolution microscopy

Super-resolution microscopy acquires its name from its ability to circumvent Abbe's diffraction limit of light: $d = \frac{\lambda}{2NA}$, where λ is the wavelength of light and NA is the numerical aperture of the objective. There are three main methods: (1) illuminating the molecules with pattern illumination, (2) making use of REversible Saturable OpticaL Fluorescence Transitions (RESOLFT) or (3) using the ability of fluorescent molecules to blink and imaging such events sequentially (Single Molecule Localisation Microscopy). Since the discovery of these methods there have been many improvements in both sample preparation and technical advancements. The following

section explores structured illumination super-resolution methods and then describe in details physical principles of different methods of Localisation microscopy, which form part of the backbone of the current project's research.

1.3.1   Illumination pattern microscopy

One of the illumination pattern super-resolution microscopies is structured illumination microscopy (SIM), a wide-field super-resolution microscopy technique. In the simplest SIM configuration excitation light is passed through spatially structured line-patterned grating before it illuminates the sample [99]. When the excitation pattern mixes with that of the sample, they produce an interference pattern called moiré fringes. The striped excitation pattern is translated and/or rotated such that multiple images with the high-resolution information of the sample contained in the different moiré fringes are generated and then post-processed (mathematically removing the known illumination pattern), resulting in the final reconstructed image with improved resolution. SIM resolution has been pushed down to 100nm laterally and about 400nm axially [100]. This can be further enhanced to have the lateral resolution of up to 50nm by using saturated SIM (SSIM). In SSIM fluorescent emission of fluorophores is saturated to the point at which it exhibits a non-linear response to increasing excitation intensities. This effectively encodes higher-order harmonics into the illumination pattern and higher resolution is achieved using 4 total harmonic orders [101].

1.3.2   Reversible saturable optical fluorescence transitions

Another way to image beyond the resolution limit is to use a "doughnut"-shaped illumination beam to deplete fluorescence around a central focal spot by stimulated emission using intense laser light. This technique is called stimulated emission depletion (STED) microscopy [102]. All molecules illuminated by the excitation beam will return to the ground state and only the fluorescence at the small focal spot will be detected. The general term for super-resolution techniques that make use of the reversible non-linear switching of the fluorophore state is Reversible saturable optical fluorescence transitions (RESOLFT). RESOLFT includes STED and Ground State Depletion microscopy (GSD). The maximal resolution achievable with STED can be described using a modified Abbe's equation:

$$d = \frac{\lambda}{2NA\sqrt{1 + \frac{I_{max}}{I_{sat}}}}, \tag{1.1}$$

where I_{max} is the intensity of the depletion beam and I_{sat} is saturation intensity.

1.3.3 Single Molecule Localisation Microscopy

Single Molecule Localisation Microscopy (SMLM) is based on high accuracy localisation of photoactivatable or reversibly photoswitchable fluorophores with nanometer-precision. Each of the imaging cycles records a diffraction limited image of multiple single emitters. The centre of each of fluorescence emissions can be calculated through a Gaussian fit of the point spread function (PSF). Since only a fraction of fluorophores are in the ON state at a given time, over multiple imaging cycles the final sub-diffraction image can be reconstructed. Photoactivated localisation microscopy (PALM) [103] and fluorescence PALM (fPALM) [104] use photoswitchable proteins (e.g. photoactivatable green fluorescent protein (PA-GFP)) to genetically tag the proteins of interest.

Both stochastic optical reconstruction microscopy (STORM) [105] and direct STORM (dSTORM) [106] use photoswitchable dyes (e.g. cyanine dyes Cy3 and Cy5 or Alexa647). STORM relies on the close proximity of the dyes for controlled switching between the ON and OFF states. In STORM red laser switches the fluorescent emitters to a stable dark state and green laser converts them back to the emitting ON state in a controlled manner. dSTORM uses the properties of the fluorophores to be reversibly switched by themselves depending on the laser power or with the help of thiol-containing reducing agents like β-mercaptoethylamine MEA [107] and an oxygen scavenging system of glucose oxidase with catalase (GLOX) [108]. STORM, PALM and fPALM use two excitation beams. PALM and fPALM use a second violet laser (usually 405nm) to activate the remaining inactive photoswitchable proteins after each cycle of excitation and bleaching. dSTORM and ground-state depletion and single-molecule return microscopy (GSDIM) [109] mostly use only one high intensity excitation beam to send all of the emitters to their triplet or another metastable dark state, such that only a subset of these emitters spontaneously return to the ground state and get excited to the singlet state, emitting a photon. But they also use some help from the second violet (405nm) laser to increase the activation rates of the fluorophores to undergo switching [110].

1.3.4 Localisation precision and resolution of SMLM

There are two important parameters used in determining the quality and the effective resolution of super-resolved images: localisation precision and resolution. Localisation precision represents the accuracy of the estimation of a position of a fluorescence emitter. The error in the position determination or localisation precision can be theoretically calculated using the equation [111] for an electron multiplying charge-coupled device EMCCD camera also used in this study:

$$\sigma_i = \sqrt{\frac{s_i^2 + \frac{a^2}{12}}{N} + \frac{8\pi s_i^4 b^2}{a^2 N^2}}, \tag{1.2}$$

where N is the number of photons, a is the pixel size of the camera, b is the standard deviation of the background (including fluorescence background and detector noise), and s_i is the width of the Gaussian. The index i refers to either the x or y direction. The first term under the square root on the right-hand side of Equation (1.2) refers to the photon noise, whereas the second term encompasses the effect of finite pixel size of the detector.

This equation clearly demonstrates the importance of the photon yield and signal-to-noise ration in the experimental method. For example, organic fluorophores can provide several thousands of photons per on-state and can be used in to more densely decorate the protein of interest (POI), compared to proteins at only hundreds of photons per on-state. Since background also plays a crucial role, most of the above mentioned methods use wide-field microscopy with total internal reflection fluorescence (TIRF) illumination. In TIRF microscopy a very thin region of sample (up to 200nm) on the surface of the coverslip is illuminated, thus avoiding any fluorescent background coming from the rest of the sample and significantly improving the signal-to-noise ratio.

Resolution of the reconstructed images is determined by using the concept of Fourier ring correlation (FRC) originated from electron microscopy [112, 113]. FRC method is considered a universal method of determining resolution in all super-resolution techniques [112, 113, 114].

1.3.5 Super-resolution microscopy of the NPCs

Due to its eight-fold symmetry and well known composition and numbers of each of the proteins, the NPC became one of the standard probes to be used in super-resolution microscopy (SRM). One of the first attempts to use SRM was to either look at relative positions of different Nups in nuclear envelope cross-sections using SIM [41] and 3D SIM [100] or to resolve the eight-fold symmetry of a ring formed by one of the structural Nups -gp210 in *Xenopus* nuclear extracts using high-resolution dual colour dSTORM imaging [115] and later STED imaging [116] or to quantitatively assess absolute stoichiometry of the NPC [117] in an integrative approach.

NPCs are also widely used as a standard labelling sample for method development and resolution assessment. Rapid progress in development of SRM allowed to improve its resolution to be applicable to resolve many protein complexes structures. One of the studies used localisation microscopy and a combination of specific antibodies against Nups and labelled anti-GFP nanobodies in stable CRISPR knock-in cell lines of Nup-GFP fusions to precisely map the individual proteins of the NPC Y-complex [118], which agreed well with the high resolution structure obtained by cryo-electron tomography [39]. Another recent study suggested using NPCs as a

universal biological standard for quantitative assessment and systematic optimi-
sation of factors like microscope settings, labelling and imaging conditions on the
quality of the final image using single molecule localisation microscopy [119]. The
standard nucleoporin was suggested to be Nup96 due to its very well characterised
architecture of two well defined eight-fold symmetry rings, each of 16 copies of the
protein [120, 118]. Authors suggest using homozygose CRISPR knock-in cell lines
expressing Nup96 tagged with either EGFP or an enzymatic label like SNAP-tag or
HaloTag as an experimental sample and provide software for optimisation of imag-
ing conditions for all types of super-resolution microscopies with a focus on single
molecule localisation microscopy methods.

All the described studies focus on structured Nups as they resemble a structure
of well-defined dimensions and composition. Nevertheless, FG-Nups of the NPC
remain poorly researched even by such a versatile method as super-resolution mi-
croscopy. Multiple studies managed to resolve the eight-fold symmetry of struc-
tured domains of some FG-Nups like Nup98 *Xenopus* [121] and Nup153 in mam-
malian cells [98, 122]. The FG-region of Nup62 was reported to be at the centre
of the NPC in super-resolution images of labelled GFP-nanobodies targeting its N-
terminally fused GFP [122] or a Nup62 specific antibody [118]. Polarisation PALM
was used to asses rotational mobility and crowding using fluorescent proteins in-
serted into different parts of the Nup98 and Nup153 chains, however this does not
direcötydirectly report on their conformational state [123]. Thus, there remains a
clear gap in the knowledge of systematic localisation and arrangement of the FG-
regions of the FG-Nups in the NPC.

1.4   A powerful fluorescence tool: FLIM-FRET

1.4.1   Fluorescence resonance energy transfer as a tool measure absolute distances in biological molecules

Fluorescence or Förster resonance energy transfer (FRET) is a radiationless mecha-
nism of energy transfer between a donor fluorophore in an excited electronic state
and an acceptor fluorophore via long-range dipole-dipole interactions. This phe-
nomenon was first quantitatively described by Theodor Förster in 1946 [124]. FRET
is one of the widely used techniques to precisely measure distances between sites
of biological molecules in a typical range of 10-90Å as the transfer rate depends on
distance and dye orientation.

In vitro the most sensitive way of studying conformational dynamics and interac-
tions of proteins with single molecule precision is using single molecule FRET (sm-
FRET). smFRET resolves the FRET signal of each individual molecule in the con-
focal volume. This method allows to precisely measure end-to-end distances and

dynamic properties of molecules from histograms of FRET efficiencies and donor lifetimes. The fluorescence lifetime is the average time that a molecule remains in the excited state before returning to the ground state by emitting a photon. smFRET is widely used to study native states of IDPs and also any changes in their structure or dynamics upon binding [125, 126, 11].

In cells FRET is mainly used as a powerful tool to study protein-protein interactions (intermolecular FRET), for example, in oligomerisation studies [127] and in protein complexes [128], as a biosensor (intramolecular FRET), for example, for Ca^{2+} ions [129]. In these types of experiments two fluorescent proteins are commonly used. There are five most commonly used approaches to measure FRET:

1. intensity-based FRET or sensitised emission: two-channel imaging with applied corrections for excitation and emission cross talk [130, 131]

2. acceptor photobleaching or donor dequenching [132, 133, 134]

3. fluorescence lifetime imaging (discussed in the next section)

4. spectral imaging [135]

Sensitised emission is one of the simplest methods of measuring FRET, yet it requires a broad set of controls. This method can be used in both wide field and confocal microscopies with two channels, two emission filters and two excitation wavelengths each corresponding to donor and acceptor fluorophores. Using this method one observes that donor emission is decreased, while acceptor emission is increased (sensitised) in a process of energy transfer. Thus, imaging involves measuring donor signal in both donor and acceptor channels and acceptor signal in both channels. Moreover, one has to measure appropriate controls like the sample that only has the donor fluorophore present and the sample with only acceptor fluorophore to appropriately correct for cross talk [136, 131]:

$$
\begin{aligned}
E_{FRET} &= \frac{FRET - D\beta - A\alpha}{A}, \\
\alpha &= \frac{A_{\lambda_D}}{A_{\lambda_A}}, \\
\beta &= \frac{A_{\lambda_D}}{D_{\lambda_D}},
\end{aligned}
\tag{1.3}
$$

where FRET=FRET channel fluorescence intensity; D=Donor only channel intensity; A=acceptor only channel intensity (subscripts indicate excitation wavelength); α is the acceptor direct excitation correction factor and β is the donor leakage correction factor.

Acceptor photobleaching is another easy to use method. It is based on the fact that in a FRET sample donor is being quenched by the acceptor. Thus, once the acceptor

is fully photobleached, the donor will dequench, resulting in an increase of donor fluorophore emission. However, this method assumes that there is no other photophysical change in the local environment other than photobleaching of the acceptor. To quantify FRET efficiency in this process one measures donor channel fluorescence intensity D_{pre} in the selected region before the photobleaching and D_{post} after the photobleaching. In this experiment FRET efficiency can be quantified as:

$$E_{FRET} = \frac{D_{post} - D_{pre}}{D_{post}} \tag{1.4}$$

Both described methods estimate the apparent FRET efficiencies in cells. Absolute FRET efficiencies can be calculated applying normalisation to account for differences between the donor and acceptor fluorophores quantum yield and detection efficiency. However, they both estimate FRET from the ensemble of all donors in the sample, failing to precisely quantify donor only species contribution. Moreover, the described techniques have too many controls and bear the uncertainty of many artifacts originating from photobleaching and differences in excitation power and fluorophore concentrations. Fluorescent lifetime imaging microscopy, on the other hand, allows for direct absolute FRET measurements and precise estimation of the donor species that do and do not undergo energy transfer.

1.4.2 Fluorescent lifetime imaging microscopy

Fluorescent lifetime imaging microscopy (FLIM) is a fluorescence technique that produces an image based on the differences in the exponential decay rate of the fluorescence from the sample. For the most common applications of fluorescence lifetime imaging the technique is based on a combination of a confocal scanning or a multiphoton (for deep tissue imaging) microscope with pulsed lasers, high time resolution detectors and special detection electronics. FLIM can be performed in either time or frequency domain. In frequency domain excitation and then the emission light are modulated at frequencies comparable to the lifetime is often measured in camera-based microscopes like widefield or TIRF.

FLIM used in this study is based on measurements in the time domain with the confocal scanning microscopy that in addition to forming a fluorescence intensity image also measures lifetime of fluorophores in each pixel. Lifetime measurement can be obtained using the time-correlated single photon counting (TCSPC) method [137, 138, 139]. TCSPC is known to deliver the best lifetime accuracy at very low photon efficiency [139, 140]. TCSPC uses the principles of photon counting statistics. Occurrence of a photon is a random process and for perfectly coherent light with a constant intensity can only be described in terms of Poisson statistics. In TCSPC the excitation laser is pulsed with a high repetition rate. Photons from the excited fluorophores are emitted and then detected with a photosensitive detector

such as an avalanche photodiode detector (APD). Data is accumulated until a sufficient number of events is recorded to extract a fluorescence decay curve. The process can also be described as a cycle of "stopwatch" events [141]. The "start" is an excitation pulse and the "stop" is the reading of the detector of photon arrival times. Thus, the stopwatch readings of photon statistics are sorted in a histogram of time bins, where the size of the bin corresponds to the time resolution of the system. The histogram can be fitted with a fluorescent decay curve that can be described by an exponential decay of counts toward later arrival times, similar to that of a radioactive decay of a certain decay constant or half-life. In fluorescence such a constant is the average time that a fluorophore spends in an excited state prior to returning to the ground state by emitting a photon - its lifetime [142].

Fluorescence lifetime is an absolute measurement which is less dependent of imaging artefacts arising from scattered light, photobleaching, intensity, light path length, or concentration, in contrast to a conventional intensity based confocal imaging. Fluorescence lifetime therefore can be used to measure changes in the environment around the dye. Factors such as temperature, pH, ion concentrations and other chemicals can all affect the lifetime of the chromophore [143]. This effect was also taken the advantage of as it FLIM was used to map changes in cellular pH [144], ions [145], oxygen [146], temperature [147, 148], viscosity [149, 150], glucose concentration [151] and other factors [152, 153]. Due to its multifunctionality, high sensitivity and use *in situ* FLIM is also a highly used technique in multiple biomedical applications like studying early stages of diseases like age-related eye macula degeneration [154] or cancer [155, 156]. Another common application of FLIM is to use the unique value of the lifetime to map different chromophores that can be spectrally indistinguishable [157]. One of the applications that is used in this thesis is to measure FRET using FLIM (FLIM-FRET).

1.4.3 Fluorescence lifetime imaging microscopy application to study FRET

One of the main advantages of using FLIM to quantify FRET in cells is that it measures changes in fluorescence lifetime that are independent of the variations of concentration of fluorophores across the sample. It is also internally calibrated meaning no additional spectral correction imaging such that for intensity based FRET imaging is required and FRET efficiencies and fraction of the donor species can be measured directly. The main advantage of FRET is that it provides an absolute measure of end-to-end distances between the two fluorophores. Quantifying changes in the lifetime upon FRET thus would lead to a very precise estimate of distances in biological molecules in their physiological environment.

Fluorescence lifetime of the donor changes when after absorbing a photon donor transfers energy to the acceptor. By definition the rate constant of energy transfer

$k_{RET}(R)$ in a dipole-dipole interaction between donor and acceptor is:

$$k_{RET}(R) = \frac{1}{\tau_D}\left(\frac{R_0}{R}\right)^6,\tag{1.5}$$

where R_0 is the Förster distance, at which FRET is 50% efficient, and τ_D is lifetime of the donor in absence of the acceptor and R is the distance between the fluorophores or end-to-end distance (R_E). The critical transfer distance R_0 is the distance at which FRET and fluorescence emission are equally likely. R_0 can be calculated from the spectral overlap integral J_F of the two dyes involved in FRET:

$$R_0^6 = \frac{9000(ln10)\kappa^2 Q_D}{128\pi^5 N_A n^4}J_F,\tag{1.6}$$

where Q_D is the quantum yield of the donor in the absence of acceptor, n is the refractive index of the medium, N_A is Avogadro's number, κ^2 is the orientation factor that can vary between 0 and 4. κ^2 is a factor that describes relative orientations of the donor to acceptor dipoles. When the dyes are freely rotating and can be considered isotropically oriented so that the orientation of their dipoles randomises prior to the energy transfer, κ^2 is $\frac{2}{3}$. Experimentally, from anisotropy measurements of unfolded proteins, also in cells, κ^2 was found to be equal to $\frac{2}{3}$ [158, 159, 143]. The rotational correlation times of dyes attached to IDPs are in the range of a few hundred picoseconds, while their fluorescence lifetimes are in the nanosecond range. In this regime the dye reorientation is fast enough for the relative orientations of the donor and acceptor dipoles to be averaged out [160]. $J_F(\lambda)$ describes the degree of spectral overlap between the donor emission ($F_D(\lambda$, normalised to unity)) and the acceptor absorption:

$$J_F(\lambda) = \int_0^\infty F_D(\lambda)\epsilon_A(\lambda)\lambda^4 d\lambda,\tag{1.7}$$

where ϵ_A is an extinction coefficient of the acceptor at wavelength λ (typically in units $M^{-1}cm^{-1}$).

From the definition of efficiency of energy transfer being the fraction of photons transferred from donor to acceptor, one can derive the equations for FRET efficiency as a ratio of transfer rate to the total fluorescence decay rate of the donor:

$$E = \frac{k_{RET}}{k_{RET} + \frac{1}{\tau_D}}\tag{1.8}$$

Combining this equation with the equation 1.5 for transfer rate gives the FRET equation in terms of distances:

$$E = \frac{R_0^6}{R_0^6 + R^6}\tag{1.9}$$

In the simplest case scenario when donor has a monoexponential decay in the absence and presence of acceptor and its quenching only occurs via FRET, FRET efficiency of such process can be expressed as:

$$E_{FRET} = 1 - \frac{\tau_{FRET}}{\tau_D}, \tag{1.10}$$

where τ_D is the donor lifetime in the absence of acceptor and τ_{FRET} is the lifetime of the quenched component of the donor lifetime upon energy transfer. τ_D is measured using a monoexponential decay fit to the intensity decay (I_D) with the equation:

$$I_D(t) = \exp(-\frac{t}{\tau_D})) \tag{1.11}$$

Biological systems exhibit one more level of complexity as monoexponential decays are rare in biology [142]. Therefore, more complex equations are needed to describe them. A multiexponential decay is sum of all exponential components (τ_{Dn}), multiplied by their preexponential factors A_n, which are normalised to unity ($\sum_{n=1}^{\infty} A_n = 1$):

$$I_D(t) = \sum_{n=1}^{\infty} A_n \exp(-\frac{t}{\tau_{Dn}}) \tag{1.12}$$

When energy transfer occurs the decay time of the donor is shorted or quenched and the decay curve is altered (I_{DA}). Implementing the transfer rate expression into the lifetime equation one obtains:

$$I_{DA}(R,t) = \sum_{n=1}^{\infty} A_n \exp(-\frac{t}{\tau_{Dn}} - k_{RET}(R)t), \tag{1.13}$$

where energy transfer rate $k_{RET}(R)$ is a function of R, as seen in the equation 1.5. Now including the equation 1.5, expression for the fluorescence decay in terms of time and distances simplifies to:

$$I_{DA}(R,t) = \sum_{n=1}^{\infty} A_n \exp(-\frac{t}{\tau_{Dn}} - \frac{t}{\tau_{Dn}}(\frac{R_0}{R})^6), \tag{1.14}$$

Experimentally labelling is almost never 100% accurate so the sample usually has some molecules that have been labelled with two donor dyes or even two acceptor dyes. Thus the best theoretically possible labelling with a 50-50 mixture of dyes would only contain 50% of the sample labelled to undergo FRET under perfect conditions, where no bleaching occurs. In the donor only labelling case the donor would not be quenched by the acceptor. Therefore, their fraction in the overall signal *d*

should also be taken into account [143]:

$$
\begin{aligned}
I_{FRET}(t) &= (1-d)I_{DA}(t) + dI_D(t) = \\
(1-d) &\sum_{n=1}^{\infty} A_n \exp(-\frac{t}{\tau_n}(1+(\frac{R_0}{R})^6)) + d\sum_{n=1}^{\infty} A_n \exp(-\frac{t}{\tau_n})
\end{aligned}
\tag{1.15}
$$

DNA can be approximated to be a rigid molecule with specified distances (static model). Therefore this equation would already well describe a molecule with defined end-to-end distances. However, in biology and especially in the field studying disordered proteins, IDPs by definition do not exist in a single, well-defined conformation. The use of FRET to study configurations of flexible molecules was pioneered by Haas and Steinberg [161, 162]. They proposed using the characteristic end-to-end distance distributions P(R) between the single donor and the single acceptor labelling sites in a flexible chain. Determination of such distributions has also been shown to be possible in time-resolved measurements using TCSPC. Implementing the end-to-end distance distribution into equation 1.15, one derives:

$$
I_{DA}(t) = \int_0^{\infty} P(R)I_{DA}(R,t)dR = \int_0^{\infty} P(R)\sum_{n=1}^{\infty} A_n \exp(-\frac{t}{\tau_{Dn}} - \frac{t}{\tau_{Dn}}(\frac{R_0}{R})^6)dR,
\tag{1.16}
$$

Therefore, the final equation for time-resolved FRET becomes:

$$
\begin{aligned}
I_{FRET}(t) &= (1-d)I_{DA}(t) + dI_D(t) = \\
(1-d) &\int_0^{\infty} P(R)\sum_{n=1}^{\infty} A_n \exp(-\frac{t}{\tau_{Dn}}(1+(\frac{R_0}{R})^6))dR + d\sum_{n=1}^{\infty} A_n \exp(-\frac{t}{\tau_{Dn}})
\end{aligned}
\tag{1.17}
$$

This is the equation that I will refer to on multiple points in the next sections as this is the main equation used in the analysis of lifetime decays of flexible proteins. This type of analysis assumes that energy transfer is the only mechanism of donor quenching, multiexponential donor rates of energy transfer scale according to the decay times of each of the decay components and makes sure that the probability distribution function P(R) is normalised by the constant $\frac{1}{\int_{R_{min}}^{R_{max}} P(R)dR}$ for $R_{min} < R < R_{max}$.

Generally, the dynamics of disordered states at the scale accessible to FRET, especially in IDPs, are much slower than the lifetimes of typically used dyes. Confocal experiments employing time correlating single photon counting (TCSPC) can be used to generate fluorescence intensity decays that can be analysed with different P(R) to identify the most suitable model. P(R) used for these equations is not arbitrary but usually contain parameters describing the physical state of the molecule and will be discussed in more details in the next section. P(R) is generally used as a Normal distribution or can be modified to introduce polymer chain statistics, for example, Gaussian chain model, worm-like chain model, self-avoiding walk model,

excluded volume chain and others. In cases of the segment length being much larger than the persistence length these models converge to provide similar results [163].

Experimentally lifetime measurements are not just the lifetime decays but the decays convoluted with an instrument response function (IRF). IRF describes the distribution of photon arrival times affected by the the excitation impulse and electronic and detection time responses of the instruments. It is essential for a correct lifetime analysis. The experimental decays can therefore be represented as:

$$I(t)_{total} = IRF \otimes I(t), \tag{1.18}$$

where I(t) is any exponential fluorescence decay coming from the sample. The usual FLIM-FRET analysis therefore is based on reconvolution fitting to fit the complete decay. Reconvolution fitting is typically based on nonlinear least squares fitting using a Jacobian matrix to reduce the computational complexity [164, 165].

1.5 Polymer scaling laws and end-to-end distance distributions

Polymers are long molecules made of multiple units called monomers, all linked together in a chain. In biology single chains of DNA, RNA and proteins are polymers. Biopolymers can be physically described in terms of two chain models: ideal and real. Ideal chain assumes monomers do not interact with each other, while real chain models interchange interactions between each other and the solute, which can be described in terms of excluded volume.

End-to-end vector is essentially a sum of all bond vectors in a chain, such that:

$$\vec{R}_n = \sum_{i=1}^{n} \vec{r}_i, \tag{1.19}$$

where $\vec{r}_i$ represents the bond vector between the monomers and n is the number of bonds. One of the simplest models of an ideal polymer is a freely joined chain model. In this model all bond vector have the same length $l = |\vec{r}_i|$ and no correlation between bond vectors direction. This model gave rise to the Kuhn length. Kuhn length is the effective bond length b of N freely-joined bonds, such that the contour length of the polymer (R_{max}) can be expressed as $R_{max} = bN$, thus mean end-to-end distance becomes

$$\langle R_E \rangle^2 = bR_{max} = Nb^2 \tag{1.20}$$

I will use the concept of Kuhn length to call all functional units (amino acids) in the proteins monomers or residues, such that a protein has N=N_{res} monomers of effective length b.

Every conformation of an ideal chain in solution can be described using statistical derivation of a random walk. Random walk essentially is essentially a path or distance travelled following a random distribution. According to the central limit theorem a random distribution of very large (infinite) steps tend to follow a Gaussian probability distribution function. Ideal chains undergoing a random walk can therefore be characterised by Gaussian relationship of their end-to-end distances [166]. In 1-dimension end -to-end distance distribution becomes:

$$P(R) = \frac{1}{\sqrt{2\pi\langle R_E\rangle^2}} exp(\frac{R^2}{2\langle R_E\rangle^2}) \tag{1.21}$$

One-dimensional case of the probability distribution is a specific case of a Normal distribution:

$$P(R) = \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{1}{2}(\frac{R - \langle R_E\rangle}{\sigma})^2 \tag{1.22}$$

Extending equation 1.21 to a 3-dimensional case, we derive what is called the Ideal Gaussian chain distribution:

$$P(R) = 4\pi R^2 (\frac{3}{2\pi\langle R_E\rangle^2})^{\frac{3}{2}} \exp -\frac{3}{2}\frac{R^2}{\langle R_E\rangle^2} \tag{1.23}$$

Both normal and Ideal Gaussian chain end-to-end distances distributions are widely used in lifetime analysis using equation 1.17. Normal distribution is used for end-to-end distances distribution approximation in time-resolved FRET and FLIM analysis for different structured and unstructured proteins [167, 168, 169], while Ideal Gaussian distribution has been applied to study intrinsically disordered proteins and even FG-Nups [170].

Flory theory applies concepts from statistical physics to mathematically describe polymers. It states that dimensionality of a polymer is independent of fine tuning of parameters such as temperature, pressure and other external fields but in all cases follows power laws [171, 166].According to the Flory theory polymer size R scales in a power law relationship with the number of monomers in a polymer chain N_{res} [172, 171, 166]:

$$R \propto N_{res}^\nu, \tag{1.24}$$

where ν is Flory scaling exponent. Flory theory of an ideal chain assumes that monomers are uniformly distributed within the volume R_E^3, have no interaction between them and undergo a three-dimensional random walk. Since ideal chains obey Gaussian statistics, Flory exponent of the ideal chain is equal to $\frac{1}{2}$. Real chains, however, do have interactions between monomers and can be described using excluded volume principles taking into account their interaction with the solvent. The interaction of a monomer essentially depends on the difference in its interaction with another monomer and the surrounding molecules. By this interpretation when

monomers do not interact with each other in the ideal chain there is a perfect balance between chain–chain and chain–solvent interactions. At this point the thermodynamic phase boundaries disappear. This state is referred to as θ - state ($\nu = \frac{1}{2}$), where θ is the critical point for polymers. When a polymer is in a good solvent, interactions between the monomers and solvent molecules are energetically more favourable, therefore the polymer appears more expanded or swollen ($\nu = \frac{3}{5}$). In a bad solvent monomers prefer to stick to each other, forming a globule, rather than interacting with the solvent ($\nu = \frac{1}{3}$), usually in a very concentrated solution [171].

Flory relationship for a chain dimension provides a universal power-relationship that is a reasonable approximation to the real chains and is often used in both polymer simulations and to describe the biological data to a good agreement [173, 174, 175, 176, 177, 170]. Nevertheless, even in *in vitro* studies of conformations of IDPs using scaling laws there are conflicting evidence reported between the values of scaling exponents of proteins measured in native conditions. Some studies show disordered proteins having a scaling exponent between 0.46 and 0.51 under physiological conditions in solution, being close to the θ - state [175, 170]. However, these values may be slightly underestimated due to the dye contribution to IDP collapse [178], which is still a controversial issue. Inspired by success of a simple theory such as scaling law in describing complex properties of IDPs *in vitro* I decided to use Flory scaling law to apply to my measurements of end-to-end distances *in situ*.

1.6 Objectives

The conformational state of FG-Nups in the NPC has been a fundamental question in the scientific community for decades. Due to technical difficulties working at nanometer range with highly dynamic and flexible disordered proteins at very high concentrations, no reliable method exists to address such a complex system in its natural environment. My main goal is to develop an experimental strategy to visualise conformational plasticity of FG-Nups inside the central channel of the NPC *in situ* using a combination of site-specific labelling and high-resolution fluorescence microscopy approaches. Specifically, I will develop two complementary imaging procedures focused on two major players in the NPC's permeability barrier, Nup62 and Nup98. I aim to quantitatively assess and compare dimensions of the FG-Nups inside cells using a polymer physics approach. My overarching goal is to describe the conformational plasticity of FG-Nups in the NPC and its relevance to the nuclear transport function and to generalise the application of the developed method onto other disordered proteins.

1.6.1 Extending on the site-specific super-resolution imaging approach to map FG-Nups conformation

My first aim is to use a combination of genetic code expansion and click-labelling to site-specifically label and image FG-Nups *in situ* in mammalian cells. I will use and develop a dual colour dSTORM approach to image different Amber sites of one of the essential components of the central channel Nups - Nup62.

1.6.2 FLIM-FRET approach to visualise the spatial organisation of FG-Nups in mammalian cells

I will advance genetic code expansion and site-specific labelling to two explicit sites in the same protein for FRET measurements with residue precision. I will then establish FLIM-FRET imaging in cells using dual colour labelling at systematically increasing distances between two sites. This method will allow me to resolve native FG-Nups conformations up to 10nm scale.

1.6.3 Polymer physics perspective to solve plasticity of FG-Nups in the NPC

From a polymer physics perspective, scaling dimensions of a single polymer follow a power-like relationship with the number of amino acids. My third aim is extract the dimensions of the FG-Nups directly in the NPC using the power-law relationship between the measured end-to-end distances and number of residues. The method that I will develop will eventually not just help solve the mystery of the FG-Nup permeability barrier functional state but can be applied to study dimensions and the functionality of different intrinsically disordered proteins *in situ*.

2 Materials and Methods

2.1 Confocal imaging to test expression and labelling of FG-Nups in mammalian cells

Confocal imaging presented in this study was done using Leica SP8 STED and Olympus FV3000 microscopes available at EMBL Advanced Light Microscopy facility. Leica SP8 microscope is equipped with with HCX PL APO 100x/1.4 oil objective (Leica, Germany), white light laser line and allows for gated detection on two hybrid HyD detectors. Olympus FV3000 is equipped with UPLXAPO60XO 60x/1.42 oil objective and four 4 Gallium arsenide phosphide GaAsP spectral detectors and OBIS lasers: 405, 488, 514, 561, 594, 640 nm. Microscopes were used to check expression and labelling of constructs mentioned in this study.

2.2 Dual colour super-resolution imaging

2.2.1 Cells preparation for super-resolution imaging

Culture media:

- Dulbecco's Modified Eagle Medium (DMEM, Life Technologies, cat. no. 41965-039) supplemented with 10% v/v fetal bovine serum (FBS, Sigma, cat. no. F7524), 1% Pen-Strep (Sigma, cat. no. P0781), 1%L-glutamine (Sigma, cat. no. 158127) and 1% sodium pyruvate (Life Technologies, cat. no. 11360).

High salt buffer(10xTB):

- 200mM HEPES, 1.1M potassium acetate (KOAc), 50mM sodium acetate (NaOAc), 20mM magnesium acetate (MgOAc), 10mM ethylene glycol-bis (β-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA), pH 7.3

HEK293T cells were cultured at 37°C in 5% CO_2 atmosphere in the culture media and passaged every 2-3 days up to maximum of 15-18 passages. For imaging cells were seeded in a 24-well plate at 110,000 cells/well with the total volume of 500μl.The next day (in 16-20 hours) cells were transfected with 1.0μg/ml of DNA in total (0.5μg/ml of a plasmid with POI containing an Amber codon and 0.5μg/ml of the plasmid with PylRS/tRNA (pcDNA3.1-NESPylRSAF-U6Full(RV) (cat. no 1533)). Transfections were done with jetPRIME reagent (Polyplus-transfection, cat. no. 114-15) according to manufacturer's protocol. Media was changed after 4-6 hours and 250μM of TCO* was added at 1:4 in 1M HEPES.The next day, approximately 24

hours after transfection, cells were washed with warm PBS, trypsinised (100μl Trypsin per well), resuspended in 800μl culture medium and re-seeded (250μl per well) at a density of $\sim 100,000$ cells/well in a prior to that poly-l-lysine (PLL, Sigma, cat. no. P5899) coated 8-well Lab-Tek II chambered coverglass (Thermo Scientific Nunc, 155411). Coating was done at least 4 hours in advance by incubating Lab-Teks with a 1:100 dilution of stock solution (1 mg/ml) in ddH2O (coated Lab-Teks were afterwards stored at 4°C for up to 2 months).After 4-5 hours, once the cells have settled,they were taken out of the incubator.

labelling was done at room temperature in several steps. Cells were:

1. once washed in 1xPBS,
2. permeabilized with digitonin (1:500 dilution of 20 mg/ml stock in dimethyl sulfoxide (DMSO)) in 10xTB for 3-5 min,
3. once washed in 10xTB,
4. incubated with 1.5μM of dye solution (here Alexa647-H-tet) for 10 min at room temperature,
5. washed twice in PBS,
6. fixed with 2% PFA for 10 min,
7. washed twice in PBS,
8. labelled with WGA-CF680 at the final concentration of 1μg/ml for 10 min,
9. finally washed twice with PBS and either imaged immediately or stored up to 2 days at 4°C.

All the UAAs mentioned in this thesis were purchased from SiChem GmbH. SCO (strained cyclooctyne lysine), TCO* (*trans*-cyclooct-2-ene-lysine) first isomer and endoBCN (here, BCN endo-bicyclo [6.1.0] nonyne - lysine). Stock solutions of all UAAs (100 mM) were prepared by dissolving them in 15% (vol/vol) DMSO/0.2 M NaOH.

2.2.2   Experimental TIRF setup

The microscope was designed to perform both in epi- and TIRF illumination modes. For dSTORM imaging of nuclear enveloped we have used TIRF illumination (about 100nm thick TIRF field) since it helps to reduce fluorescent background of out-of-focus light and to improve signal-to-noise ratio. Details of all optical parts described further can be found in the Materials list in section 2.3.

The TIRF microscope was equipped with four lasers (detailed parameters can be found in the table in section 2.4): Omicron 405 nm, Omicron 488nm, Sapphire 568nm and Omicron 660nm, each supplied with clean–up filters: 405/10 (405), 475/23 (488), 572/15(568) and 661/11(660). In the excitation path the lasers were expanded with 10x beam expanders (Qioptiq) for the 660, 594 and 488 lasers and a 7x beam expander (Qioptiq) for the 405 laser. The lasers were then combined using dichroic mirrors: 458 for the 405nm laser, 495 for the 488 laser, 605 for the 568nm laser.

The lasers polarisation was set to circular to avoid any polarisation effects using a quarter-wave plate (G362021491,Qioptiq). To achieve TIRF a mirror mounted on a linear stage (PI, M505) was used to shift the beam with respect to the the back-pupil of a UAPON 100X TIRF, NA 1.49 objective (Olympus). The beam was focused to the objective by a 300 mm focal distance lens (Thorlabs). Then the excitation light was separated from the emission light by a quad-edge dichroic beamsplitter (Semrock). In the emission part there were three consecutive lenses (two 180 mm and one 120 mm focal length) installed and optimised to minimise aberrations. Notch filters (568/647 and 658) were used to filter out any residual excitation light. An AHF 690 dichroic mirror then split the emission light onto two EMCCD cameras (Andor Technology, DU-897 and DU-897D, iXon X3 EMCCD), equipped with ET 700/100 emission filters for dual colour super-resolution imaging. To check expression in the GFP channel emission filters were switched such that, ET 525/50 filter and 488 laser excitation allowed for GFP imaging, while ET 700/75 filter and 660 laser in the same configuration allowed for the labelling channel imaging (Alexa647 and/or CF680). This configuration was used to find appropriate cells for dual colour dSTORM (subsection 2.1.3)

Since super-resolution imaging acquisition takes a long time and uses high power of excitation lasers sample is in a constant risk of drift, which might significantly impair the final results. Therefore, a focus stabilisation system was installed. The system operates by an electronic feedback loop that helps to constantly monitor the position of the objective with respect to the bottom of the Lab-Tek and move a piezo-driven stage accordingly to compensate for the drift in z-focus. It operates by following a displacement of a totally reflected beam of a near-infra red laser (Coherent Inc., cat. no. LuxX 905), focused at the side of the back-pupil of the objective to the bottom of the Lab-Tek. Any z-displacement of the Lab-Tek then resulted in the lateral displacement of the totally reflected beam on the photodiode (Thorlabs Inc., cat. no. PDP90A), triggering the feedback loop.

All the components of the microscope were controlled by a custom-written LabView (National Instrument) software.

2.2.3 Dual colour dSTORM

Composition of GLOX-MEA blinking buffer:

- 50mM Tris (pH 8.0), 10mM sodium chloride (NaCl), 0.5mg/ml glucose oxidase, 40g/ml catalase , 10 %(w/v) glucose and 10mM Mercaptoethylamine (MEA) (buffer was prepared freshly before each measurement).

One of the very well established methods to use dSTORM for multicolour imaging of the dye pair is to apply spectral demixing[179]. This method involves using one laser to excite two spectrally close fluorophores. The choice of fluorophores should also take into account their brightness and similar duty cycles in blinking buffer. The

emissions are then split by a dichroic mirror onto two EMCCDs cameras. Each of the cameras will record all the blinking events, which can then be computationally split according to their relative brightness. We used Alexa647 and CF680 fluorophores that in our experimental setup give 45/55 and 30/70 ratios of brightness between the cameras respectively[122].

dSTORM imaging was done in blinking buffer, which was refreshed every 60-90 minutes of imaging. I selected cells based on the GFP signal, taking into account its brightness as well as the shape of the nuclear footprint. Imaging was started with bringing the fluorophores onto the dark state by using the maximum power of the red laser at 660nm ($\sim 1kW/cm^2$) until blinking is observed (10-30s). Then the laser power was reduced to about a half ($\sim 0.5kW/cm^2$) and the final movie was acquired for $\sim$30,000 frames with the exposure time of 30ms. Increasing the intensity pumping with the 405nm laser was then started at around 10,000-15,000 frames to keep the number of blinking events per frame constant (50-100). Resulting images were reconstructed using Igor Pro 7 software with the Localizer package [180] and demixed with a custom written code by Dr Giulia Paci. Resolution was estimated using the FRC method and a custom-written code in Igor Pro 7 software [112].

2.2.4   GFP-nanobody: expression, purification and labelling

Plasmid pETM20-Trx-6His-TEV-GFP(Q14C) (cat no. 1709) encoding a cysteine-modified GFP-nanobody was expressed in *E.coli* BL21(DE3)AI cells in terrific broth medium (TB). Expression was induced with 1mM IPTG and 0.02% arabinose at O.D.=0.6 and left at 25°C for 16 hours shacking. Cells were harvested by centrifugation in a Beckmann centrifuge, rotor JLA 8.100 at 4000g for 20 minutes. The bacterial pellet was resuspended in lysis buffer (20 mM HEPES, 250mM NaCl, 20mM imidazole (Sigma (I5513)), 1mM phenylmethylsulfonylfluorid (PMSF) (Biochemica A0999.0025), 0.2mM Tris(2-carboxyethyl)-Phosphine (TCEP) pH 7.5) and sonicated. The lysate was centrifuged in SS34 tubes in a Beckmann centrifuge (rotor JA 25.50) at 4°C, 18,000g for 1 hour. The clear lysate was then incubated with Ni–beads (1 ml/l expression) for 1-2 hours at 4°C under gentle rotation. Using polypropylene (PP) columns Ni–beads were separated from the flow through (FT). Then Ni–beads were washed three times with lysis buffer and the bound protein was eluted with 5 ml of elution buffer (20mM HEPES, 250mM NaCl, 0.2mM TCEP, 300mM Imidazole, pH7.5). Eluted protein was first dialised against lysis buffer (at least 8 hours of 1l per 1l of expression) and then incubated for 16 hours at room temperature with tobacco etch virus (TEV) protease (1ml of 1mg/ml TEV per 1l of expression) slowly shaking or rotating. TEV efficienctly cleaves off Txr-6His tag from the nanobody, yet it also has a His tag so bot the cleaved Trx and TEV can be separated from the nanobody using Ni-beads. Cleaved protein solution was again incubated with Ni-beads for 2 hours and then separated into flow through, washes and elution with

the elution buffer were collected. FT and the first wash were combined and concentrated in a centrifugal concentrator filter (AmiconUltra, Millipore) with 5 kDa cutoff. Finally, the sample was purified by gel–filtration chromatography using a Superdex 200 (S200) column. For long–term storage at -20°C, the protein was flash frozen in liquid nitrogen after adding 25% glycerol. Otherwise, the protein was directly labelled using maleimide chemistry. Protein was dissolved in a lysis buffer (usually 1 reaction was set for 25nmoles of protein) with 10mM DTT to reduce disulfide bonds. Protein then was washed in lysis buffer at least 5 times using an amicon concentrator. 5-10x excess of maleimide-conjugated dye was dissolved in DMSO and slowly added while vortexing the protein. The reaction was left to complete at room temperature for at least 2 hours in the dark or overnight at 4°C. After completion the reaction was quenched with 10mM DTT. labelled protein was washed at least 5 times using an amicon concentrator and then loaded on a Superdex 200 (S200) column. Fractions were collected, concentrated flash frozen in liquid nitrogen after adding 25% glycerol and stored at -20°C.

2.2.5 Immunostaining

COS-7 cells were cultured in the same way as HEK293T as described in subsection 2.2.1. COS-7 cells were seeded directly into a 4-well Lab-Tek II chambered coverglass (Thermo Scientific Nunc, 155383) at 50,000 cells/well with the total volume of 500μl. The next day after seeding (in 24 hours) cells were prepared for immunostaining. Cells were washed once in PBS, then fixed with 2% PFA for 10min and then permeabilised with 0.5% Triton-x 100 (AppliChem, A1388.0500)in PBS for 10-15min at room temperature. Cells were then washed 2-3 times with PBS and then blocked with 5% BSA (Sigma Aldrich, A7030) in PBS for 60-90 min at room temperature. The blocking solution was removed and the cells were incubated overnight at 4∘C with the primary antibody. The next morning, cells were rinsed and incubated with the secondary antibody for 1 hour at room temperature. To check intactness of the NPC after high salt buffer cells were prepared according to the labelling protocol in subsection 2.2.1, except no dye was added in step 4. After step 9, cells were prepared the same way as the control.

2.3 FLIM imaging

2.3.1 Cell preparation for FLIM

COS-7 cells were cultured in the same way as HEK293T as described in subsection 2.2.1. For imaging, COS-7 cells were seeded directly into a prior to that poly-l-lysine coated 4-well Lab-Tek II chambered coverglass at 35,000 cells/well with the total volume of 500μl. The next day after seeding (in 16-20 hours) cells were transfected with 1.2μg/ml of DNA in total (0.4μg/ml of a plasmid with POI, 0.4μg/ml of a plasmid with the PKR (pIEX-PKR(1-170) (cat no. 1702)) and 0.4μg/ml of the plasmid with

PylRS/tRNA). PKR is protein kinase R, an interferon inducible serine/threonine protein, which plays multiple roles in apoptosis, immune responce to viral infections, cell proliferation and differentiation[181]. It is also know that PKR is a major regulator of the eukaryotic translation initiation factor eIF2α, responsible for the initiation phase of mRNA translation [181]. In context of Amber suppression we have observed that PKR helps to significantly increase the levels of Amber-encoded protein expression (data not shown). Media was changed after 4-6hours and 250μM of BCN was added at 1:4 in 1M HEPES. Buffer was changed again 24 hours after transfection. labelling was started after 48hours.

Labelling was done in the same way as in subsection 2.2.1 except step 4 and eliminating step 8. Instead, in step 4, a 50:50 mixture of 1.5μM Atto532-H-tet (Donor) and 1.5μM LD650-H-tet (Acceptor) in 10xTB was left for 15min at room temperature for samples used for FRET measurements or only 1.5μM of one of the dyes for samples used for Donor only or Acceptor only controls.

2.3.2   Experimental FLIM setup
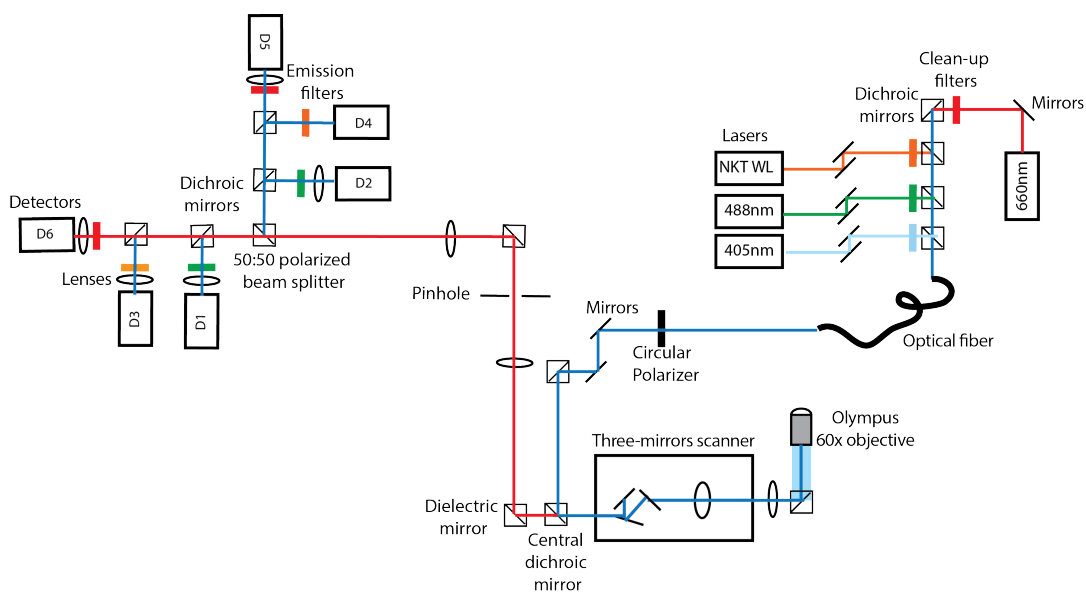
FIGURE 2.1: Schematic drawing of the scanning confocal FLIM setup. On the top right, the lasers are combined to a single beam in the excitation path. It then travels through the optical fibre and after the polariser and the central dichroic mirror goes through the scanner and reaches the objective. The emitted light travels back the same way until it passes through the central dichroic mirror and gets split 50/50 by a beam splitter onto two identical detection paths. Light is finally filtered and focused on the detectors.

The setup is based on a regular scanning confocal microscope with a custom-designed and built three-mirrors scanner and an additional Time-Correlated Single Photon Counting (TCSPC) modules implemented to measure fluorescent lifetime at high precision. Details of all optical parts described further can be found in the Materials list in section 2.3.

The schematic drawing of the microscope can be seen in Figure2.1. The excitation part of the microscope consists of 4 pulse lasers (detailed parameters can be found in the table in section 2.4): Omicron 660nm, a white light laser from NKT Photonics that we normally operate at 532nm or 561nm at repetition rate of 40MHz, Omicron 488nm and Omicron 405nm. Pulsed interleaved excitation was used to avoid any cross-talk between the lasers. All lasers have a corresponding clean-up filter (661/11, 572/15 or 525/20, 488/10 and 405/10 respectively) and are then combined by dichroic mirrors (605, 488 and 455 respectively) and coupled into an optical fibre (Qioptiq).

From the optical fibre the light is split by a pellicle beamsplitter (Thorlabs) such that 8% is used for a constant power monitoring with a power meter (FiledMax II, Coherent) and 92% get circularly polarised with a quarter-wave plate (G362021491,Qioptiq) before being reflected with a central multi-band dichroic mirror (ZT405/488/532/640 in 532nm configuration and ZT405/488/561/660/905 in 561 configuration) to go through the FLIMbee (PicoQuant) controlled mirror scanner onto the objective (Olympus 60x UPlanSApo NA1.2). Sample gets excited by the confocal focused spot, and the light travels back through the same optical path: the scanner and the dichroic mirror onto a custom-designed piezo-controlled pinhole plate. After the pinhole light travels through the polarising beamsplitter (Thorlabs) and gets split equally into two identical emission paths. Ideally each of the emission paths would be used for detection of parallel and perpendicular polarised light when I linearly polarise it after the fibre. Such configuration would allow us to additionally measure anisotropy. Since I didn't want to over-complicate the experimental system, I minimised all polarisation effects by keeping light circularly polarised. The two emission paths are equiped with three detectors: for red, for orange and for green. Emission light gets split onto the detectors by dichroic mirrors (708/75 for red, for orange: 579/59 in 532 configuration or 609/57 in 561 and 579/59 for green). Additionally each of the detectors has an emission filter in front: 708/75, 579/59 or 612/65 and 579/59 respectively.

Pinhole gets positioned with the nanometer precision using Precision Tool Commander 2 software such that the signal is maximum intensity and the point spread function (PSF) is of appropriate quality. Pinhole is aligned every day before each experiment. To do so a Lab-Tek containing a dilute solution of a fluorescent dye (typically fluorescein) is placed on the microscope. Confocal spot focus in the solution is found by mechanically switching a mirror to reflect a back reflection plane of the objective to the camera (Imagesource). Using the next mechanically switchable mirror the light passing through the pinhole plate is reflected onto the PMT detector (Hamamatsu). A sensitive piezo-motor operated by the software automatically find the best position for the pinhole according to above mentioned criteria.

In addition to optical components the microscope is also equipped with additional electronic components: a galvo scanner FLIMbee(PicoQuant) and a Multichannel Picosecond Event Timer TCSPC Module HydraHarp 400(PicoQuant). Electronic components serve as temporal controllers and triggers to all essential devices like lasers and detectors. All components were synchronised in time in such a way that there are no delays between the identical detectors on the two arms receiving the signal and that the detectors used for donor/FRET channel lifetime data acquisition (MPD PDM for DNA measurements and PMA Hybrid40 for cells) are detecting all the photons at the 40MHz frequency and there is no interference lifetime signal coming from the acceptor lifetime. All was done manually using PicoQuant SymPhoTime 64 software. In such configuration the set up was used for all of the subsequent FRET and FLIM-FRET measurements.

2.3.3   Tetrazine dye reactions with UAAs for lifetime measurements

Here I have used Atto532 dye functionalised with an azide(Atto-tec) and H-tetrazine(Jena Bioscience), Tetrazine-Amine HCl salt (Jena Biosciences). Functionalised dyes and tetrazine were used as reference measurements. Highly concentrated reactions were set up to quickly and completely react tetrazine group with TCO*, BCN and SCO unnatural amino acids (1mM dye with 1.2mM TCO*,BCN or SCO in 10μl in PBS) for 15 minutes at room temperature. Reaction product was then measured at 1μM in PBS in an 8-well LabTek up to 10,000 counts. IRF was measured on the same day using 1μM quenched with KI Atto532-azide solution. Data was analysed using SimPhoTime software with mono and double-exponential decay fits.

2.3.4   DNA titration

Single stranded DNA oligos with donor and acceptor dyes respectively (see the full list in section 2.4) were hybridised using polymerase chain reaction (PCR). The PCR reaction mix contained annealing buffer (10 mM Tris, pH 7.5–8.0, 50 mM NaCl, 1 mM EDTA) and 10μM of each of DNA strands (100μM stock) in 50μl of reaction volume. The reaction underwent 3 cycles: 95°C for 2 minutes; ramp cool to 25°C over a period of 45 minutes; storage at 4°C. Out of the PCR machine samples were ready to use and were stored for longer times at -20°C.

FLIM measurements were done using 561 configuration (see subsection 2.2.1). For Donor and FRET samples FLIM was recorded using 488nm laser excitation/MPD PDM detector up to 10,000 counts. Acceptor channel was recorded using 561nm laser excitation and PMA Hybrid40 detector. IRF was recorded by measuring 200μl of saturated solution of KI with fluorescein in PBS, up to 10,000 counts. The concentration of the DNA was 10nM dsDNA in 200μl of PBS. Laser power of 488nm was 50% - 0.84uW, dwell time 2.5ns.

2.3.5 Dimerisation activated FRET sensor

HEK293T cells were cultured the same way as in subsection 2.1.1. Cells were seeded in onto 8-well Lab-Tek II chambered cover glass at 110,000 cells/well, total volume 250μl. The next day (in 16-20 hours) cells were transfected with 0.5-1μg/ml of DNA in total (0.5μg/ml for controls and 1.0μg/ml for FRET samples). Transfections were done with jetPRIME reagent (Polyplus-transfection, cat. no. 114-15) according to manufacturer's protocol. Media was changed after 4-6 hours and 100nM rapamycin (100μM stock in DMSO) was added. In about 24 hours after adding rapamycin cells were taken out of the incubator, washed with PBS and fixed with 2% PFA (steps 1,6,7 in subsection 2.1.1). Cells were then imaged on the FLIM set up using 561 configuration.

Dimerisation FRET sensor plasmids:

Plasmid	Internal no.
pAAV-FKBP-SCF2-LD6-mCherry	2223
pAAV-FRB-SCF2-LD6-EGFP	2225

For intensity FRET corrections separate images of Donor only and Acceptor only controls were taken at both 488nm and 561nm laser excitations. Images were recorded on all detectors simultaneously, final images were summed from identical arms, rescaled to the maximum number of counts and exported as .bmp file from SymPhoTime 64 software. Using ImageJ software and FRET and Colocalisation Analyser plugin 8-bit control images were used to calculate the FRET correction factors α and β in the equation 1.3. The average values were taken to apply to correct for resulting FRET images and for E$_{FRET}$ calculations. Resulting FRET images were calculated using Image Calculator tool in ImageJ.

SymPhoTime 64 software (PicoQuant) was used to fit resulting lifetime curves with mono-exponential (Donor only) and double-exponential (FRET channel) decays to calculate lifetimes. Equation 1.10 was used to quantify E_{FRET} and directly compare to the results from intensity measurements.

2.3.6 FLIM-FRET imaging of FG-Nups *in situ*

To establish best experimental settings I have acquired images from one of the FRET samples (Nup62-119-181), labelled with Atto532-LD650 (H-tet) in 532nm configuration (see subsection 2.2.1). I have varied the image size and the laser power, keeping dwell time and pixel size the same. Both Donor and Acceptor channels were recorded simultaneously. Final images were summed from the detectors. Individual frames (1-5) and (95-100) were summed up respectively and exported as .tiff files. Resulting images were opened in ImageJ software, where I compared their relative total intensities and intensities in manually selected rims. Settings that resulted in less than 20% drop in intensity, both total and in the rim in both Donor and Acceptor

channels at 532nm excitation, were taken for subsequent measurements of all mutants (Table 3.2).

Image setting are summarised in table 3.2. Using those settings I acquired all FLIM-FRET images, shown in the Appendix. I have also acquired the Hoechst and the acceptor channels to be used in automatic segmentation that is under development and for reference of successful dual-colour labelling in the latter case.

2.3.7   FLIM-FRET data analysis and simulations

The general description for the choice of analysis software used is presented in section 3.2.4. Here I show more details of the analysis pipeline and a scheme for a better visualisation 2.2. All images were segmented using drawing tools available in SymPhoTime since the software does not allow for import of ROIs from any other sources. Loading imaging files directly into PAM is possible yet time-consuming as generated files are of a large size. The solution decided upon was to directly load the fluorescence decay curves from SimPhoTime into TauFit component of PAM for reconvolution fitting. Reconvolution fitting in TauFit allows to optimise non-linear least-squares fit for background, scattering (experimental) and IRF shift to find minimal χ^2. Global fitting was used in SimPhoTime to determine lifetime and relative fraction of the two exponential components of the donor fluorescence decay to be fixed in the final fit. In the future I plan to implement automatic image segmentation and global fitting into PAM for a more robust, potentially, automated analysis of all data files. The final changes to PAM software will be shared with the scientific community on GitHub.
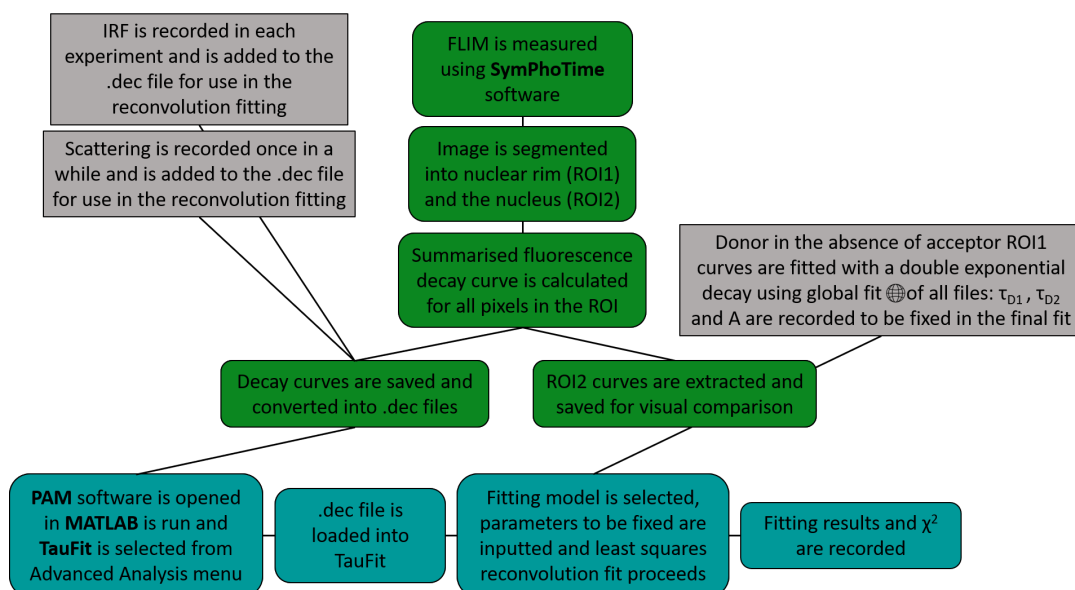
FIGURE 2.2: Data analysis pipeline used in this study. Two software SymPhoTime and PAM were used for different purposes. SymPhoTime for data acquisition, image segmentation and parameter determination and PAM for the reconvolution fitting with end-to-end distances model P(R).

2.4 Materials used in this study

List of plasmids for mammalian expression of FRET mutants:
All Amber STOP codons TAG were introduced into the DNA sequence by site-specific quick change mutagenesis using PCR

Plasmid	Internal no.	N_{res}
pCI-Nup62(119TAG)-mCerulean3	1791	donor
pCI-Nup62(119TAG-153TAG)-mCerulean3	2695	34
pCI-Nup62(119TAG-161TAG)-mCerulean3	2696	42
pCI-Nup62(119TAG-181TAG)-mCerulean3	2698	62
pCI-Nup62(119TAG-193TAG)-mCerulean3	2702	74
pCI-Nup62(119TAG-207TAG)-mCerulean3	2703	88
pCI-Nup62(119TAG-226TAG)-mCerulean3	2705	107
pCI-Nup62(119TAG-231TAG)-mCerulean3	2700	112
pCI-Nup62(119TAG-245TAG)-mCerulean3	2704	126
pCI-Nup62(119TAG-257TAG)-mCerulean3	2701	138
pCI-Nup62(119TAG-292TAG)-mCerulean3	2707	173

Plasmid	Internal no.	N_{res}
pCI-Nup62(13TAG)-mCerulean3	2686	donor
pCI-Nup62(13TAG-61TAG)-mCerulean3	2688	48
pCI-Nup62(13TAG-71TAG)-mCerulean3	2690	58
pCI-Nup62(13TAG-81TAG)-mCerulean3	2691	68
pCI-Nup62(13TAG-119TAG)-mCerulean3	2689	106
pCI-Nup62(13TAG-181TAG)-mCerulean3	2693	88
pCI-Nup62(13TAG-231TAG)-mCerulean3	2692	107

Plasmid	Internal no.	N_{res}
pCI-mCerulean3-Nup98(38TAG)	2709	donor
pCI-mCerulean3-Nup98(38TAG-101TAG)	2710	63
pCI-mCerulean3-Nup98(38TAG-138TAG)	2711	100
pCI-mCerulean3-Nup98(38TAG-153TAG)	2712	115

List of synthetic dyes:

Name	Reactivity	Company / cat.no.
Atto 532	H-tetrazine	Jena Bioscience CLK-011-02
Alexa Fluor 647	H-tetrazine	Invitrogen A20347
LD 650	H-tetrazine	Lumidyne Technologies
SiR	H-tetrazine	Spirochrome SC 008
TAMRA	H-tetrazine	Jena Bioscience CLK-017-05
Fluorescein	-	Sigma 46960-25G-F

List of used antibodies and lectins:

Name and Type	Target	Produced in	Modification	Company /cat.no.	Dilution
mAB414 primary	FG-Nups	mouse	-	BioLegend, 902901	1:2000
anti-Mouse IgG secondary	mouse	goat	Alexa647	ThermoFisher, A32728	1:1000
WGA lectin	O-GlcNAc	-	CF680	Biotium, BTIU29029-1	1:1000

List of DNA oligos:

Name	Dye modification	Sequence
Donor strand	Alexa488	*TAAATCTAAAATGTAACTAAGTAACATAACGGTAAGTCCA
Acceptor strand 1	Alexa594	TGGACTTACCGTATGTTACTTAGTTACATTT*TAGATTTA
Acceptor strand 2	Alexa594	TGGACTTACCGTTATGTTACTTAGT*TACATTTTAGATTTA
Acceptor strand 3	Alexa594	TGGACTTACCGTTATGTTACT*TAGTTACATTTTAGATTTA

symbolises position of the C6 linker to the dye

List of frequently used reagents:

Name	Company/cat.no.
EDTA	Ambion, AM9261
glucose oxidase	Sigma, G0543
glucose	Sigma, D9559
catalase	Sigma, C3155
MEA	Sigma, M9768
KI	Sigma, 60399-100G-F
KOAc	Merck, 1.05819.0250
NaOAc	Merck, 1.06268.1000
MgOAc	Sigma-Aldrich, M2545
PFA	Sigma, 158127

List of TIRF microscope components:

Lasers:

Name	Wavelength	Company
Omicron LuxX 405-60	405 nm	Omicron
Omicron LuxX 488-200	488 nm	Omicron
Sapphire 568-100	568 nm	Coherent Inc.
Omicron LuxX 660	660 nm	Omicron
Coherent LuxX 905	905nm	Coherent Inc.

Filters:

Type	Name	Company / cat.no.
Clean–up	Brightline HC 475/23	Semrock FF01-475/23
Clean–up	Brightline HC 572/15	Semrock FF01-572/15
Clean–up	Brightline HC 661/11	Semrock FF01-661/11
Emission	525/50	Semrock FF01-525250-25
Emission	700/75	Chroma ET 700/75 M
Emission	700/100	AHF ET700/100
Notch	StopLine 568/647	Semrock NF01-568/647
Notch	Single notch 658	Semrock NF03-658E

Dichroic mirrors:

Name	Company / cat.no.
Single Band Dichroic Filter 458	Semrock FF458-Di02
Single Band Dichroic Filter 495	Semrock FF495-Di03
Single Band Dichroic Filter 605	Semrock FF605-Di02
Multi Band dichroic	Chroma ZT405-488/568/660 RPC
Single Band Dichroic Filter 643	Chroma F43-T08 643 RDCXT
Single Band Dichroic Filter 690	AHF 690 DCXR

Lenses:

Focal length	Company / cat.no.
300 mm	Thorlabs AC508-300-A-ML
180 mm	Linos G322-246-000
120 mm	Linos G322-388-000

List of FLIM microscope components:

Lasers:

Name	Wavelength	Company
Omicron QuixX 405-120	405nm	Omicron
Omicron BrixX 488-200PS	488nm	Omicron
NKT SuperK VARIA	whitelight	NKT Photonics
Omicron QuixX 660-130PS	660 nm	Omicron

Filters:

Type	Name	Company / cat.no.	Quantity
Clean–up	Brightline HC 488/10	Semrock F39-489	1
Clean–up	Chroma ET 525/20 Bandpass (532nm laser)	Chroma ET F47-528	1
Clean–up	Brightline 572/15 (561nm laser)	Semrock FF01-572/15	1
Clean–up	Chroma ET 405/10 Bandpass	Chroma ET F49-401	1
Clean–up	Brightline HC 661/11	Semrock F37-661	1
Emission	BrightLine 520/50	Semrock F37-516	2
Emission	Brightline HC 575/59 (532nm config.)	Semrock F39-075	2
Emission	Brightline 612/69 (561nm config.)	AHF 612/69 HC	2
Emission	Brightline HC 708/75	Semrock F39-708	2
Emission	Chroma HQ 530/50m	Chroma HQ530/50m	1

Dichroic mirrors:

Name	Company / cat.no.	Quantity
Single Band Dichroic Filter 455	Zeiss FT455	1
Single Band Dichroic Filter 488	Semrock Di 01 R488 25x36	1
Single Band Dichroic Filter 605	Semrock FF605-Di02	1
Multi Band Dichroic (561 config.)	Chroma ZT405/488/561/660/905 RPC-UF3	1
Multi Band Dichroic (532 config.)	Chroma ZT405/488/532/640 RPC-XT-UF3	1
Single Band Dichroic Filter 555	Semrock HC BS555	2
Single Band Dichroic Filter 647	Chroma F48-647 T3 25.5x36x3mm	2
Single Band Dichroic FIlter 484	AHF F38-484 HC BS 484i	2

Detectors and cameras:

Item	Company / cat.no.	Quantity
Detector	Perlin Elmer SPCM-AQRH-13	2
Detector	MPD PDM series	2
Detector	PicoQuant PMA Hybrid 40	2
Camera	Imagesource DMK 23UX174	3
Detector	Hamamatsu PMT H12386-110	1

Other optical parts:

Item	Company / cat.no.	Quantity
Optical Fibre	Qioptiq	1
Pellicle Beamsplitter 8:92	Thorlabs	1
U-TLU tube lens F=180 mm	Olympus	1
FLIMBee optics (one scan lens +3 scanning mirrors)	custom made	1
Achr. lens 160/25	Linos G063-236-000	2
f=50mm lens	Thorlabs LA1131-A-ML	4
f=50mm lens	Thorlabs LA1131-B-ML	2
Broadband dielectric Elliptical mirror	Thorlabs BBE1-E02	2
Broadband dielectric mirror	Thorlabs BB1-E02	11
Non-Pol BS cube 50:50 400-700nm	Thorlabs CCM1-BS013/M	1

3 Results

Technological challenges render the conformational state of the FG-Nups one of the central remaining mysteries of the NPC. Solving the configuration of the FG-Nups will provide the foundations for further advances regarding the plasticity and functionality of the NPC. To develop a new approach to visualise spatial configuration of FG-Nups inside the NPC *in situ*, I used a combination of site-specific labelling and high-resolution fluorescence microscopy approaches. In section 3.1, I will present the first experimental strategy: super-resolution localisation microscopy super-resolution localisation microscopy of residue-specially labelled FG-Nups, which I used to map distances >10nm. In section 3.2, I will show the development and validation of a complimentary fluorescence microscopy approach: FLIM-FRET, which can measure very short distances up to 10nm. In the final section 3.3, I will show the results of using the FLIM-FRET approach to quantitatively determine the spatial configuration of FG-Nups Nup62 and Nup98 in the NPC and Nup98 condensates - GLFG bodies. I will implement the use of polymer scaling laws to find a descriptor that provides the quantitative measurement of the conformational state of the FG-Nups.
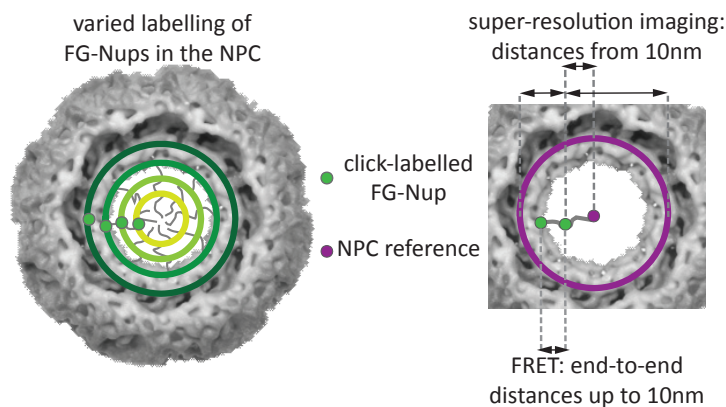
FIGURE 3.1: Schematic representation of the overall experimental design. Left picture represents different Amber suppression encoded labelling sites ranging across the full disordered sequence of the FG-Nup. Right picture shows that super-resolution microscopy is intended to be used to measure distances above 10nm from the labelled position to the reference in the NPC, either at the centre or in the scaffold. FRET will be used to measure absolute distances below 10nm between the two positions in the FG-Nups.

3.1   Site-specific super-resolution imaging approach

In this section I will present my results from the dual colour super-resolution imaging approach to map FG-Nups *in situ*. This approach measures distances in reconstructed single molecule localisation images between a labelled position in the FG-Nups and a reference in the NPC (Figure 3.1). The reference is key as click-labelling technique does not result in perfect labelling and incorporation of all targeted FG-Nups. Since FG-Nups do not posses any defined structure or arrangement, resulting images on their own would not be able to provide necessary information about the labelled position in the NPC. The distances are measured in two-dimensional images representing a cross-section of the real three-dimensional spatial separation. From these distances one can quantitatively estimate the conformational state of FG-Nups. Site-specific labelling of the FG-Nups is achieved using Amber suppression, which allows for incorporation of an UAA with a functional group to perform click-labelling. The reference for the NPC is either wheat germ agglutinin (WGA), which marks the centre of the NPC, or GFP-nanobody labelling the EGFP as a terminal marker of recombinant FG-Nup expression. Measuring distances to WGA in the centre can only precisely estimate distances that are more than 20nm apart as predicted by the simulation [182]. The advantage of using the GFP-nanobody instead is that localisations of click-labelled positions can be mapped to several well-defined positions in the NPC, compared to one localisation in the centre as in the case of WGA.

In this section I focus specifically on Nup62. Firstly, I demonstrate experimental design and labelling (section 3.1.1) and then show super-resolution imaging results of Nup62 Amber mutants around the structural region of Nup62 (section 3.1.2). I will then demonstrate a new dual-colour super-resolution mapping approach using a GFP-nanobody for a reference of the scaffold and show reconstructed dual-colour super-resolution images of two N-terminal Nup62 Amber mutants.[1]

3.1.1   FG-Nups expression and labelling

Nup62 Amber mutants here were designed to contain one Amber codon (TAG) in the Nup62 DNA sequence (number indicates the mutated amino acid) and a C-terminally tagged via a short flexible linker EGFP. EGFP is necessary as a readout of Amber suppression success. When the UAA is incorporated, the full protein gets translated and EGFP is observed, while in the ideal case scenario, when no UAA is incorporated, protein gets degraded. Amber mutants were chosen such that there is minimal perturbation to the FG-Nup functionality (STOP codon was used to replace

---

[1]labelling and super-resolution microscopy methods are based on the protocols established by Dr Jun Hee Kang [182].

amino acids in between the functional FG repeats).TCO* was used as the best performing UAA to date in terms of reaction speed, stability and yield [95, 183].

The central component of super-resolution microscopy is a well-labelled sample. Both labelling the NPC with WGA and GFP-nanobody are well established methods [115, 23, 184, 119], while click-labelling of FG-Nups with synthetic dyes *in situ* is new and is yet unpublished. The most advanced methods established in the group to click-label FG-Nups is to use a high salt buffer (10xTB). The high salt composition of the buffer helps to wash out undesired UAAs and doesn't fully lyse the nuclei [185]. Using this method I validated specific labelling of all Nup62 Amber mutants, shown in Figure 3.2. Figure 3.2 demonstrates specific labelling of four Amber mutants of Nup62. I imaged the two N-terminal Amber mutants 13 and 61 using TIRF microscope to better see the contrast between added TCO* and no UAA added case. This was necessary since one can still see the expression of C-terminally tagged EGFP, even when no UAA is added. In these mutants the Amber codon is at the beginning of mRNA sequence, it might be bypassed by the ribosome hopping over it [186, 187] and starting translation later in the sequence. I also mutated a native N-terminal methionine at position 108 to a leucine to prevent such a tentative translation initiation at this site. Even though this modification did not reduce the expression of GFP in the N-terminal mutants of Nup62, there was a clear difference between labelled NPCs in the nuclear footprints in cells with UAA and without (Figure 3.2). Therefore, these mutants are still suitable for further experiments, since they exhibit clear and specific labelling upon addition of the UAA. Mutants 257 and 515 also display specific labelling and a clear difference between addition and no addition of UAA.

I also checked the intactness of the NPCs using 10xTB buffer in the dual colour dSTORM pipeline. Images on Figure 3.3 show cells labelled with anti-MAb414 antibody (and anti-mouse secondary antibody tagged with Alexa647) that specifically binds FG-repeats together with wheat germ agglutinin (WGA), attached to CF680 dye, as the central channel marker. WGA is a lectin that specifically binds N-acetylglucosamine (GlcNAc) sugar residues, a very common post-translational modification of FG-Nups in metazoan, inhibiting transport through the NPC [188, 189]. Figure 3.3 shows that at this resolution there are no clear structural changes observed at the NPCs treated with high salt buffer 10xTB compared to intact cells. Taking this observation and the fact that it is the best performing methods available, high salt buffer was used in all experiments involving click-labelling of FG-Nups.
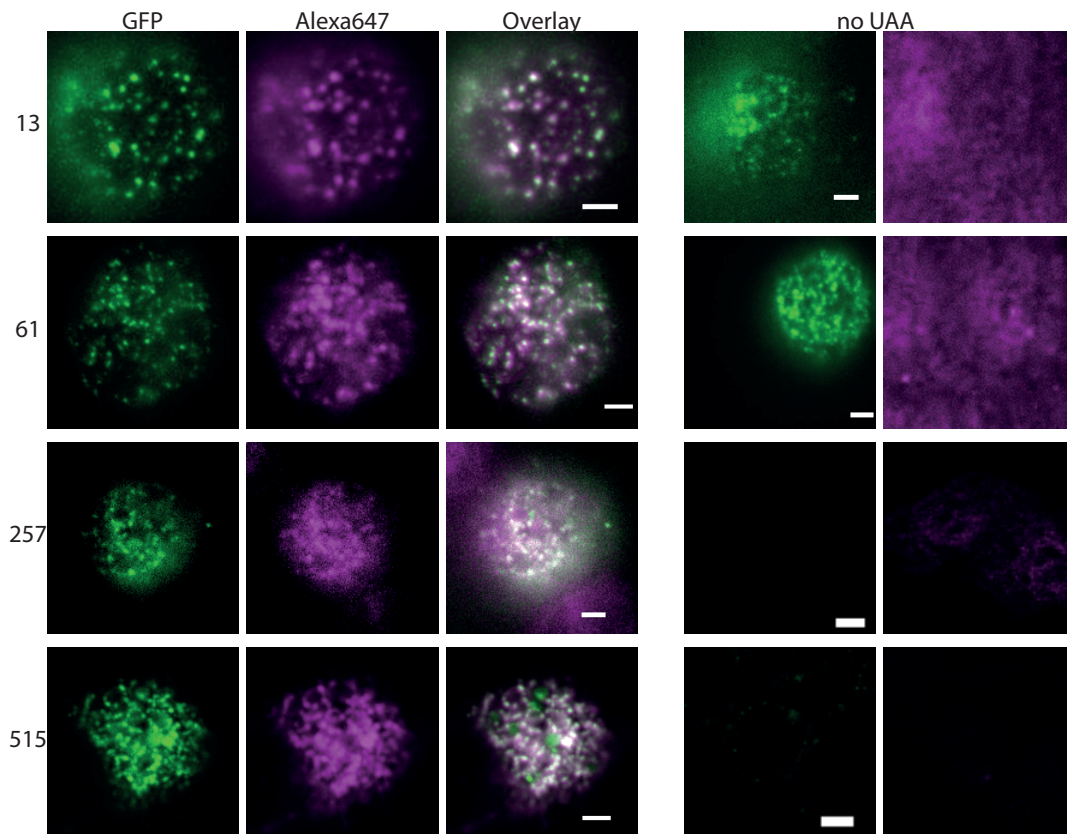
FIGURE 3.2: Representative images of nuclei footprint of HEK293T cells expressing Nup62 mutants for super-resolution studies. Two columns of image sets show representative expression (GFP) and labelling (Alexa647) channels for cells with added UAA (in this case TCO*) and with no added UAA (no UAA). Rows represent some of the Amber mutants used in this study, where the number indicates the position of the amino acid modified with an Amber codon. Overlay channel show colocalisation between the expression (expected NPC position) and specific labelling. Scale bars are 2μm, the last two images in the last column is 5μm.
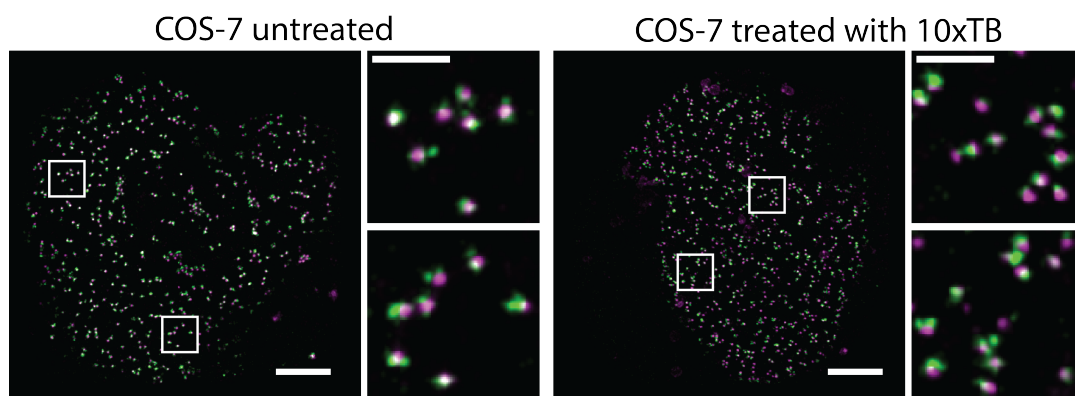
FIGURE 3.3: Side-by-side comparison of COS-7 cells in native and 10xTB conditions stained with an mAb414 antibody targeting FG-Nups. Reconstructed dual colour super-resolution images of nuclear footprints of COS-7 cells (scale bar 2um) with Alexa647 in green and CF680 in magenta. White boxes mark respectively upper and lower positions of the zoomed images (scale bar 500nm) to better display the individual NPCs.

3.1.2 Dual-colour dSTORM imaging of Amber mutants of Nup62 in HEK293T cells

Both colours of the click-labelled FG-Nup and the reference were imaged with dual-colour spectral demixing dSTORM. This method allows simultaneous imaging of two spectrally close to each other dyes like Alexa647 and CF680 using the same laser excitation, given the dyes have similar blinking kinetics [122]. Before starting with the dual colour dSTORM imaging for mapping Nup62 I checked whether Nup62 constructs with Amber codons around the known structured $\alpha$-helical region would reliably produce a super-resolved NPC ring-like structure. Indeed, after reconstructing localisation images of both mutants (schematic representation of Amber codon positions and consequently labelling site in top row of Figure 3.19) show clear ring-like structures of expected dimension around 40nm in diameter (Figure 3.19).

One of the limitations of dual colour dSTORM imaging using WGA was that imag-
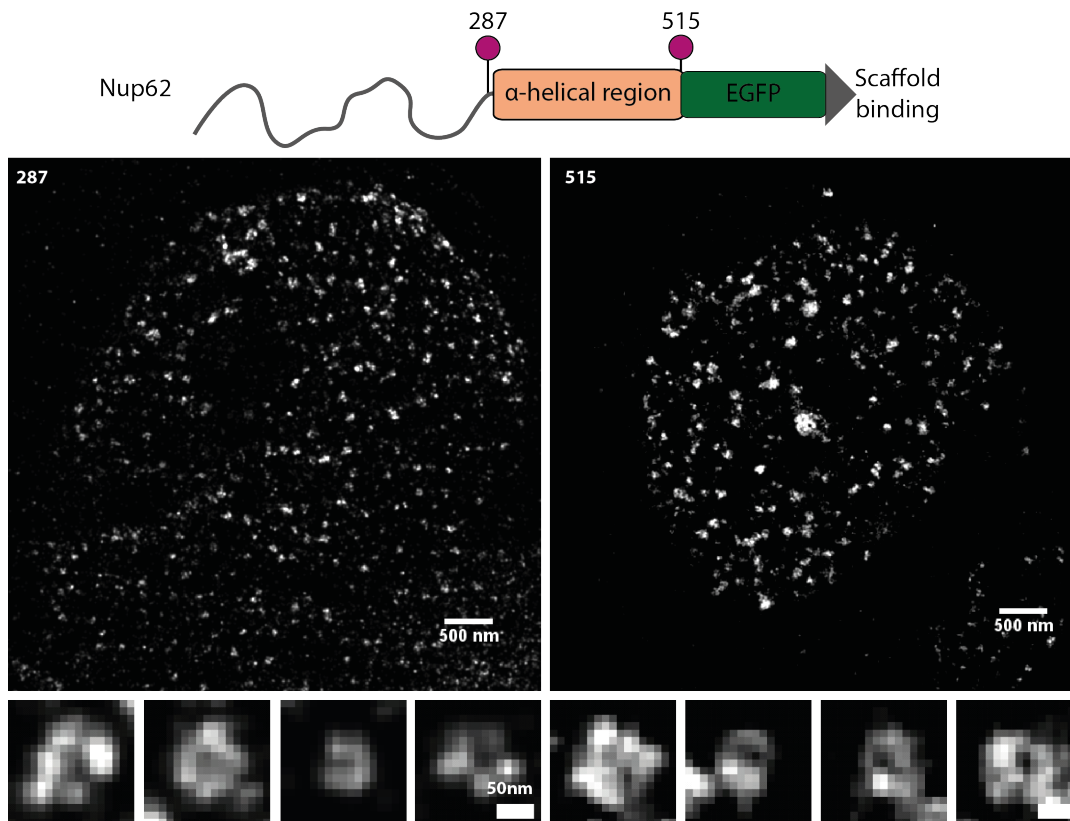
FIGURE 3.4: Super-resolution images of Nup62-287 and 515 Amber mutants. The figure shows a schematic representation of Nup62 domains, where grey is the disordered part, beige is the α-helical domain, that targets Nup62 to the pore and an engineered EGFP in green. Coloured circles with a number corresponding to Amber codon positions in the sequence of the protein and also the dye position, which is introduced as a result of UAA incorporation and subsequent click-labelling. Second row are the full nuclear footprint reconstructed images and the last row are zooms on individual pores. Resolution is 31 and 42nm respectively.

ing distances to the centre of the pore around 20nm become more and more unreliable with the localisation precision of 30nm of the TIRF setup used [182]. Since the diameter of the central ring of the NPC containing the Nup62 structural domain is around 40nm as predicted by cryo-EM [14], it becomes clear that the same procedure cannot be used reliably. Therefore, I aimed to use a different reference structure in the scaffold of the NPC to measure distances from Amber localisations to the localisations of the Nup62 structured domain in the NPC instead, thereby solving the aforementioned problem. The most straight forward and unbiased method was to directly target the already encoded EGFP [82], which should localise to the same site as the α-helical domain of Nup62, with labelled GFP-nanobody. Using the same plasmid I could image two sites of the same protein - the Amber position in the flexible FG-domain and the scaffold reference tagged by the GFP-nanobody (Figure 3.5 shows a schematic representation). Using the same imaging pipeline I reconstructed super-resolved images of Nup62 labelled at 515 Amber site with Alexa647-H-tet dye and the position of the EGFP labelled with GFP-nanobody-CF680. From the overlay image in Figure 3.5) it displays that both images do colocalise well, even though the 515 Amber targeted position does not exhibit a clear ring-like structure like in Figure 3.4.
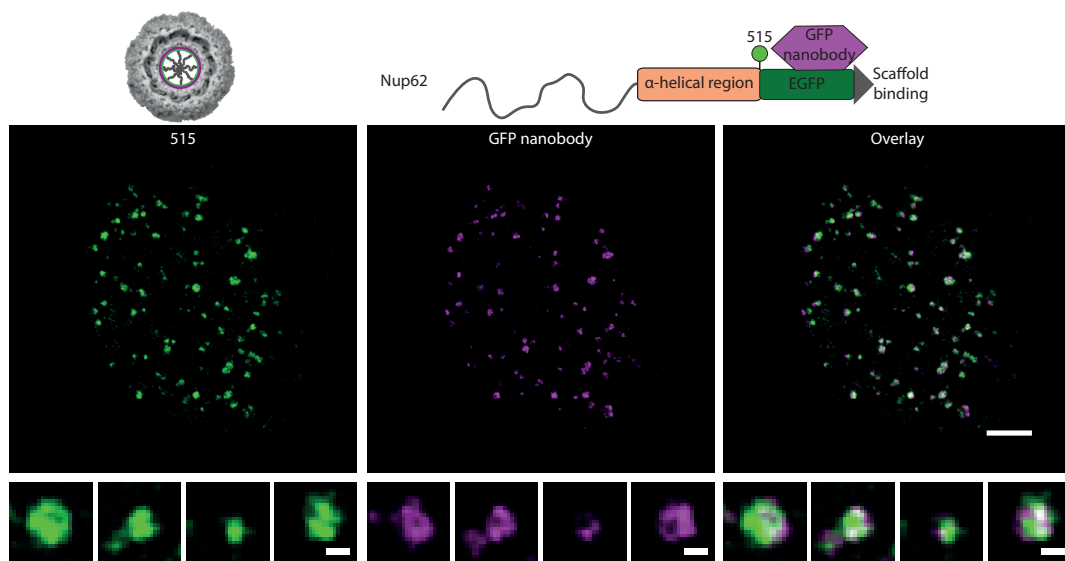
FIGURE 3.5: Dual colour super-resolution images of Nup62 labelled with both click-chemistry and GFP nanobody. Green channel represents Alexa647-H-tet click-labelled Amber site, magenta channel shows GFP-nanobody-CF680 and overlay channel shows colocalisation of both. Scale bars are 500nm and 50nm respectively and resolution is 42nm.

After seeing the convincing evidence that Amber suppression works well for the chosen sites and can be combined with a GFP-nanobody as a scaffold reference in the same experiment, I created more mutants in the FG-region of the N-terminus of

Nup62 and checked that they could be labelled (Figure 3.2). Checking for colocalisation expected between WGA and the N-terminal Amber sites of Nup62 I observed that there is a lot of fluorescent background in the labelling channel which significantly prevents from reliable validation of successful labelling (Figure 3.6). One of the reasons for this discrepancy might be the observed expression of the EGFP even in absence of UAA. This would mean that the ribosome may hop over the N-terminal Amber codon of the Nup62 and translate it past this site, resulting in a lot of Nup62 without an incorporated UAA being targeted to the NPC. Despite that, high background was also noted for other more C-terminal mutants, that do not have this problem. Therefore, it more likely that the high background fluorescence is specific to Nup62, potentially because it performs a variety of functions in cells, outside the NPC [34, 30, 31].

In summary, this section shows the development and establishment of a dual colour super-resolution approach which combines site specific click-labelling of FG-Nups and a GFP-nanobody targeted NPC scaffold reference. It details successful super-resolution imaging of the structured domain of the Nup62 and less successful imaging of N-terminus Nup62 mutants. Very high unspecific labelling background led me to focus instead on FLIM-FRET, which I describe in the following section.
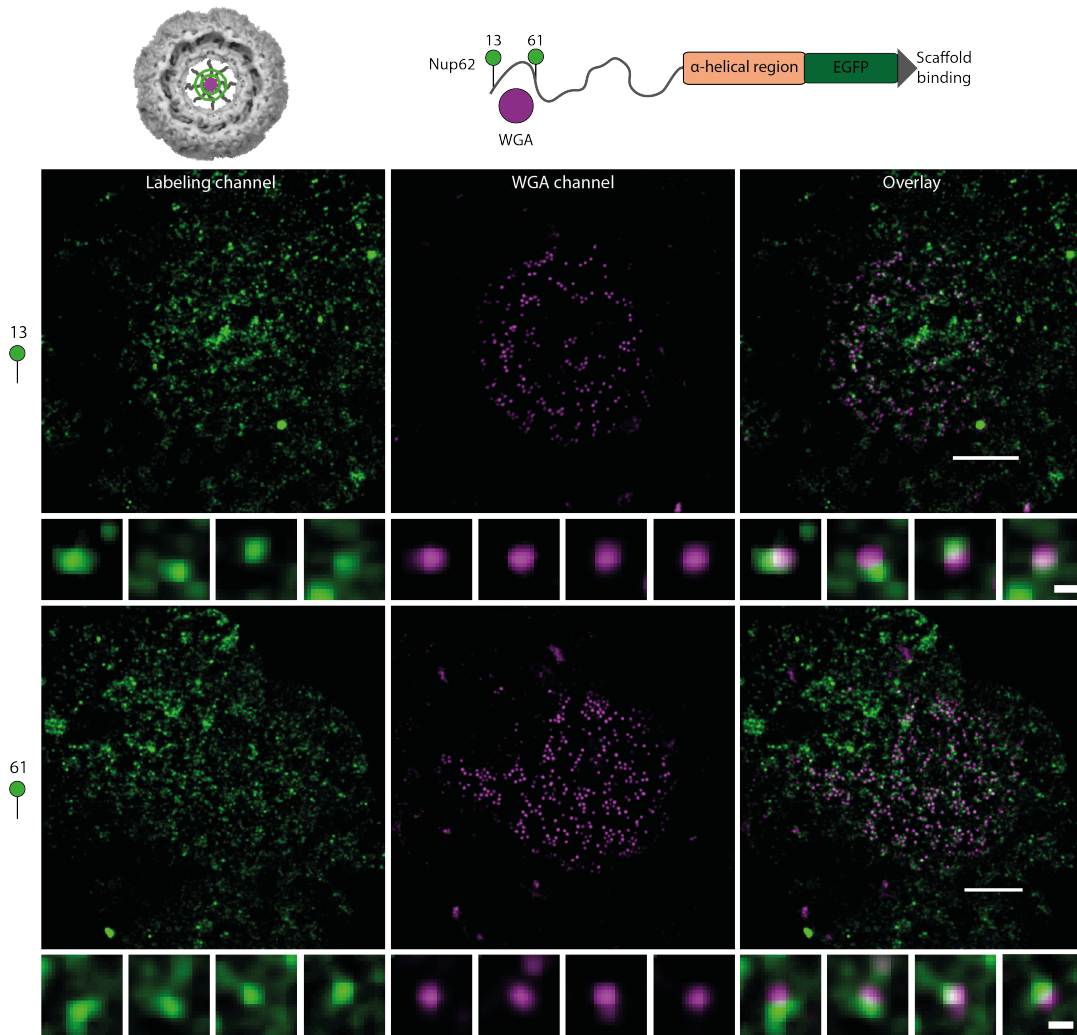
FIGURE 3.6: Dual colour super-resolution images of Nup62 Amber mutants 13 and 61, labelled with both click-chemistry and WGA. Schematic representation on the top shows the expected distribution of labels around WGA and the schematic representation of experimental system consisting of Nup62 and WGA.Green channel represents Alexa647-H-tet click-labelled Amber site, magenta channel shows WGA-CF680 and overlay channel shows colocalisation of both. Scale bars are 500nm for the nuclear footprint image and 50nm for individual pores. Resolution is 33 and 24nm respectively.

3.2 Establishing the FLIM-FRET approach

FRET is a highly sensitive mechanism that provides quantitative measurement of end-to-end distances between the two dyes in close proximity to each other, usually up to 10nm. FLIM is a highly sensitive fluorescence technique to sense FRET or more generally speaking, any changes to the surrounding dyes by monitoring their fluorescence decay. If the dye transfers energy to another dye using FRET, this process changes the donor fluorescence decay and can be quantified using exponential decay models. FLIM is an imaging technique which can record fluorescence decays in different pixels. The use of FLIM to measure FRET therefore provides a powerful tool to measure distances between fluorophores with high precision in cells. Inspired by the capabilities of this method, having developed a site-specific click-labelling method in FG-Nups I proceeded to establish, develop and validate this method to study FG-Nups *in situ*. The key requirement was the ability to label a single protein in two sites *in situ* to measure intramolecular FRET. Once my preliminary experiments made it possible, this method became the prime focus of my work.

In this section, I will show the experiments by which I established the FLIM-FRET approach for FG-Nups. I will begin with the experimental segment, before detailing its analytic validation. In the experimental part, I will show in section 3.2.1 how I selected the dyes and the UAA to be used, followed by the experimental settings I established. In section 3.2.3, I will display the results from an in vivo FRET sensor that validate use of FLIM-FRET in cells and quantitatively agree with an intensity-based FRET imaging method. Finally, I will demonstrate the experiments I performed using DNA as a FRET nanoruler to investigate the precision of calculating end-to-end distances and the effects of fraction of the donor species on these measurements. Having validated the use of the experimental setup, I will also show the validation of the analysis pipeline using simulations of both experimental results I obtained with DNA nanorulers and the prediction for Ideal Gaussian chain model implementation (section 3.2.4). [2]

3.2.1 Searching for a suitable dye pair

It was important to first find a dye pair that would be most suitable for use inside cells and perform well in FRET. To choose best-suited dyes for FLIM-FRET I considered several criteria:

1. Commercial availability with 1,2,4,5-tetrazine group (H-tet).

2. Long lifetime and high quantum yield of the donor.

---

[2]confocal image scanning FLIM setup was built by Dr Piau Siong Tan. Experiment in figure 3.7 was done together, Nup98 constructs were developed and PAM end-to-end distribution analysis was implemented by Dr Piau Siong Tan.

3. Large spectral overlap between donor and acceptor dyes.

4. Minimal sticking in cells.

5. Monoexponential decay after click-reaction.

6. Photostability of the dyes.

To address the first four criteria I tried two dye combinations: a well established FRET pair Cy3 and Cy5 and another easily available good candidate - Atto532 and Alexa647 pair [190]. Alexa488 and Alexa594 would have also been a suitable candidate and are the best dye pair regularly used in *in vitro* smFRET measurements. However, since Alexa488 emits in the green range of the spectrum, I would have faced more problems with demixing the real signal from autofluorescence background. The autofluorescence in cells is often connected to their metabolic state and flavins, nicotinamide-adenine dinucleotide (NAD) and aromatic amino acids that emit in this range of the spectra upon appropriate excitation [191, 192]. Therefore, my goal was to find a similarly performing dye pair in the orange/red part of the spectrum.

Dye pair	Donor lifetime [ns]	Q_D	$J(\lambda) [nm^4 \cdot l \cdot mol^{-1}]$	R_0[nm]
Cy3 - Cy5	0.3	0.15	7.4×10^{15}	5.4
Atto532 - Alexa647	3.8	0.9	3.5×10^{15}	6.1
Atto532 - LD650	3.8	0.9	3.7×10^{15}	6.1
Alexa488 - Alexa594	4.1	0.92	2.9×10^{14}	6

TABLE 3.1: Physical parameters of chosen dye pairs compared to the standard Alexa488 and Alexa594. Overview of donor's quantum yields (Q_D), spectral overlap integral (J) and Förster radii (R0) for the different FRET pairs.

The first pair was observed to stick more in the cellular membranes (Figure 3.7) and the lifetime of Cy3 was too short for the purposes of this study, while Atto532 and Alexa647 satisfied the first four criteria. In the end, COS-7 cells were selected as they are much larger, and therefore have a larger nucleus and easily visible cytoplasm. HEK293T cells tend to collapse more, making it harder to segment the nuclear rim. Since both cell lines express the double Amber codon recombinant protein equally well, COS-7 cells were selected. To examine whether this dye pair satisfies the fifth criterion, I performed FLIM experiments on Atto532 dye modified with an azide group as a control and Atto532 modified with H-tet for click-labelling. Figure 3.8(a) shows the dye control is monoexponential, while H-tet modified dye has a second, shorter component (0.34ns) most likely coming from tetrazine (Figure 3.8(b,c)). After this experiment, I studied whether upon click-reaction with TCO*, BCN and SCO I could eliminate the second component. I selected those three unnatural amino acids and SPIEDAC types of reactions as the three performing best in terms of reactivity, speed (except SCO) and quality of labelling [95]. The final product of click-reaction is
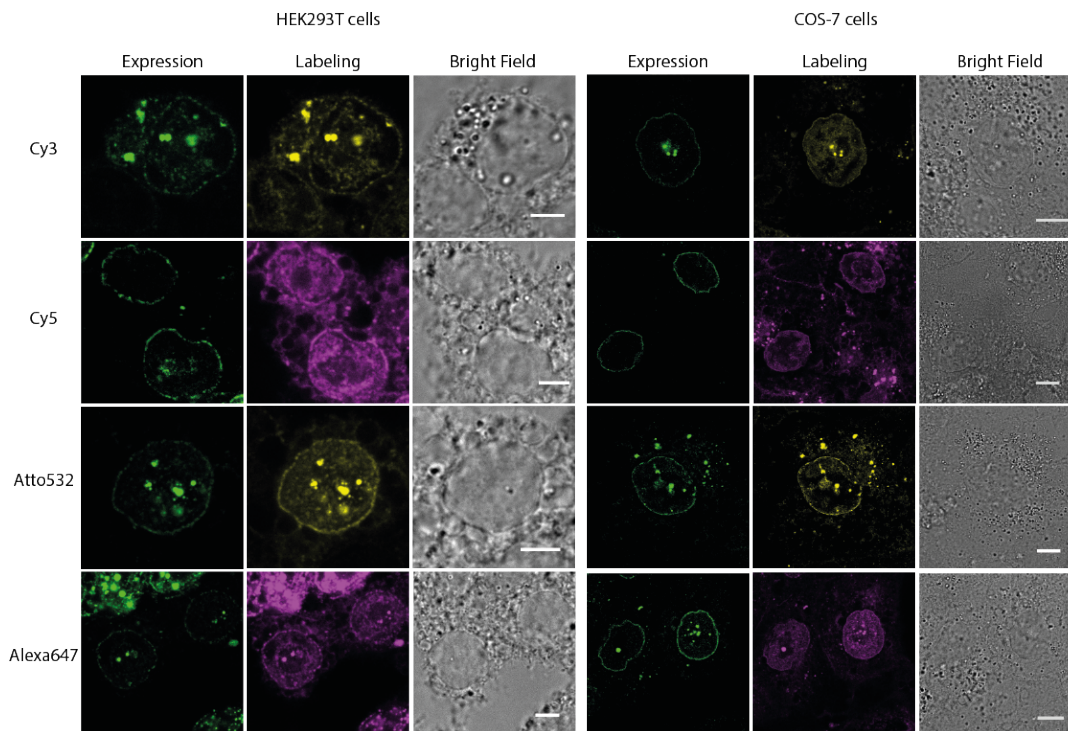
FIGURE 3.7: labelling HEK293T and COS-7 cells expressing GFP-Nup98(38TAG) construct. The first three columns (from left to right: GFP channel, labelling channel, bright field view) show confocal images of nuclear rims of HEK293T cells (scale bar 5um) and the second three columns COS-7 cells(scale bar 10um).

monoexponential only with BCN and SCO (Figure 3.8(e,f)) and double exponential with TCO* (Figure 3.8(d)), which was the prime choice for super-resolution imaging. This is in line with the fact that the reaction of TCO* with a tetrazine can lead to two isomers, while BCN and SCO lead only to a single isomer [193]. Therefore, I decided to continue using H-tet-Atto532 as the donor to click-label BCN-incorporated Nups. The last criterion on the list is critical for long measurements. Alexa647 is a very well suited dye for super-resolution imaging, but it is not a very good acceptor as I noticed strong bleaching by donor laser excitation and FRET over time. Consequently, I selected the better acceptor - LD650 dye [194]. It is a 1,3,5,7-cyclooctatetraene (COT) modified Cy5 dye which is known for its increased photostability due to so called self-healing effect. LD650 satisfied all of the first five criteria and was also very stable in the imaging conditions in PBS. Atto532-LD650 dye pair was therefore selected and used in all further experiments in cells.

3.2.2 Establishing imaging settings

To find the best FLIM settings to obtain the maximum valid signal out of our regions of interest (ROIs) I needed to consider several things:

1. The number of photons hitting the detector should be maximised at minimum excitation power to prevent bleaching and maximise resolution, therefore a
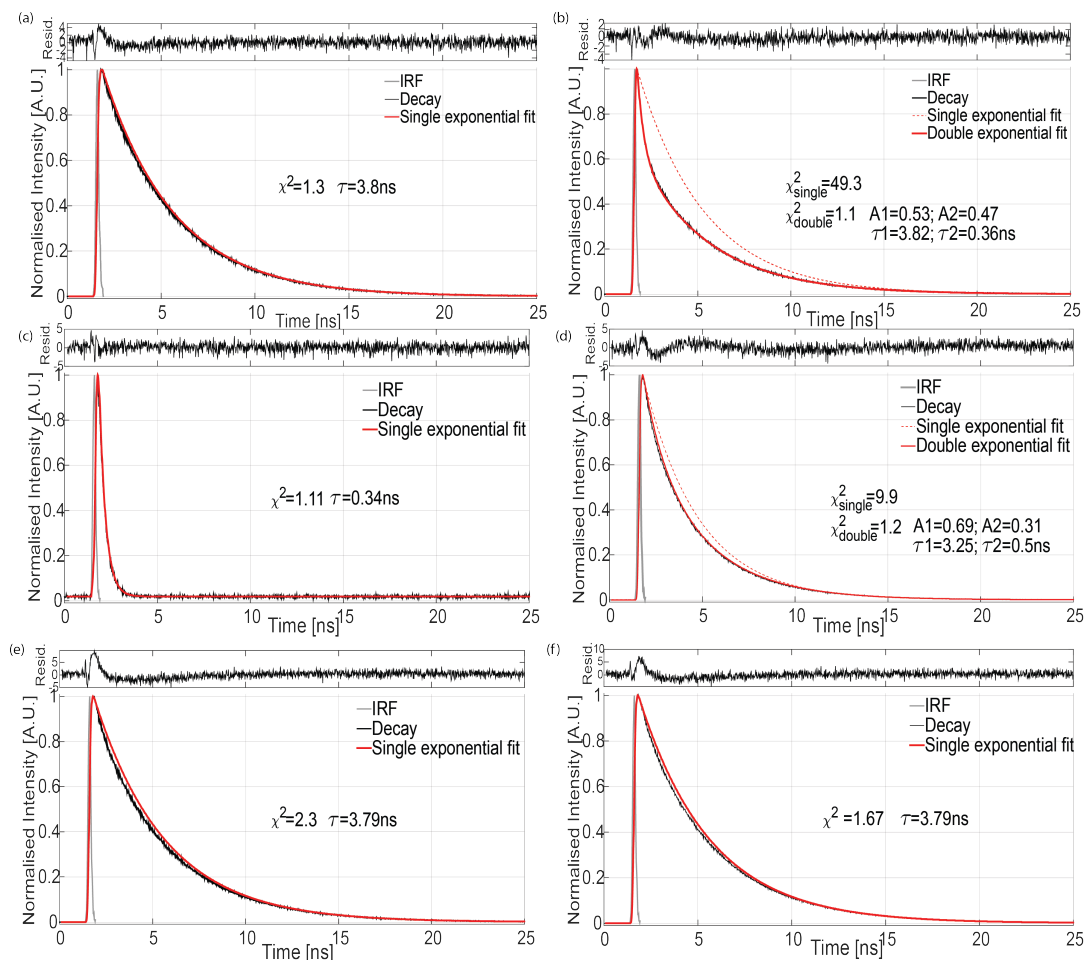
FIGURE 3.8: Lifetime analysis of Atto532 - tetrazine conjugation and in vitro click-chemistry reactions. Grey curve is the measured IRF, black is the lifetime decay data and red is an exponential fit. The graph on top of each represents the residual standard deviation curves that are an indication for the goodness of the fit. (a) azide-Atto532 dye lifetime decay fitted with a monoexponential function; (b) H-tet-Atto532 dye lifetime decay fitted with both mono- and double exponential functions; (c)tetrazine lifetime decay fitted with a monoexponential function; (d-f) reacted H-tet-Atto532 with TCO*, BCN and SCO respectively. All fitted with a monoexponential decay, while the reaction with TCO* is also fitted with double exponential function due to a poor fit of the former.

pinhole size was chosen to be 50 μm.

2. Total photon count rate in the pixel should be well under 10% for single-photon statistics to prevent the pile-up effect arising from the dead time of the detector [142, 139].

3. Total number of photons in the ROIs should be at least 30000 counts total or 650 counts/pixel to have a good photon statistics for reliable analysis of multiexponential curves [139, 195, 196].

4. At the laser powers used bleaching of both donor and acceptor should not exceed 20%.

5. Pixel size in the image should ideally conform to Nyquist criterion (maximum pixel size should be around $\frac{1AU}{2.3} = \frac{226nm}{2.3} = 98nm$, where 2.3 is the value of pixels per minimum resolved feature.

6. Dwell time and acquisition length should be minimal to minimise bleaching.

Taking all these parameters into account I found the most optimal settings for laser power, dwell time, image size and acquisition time (Table 3.2). I monitored both total image intensities and the intensities in the rim of the Donor and the Acceptor images at the beginning compared to the end of the measurement. The resulting parameters that meet all the above criteria were then used to image nuclear envelopes and slightly adjusted parameters (laser power at 6μW, number of frames at 200) were used for very bright GLFG bodies in Nup98 expressing cells) 3.2.

Pinhole	50 μm
Laser power	12μW
Pixel size	0.091μm
Image size	90μm 1024pixels
Number of frames	150
Dwell time	2.5μs
Detector count rate	40MHz
660nm laser power	12μW
405nm laser power	12μW

TABLE 3.2: Imaging settings applied to FLIM measurements in fixed cells.

3.2.3 Validation of FLIM-FRET imaging using custom-designed dimerisation activated FRET

In the need of an established *in vivo* FRET sensor control I have adapted the dimerisation-activated FRET sensor [197] to be used with a in mammalian cells. The goal of this experiment was to directly compare FRET efficiencies calculated using intensity-based FRET and lifetime measurements on the same microscope. Based on the cytoplasmic proteins FK506-binding protein 12 (FKBP) and FKBP12-rapamycin binding domain (FRB), dimerisation of which can be induced upon addition of rapamycin [198], I designed a simple GFP-mCherry FRET sensor [197]. GFP was fused to FKBP and mCherry to FRB (Figure 3.9(a)). Upon addition of rapamycin, I saw dramatic changes in the FRET channel. Firstly, I saw changes in the Donor lifetime curve. As expected upon addition of rapamycin the curve appears more quenched compared to controls, where the Donor is not participating in energy transfer: Donor only (GFP) and Donor in the presence of Acceptor before dimerisation (GFP and mCherry) (Figure 3.9 (b)). Secondly, after all images were thoroughly analysed using equation 1.3, corrected FRET images were reconstructed and again showed clear differences between the cells with added rapamycin and those without (Figure 3.9

(d)). The combination of these two observations clearly indicated there is FRET taking place only upon dimerisation of the proteins. To validate use of FLIM imaging for quantifying FRET *in vivo*, FRET efficiencies were analysed using equations 1.10 and 1.3. The bar plot in Figure 3.9(c) shows that the calculated FRET efficiencies between intensity-based FRET and FLIM match very well. The differences observed in the "- rapamycin" case occur since donor lifetime changes slightly in the presence of acceptor, thus altering FRET efficiency from equation 1.10. I therefore concluded that our FLIM setup was sensitive enough to study FRET *in vivo*.
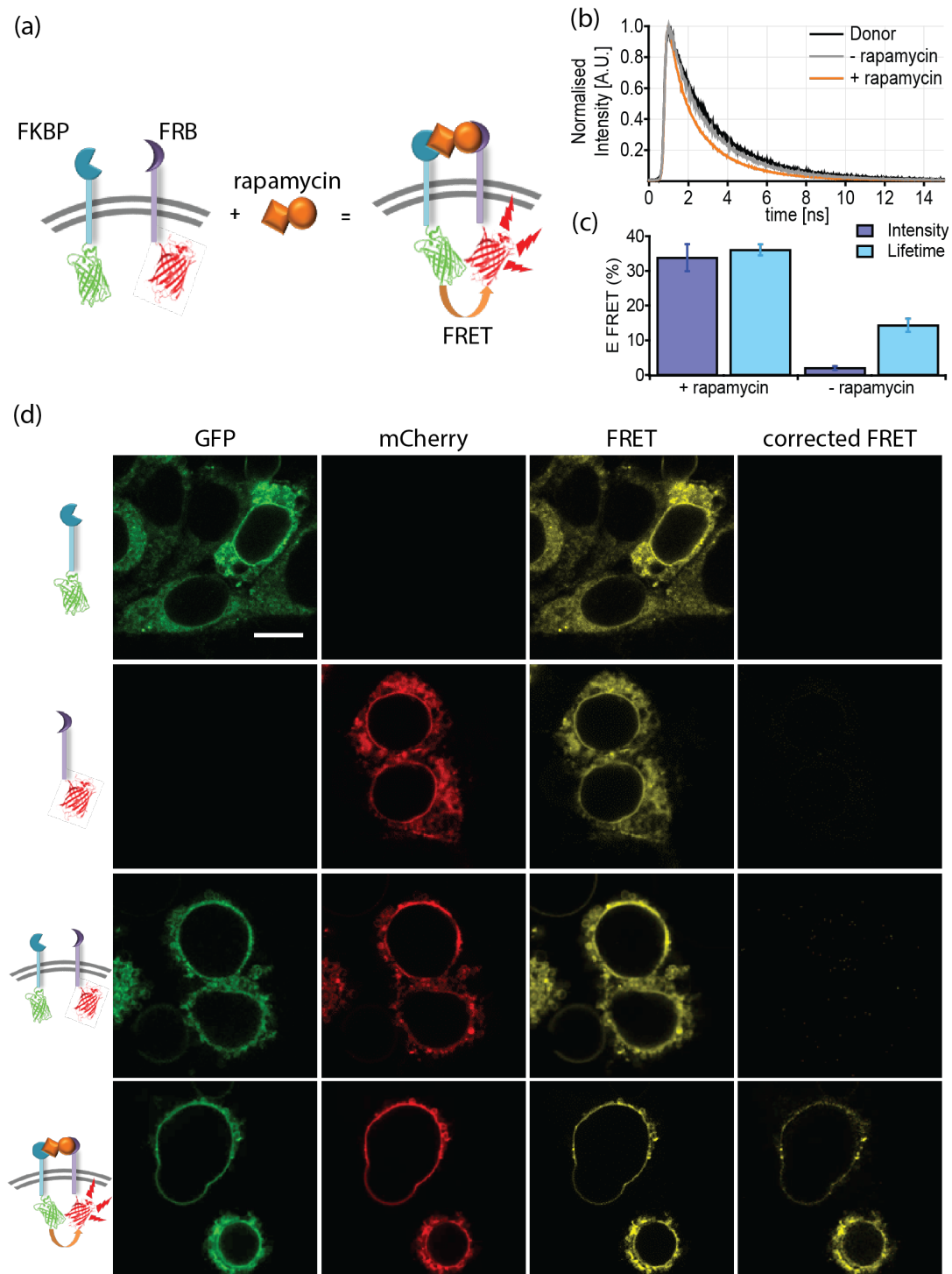
FIGURE 3.9: Validation of lifetime FRET compared to intensity based FRET on custom-designed GFP-mCherry FRET sensor *in vivo*. (a) schematic representation of the sensor; (b) overlaid summary lifetime curves with Donor (GFP only) in black, no rapamycin added (both GFP and mCherry present but do not interact) in grey and with rapamycin added in orange; (c) bar plot of FRET efficiencies; (d) corresponding images with a cartoon plot on the left (scale bar is 10 µm). From left to right there are 3 channels monitored: 488 excitation, 488 detection (GFP); 561 excitation, 561 detection (mCherry); 488 excitation, 561 detection (FRET), last image is the corrected FRET channel according to equation 1.3. From top to bottom there are 4 separate experiments: cells expressing only GFP, cells expressing only mCherry, cells expressing the fully functional system but with no rapamycin added and cells expressing the fully functional system and added rapamycin.

3.2.4   Data analysis and validation of FLIM-FRET with custom-designed DNA FRET-nanorulers simulations

Finding the best way to analyse the data from segmented nuclear rims of cells is a challenge in itself. In our custom built microscope many parts are from PicoQuant who also provide the acquisition and analysis software SymPhoTime 64. It is a useful tool for multiexponential analysis of time-resolved data and provides an option for reconvolution fit with an IRF. It also allows for global fitting between different experiments and for manual selection of regions of interest (ROIs) in the image. Test data such as DNA, organic dyes, dimerisation activated FRET sensor and donor samples were all analysed using SymPhoTime. However, adding the essential option of end-to-end distances distribution function calculation to the software or an automatic segmentation pipeline was not easily possible.

Pulsed interleaved excitation (PIE) Analysis with MATLAB (PAM) is an open source GUI-based software package for the analysis of fluorescence ensemble, single-molecule and imaging data https://gitlab.com/PAM-PIE/PAM [199, 200, 165]. PAM has multiple modules for different purposes and is easy to edit. It also has an active github community constantly updating the interface and solving any issues. I mainly used the TauFit module - the TCSPC module of PAM, dedicated to the analysis of fluorescence intensity decays and time-resolved anisotropy. I used reconvolution fitting of fluorescence decays procedure in MATLAB with the newly introduced to TauFit distribution of donor-acceptor distances. TauFit uses MATLAB (MATLAB R2017a V5-2) non-linear least-squares solver to find the best parameters to fit the data curve. The equation that the data was fitted to is derived in section 1.4 of the Introduction (equations 1.18 and 1.17). All parameters that relate to the donor characteristics were kept constant, while others were fitted using the least squares fit procedure with an appropriate model (P(R)) (equation1.23). From the double exponential fit of all donor curves from the nuclear rims τ_{D1}, τ_{D2} and A (the normalised ratio between the two exponential components) were calculated and fixed (using equation 1.12). R_0 was fixed at 61Å and was calculated from the overlap spectra of Atto532 and LD650 according to equation 1.6. The fitted parameters included the end-to-end distance R[Å], the fraction of donor only species d, scattering and background components and the IRF shift. The best fitting parameters were found by reducing χ^2 - the maximum likelihood estimate for Gaussian errors as a measure of the goodness of the fit.

DNA oligos or FRET nanorulers (see Methods Section2.4) were designed to test the performance of PAM in end-to-end distances calculation and also the effects of multiple fitting parameters like the donor fraction both experimentally and analytically. Upon hybridisation of two complementary oligos a double-stranded DNA (dsDNA) was formed with 3 varying end-to-end distances (R_E) between the dyes (see List of DNA oligos). I measured double-stranded Donor DNA (dsDonor), FRET1 (Donor strand + Acceptor strand 1), FRET2 (Donor strand + Acceptor strand 2), FRET3

(Donor strand + Acceptor strand 3), linearly varying the concentration of dsDonor to study the effects of donor fraction on R_E. From lifetime fitting using SymPho-Time and fitting the resulted curves using the static model (Equation 1.15) in TauFit I could monitor the effect of titrating the amounts of donor on the FRET population. Static model was fit keeping $\tau = 3.88 ns$ (measured from dsDonor), $R_0 = 60 Å$ (Table 3.1) fixed. Adding more donor fraction significantly alters the curves which become less quenched and "move" towards the dsDonor control as shown in Figure 3.10(c) and (d). Moreover, it becomes apparent that the sample with higher FRET (FRET1) is more sensitive to the added dsDonor (slope in Figure Figure 3.10(b)).
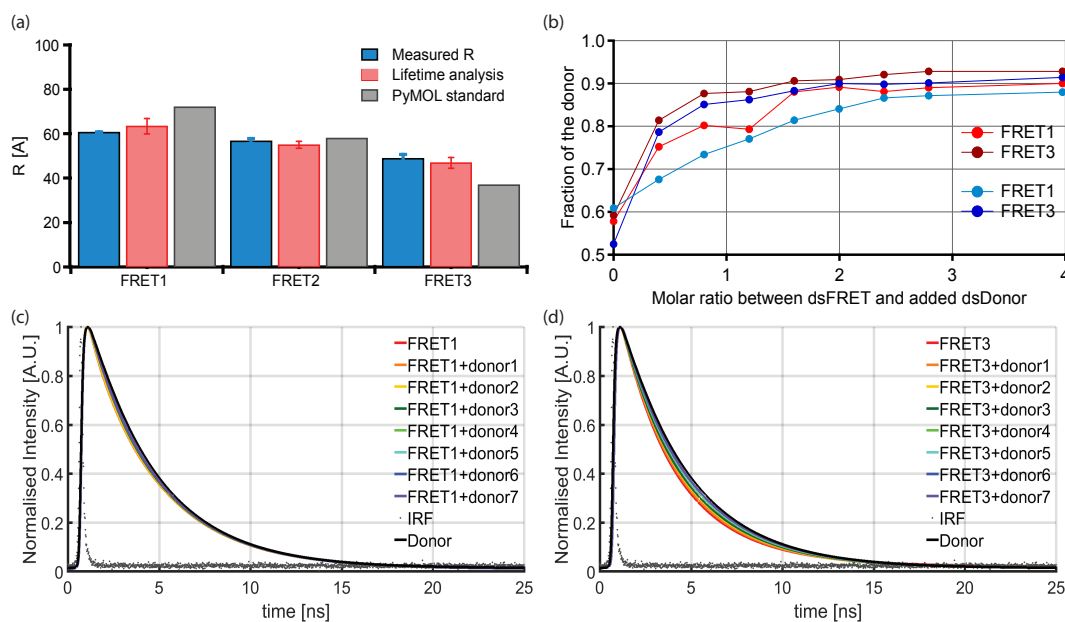
FIGURE 3.10: DNA titration FLIM-FRET experiment. (a)Bar plot of calculated R_E values from the static model fit and end-to-end distances measured using PyMOL; (b) Donor fraction amplitudes plot versus molar ratios; (c) and (d) are overlay plots of all fitted curves with increasing donor only concentration in solution.

End-to-end distances R_E from the model fit can also be verified by calculating R_E from E_{FRET} using lifetime equation (equations 1.9, 1.10). This experiment shows that I can reliably extract end-to-end distances using our pipeline (Figure 2.2). As shown in Figure 3.10(a) error bars around R are very small in measured R_E case (using the static model) and exactly match those calculated from E_{FRET}, even though the measurements do not fully agree with values measured in PyMOL, where flexibility of linkers to the dyes was not taken into account. Despite that, the relative trend is clearly the same. Secondly, the measurements of end-to-end distances are independent of the donor fraction, while the fraction of the donor d in the FRET ensemble measurements follows almost a straight line for the low FRET samples (FRET3) up to a saturation around 90% or at 1-2 times excess of the donor and is much more sensitive in measuring d in high FRET samples (FRET1) (Figure 3.3 (c)). Interestingly, estimating the donor fraction using both PAM (directly calculated R_E)

and SimPhoTime (using double exponential fit, where d is the fraction of the first exponent, related to the donor) fully agree with each other. This is evident from Figure 3.10(b), where red colours correspond to SymPhoTime and the lifetime analysis and blue correspond to PAM and measured R_E. In conclusion, it is possible to reliably analyse lifetime data using the newly established FLIM pipeline to extract end-to-end distances (at least in case of a monoexponential donor decay and using a static model) and reliably identify the amount of donor species of up to 1 molar excess.
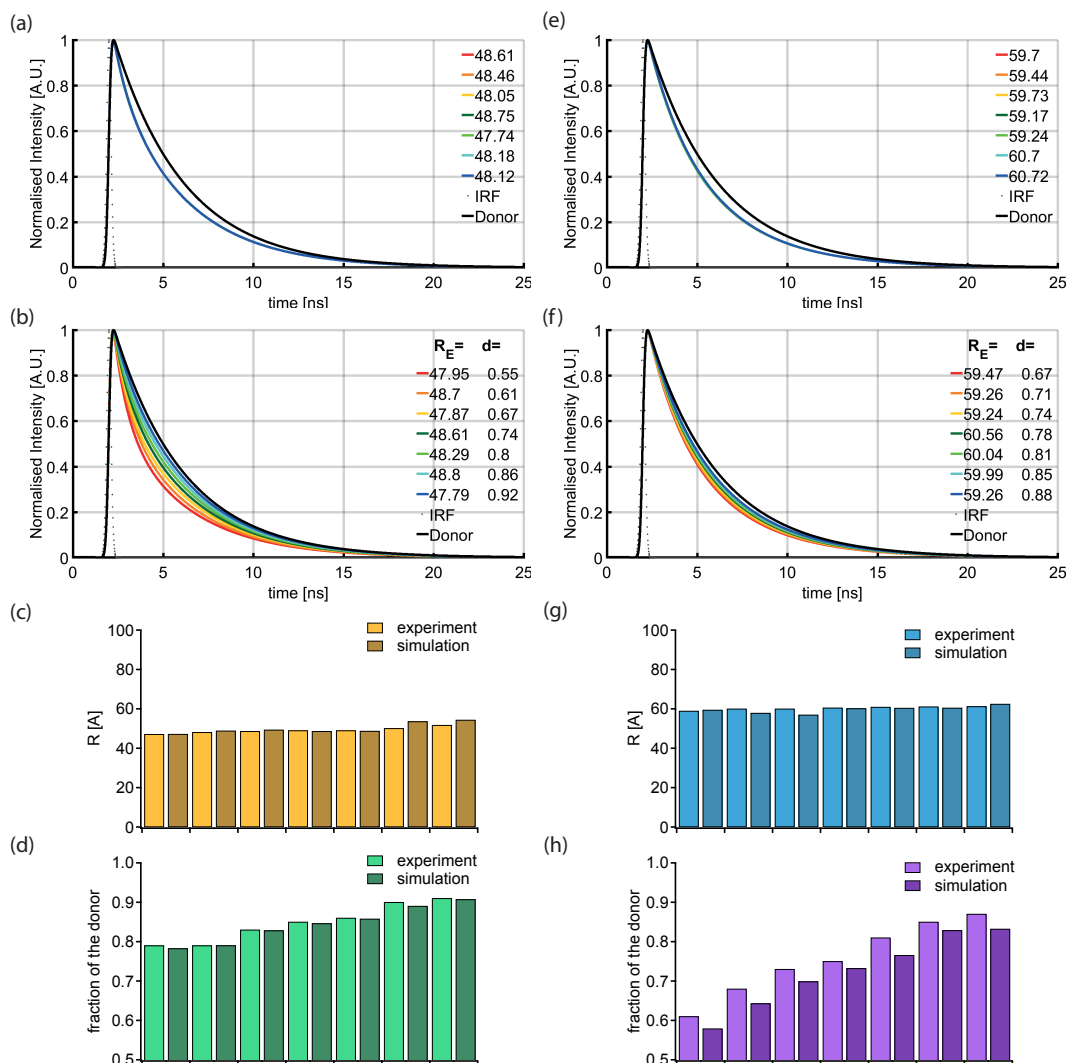
FIGURE 3.11: DNA titration FLIM-FRET simulation. Simulation is based on the static model for DNA R_E distances analysis. (a-d) are the results of fluorescence decay curves of FRET3 DNA sample with a fixed fraction of the donor d (a), randomly varying fraction representative of the experimental values (b), bar plot of the results comparing back calculated values from the simulation for R_E (c) and d (d). (e-h) are the results of FRET1 simulated DNA sample in the same order.

To test how well PAM performs in experimental data analysis, I simulated fluorescence decay curves with the fitting parameters results I obtained for the DNA. The simulation was also done in MATLAB using the exact same code also for the static

model in the equation 1.15. The first aim of the DNA experiment was to test the effect of the donor species fraction d on the end-to-end distances (R_E) calculation and also on the appearance of the fluorescence decay curve. Figure 3.11(a,b,e,f) shows simulated curves of same values of end-to-end distances of two different FRET efficiency values FRET3 (first column) and FRET1(second column), within their calculated error (exact values are shown in the legend). It is clear that such small standard deviation in the R_E would not in any way affect the fluorescence decay curves. Interestingly, varying the fraction of the donor species d 3.11(b,f) significantly alters the curves in the same way as observed in the experiment in Figure 3.10. This confirms that the fraction of donor species has a very pronounced effect on the appearance of the decay curve, yet has no effect on the measurement of R_E, suggesting R_E could be very well accounted for. I also used the curves I created in the simulation with exactly same values as in the experiment to be analysed with PAM. As Figure 3.11 (c,d,g,h) shows, I can perfectly recover all the values expected. It is potentially more difficult to get the same estimation of the fraction of the donor species at larger R_E (Figure 3.11 (h)) since I am underestimating d.
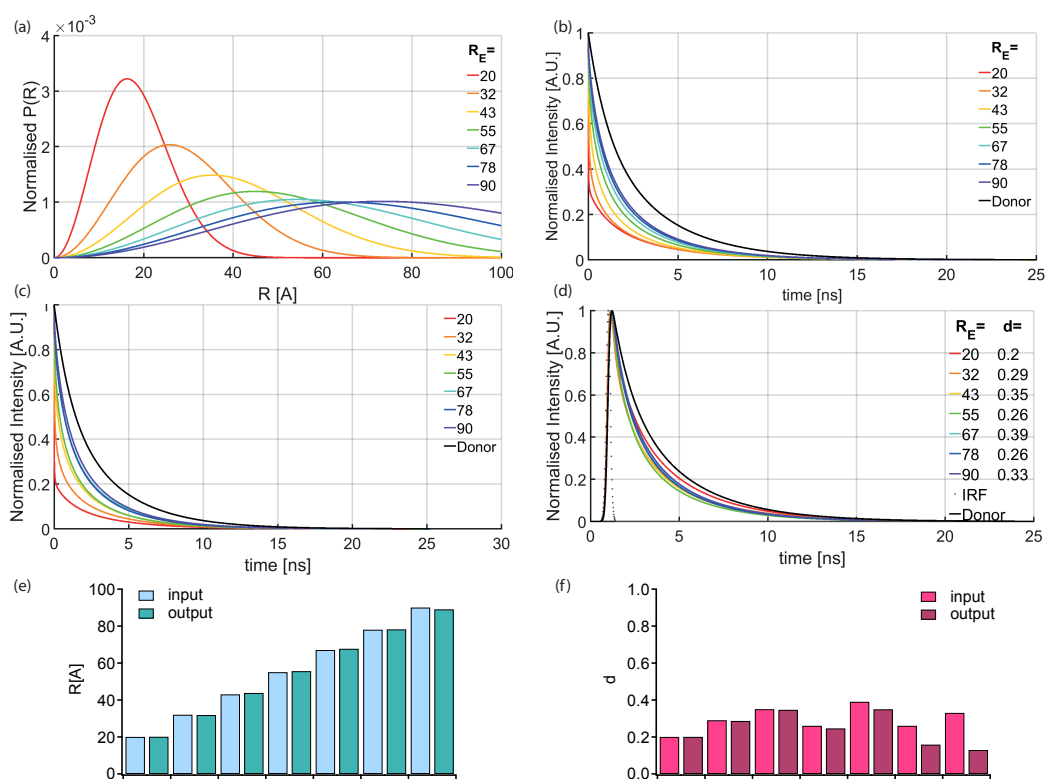
FIGURE 3.12: Simulation of Ideal Gaussian distribution in FLIM-FRET. (a) is end-to-end distances distribution P(R) in terms of R with varying R_E from 20 to 90 Å; (b) are fluorescence decay curves of varying R_E at the fixed d=0.3; (c) same fluorescence curves at randomly varying d from 0.2 to 0.45; (d) convoluted with a simulated IRF curves from (c); (e) and (f) are the results obtained from back calculating R_E and d, respectively

After seeing that the results of the DNA experiment and the simulation fully agree I simulated the fluorescence decay curves with the end-to-end distance distribution

P(R) (equation 1.17), which was chosen to be the Ideal Gaussian (equation 1.23). This model was chosen to study FG-Nups as it is most commonly applied to IDPs in general [201] and FG-Nups in particular [170] in *in vitro* studies. My simulation was based on the predicted range of end-to-end distances (Figure 3.13) and the double exponential decay parameters measured in the donor only sample (Figure 3.15), which will be discussed in details in the next section. I took those parameters along with d varying between 0.2 and 0.45 for the ideal labelling case to reconstruct the curves. Figure 3.12 (a) shows the probability distribution itself, (b) the fluorescence lifetime curves resulted from integration of the distribution and (c) demonstrates the effect of the fraction of the donor *d*, known to significantly alter curves appearance from the simulation of DNA case. In the case of the Ideal Gaussian distribution it is also noticeable that the fraction of donor species has a significant effect on the curves, especially in relation to each other. Figure 3.12 (d) shows curves in (c) after convolution with the simulated by a Gaussian function IRF. It is important to note that convolution with the IRF has a pronounces effect on the relative pattern in decay curves. The red curve corresponding to the smallest R_E has a very sharp slope at the beginning and is the furthest from the double exponential donor curve in black. After the convolution the curve's tail actually becomes closest to the donor, even though the initial slope of the long lifetime component is still the furthest (Figure 3.12(d)). I give those details to emphasise that in the region around R_E=20Å E_{FRET} measurements are not easily distinguishable (Figure 3.13) and the trend that one expects to see of the curves appearing most quenched at smallest R_E might be different in this particular range up to 31Å. Despite that, computationally the analysis pipeline can easily distinguish between all different R_E and d used in this simulation as Figure 3.12 (e,f) demonstrates. Although *d* is not recovered as well at large R_E above 70 Å, it again doesn't impair a precise evaluation of R_E. Thus, the Ideal Gaussian chain model implementation into my analysis pipeline was concluded to be successful and applicable to analyse *in situ* data.

In summary, this section shows that I successfully established both the experimental conditions and the analysis for the direct measurements of end-to-end distances using FLIM-FRET. I used a rapamycin-induced FRET sensor *in vivo* to demonstrate that the measurements of FRET efficiencies between FLIM-FRET and intensity-based FRET agree. I also showed measurements of R_E directly in DNA nanorulers and compared the results calculated using the classical FRET efficiency analysis for monoexponential donors and the developed PAM software analysis using the static model for direct R_E calculation. I further simulated the experiment and showed that in all cases the results obtained for R_E and the donor fraction agree well. Moreover, I demonstrated that even though the donor fraction *d* has a pronounced effect on the fluorescence decay curves, it does not prevent the precise estimation of R_E. d itself can be precisely estimated up to 90% or 1 molar excess of the donor only dye in the FRET sample. Simulations of the later experiments in which I directly measure R_E

in FG-Nups yielded similar conclusions to the simulation of DNA. I demonstrated that despite using end-to-end distance distribution such as the Ideal Gaussian it is possible to precisely estimate R_E and d. d was estimated least precisely at largest distances, which may be due to the large width of the distribution. I therefore concluded that both the experimental and the analysis pipelines should be well-suited to quantitatively measure end-to-end distances in FG-Nups.

3.3 FLIM-FRET imaging of FG-Nups *in situ* and applying polymer scaling laws

After validating the use of FLIM-FRET in the confocal scanning setup built in-house and establishing the best analysis pipeline, I now moved to imaging FG-Nups *in situ*. The aim of the experiments presented here was to fit precise end-to-end distances from FLIM-FRET to polymer scaling laws to determine how expanded or collapsed FG-Nups are in the NPC.

In section 3.3.1, I will present the experimental design and in section 3.3.2 the initial experimental evidence of scaling using the direct comparison of FRET efficiencies of the click-labelled FRET Amber mutants of Nup62 after acceptor photobleaching. In section 3.3.3 I will show how I determined the essential fitting parameters for the exponential fit models using donor lifetimes. In section 3.3.4, I will show all lifetime curves from the Amber FRET mutants side by side for a qualitative assessments of the data and in the final section 3.3.5, I will show quantitative results from the exponential fits. The results from the fits are end-to-end distances, which I will then plot against the number of residues between the two dyes to calculate the scaling exponent using scaling laws. The resulting exponent will be used to derive the conformational state of the FG-Nups *in situ*.

3.3.1 Experimental design

A systematic design of FRET mutants involved one Amber site being always kept the same and one varying with increasing spacing between the two. In the current system labelling is random, meaning I do not have control of which Amber site has which dye, since both sites have the same Amber codon and rely on UAA incorporation. For simplicity I graphically depicted what the sample looks like in Figure 3.13 (a). In FLIM-FRET I first measure the donor samples, labelled with only the donor dye (Atto532-H-tet). This is done to determine the necessary parameters for fitting. I then measure FLIM-FRET of all Amber FRET samples with the varying number of amino acids between the Amber sites (N_{res}aa) and calculate the mean of the end-to-end distance distribution from the fit (R_E). Taking calculated R_E for all samples I plot them in terms of N_{res} and fit the trend with a power law (Figure 3.13 (b)). The
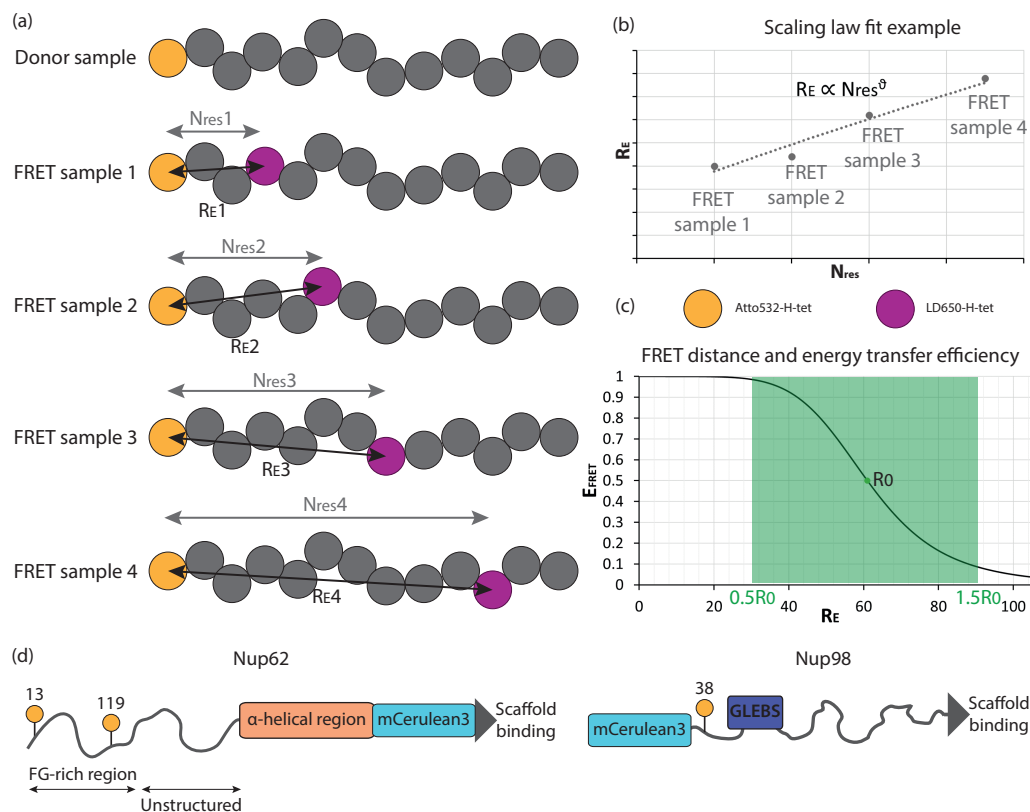
FIGURE 3.13: Experimental design. (a) schematic view of polypeptide chains with grey spheres representing native amino acids and coloured spheres - Amber sites with integrated UAAs and labelled with Atto532-H-tet(yellow) and LD650-H-tet(magenta) dyes. Gray arrows represent relative number of amino acids between different FRET samples (with two Amber sites) and black arrows show relative-end-to-end distances; (b) schematic scaling law fit of mean R_E vs N_{res}; (c) Plot of FRET efficiency vs R_E with green spot being $R_E = R_0$ at $E_{FRET} = 0.5$ and green area showing the range for the reliable range of end-to-end distances, where $0.5R_0 < R_E < 1.5R_0$; (d) schematic representation of the Nup62 and Nup98 proteins tagged with mCerulean3 at the site of scaffold integration for Nup62 and at the N-terminus for Nup98 with yellow spheres showing the starting sites for other FRET samples. GLEBS structured domain of Nup98 is shown in blue and α-helical domain of Nup62 in orange.

power law then gives us a value for an exponent ν, which represents the conformational state of the polymer according to the scaling law theory.

In all experiments presented here, I used Atto532-H-tet and LD650-H-tet dyes. R_0 calculated by equation 1.6 is the main determinant of the range of distances that these dyes can precisely estimate. From the FRET efficiency to end-to-end distance plot (Figure 3.13(c)) for these dyes I can quantify this range (green area) to be in the limits of 0.5 and 1.5 x R_0. I, therefore, expect to reliably measure distances between 3 and 9nm with these two dyes.

Since I saw that EGFP gets excited by the 532nm laser and leaks into the FRET channel I decided to replace it with another more blue-shifted fluorescent protein mCerulean3. I have since observed no visible contribution of the fluorescent protein to our lifetime measurements. all the FRET mutants I made are therefore tagged with mCerulean3. Figure 3.13(d) represents the schematic view of the proteins Nup62

and Nup98 used in this study. Yellow spheres indicate sites which I picked to be 'the Donor site', from which I varied the Amber positions. In Nup62 I picked two sites as for site 13 I have previously observed some signal in the -UAA control, even though labelling was clearly specific (Figure 3.2). Consequently, I measured distances in two different regions in Nup62 and in one region but two different cellular localisations (nuclear rim NPCs and GLFG bodies) in Nup98. The two regions in the Nup62 are marked in Figure 3.13(d) I will be referring to as FG-rich region (13+) containing fifteen FG repeats, of which five are FxFG, and unstructured region (119+) containing only five FG-repeats, of which one is FxFG (Figure 1.2). Nup98 region (38+) will be the FG-rich region up to the GLEBS domain (Figures 3.13(d) and 1.2).

3.3.2 Acceptor photobleaching FRET experiments: the first indication of scaling law in the NPC

Before proceeding with the complex pipeline of using FLIM-FRET to measure end-to-end distances to ultimately fit them with scaling laws, I decided to check whether the experimental system produces visible FRET changes between different FRET mutants of Nup62 (unstructured region). To test this, I used the double Amber codon FRET constructs, expressed and labelled in both HEK293T and COS-7 cell lines. Using the FRET AB module of the confocal Leica SP8 (LAS AF) microscope I imaged the samples before photobleaching (both donor and acceptor channels (Figure 3.14 top row)), selected rim areas (2 to 3 per image) using the hand draw tool (Figure 3.14 top row, last image), photobleached the acceptor channel using maximum power of the 660nm laser for 2min and acquired the image again. In the resulting image of the donor channel (Figure 3.14 second row), one can clearly see signal increase exactly in the areas selected for photobleaching, indicating that donor fluorescence was quenched by the acceptor in FRET and is now unquenched due to the absence of acceptor. Figure 3.14 (second row, last image) shows a resulting FRET efficiency image with FRET clearly only taking place in the nuclear rim.

Repeating the experiment multiple times and using different mutants of variable spacing between the two Amber sites (N_{res}) I observed a clear decrease in mean FRET efficiency (E_{FRET}) after about 88 amino acid (aa) spacing in the Nup62 (Figure 3.14 last two rows). Here I show a bar plot and a violin plot side by side as in bar plot it is perhaps easier to see the differences in the mean, although violin plot is a statistically preferred way of plotting data. The coloured area of the violin plot shows the density of the data distribution, individual points are data points calculated per cell, the grey bar shows the interquartile range, the white circle shows the median of the data and the coloured bar the mean. In the next figures in this section all comparative plots of end-to-end distances will be violin plots. Note that here I am looking at single values of FRET efficiency calculated according to equation 1.4. However, it is known that in a flexible polymer like Nup a distribution of end-to-end distances

and FRET efficiencies exists. The comparison between the FRET efficiencies is therefore rather qualitative in this experiment. Another interesting feature apparent from the plots is that mutants with spacing less than 88aa also show a decrease or similar signal compared to N_{res}=88. This might either be explained by an actual structural feature such as a loop, or by the insensitivity of FRET in this regime. Nonetheless, this was the first evidence that FRET differences with the varying number of amino acids between the two Amber sites can be distinguished. I therefore concluded that with more sensitive measurements I could directly compare end-to-end distances to fit into scaling law.

3.3.3   Calculating fitting parameters from donor lifetimes

Before beginning experiments I verified the nature of the exponential fluorescence decay of Donor dye in the cellular environment. I noted that dye reacted with BCN in contrast to being in solution (Figure 3.8) is clearly double exponential in the cell (Figure 3.15). Moreover, there is no difference between the donor lifetimes of the dye specifically labelling the FG-Nups and unspecifically sticking in the nucleus. Therefore the second component most likely comes from the environment the dyes are exposed to rather than scattering or autofluorescence in the cell, since those generate different lifetime values compared to around 1ns. Furthermore, judging from the relative amplitudes (A) between the two components it is also apparent that the second component has a large enough contribution (40%) to be considered coming from the dye itself. Experiments on different days with different samples also always yield the same values. Therefore, I decided to use the two components with their relative contributions for the final data analysis. Figure 3.15 shows the double exponential fits created using SymPhoTime between different samples using global fitting and the resulting values of the two exponential components $\tau 1$ and $\tau 2$ and their ratio A1.

3.3.4   Qualitative check of individual lifetime curves

Before quantifying all the collected fluorescence decays I wanted to first qualitatively assess whether there are any visible differences. This would give an indication whether the whole established experimental pipeline is sensitive enough to FRET changes. From the data I can see that all the 'donor in presence of acceptor' curves are slightly under the donor only curve, meaning the presence of acceptor dye in the vicinity of the donor dye does have a small effect on the donor environment. The decay curves I get from nuclear envelopes (Figures 3.19,3.17) are visually significantly more quenched than the curves collected from inside the nuclei, where dyes tend to stick (Figure 3.18). In the nuclei all curves, independent of the FRET mutants almost exactly overlap with the 'donor in presence of acceptor' curve, which is expected if
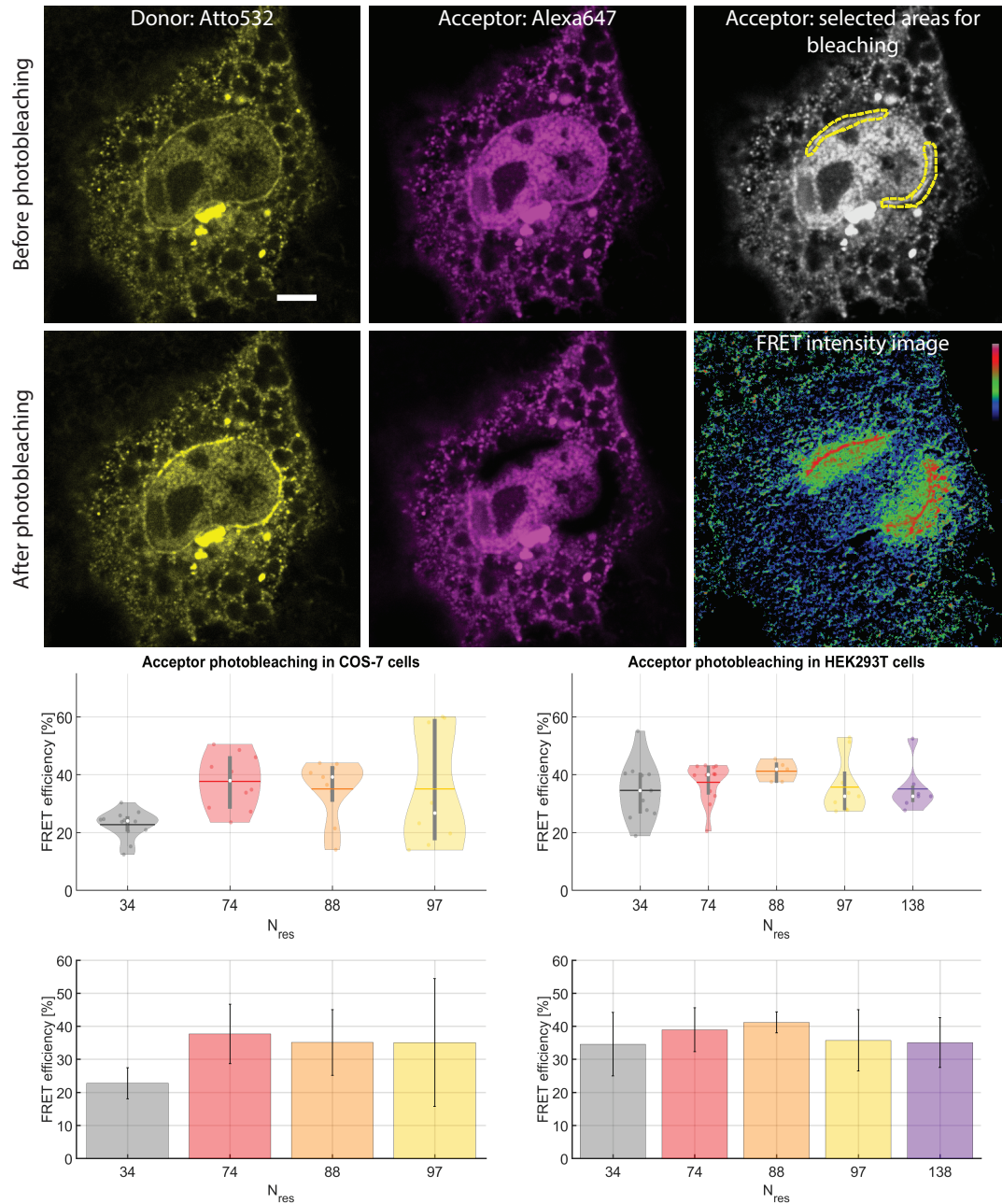
FIGURE 3.14: Acceptor photobleaching of FG-Nups in cells. The two rows of images demonstrate representative intensity confocal images of the donor channel (Atto532) and the acceptor channel (Alexa647) before (first row) and after (second row) the photobleaching in COS-7 cells. The last column shows selected are in the rim for photobleaching and the resulting FRET intensity image calculated by the Leica SP8 software. The last two rows of the figure shows all calculated FRET efficiencies colour-coded according to the mutant (with different N_{res}) summarised in two violin and bar plots for COS-7 cells and HEK293T cells respectively. Scale bar is $5\mu m$

no FRET is taking place. This small observation already confirms that the measurements in the nuclear rims are most likely coming from the specific labelling of the targeted Nup.

Individual fluorescence decay curves of Nup98 in Figure 3.16 and 3.17 show the
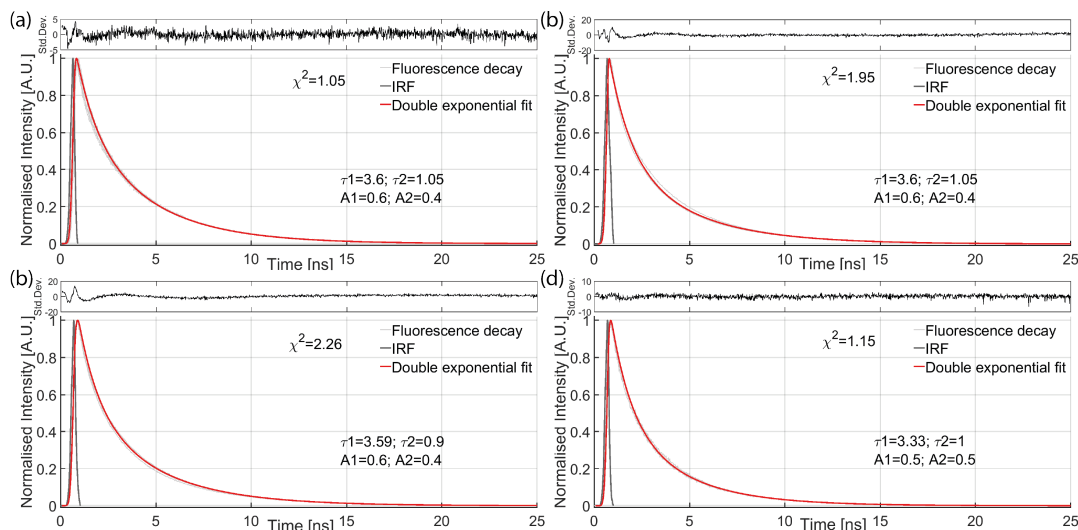
FIGURE 3.15: Fitted fluorescence decay curves of Nup62 and Nup98, labelled only with the donor dye. (a) shows Nup62 Amber mutant 119; (b) Nup62 Amber mutant 13; (c) Nup98 Amber mutant 38 measured in the nuclear rim; (d) Nup98 Amber mutant 38 measured in the GLFG bodies

FIGURE 3.16: Averaged fluorescence decay curves from nuclear rims of COS-7 cells expressing FRET mutants of Nup98. FRET mutants are colour-coded in agreement with Figure 3.22.

expected trend, where the curve of the highest FRET or smallest spacing ($N_{res} = 63$) being most quenched or most distant from the donor curve and the curve of smallest FRET or largest spacing ($N_{res} = 155$) being least quenched and least distant from the donor. One important thing to mention in this assessment is the standard deviation.

FIGURE 3.17: Averaged fluorescence decay curves from GLFG bodies of COS-7 cells expressing FRET mutants of Nup98. FRET mutants are colour-coded in agreement with Figure 3.23

Here the bold curves are the averaged curves over multiple experiments with the shaded area representing the standard deviation. It is apparent that the standard deviation for all fluorescence decay curves is very large and will also be reflected in the quantitative measurements of end-to-end distances.

In Figure 3.19 there is a clear trend of decreasing quenching from the FRET mutant with the smallest spacing ($N_{res} = 62$) to the FRET mutant with 64 more amino acids ($N_{res} = 126$), however after that the most distant mutants ($N_{res} = 138, 173$) still appear quenched. This is again comparing mean curves directly, the differences in general are much harder to see taking into account the large standard deviation.

Overall, the direct comparison of the fluorescence decay curves indicates that I will see a trend of differences in FRET signal with varying spacing between the donor and acceptor dyes but the variance of the measurements is expected to be very large. According to the simulation in Figure 3.12 the large variance and the curves outside the expected trend might be due to variations in the fraction of the donor species per sample, however it should not interfere dramatically with a quantitative measurement of R_E.
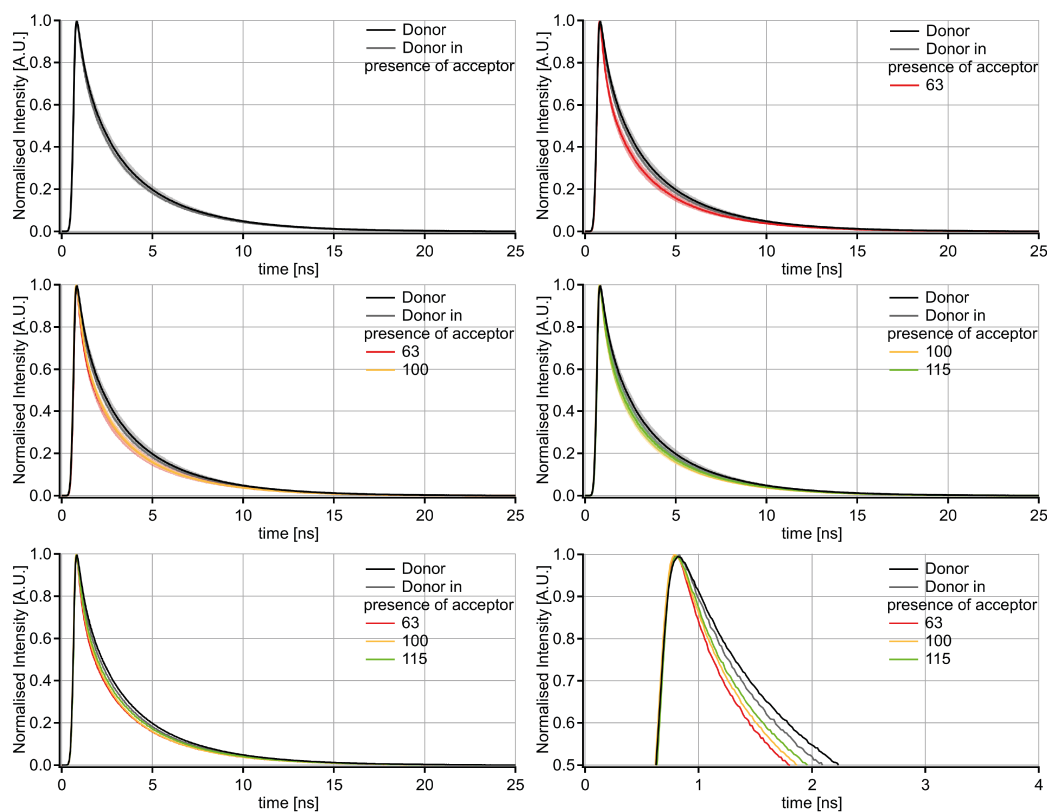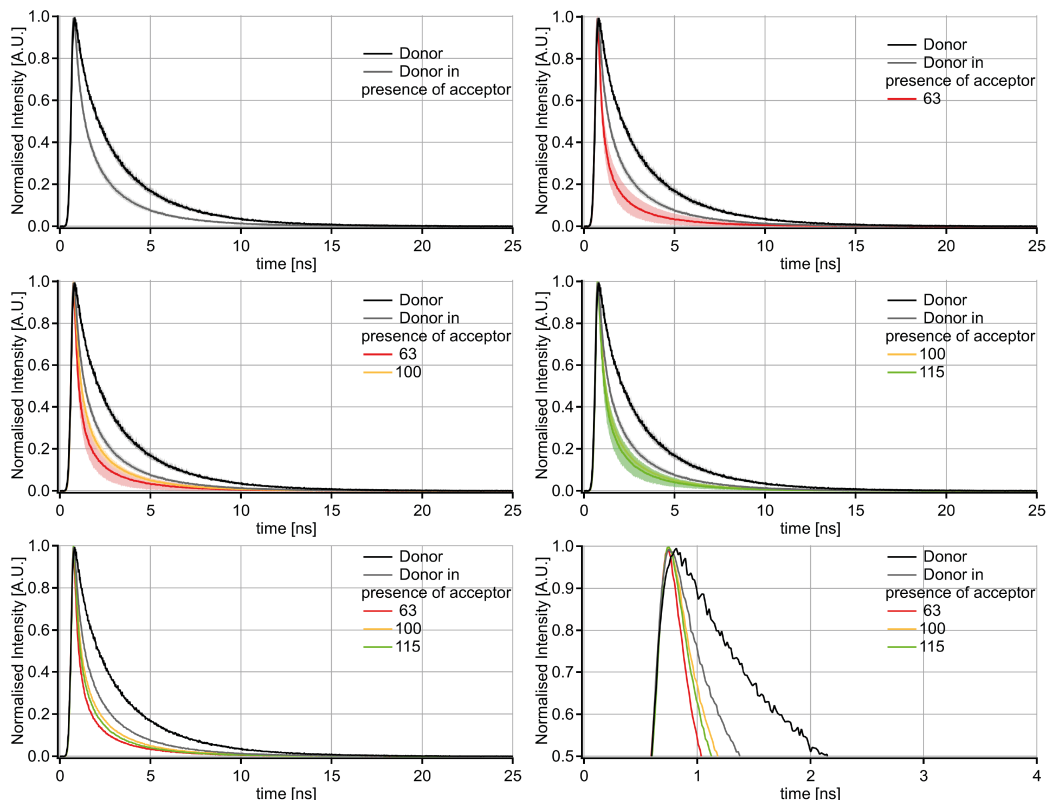
FIGURE 3.18: Averaged fluorescence decay curves from nuclei of COS-7 cells expressing FRET mutants of Nup62. FRET mutants are colour-coded in agreement with Figure 3.20.

3.3.5 Ideal Gaussian chain model fit of lifetime data

Nup62

The most extensive studies for the FLIM-FRET were done with Nup62 unstructured region. As mentioned before in section 3.1.1 this is due to the bypassing of the Amber site by the ribosome at earlier occurring amino acid sites. Site 119, however, provided a very clean + and - UAA comparison. I created 10 mutants from 34 to 173 amino acid spacing up to the structured region of Nup62. I then sequentially imaged all mutants with the selected pair of dyes (Atto532 and LD650) and recorded their TCSPC signals. I then fitted all resulting fluorescence lifetime curves for each of the nuclear rims using PAM software to directly calculate end-to-end distances using the Ideal Gaussian chain distribution fit (equations are presented in section 1.4.3. The results are summarised in Figure 3.20. Means of the distribution of all data (coloured bars in Figure 3.20(a)) were then taken for the scaling law fitting with the equation 1.24. The fitting itself, even though looks good in Figure 3.20(b), is not very robust as the coefficient of determination R^2 (R^2 is the square of the correlation between the real and the predicted values and is from 0 to 1, where 1 indicates a greater proportion of variance is accounted by the model) is only 0.57. However, one can also note that our *in vivo* measurements are too widely distributed (coloured areas in Figure 3.20(a)). This can be most likely explained by the biological variability between cells, labelling quality between the samples, high background levels (e.g.
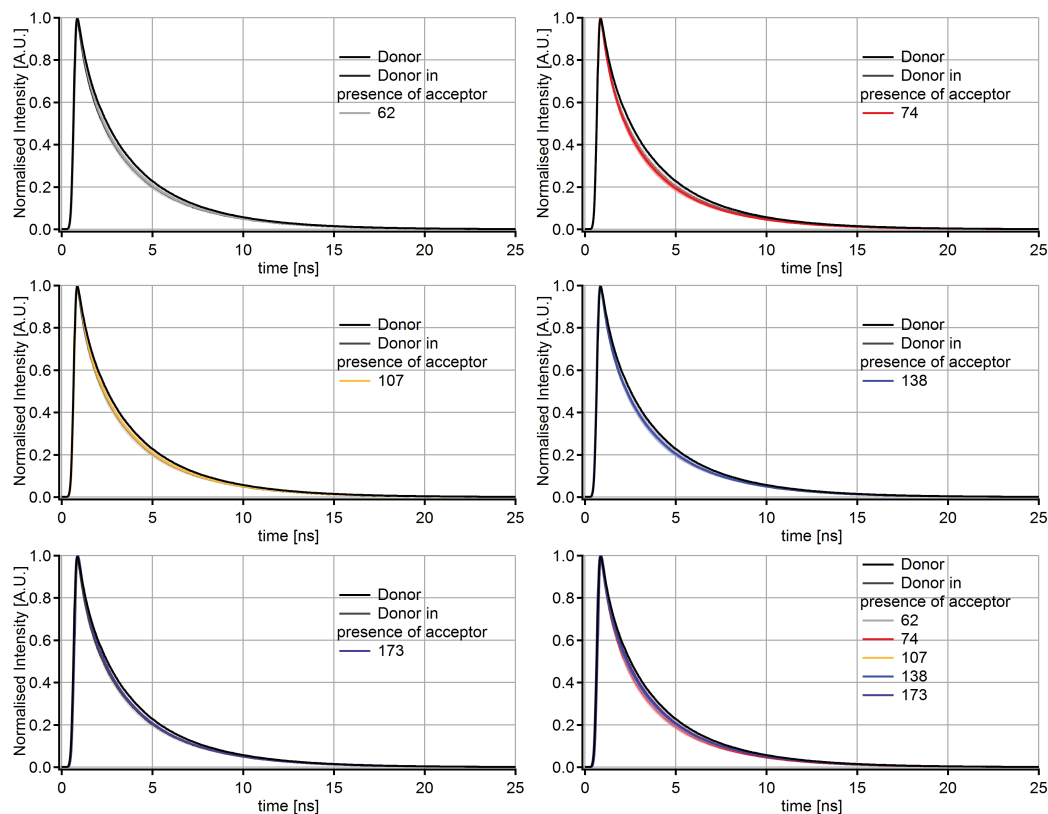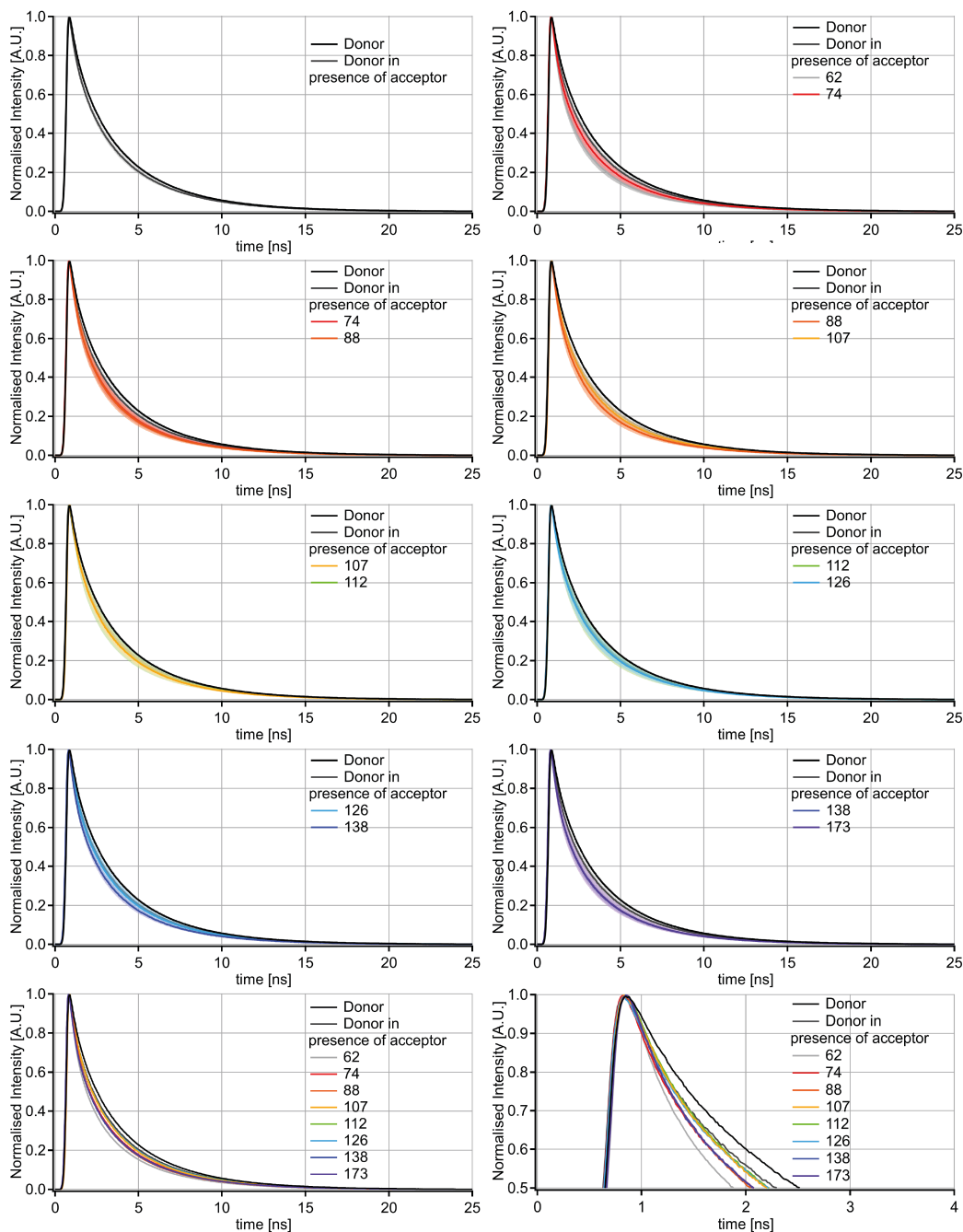
FIGURE 3.19: Averaged fluorescence decay curves from nuclear rims of COS-7 cells expressing FRET mutants of Nup62. FRET mutants are colour-coded in agreement with Figure 3.20.

dye sticking) or potentially high variability of FG-Nups conformations between different NPCs (transporting large cargoes like ribosomes *versus* "inactive" NPCs). To narrow down the distribution and more precisely determine the valid end-to-end distance per mutant one would require a larger data set or a more direct way of measuring FRET in the NPC, for example, by only imaging the bottom of the nuclei with clear foci indicating individual NPCs. Nonetheless, the scaling law fit indicates in what range I expect the Flory exponent to be. For Nup62 unstructured region, the exponent is 0.45, which suggests it is resembles a collapsed coil.
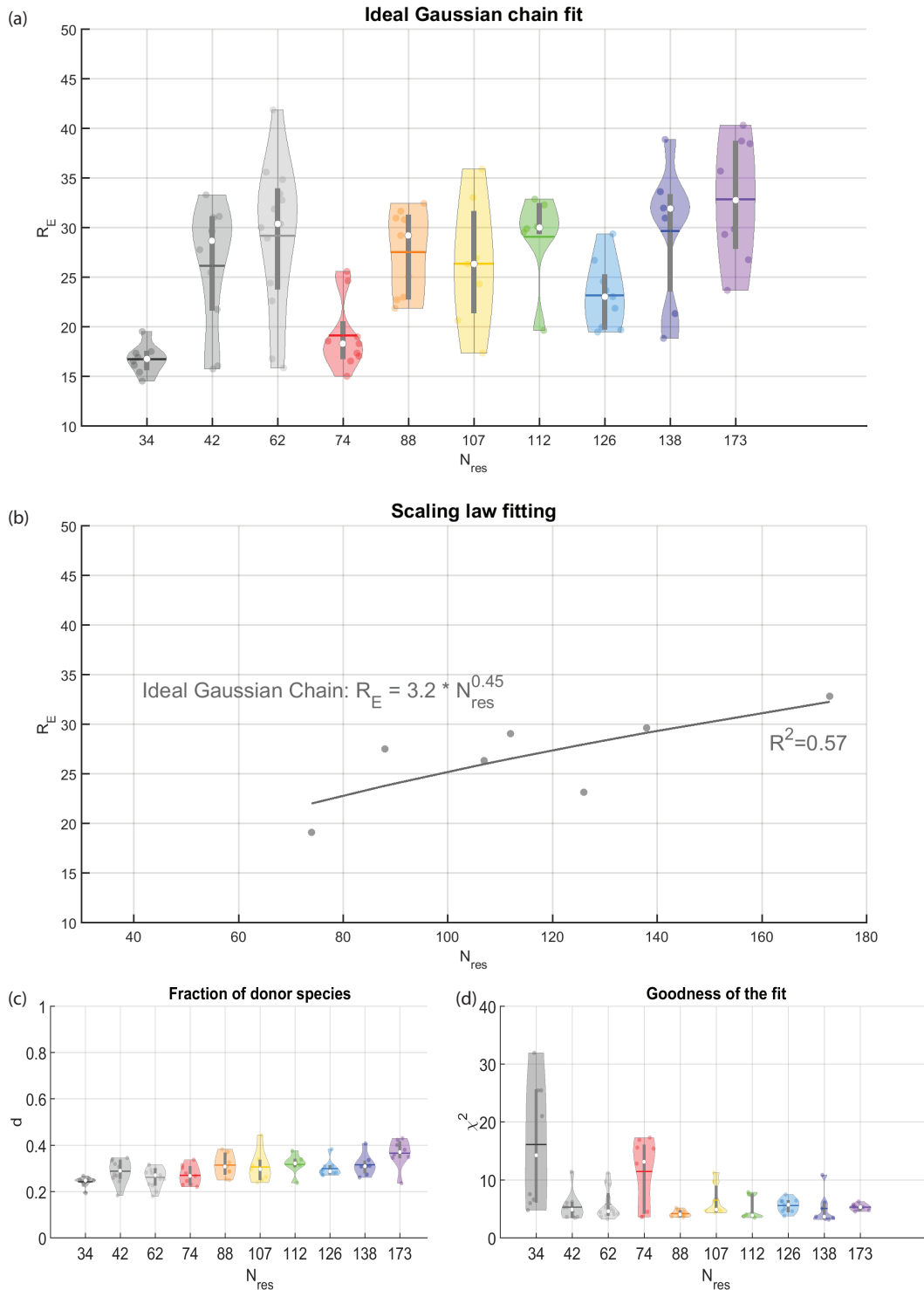
FIGURE 3.20: Nup62: unstructured region. (a) violin plots of all calculated end-to-end distances colour-coded according to the mutant (with different N_{res}). The coloured area of the violin plot shows the density of the data distribution, individual points are data points calculated per cell, grey bar shows interquartile range, the white circle shows the median of the data and the coloured bar the mean. (b) Fitting of the mean end-to-end distances with scaling law. R^2 demonstrates the goodness of the fit and the equation shows fitting parameters. (c) and (d) are violin plots of donor fraction d and goodness of the end-to-end distances fit χ^2 distributions, colour-coded according to N_{res}.

FIGURE 3.21: Nup62: FG-rich region. (a) violin plots of all calculated end-to-end distances colour-coded according to the mutant (with different N_{res}); (b) fitting of the mean end-to-end distances with scaling law. R^2 demonstrates the goodness of the fit and the equation shows fitting parameters; (c) and (d) violin plots of donor fraction d and goodness of the end-to-end distances fit χ^2 distributions, colour-coded according to N_{res}.

Fitting in general works well for my data set with the exception of mutants that have $N_{res} < 88$, as previously observed in acceptor photobleaching experiments (section 3.3.2). This is most likely due to the poor fitting in this regime as indicated in Figure 3.20(d) and also in the width of the distribution in Figure 3.20(a). I hypothesise that in this range I have very high FRET that is undetectable and unfittable in *in vivo* studies. Fluorescence decay curve of FRET mutant with $N_{res} = 62$ in Figure 3.19) support this hypothesis as the fluorescence decay curve of this mutant is expected tho correspond to the highest FRET signal as it appears most quenched. This is unexpected as many *in vitro* studies using smFRET have used such short fragments and reported FRET efficiencies upwards of 70% [175, 170], so the high FRET in *in vitro* studies is clearly not problematic. However, NPC is a highly crowded environment and it could be that in such conditions one can not quantify FRET in the high-efficiency regime.

To complement and extend the spacing between the FRET sites I also started to work on the FG-rich region of Nup62. Since I have previously shown specific labelling at the site 13 Amber mutant (Figure 3.2), I proceeded to use this site for varying spacing to a larger range than in the unstructured region, limited by the presence of the α-helical region past residue 300. Unfortunately, there are even fewer data points collected for this region but it becomes apparent that the trend is very similar to

the unstructured region. Again, in the high FRET regime ($N_{res} < 88$) the distribution width (Figure 3.21(a)) and the χ^2 (Figure 3.21(d)) are very large but there is a clear increase in mean end-to-end distance of mutants with longer spacing between the FRET sites (Figure 3.21(a)). Donor only fraction in these experiments also remains around 30% on average, though it is higher for the two low FRET mutants $N_{res} = 168$ and 218 (Figure 3.21). Since in these experiments I don't have many points and the first three points are most likely invalid, fitting with scaling laws delivers a significantly lower exponent value 0.32, even though the fit converges better than for the unstructured region (Figure 3.21(b)). Surprisingly, I also do not measure much higher values of end-to-end distance in the $N_{res} = 218$ mutant compared to the $N_{res} = 173$ mutant in the unstructured region. I would therefore need to validate that the Amber suppression system works properly in the FRET mutants in this range.

Nup98

Extending from Nup62 and seeing a clear trend of end-to-end distances increasing with increasing spacing between the dyes, I began working on Nup98. I used three mutants with $N_{res} = 63, 100, 115$ to measure end-to-end distances in both nuclear envelopes and the GLFG bodies (which I distinguished using image segmentation). I hypothesised that NPC's FG-phase might be similar to that of a liquid-like or gel-like structure such as the GLFG body, state of which is currently underresearched. Moreover, I wondered whether I would see quantitative differences in the conformational state of the Nup98 compared to Nup62.

Figures 3.22 and 3.23 summarise the findings. Once again, I suffer from large data variance and the small number of measurements. Yet in both cases (Nup98 nuclear rim and Nup98 GLFG bodies) there is again a clear increase in mean end-to-end distances with increasing spacing between the amino acids. I also fit both data sets with scaling law to retrieve preliminary scaling exponents of 0.37 and 0.46 for the nuclear rims and GLFG bodies respectively (Figures 3.22(b) and 3.23(b)). Fractions of donor species are again around 0.3, though for GLFG bodies data they are significantly lower - around 0.1. The curves are also significantly different even for donor only species as shown previously (Figures 3.15). GLFG bodies show a larger scaling exponent of 0.46, suggesting that the FG-Nups in such a liquid-like state may be more expanded than in the NPC.

In summary, my research results applying the FLIM-FRET approach and polymer scaling laws to derive the inner structure of the FG-Nups *in situ* demonstrate that FG-Nups accept a more collapsed conformation inside the NPC. My *in situ* measurements are sensitive enough to resolve different end-to-end distances in the nuclear rims and the FRET signal is, as expected, specifically detectable from the rims
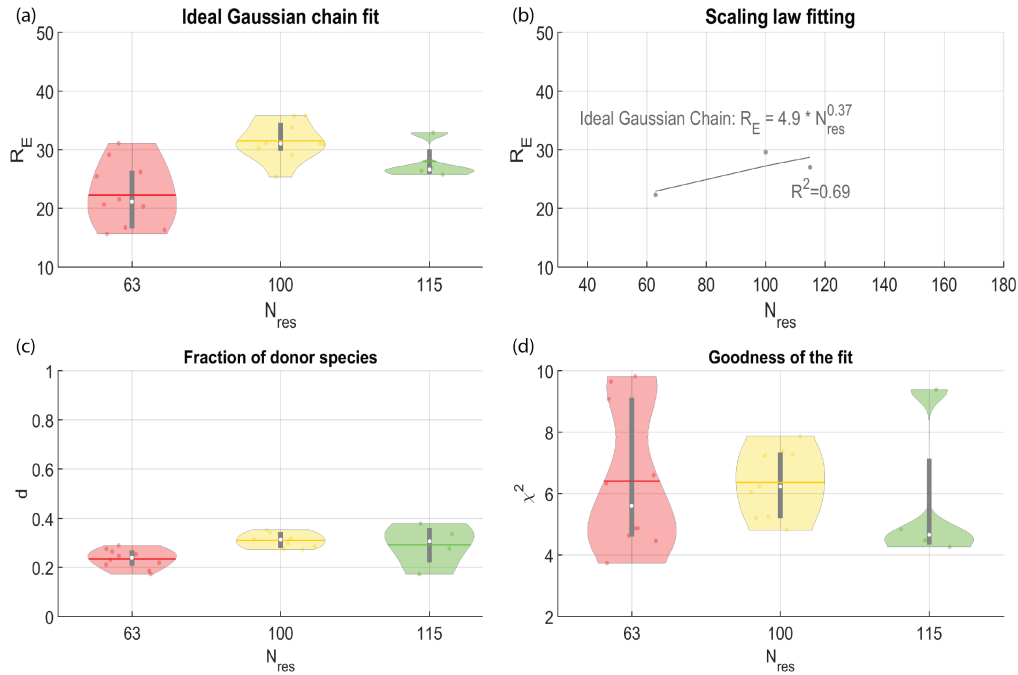
FIGURE 3.22: Nup98: nuclear rims analysis.(a) violin plots of all calculated end-to-end distances colour-coded according to the mutant (with different N_{res}); (b) fitting of the mean end-to-end distances with scaling law. R^2 demonstrates the goodness of the fit and the equation shows fitting parameters; (c) and (d) violin plots of donor fraction d and goodness of the end-to-end distances fit χ^2 distributions, colour-coded according to N_{res}.
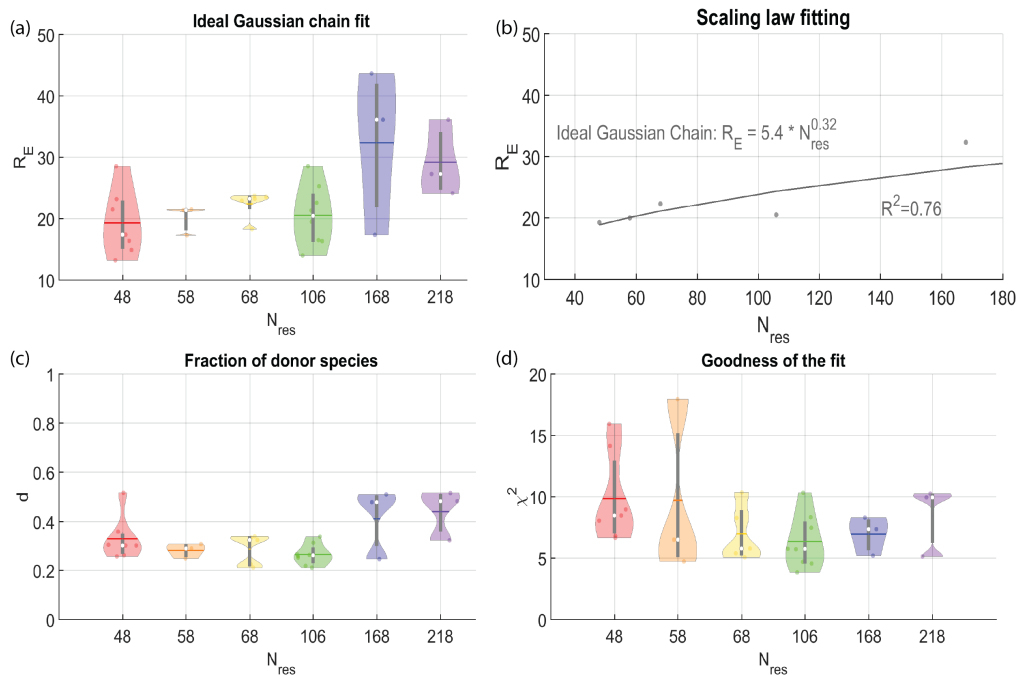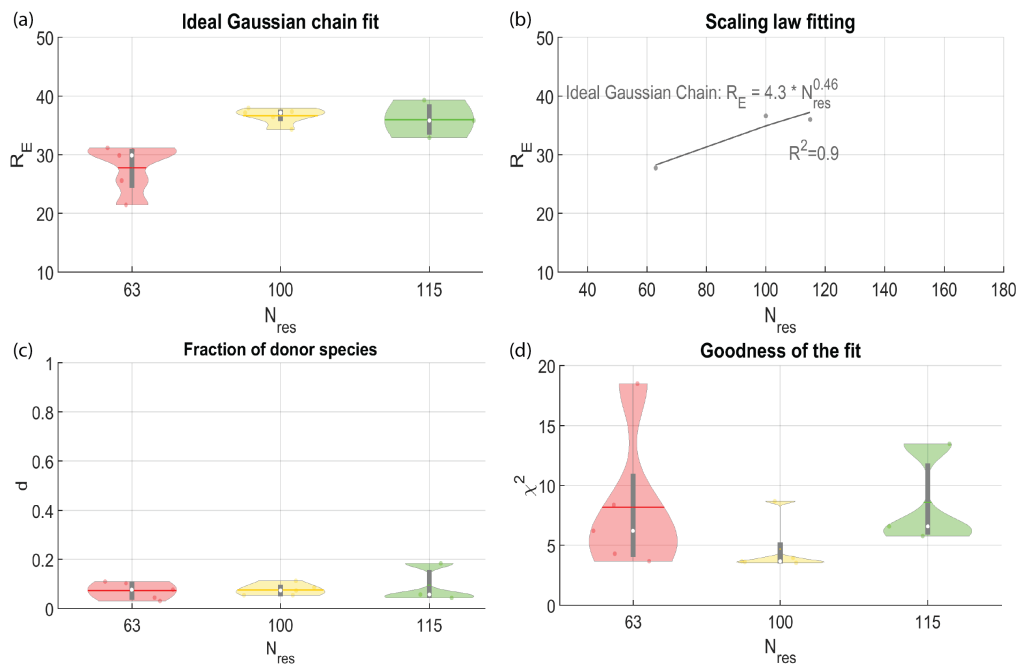

FIGURE 3.23: Nup98: GLFG bodies analysis.(a) violin plots of all calculated end-to-end distances colour-coded according to the mutant (with different N_{res}); (b) fitting of the mean end-to-end distances with scaling law. R^2 demonstrates the goodness of the fit and the equation shows fitting parameters; (c) and (d) violin plots of donor fraction d and goodness of the end-to-end distances fit χ^2 distributions, colour-coded according to N_{res}.

and not the nucleus. I also show that I can extend my measurements from the unstructured region of Nup62 to the FG-region and also to a different Nup - Nup98. Moreover, I attempt to describe the differences in Nup98 conformation in the NPC in contrast to that in a liquid-like GLFG body, in which I see it conform to a more extended configuration. Overall, my results prove that scaling of FG-Nups, with an exponent between 0.32 and 0.46, can be detected *in situ* inside the NPCs. Further measurements are required for more quantitative and precise estimation of the scaling exponent.

4 Discussion

FG-Nups are the essential components that form the selective permeability barrier of the NPC. Nonetheless, they are among the least well understood protein assemblies of the nucleocytoplasmic transport machinery. Being intrinsically disordered, FG-Nups are highly dynamic and flexible. Multiple models attempt to describe the assembly and collective behaviour of FG-Nups but still do not converge on a solution. Since IDPs are heteropolymers, most models consider the two extreme behaviours of polymers. The polymer brush model assumes the FG-Nups to be mostly governed by entropic forces, therefore being expanded in solution, while the hydrogel model posits that the FG-Nups interact with each forming a sieve-like structure.

FG-Nups can be considered polymer-like and therefore described by principles of polymer physics. In simple terms, polymer physics considers proteins as a set of functional units or amino acids connected in a chain that behave according to physical laws derived from statistical mechanics and thermodynamics. One of the general principles derived by Flory suggests that there is universal scaling between the conformational state of a chain and the number of functional units (scaling law). In the highly dynamic structural assemblies typical for disordered proteins, scaling behaviour is a simple principle that has been successfully applied to explain complex properties and behaviour of IDPs in isolation and to predict their behaviour in supramolecular assembly states. Nonetheless, even *in vitro*, the scaling behaviour is still hard to measure precisely and debated in many cases [178].

The main goal of my research was to develop fluorescence microscopy imaging approaches to study the conformational state of FG-Nups *in situ* from a polymer perspective. I focused on two techniques. The first was a dual-colour super-resolution microscopy approach to map different positions in the FG-domain of Nup62. This super-resolution microscopy approach was largely based on Amber suppression and click-labelling. In section 4.0.1, I will discuss both the advantages and drawbacks of this system. The second technique was the FLIM-FRET approach that covers a complimentary higher resolution range. It is also based on Amber suppression and click-labelling, but I extended it to two Amber site dual labelling with donor and acceptor dyes which provided an intramolecular readout of dye-to-dye distance. In section 4.0.2, I will summarise and discuss the results and advantages of this methodology. My last aim was to use general principles of polymer physics and, more specifically,

scaling concepts to apply these principles to derive information about the conformational state of FG-Nups based on the measurements from FLIM-FRET. Finally, in section 4.0.3, I will discuss my conclusions and provide an outlook on further applications.

4.0.1   Site-specific super-resolution imaging approach to mapping FG-Nups conformation

The nuclear pore complex is a well defined structure of eight-fold symmetry. Its dimensions and composition are very well known from cryo-electron microscopy, X-ray crystallography and proteomics [202, 13]. The eight fold symmetry and ring-like structure cannot be resolved by conventional light microscopy, but falls into the resolution range of most modern super-resolution microscopy techniques. It therefore became a model system for validation of labelling technologies and experimental super-resolution microscopy conditions. There are multiple labelling techniques available ranging from specific antibodies to stable cell lines with genetically encoded GFP or SNAP-tag at termini of multiple structured nucleoporins [118, 119]. In combination with high resolution structural approaches, this makes it possible to resolve the positions of the structured regions of the FG-Nups in the scaffold of the NPC. The disordered regions in between the termini of FG-Nups, however, remain largely unaddressed.

Amber suppression combined with click-labelling are unique methods that provide an opportunity to site-specifically label the FG-Nups at an exact position in the protein with an organic fluorophore. Since the exact localisation of FG-domains within the NPC is unclear and the site-specific click-labelling methodology does not allow for complete labelling and incorporation of the POI, it is essential to use a reference in the NPC to map the relative localisation of the labelled position. Therefore a dual-colour super-resolution imaging approach was developed. The method was first established in our research group by Dr Jun Hee Kang [182]. The method involved labelling the centre of the NPC with WGA and then measuring relative distances of the labelled position on the FG-Nup in another colour. Motivated by the results of mapping Nup98 using this approach, I extended this method to another major player in the central channel - Nup62, and developed an improved approach to mapping the FG positions with respect to the scaffold.

Amber suppression is the most versatile for site-specific protein labelling in mammalian cells. The two main limitations of this technique are the need for transient expression of the Amber suppression system in mammalian cells for high yielding expression [203] and the presence of truncated proteins. The truncated proteins are produced in either of two scenarios: (1) when the natural termination machinery wins against the suppressor tRNA so that the protein is terminated at the STOP codon (N-terminal truncation) or (2) when the ribosome re-initiates after the STOP

codon of choice, at an alternative a 3′ start codon after the targeted Amber site (C-terminal truncation). The poor understanding of this mechanism prevented improvements in the experimental design but was at least easily detected experimentally so that I could develop more suitable mutants, minimising above mentioned limitations.

Working with a transiently expressed system on one hand provides the benefit of having plenty of protein containing an UAA to replace the endogenous protein and therefore provide more labelling sites in the NPC. On the other hand, too much expression of in particular FG-Nups also leads to high background and higher variability between the samples. Creation of truncated protein only limits the choice of the protein. For example, FG-Nups that contain a structured NPC-targeting domain on the N-terminus and the Amber suppression site in the FG-region in the C-terminus might not all be degraded and some might incorporate into the NPC, thus blocking the successfully Amber suppressed full length FG-Nup from association with the pore. This problem is solved by primarily focusing on the FG-Nups that have the pore-targeting domain in the C-terminus and would not be incorporated unless Amber suppression is successful (in this study Nup62 and Nup98).

Taking this limitation into account I choose from C-terminally targeted permeability barrier relevant FG-Nups and focused on Nup98 and Nup62. Nup98 has the limitation that a GFP fusion could only be attached at the N-terminus. Typically such a GFP introduced at the C-terminus of the POI serves as a marker of success of Amber suppression incorporation of UAA and is of practical help when identifying proper cells under the microscope. Therefore, in a transiently transfected culture this made proper identification of cells containing the UAA in Nup98 much harder. Nup62 can be modified at the C-terminus and Amber suppression worked at almost all chosen sites. It therefore became my primary choice for the super-resolution microscopy approach. I first successfully resolved the two C-termini Amber mutants around the structural domain of Nup62 (287 and 515). This gave me confidence that Nup62 is a good choice of the Nup and that it is presumably incorporated into the NPC. My results of using one of these mutants together with GFP-nanobody showed the co-localisation of the two structures. This proved the success of Amber suppression at that site. To my surprise, my results of super-resolution images of the two N-termini labels of Nup62 (13 and 61) relative to the central marker WGA even under such high background showed that there is no perfect co-localisation of the Nup62 and WGA signals. I would have expected the N-termini of Nup62 to be exactly at the centre [118, 122] and also to be O-GlcNAcylated, therefore labelled by WGA. Comparing my results of the N-termini labelled Nup62 to the unpublished results from Nup98 (N-termini mutants 6 and 38) [182], I also observe that N-termini of Nup98 do not colocalise well with the WGA channel. The distance measured for Nup98 was

20nm, which for Nup62 would mean its FG-domains are more collapsed and concentrated on the peripheries of the central channel. Since a good quality dual-colour super-resolution image with the GFP-nanobody required for a precise measurement was not achieved, this hypothesis remains open. From comparison of the images to Nup98 I would expect that the N-terminus of Nup62 would also be around 20nm away from the centre of the NPC.

Nup62 exists in two subcomplexes in mammalian cells - Nup62 in the central channel and Nup214 on the cytoplasmic side. Nup214 and Nup62 subcomplexes are heterogeneously distributed in different NPC populations. Additionally, the Nup214 structured part of the subcomplex has never been accurately resolved using high resolution structural microscopy methods [204, 13]. This heterogeneity may be regulated by O-GlcNAcylation, as Nup88, which is a member of the Nup214 subcomplex, only associates with O-GlcNAcylated Nup62 [27]. From my preliminary results from super-resolution microscopy most of the NPCs I imaged do not show co-localisation of the Amber suppression site at the N-terminus and the WGA signal. The Nup62 Amber mutants therefore most likely localise within the Nup62 subcomplex. More localisations and improved super-resolution images of Nup62 could provide a unique opportunity to classify NPCs based on Nup62 localisation and compare its conformation. Having successfully designed, labelled and imaged Nup62 using two-dimensional super-resolution microscopy the next step in development of this methodology would be to use three-dimensional super-resolution microscopy. Using another structure in the NPC for a reference one could this way potentially better resolve the localisation of recombinant Nup62 and even map its position in the Nup214 subcomplex.

Improvements of the method

My results from the super-resolution microscopy method indicate high unspecific background being the major drawback limiting the precise mapping of distances of Nup62. There are two main developments that can solve this issue: improvements in labelling strategy and improvement of Amber suppression efficiency and specificity.

The click-labelling strategy I used is based on a specific reaction between the cyclooctene in TCO* and a tetrazine conjugated to the dye. The labelling procedure I used is based on incubating permeabilised cells in a concentrated dye solution before fixing. I noticed that despite the washes after incubation with the dye high background persists, which has an severe effect on my measurements. This may originate in the tetrazine non-specifically reacting to other molecules present in cells, in contrast to e.g. potentially less reactive O6-benzylguanine or O6-benzylcytosine groups used in labelling SNAP- and CLIP-tags respectively. The main location of

dye sticking in cells is also, unfortunately, the nucleus. My results from super-resolution imaging of Nup62 clearly show that the background hugely contributes to the overall image (Figure 3.6). I tried to eliminate the background by using a DNA points accumulation for imaging in nanoscale topography (DNA-PAINT) technique instead. This technique uses DNA strands as complementary labels that can transiently bind and unbind, resulting in blinking necessary for localisation microscopy without being affected by photobleaching, therefore allowing for even longer image acquisitions [205, 206]. We extended the use of this technique to Amber suppression technology and click-labelling by demonstrating successful super-resolution images of Vimentin-mOrange(N116TAG) and GFP(N149TAG)-Nup153 [98]. Even though this technique worked well, it never worked in combination with immunostaining with DNA-functionalised nanobodies (via maleimide-labelling). This is potentially due to the fact that in DNA-PAINT it is impossible to check for labelling success before super-resolution imaging and we were never sure of the stability of the chemical association between the tetrazine group and the DNA that was commercially prepared.

Another way to improve the experimental conditions for super-resolution microscopy is to knock-out the endogenous nucleoporins. Unfortunately, both Nup98 and Nup62 are essential for cells, cannot be knocked out and are homozygosly lethal to the organism. Instead, I tried a knock down of Nup98 with shRNA lentivirus vectors, which was unsuccessful. I did not try different knock down methods available for Nup62, which could have significantly helped [204, 27, 207]. A company tasked with this reported that after six months of trying that they had to cancel the order because they repeatedly failed to achieve the desired result.

In the future this method can be significantly improved with rapid developments in new cell permeable dyes and Amber suppression technology. Even smaller tags will also benefit the precision in this project with GFP being replaced by an even smaller peptide tag, targeted by a specific nanobody, to also address any incorporation concerns [88]. Extending this method to three dimensions would also provide greater insights into the relative localisations of the FG-Nups in the NPC. This method also inspired me to undergo a different approach that is less sensitive to high background and led into the development of FLIM-FRET applied to studying conformations of disordered FG-Nups in the NPC.

4.0.2   Establishing the FLIM-FRET approach

Once it became clear that with the current experimental approach and available dyes it was not practical to obtain sufficiently reliable mapped positions from super-resolution studies, I turned to a different approach. The idea was to use FRET to directly measure distances between two sites of the FG-Nup rather than map dual-colour super-resolved positions. FRET is a fluorescence phenomena that relies on

proximity of the fluorescent labels and can precisely estimate the distance between them. FLIM is a very precise methodology that allows for a combination of quantitative FRET with imaging. The use of FLIM to measure FRET therefore seemed suitable.

Since the technology was not yet available in the research group I developed many parts of it myself. One of the major parts of my work went into establishing and validating experimental protocols for correct interpretation and analysis. To further progress, it was essential to establish factors such as finding a suitable dye pair (Section 3.2.1) and establishing suitable imaging conditions (Section 3.2.2), comparing the results from lifetime FRET measurements to "classic" intensity-based FRET (Section 3.2.3) and validating end-to-end distance measurements' precision and the analysis pipeline (Section 3.2.4). The dye pair is essential as the R_0 is defines the range of experimental end-to-end distances that can be probed reliably (discussed in Figure 3.13). Comparing lifetime measurements to intensity-based FRET measurements in cells provides a good reference for the functionality of the overall FLIM-FRET setup. Validation of end-to-end distance measurements was performed with a well known biological control - DNA nanorulers. This sample has exact dimensions and designed end-to-end distances and provides direct access to the ground truth. This sample was therefore the best candidate to validate the performance of the analysis software in measuring end-to-end distances directly.

The experiments performed in section 3.2 of the Results chapter led to the establishment of the ideal experimental conditions achievable for the conditions from the beginning (discussed in section 3.2.2) and verified in both the use of the set-up built in-house and the analysis strategy and procedure developed based on PAM software [165]. It is essential that the dye pair is stable under the imaging conditions to eliminate photophysical effects other than FRET (i.e. bleaching) and to determine the range of distances to be measured (from R_0 and discussed in Figure 3.13). Moreover, since in biology most of the donors are not mono exponential, I wanted to eliminate any stereochemical effects on the lifetime of the donor arising from the reaction of the UAA with the tetrazine (Figure 3.8) to make sure the dye itself is at least monoexponential. This would significantly reduce the complexity of the analysis as all lifetime components from the donor would most likely contribute to FRET.

Experimental validation using intermolecular FRET *in vivo*
The next logical step was to test the performance of the FLIM set up in measuring FRET in an established system *in vivo*. Since most FRET sensors used in cells are designed based on fluorescent proteins I also used one of the established systems based on rapamycin inducible dimerisation to measure intermolecular FRET. One of the simplest systems to implement was FKBP12-FRB dimerisation upon introduction of rapamycin, where one of them is anchored in a membrane. Such a system with

mTurquoise and yellow fluorescent protein was demonstrated in *Nicotiana benthami-
ana* plants, where it was also studied using FLIM-FRET [197]. I modified the system
for use in mammalian cells and used it to validate whether my fluorescence life-
time measurements *in vivo* would agree with intensity-based FRET measurements
in the same setup. The system was designed to measure FRET between GFP and
mCherry only upon dimerisation of the two transmembrane domains after addition
of rapamycin. FRET efficiency standard measured *in vitro* gives $E_{FRET} = 35\%$ [208],
which compares well to the measurements I performed in cells. More importantly,
the results from measurements of lifetime (equation 1.10) fully agreed with mea-
surements of intensity-based FRET (equation 1.3). This gave me confidence that the
experimental setup was sufficiently sensitive to detect FRET changes *in vivo*.

Experimental and analytical validation using DNA nanorulers
The final section 3.2.4 is the most important for the total validation of the experi-
mental and analytical procedures used in my research. The experiment I performed
is based on using pre-designed nanorulers of defined dimensions hybridised to con-
tain a donor and an acceptor on one piece of DNA [209, 210]. In such a sample,
after the hybridisation, FRET signal would only come from the formed DNA but
the freely diffusing donor-only sample (unhybridised donor oligo) would still be
present in solution. Since in a real sample even under perfect labelling conditions
there is always a number of donor dyes that do not undergo energy transfer, it is
important to distinguish this fraction from the FRET contribution. This experiment
was designed to test two things: to what extent this fraction affects the precise esti-
mation of end-to-end distances and whether the current analysis pipeline is sensitive
enough to precisely estimate it. From both my experiment and simulation I saw that
the fraction has an observable effect on the fluorescence lifetime curve (Figures 3.10,
3.11). Reassuringly, I could still directly measure end-to-end distances with high
precision. I also investigated the upper limit of the molar excess of the donor-only
fraction so that FRET can still be distinguished and found that even when 90% of
the donor species are present in the FRET sample one can still reliably measure end-
to-end distance and precisely estimate the fraction of the FRET-free donor (Figures
3.10, 3.11). This would translate statistically to more than three quarters of the la-
belling sites occupied by donor dyes. This experiment demonstrates that even in a
cell it is likely that the fraction of FG-Nups labelled only with the donor dye and not
undergoing FRET will not affect the absolute measurement of end-to-end distances.
Even though DNA was a great control in validating both the experimental and the
analysis procedures it is not an IDP and therefore does not account for the whole
complexity of the final experiment, which on top of that was also done *in situ*.

Analysis pipeline and validation
PAM was chosen for the analysis software as it is a large integrative platform that
not only has a lifetime analysis module but also an image analysis platform, which

I can use in the future for automated image segmentation and analysis. Moreover, it is open-source, written in easily accessible MATLAB software and has a vibrant github community helping to identify any issues and finding solutions for specific applications. It also already has simple Normal distribution implementation into the lifetime analysis for the monoexponential donor decay. We therefore decided to build the analysis pipeline for the purposes of the *in situ* experiment on this platform.

Before the final analysis of the images in the *in situ* experiments, I simulated experimental results I expected to find in FG-Nups and analysed them using a newly established PAM pipeline with integrated Ideal Gaussian chain distribution for a double exponential donor (equation 1.17, Figure 3.12). Similar to what I observed in the simulation of the DNA data, I noticed that the fraction d also has a pronounced effect on the lifetime curves' relative appearance and can therefore potentially mislead the conclusions about R_E made based on the curves' relative trend. More importantly, using the end-to-end distribution I could also recover the same input parameters for end-to-end distances and the fraction of the donor. This simulation therefore proved that the developed analysis pipeline is robust at direct measurements of end-to-end distances and could be applied to analyse *in situ* data.

Improvements in data analysis

The current data analysis yields satisfying results, but still has room for improvement. For example, PAM does not involve global fitting. I believe Global fitting will help this project. Global analysis permits simultaneous fitting of lifetimes and other contributors from all selected pixels in the image [211, 164]. In combination with automation of segmenting the nuclear rims, nuclei and the cytoplasm out, which are currently performed stand alone and not directly integrated into PAM, global analysis could analyse all files simultaneously issuing a more precise measure of the end-to-end distances and better signal-to-noise ratio. Such software tools already exist and can be applied [164, 212].

On the other hand, PAM also allows for phasor type of the analysis. Unlike classical FLIM that creates a colour-mapped image of different lifetimes, the phasor approach forms a polar coordinates plot in which each decay is represented by the two coordinates [213]. Phasor analysis of FLIM images helps to avoid potential problems with exponential analysis (for example, it is irrelevant how many exponents donor decay has) and is often used to quantify FRET efficiencies by localising the position of clusters of phasor rather than by analysing the decay in exponential components [213, 214, 215]. It would be beneficial to begin by directly comparing mean FRET efficiency values extracted from a model-free phasor approach to quantitative values from the fit with a polymer model P(R). In my analysis I can calculate mean FRET efficiency E by applying a central limit theorem derived for single molecule FRET

experiments [216, 217]:

$$\langle E \rangle = \int_0^\infty P(R)(\frac{1}{1 + \frac{R}{R_0}^6})dR \tag{4.1}$$

This expression is limited to cases when fluctuations of the energy transfer are significantly slower than the donor lifetime. Even though phasor analysis was not used in this study, in combination with my automatic segmentation routine it could be a powerful tool to quantify end-to-end distances from FRET efficiencies using a potentially 'cleaner' analysis with phasors that are also able to precisely distinguish the donor species fractions faster and are more suitable for the analysis of large datasets [213].

4.0.3 FLIM-FRET imaging of FG-Nups in mammalian cells and applying polymer scaling laws

One of the aims of my research is to use the FLIM-FRET approach to quantitatively determine the conformational state of the FG-Nups in the NPC. I focused on two major players in the permeability barrier - Nup98 and Nup62. Surprisingly little is known about their conformational state. Research into these FG-Nups is complicated by the fact that both of them are subject to multiple post-translation modifications which alter their functions and potentially their localisation in case of Nup62 as discussed in section 4.0.1. In this section I will discuss the results I obtained from end-to-end distance measurements *in situ* in the context of what was measured *in vitro* and predicted in simulations and the implications they have in the development of the permeability barrier model.

FLIM-FRET measurements of FG-Nups in mammalian cells
One of the first experimental validations of the dual-colour labelling of double Amber FRET mutants of Nup62 was to measure FRET efficiencies using photobleaching (Figure 3.14). It became clear from this that it is possible to distinguish between different mutants based on FRET, and to observe an approximate expected trend potentially following scaling laws. Measurement of FRET by photobleaching is potentially not sensitive enough and only measures apparent FRET efficiency proportional to the fraction of donor molecules that interact with the acceptor. FLIM, on the other hand, can measure absolute FRET and quantitatively estimate the fraction of donor species. Overall, this was the first experimental proof of principle of the experimental design of FRET mutants and integrating Amber suppression and click-labelling techniques.

Using the established FLIM-FRET procedure I measured end-to-end distances of FG-Nups *in situ*. There is a clear trend of increasing end-to-end distances with increasing number of amino acids in line with scaling laws for polymer chains. That

is the first time this relationship has been shown in a supramolecular assembly such as the NPC. I chose to use the Ideal Gaussian chain polymer model in the analysis as it is often used to describe IDPs in FRET experiments *in vitro* [175, 170]. The results of my measurements yield scaling constants ranging between 0.32 and 0.46, which according to Flory theory correspond to a more collapsed polymer. *In vitro* disordered proteins were reported to have scaling exponents of around 0.45 and 0.51 in native buffer and between 0.55 and 0.62 in high denaturing conditions in FRET experiments [175, 170]. However, some evidence suggest IDPs are more expanded in native buffer conditions (scaling exponents > 0.5) in label-free small-angle X-ray scattering (SAXS) measurements. Whether dyes in FRET measurements affect protein conformation is still debatable in a dynamic sample [178], however in a fixed cell in my measurements concerns due to dye labels are minimal.

I acknowledge that more measurements and more refined corrections might affect the absolute values of my scaling exponents. However, I will discuss here the implications of the exponents I measured and their relation to different polymer states predicted for FG-Nups. Nup62 unstructured region showed a scaling exponent of 0.46, which indicates that in this part of Nup62 it behaves almost like an ideal chain and the surrounding NPC environment is close to the effective θ-condition. Sequence-wise it is actually quite similar to the FG-rich region and has a large number of hydrophobic (A,P,G) and polar (T) amino acids, common for many IDPs. The FG-rich region shows more compaction with a scaling exponent of 0.32, which may relate to it is more cohesive.

My Nup98 results also show two scaling exponents calculated for two different states in the cell. I imaged the same cells to collect fluorescence decays from both the GLFG bodies and the nuclear rims for direct comparison. Interestingly, in GLFG bodies the exponent is larger, 0.46 compared to 0.37 in the rim. This may be expected that Nup98 in a liquid droplet would be more extended compared to the NPC. This can be assumed if one considers a liquid droplet as a polymer melt. Polymer melt is essentially a concentrated polymer solution with its chains completely filling up the space available, pushing the solvent out [171]. According to Flory theory, the individual chains in a polymer melt are ideal and can move freely [218, 171]. A liquid-like droplet or, for instance, the NPC both are predicted to contain a large number of small molecules and water inside them so most probably are not polymer melts but may still follow an ideal chain behaviour if θ-condition is reached. Under this condition FG-Nups will be in a perfect balance between chain–chain and chain–solvent interactions. My exponents suggest more that FG-Nups prefer to interact with each other than the solvent and tend towards a globule. Moreover, my measurements suggest that there are clear differences in chemical microenvironments in GLFG bodies and inside the NPC.

Having identified the preliminary values of scaling exponents of the two FG-Nups in the NPC I attempt to relate them to the current permeability barrier models? Most permeability barrier models to date consider the two extreme scenarios of polymer behaviour. The virtual gate/polymer brush model is based on grafted polymer behaviour, such that the individual brushes extend further with increasing grafting density, forming an entropic barrier [52, 54]. In Flory terms this would contribute to $\nu > 0.5$. The selective phase/hydrogel model, in contrast, suggests that the polymers are highly cohesive and associate with each other to form a sieve-like tough hydrogel [50, 58]. In polymer physics gel can be considered as a collection of adjacent blobs in chain. Blobs in Flory terms would correspond to ν around 0.3.

Taking cohesiveness into account, my experimental results agree with many simulations involving metazoan Nup98 and the yeast homologues of Nup62 and Nup98, Nsp1 and Nup145N respectively. The cohesiveness strength in these models is determined by charge-to-hydrophobicity ratios. According to Yamada *et al.* Nsp1 (fragment of 172aa) and Nup145N (fragment of 242aa) were experimentally shown to accept premolten globule (Stokes radius of around 3nm) and molten globule (Stokes radius of around 2.6nm) configurations respectively [75]. Stokes theory considers a polymer coil to behave like a solid sphere of a certain Stokes or hydrodynamic radius in solution. Taking their measurements and calculated cohesive strength into account, the authors suggested a forest model, in which Nsp1 appears as a 'tree' with a long relaxed coil domain grafted at the NPC (12nm) and a collapsed coil domain at the N-terminus (6nm), while Nup145N appears as a 'shrub' with two collapsed coil domains differentiated by relative charge content of 5 and 6 nm respectively [75]. Thus, the radius of the collapsed FG-Nup according to this model is roughly 5nm. If I extrapolate my measurements for both Nup62 and Nup98 in the two extreme scaling exponents I measured, these FG-Nups will have dimensions around 3nm with $\nu = 0.32$ or 6.3 and 7.6nm respectively with $\nu = 0.46$. This already shows that my measurements are within the range predicted for yeast FG-Nups and that most likely both Nup62 and Nup98 behave more like 'shrubs' (between premolten and molten globules) according to the forest model. Extending from the forest model into the phase separation field, it is also known that Nup98 FG-domain is more cohesive than that of Nsp1 and therefore is more collapsed in solution [77]. This agrees with my findings that I see a smaller ν for Nup98 than that of Nup62 unstructured domain. Independently of the sequence specificities between the FG-Nups, there is also evidence that all FG-Nups' Stokes radii scale with amino acids number following an ideal Gaussian model [80] with the scaling exponent equal to 0.5. My preliminary results indicate that *in situ* FG-Nups are more collapsed than ideal.

It is also important to mention that the dimensions of FG-Nups that I measured are the dimensions in the NPC under permeabilised cell conditions: with most other FG-Nups, some NTRs, small molecules and cargoes still present in the channel [65,

66, 67]. My measurements therefore capture the heterogeneous nature of all states of FG-Nups within and between different NPCs. Since these measurements only show the 'average' state it could also be that I am capturing the state between the phase transitions FG-Nups might undergo in the NPC from the extended to the collapsed coil [77] ('coil-globule' transitions in polymer physics), with a preferential tendency for the latter.

There is still a huge variance observed in my measurements. Firstly, this is due to using the ensemble average of the FRET signal in all pixels. The biggest challenge in this type of analysis is to make sure that the donor species can be efficiently identified in the analysis and their contribution so as not to have a pronounced effect on end-to-end distance estimation. According to the simulations in Figure 3.11 and 3.12 this should not be an issue but this was not yet properly tested in cells. Secondly, there are many experimental differences between the individual experiments as well as individual images. These differences arise from expression, labelling, buffers and other heterogeneities that would also influence autofluorescent background in the sample. Thirdly, it is believed that individual NPCs are in fact heterogeneous structures that would potentially also have multiple different conformations of the FG-Nups depending on their state. In summary, all these factors would contribute to the large heterogeneity in my measurements.

To address the factors mentioned above in the future, I suggest the following optimisations of experimental conditions and the analysis pipeline: Experimentally I would address the effect of the labelling on the physiological state of FG-Nups. In the current experimental conditions I am using high salt buffer to label permeabilised cells before fixing. The effect this labelling method has on the state of the FG-Nups is yet not clear. It would be beneficial to directly compare the measurements done in high salt buffer and measurements in live labelled samples, where cells are not even permeabilised. To address this improvement I managed to successfully label both Nup98 and Nup62 live with TAMRA and SiR H-tetrazine modified cell permeable dyes (Figure 4.1). Unfortunately TAMRA has a very short lifetime of 2.3ns and is not a good donor candidate but cell permeable Janelia Fluor dye JF549 with a lifetime of 3.84ns [219] exists, which would potentially be a better candidate for a donor dye. Silicon rhodamine can still be used as the acceptor dye in this scenario.

One of the most puzzling results is the range of end-to-end distances that I calculate using the Ideal Gaussian model. The range is 1.5-4nm for up to 218 amino acids in between the dyes. One of the direct references to compare my results from the NPC to the *in vitro* configuration is using measurements by Fuertes *et al.* [170]. This study directly compares measurements obtained using small angle X-ray scattering

(SAXS) and smFRET. The authors used different FG-Nups and other IDPs to determine their dimensions using scaling laws. Even using the Ideal Gaussian chain model R_E of Nup98 fragment of 150aa is 5.3nm compared to around 3.3nm obtained by extrapolating my results in Figure 3.22. Also the scaling exponent in the published results is 0.5 compared to 0.37 in my results. This either indicates that the protein in the NPC is either far more collapsed or a systematic technical issue, such as a limited dynamic range to probe such short distances of the chosen dye pair. This is apparent from the FRET efficiency curve in figure 3.13(c), the shape of which is dictated by R and R_0 (equation 1.9. In the part of the curve where distances are between 1.5 and 4nm, FRET efficiency is so large that it is hard to distinguish E_{FRET} values from each other. The best way to verify these distances for the current set of mutants would be to take a dye pair with R_0 around 3-4nm and remeasure, but it is unfeasible to find such a dye pair as most dyes have R_0 around 5.5-7nm.

To increase the precision of measurements I would also consider using linearly polarised light instead of circular. Circularly polarised light provides equal excitation to all fluorophores irrespective of their dipole orientation. Despite this, the spot for the electric field component parallel to the direction of polarisation of the incident light going through the high NA objective is elliptic [220]. The splitting into two channels by the polarised beam splitter might not actually be creating two identical lifetime fluorescence decay curves. By combining the two curves I might in fact alter them, potentially also affecting the fitting. Measuring using linearly polarised light and applying relevant corrections before combining the curves from the two channels could further improve the quality of my measurements and potentially also the fitting. Moreover, using linear polarised light would enable me to also measure anisotropy. This is a measure of the rotational correlation time of the fluorescent probe, which provides direct information on the dynamics of the probe, which is especially relevant in the context of IDPs.

In summary, I described the conformations of Nup62 and Nup98 *in situ* using a scaling law. This appears to be the first time the scaling of an IDP has been quantitatively probed in any cellular machine. The current data suggests that the Nups are in a state between being more collapsed and ideal inside the NPC in cells. This agrees with the very few previous measurements *in vitro* of IDPs and will be complemented by *in vitro* measurements of the same proteins. In the future this method can be applied to reveal the structural rearrangement of FG-Nups upon transport of large cargoes such as viral capsids [221], especially the HIV capsid discussed in the Perspectives section. The established method would also benefit the community studying IDPs in general such as in phase separation in cells. More than 30% of human proteome are IDRs, found to play an essential role in the regulation of signalling pathways and other crucial cellular processes. My new method of studying conformation of IDP *in situ* potentially has a significant impact on deciphering many of the

functions of IDPs in these hubs of protein interaction networks. Moreover, it may go some way towards revealing the structural assembly of many phase separated membraneless cellular assemblies such as stress granules and shine light on the assembly of amyloids at early stages, crucial in understanding disease progression.
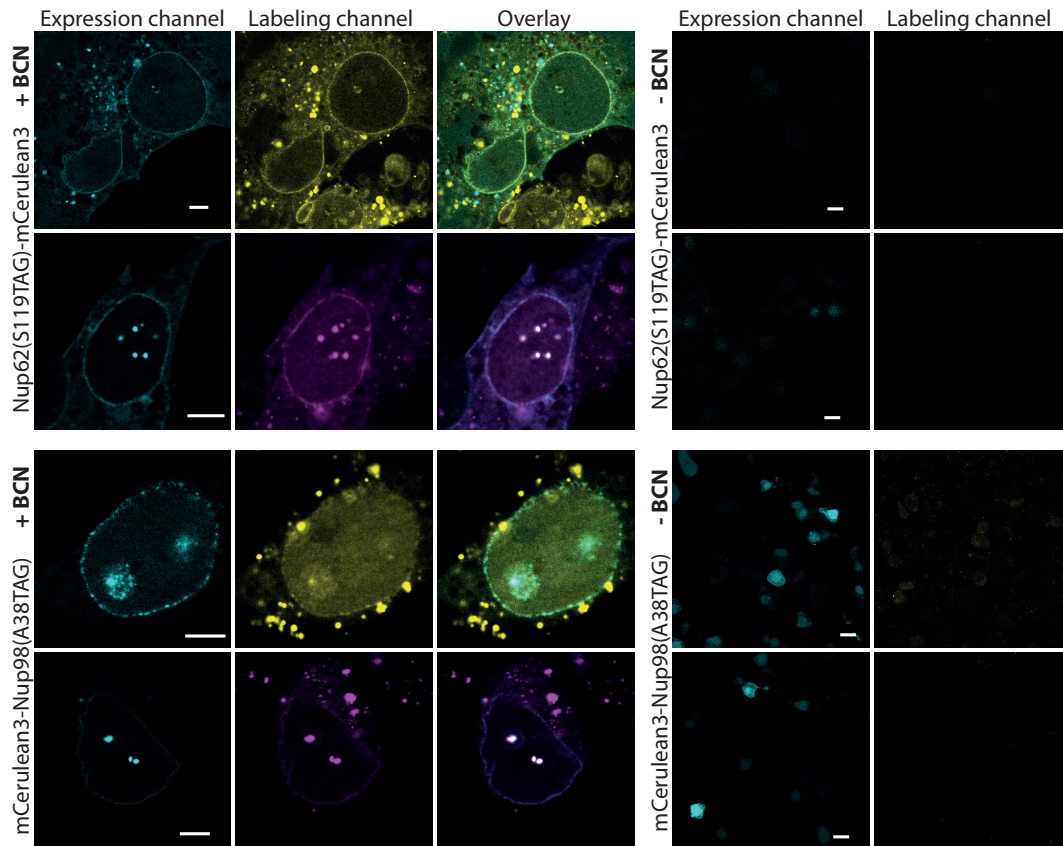
FIGURE 4.1: Live labelling of Nup62 and Nup98 Amber mutants in COS-7 cells. Cyan represents the mCerulean3 expression channel, yellow is TAMRA-H-tetrazine (Jena Biosciences) dye and magenta is SiR-H-tetrazine (Spirochrome) dye. Labelling was done according to the method published by Uttamapinant *et al.* [97] with extended fluorophore incubation time up to 1 hour. Columns represent different FG-Nups with Nup62 on top and Nup98 in the bottom. Expression plasmids are indicated on the side. Image groups on the left are the fully working Amber suppression system with added BCN (scale bars are at 5μm) and columns on the right are the ones where codon should not be suppressed as no BCN is added, therefore no labelling is observed (scale bars are 10μm).

5 Perspective: HIV-1 in the context of nuclear pore and nucleocytoplasmic transport

This outlook focuses on a different yet related topic that I have spent time working on. The long-term goal of this project is to visualise protein flexibility and protein-protein interactions between the human immunodeficiency virus 1 (HIV-1) capsid and the NPC with high spatial and temporal resolution using the methods developed in this thesis. In this section I will introduce the role the NPC and particularly FG-Nups play in the lifecycle of HIV-1 in cells and show the results of the experimental work I have done to study HIV-1 capsid passage through the NPC.

5.1 Introduction to HIV-1 in the context of the NPC and nucleocytoplasmic transport

Many viruses such a Hepatitis B (HBV), adenoviruses, herpesviruses, influenza, retroviruses and others are known to hijack the NPC to traffic essential viral proteins and their genome in and out of the nucleus of the infected cells [222]. HIV-1 is a human immunodeficiency virus of *Lentivirus* genus of retroviruses that can cause acquired immunodeficiency syndrome (AIDS). HIV-1 is an RNA virus, that rapidly reversely transcribes its genomic RNA to DNA in the cytoplasm as soon as it enters the cell. The complex of viral DNA and viral proteins such as integrase, capsid (CA), nucleocapsid (NC), viral protein R (Vpr) and some host protein is called the pre-integration complex (PIC). PIC is believed to use the cellular transport machinery to move along microtubules towards the NPCs and then hijack the NPCs' transport machinery to get into the nucleus and, as a result, integrate its genome to encode more virus production [223, 224]. This ability of HIV-1 to transport essential viral components through the NPC means it can infect both dividing and non-dividing cells. The prime targets of HIV-1 are cells of the immune system, CD4$^+$ T-cells and macrophages.

What is still unclear is how the PIC can cross the NPC as its size is estimated to be around 56nm in diameter [225, 226] in contrast to around 40nm diameter of the central channel of the human NPC [14]. Moreover, it has recently been shown that the upper limit of size of the cargo transported into the nucleus is more likely to be

even less than 39nm as previously thought for the HBV capsid [227]. In a new study by Paci *et al.* the HBV capsid was observed to accumulate at the nuclear envelope rather than successfully imported [221]. But HIV-1 is a very different scenario. First of all, it is not just the capsid that presumably is transported through the NPC but also multiple other components of the PIC. Secondly, HIV-1 capsid does not possess or acquire any localisation signals. Instead, the capsid has been shown to specifically interact with multiple FG-Nups directly. In particular, Nup358/RanBP2 (component of potential cytoplasmic filaments) and Nup153 (FG-Nup component of the nuclear basket) have been co-crystallised with capsid hexamers [228, 229]. The mature HIV-1 capsid is a fullerene cone made of about 100-1500 CA proteins, arranged in approximately 200-250 hexamers and exactly 12 pentamers [230, 225, 231]. CA protein consists of two alpha-helical N- and C-terminal (NTD and CTD) domains connected by a flexible linker. The C-terminal cyclophilin domain of Nup358 binds CA at the cyclophilin A (CypA) binding loop between helices 4 and 5 of the NTD [228]. CypA is a *cis-trans* prolyl isomerase that catalises proline isomerisation in proteins and may regulate both capsid stability and uncoating [232, 233, 234]. Nup153 potentially binds the capsid differently, via its FG-domain. Recently Buffone *et al.* showed that Nup153s' FG-motifs bind the capsid at the NTD/CTD interface between adjacent subunits, interacting specifically with residues P34,I37,P38,N53,L56,N57,V59,V142 and Q176 [235]. Moreover, Nup153, Nup358/RanBP2, Nup98, Nup62 and Nup88 were found to tightly associate with *in vitro* assembled capsid-nucleocapsid (CANC) tubes and affect viral infection in cells [236, 237, 238, 25].

It is still unclear whether the capsids in the PIC are uncoated before reaching the NPC, in or by the NPC or after passing the NPC. It was previously observed that capsid proteins potentially dock and uncoat at the NPC are retained in the nucleus after passing the NPC in non-dividing cell [239, 240]. Based on the evidences for specific interaction of the capsid with FG-Nups and observations of capsids associated with the NPCs I proceeded to establish a fluorescence microscopy assay to visualise capsids at the NPCs and an *in vitro* single-molecule fluorescence approach to decipher binding mechanism between FG-Nups and capsid hexamers.

5.2   Results of click-labelled HIV-1 capsids interactions with the NPC

In this section, I present experimental results of observing *in vitro* assembled click-labelled capsids accumulating at the nuclear envelope in permeabilised HeLa cells. I also demonstrate a strategy to deduce the specific mechanism of binding between the FG-residues of Nup153 and capsid hexamers. Firstly, I will present the experimental procedure we developed in close collaboration with Dr Martin Obr in Prof. Hans-Georg Kräusslich's research group at the University of Heidelberg. Secondly, I will show results of assembled and click-labelled capsids in permeabilised HeLa

cells and finally present preliminary results studying binding between capsid hexamers and FG-fragments of Nup153.

Dr Obr successfully identified a site in the capsid monomer that can be modified with an Amber stop codon to incorporate an UAA (I91). He then established an *in vitro* assembly strategy to form fullerene capsid cones. I have used his approach to specifically label the capsids with tetrazine-modified dyes and imaged their localisation in permeabilised cells. Initially we started working with CA(A204C) mutant, which was previously reported to assemble into closed cones similar to authentic cores [226]. We have then moved to a mixture of CANC cores of WT and R18L mutant, which was shown to form spheres. This mixture did also assembled equally well into expected cones and had the advantage that NC binds to DNA, meaning we could introduce another label on the DNA. In this section capsids and CA refer to the WT and R18L mutant mixture and are co-assembled with a DNA oligo. The additional advantage of this system was that I could use a labelled DNA oligo as an additional colour, which is expected to dissociate from the capsid protein signal only upon disassembly of the cone.
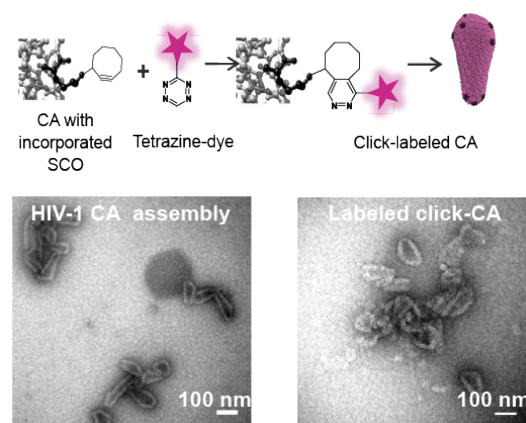
FIGURE 5.1: Negatively stained electron microscopy images of assembled HIV-1 capsids. Top picture is a schematic representation of a click-chemistry reaction between the incorporated SCO and tetrazine-dye (magenta). Below are the EM images of *in vitro* assembled capsids before labelling and after labelling with the dye.

Initially I wanted to check that the dye coupled using click-chemistry does not disrupt the capsid assembly. The negative stain images in figure5.1 clearly show that the labelled capsid cores are still assembled similarly to the ones without the dye. I then introduced the click-labelled sample to the permeabilised HeLa cells in transport buffer (TB) with the essential transport protein energy mix [241, 242, 243]. Using the established reconstructed nucleocytoplasmic transport protocol I imaged accumulation of assembled capsids. Figure 5.2(a) shows that the final assembled wild-type (WT) and R18L mutant assembled capsids sample has a lot of proteins that potentially didn't associate into the cores and, therefore, stick to all membranes in

cells without any specific accumulation at the NPC. It was striking to see both anti-capsid and DNA signal in the nucleus of one of the cells. This was the first evidence that even in HeLa cells I could potentially see some capsid assemblies transported through the NPC into the nucleus. Figure 5.2(b) shows the WT, R18L and I91^{SCO}-click-labelled sample. Here it becomes apparent that either even after several washes to remove free dye some remains or that there are a lot of labelled but unincorporated capsid proteins in the mix. I conclude this from the signal from the click-dye inside the nucleus and cytoplasm (magenta). Interestingly, capsid signal in the nucleus in this experiment did not have a corresponding signal from the DNA (yellow). This may be because the capsids were uncoated at the NPC. Figure 5.2(c) shows the dramatic effect I observed on specific capsids accumulation at the nuclear envelope by using cytoplasmic extract instead of reconstructed transport buffer condition. The two very clear nuclear rims are observed in both the click-labelled capsid channel and the DNA channel. Also here the signal from the click-labelled capsids in the nucleus does not have a corresponding signal from the DNA, yet at the nuclear envelope they clearly specifically colocalise, indicating that the capsid cone was disassembled only in the nucleus.

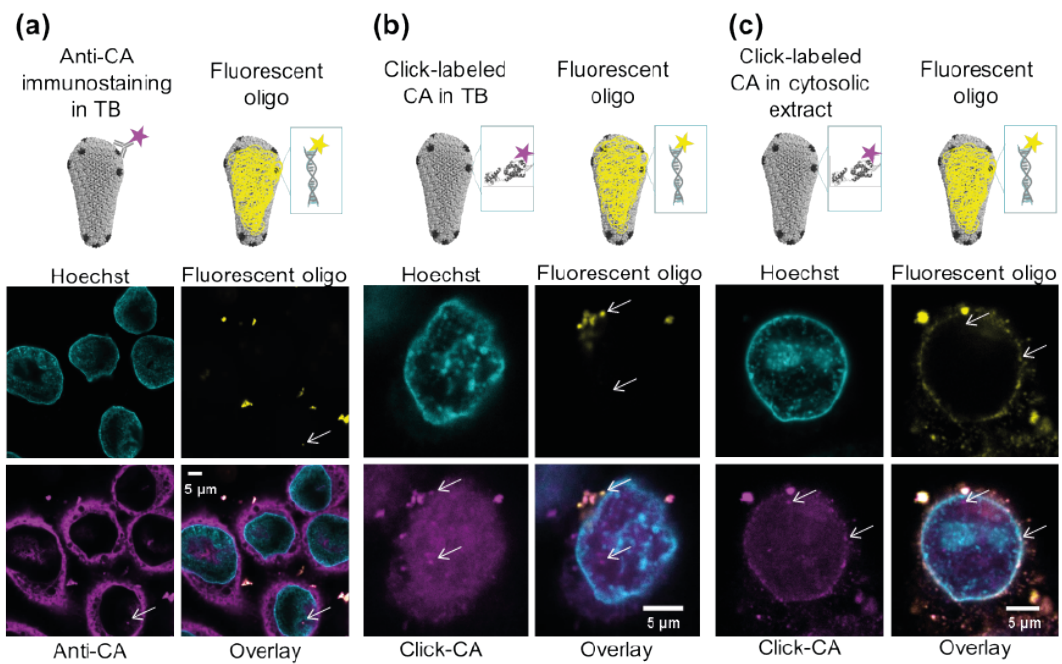FIGURE 5.2: Confocal images of assembled and labelled HIV-1 capsids in permeabilised cells. (a) immunostaining of the capsids in the reconstituted transport assay; (b) click-labelled capsids in the same assay; (c) click-labelled capsids in HeLa cytoplasmic extract. Cyan channel represents Hoechst staining of the nucleus, yellow is the fluorescent oligo channel, magenta represents the capsid channel and overlay shows all.

5.2.1 Capsid hexamers interactions with the NPC

Despite these promising results the heterogeneity of the sample's assembly was problematic. Both negative stain EM and fluorescence imaging show large clusters of capsids, which not only have the cones but also tubular assemblies. I tried to make the sample more homogeneous by using CypA at different ratios [244, 234] but I did not observe any differences in the assemblies after checking them in the EM. Unfortunately, for super-resolution imaging this means I would be looking for 'a needle in a haystack' to find assembled conical capsids in the vicinity of the NPC even in the cytoplasmic extract. Since I was unable to use the full capsids to decipher the nature of the interaction between the CA and the FG-Nups, I started working on a simpler model system of lower assembly level of the capsid proteins - capsid hexamers. I managed to successfully purify the hexamer and test it in a fluorescence anisotropy assay. Fluorescence anisotropy is a method to study binding constants and kinetics of reactions or protein binding which affect rotational time of the molecule. I used a Cy3b-labeled FG-fragment of Nup153 [11, 245] to monitor changes in its anisotropy upon binding an NTR - importin-β with a very high binding affinity (K_d) in a range of nano- to micro- molar [246] compared to the capsid hexamer with a very low K_d, around 100 μM [247]. This difference is recaptured in my anisotropy measurements in figure 5.3(c,d).

5.3 Discussion and outlook of of click-labelled HIV-1 capsids interactions with the NPC

In the course of my PhD research I successfully visualised labelled HIV-1 capsid proteins specifically associating with the nuclear envelopes in permeabilised HeLa cells supplemented with cytosolic extract and also demonstrated direct binding between an FG-fragment of Nup153 and HIV-1 hexamers. My results from permeabilised cells indicate that there are still multiple essential proteins in the cytoplasm that are crucial for successful capsid association with the NPCs. Since during cellular membrane permeabilisation one washes the cytosol out, if not supplemented with ions and ATP, this method may stall molecular motors. One of the reasons cytosolic extract might restore active capsid targeting to the NPC is that it supplies energy to activate molecular motors like kinesin and dynein (specifically dynein heavy chain and kinesin-1 and its' FEZ1 adapter) that are suspected to actively move capsids along microtubules to the NPCs [248, 249, 250, 251]. Another apparent candidate missing in the standard transport proteins mix is transportin-3 (TNPO3). TNPO3 is a β-karyopherin that has previously been shown to interact directly with the integrase and CANC complexes [252, 253, 254]. Adding these components or making
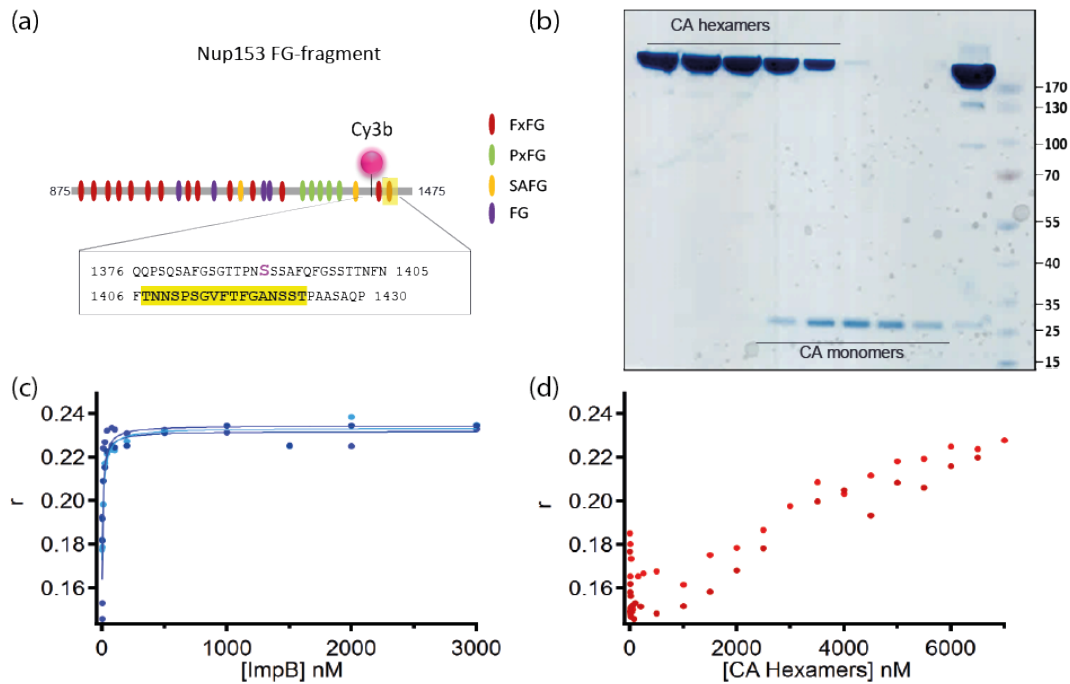
FIGURE 5.3: Association of capsid hexamers with FG-fragment of Nup153. (a) shows a schematic representation of the $Nup153_{875-1475}$ FG-fragment labelled with Cy3b at position S1391 and the peptide sequence used to study interaction between the two, highlighted in yellow [247]; (b) SD-PAGE gel showing successful purification of HIV-1 capsid hexamers (upper band), lower band shows monomers that did not cross-link; (c) anisotropy (r) measurement of Nup153 with importin-β; (d) anisotropy measurement of Nup153 with the hexamers. Different colours demonstrate different replicates.

sure they are actively functioning could recapitulate the same effect observed in using the cytoplasmic extract in a simple transport mix assay.

The main difficulty I faced using the *in vitro* reconstructed click-labelled HIV-1 CANC particles was that they form a highly heterogeneous assembly and prefer to stick together. Consequently, most of the CANC labelled sample I had were not individual fully assembled capsids. This makes it more challenging to spot a rare event of CA interaction with the NPC. It currently seems more promising to further explore the interaction of HIV-1 capsids' functional subunits: hexamers and pentamers *in vitro* and potentially in cells. First of all, hexamers sample looks very pure on the gel and, since no precipitation was observed, is not impaired by sticking like the assembled capsids. I predict that *in vitro* cross-linked pentamers would behave the same way [255]. Thus, it would be interesting to first, directly compare their binding kinetics to the Nup153 (for example using fluorescence anisotropy). This will potentially reveal aspects beyond the structural function of pentamers in the capsid cone. Secondly, it is still not clear what the binding mechanism of the capsid hexamers/pentamers to the FGs is and what effect it has on their conformational state. Direct comparison of known binding affinities of FGs to importin-β (K_d between around 1 and 1500 nM [256, 257]) and the hexamer (K_d around 100 μM [247]) indicates a 100 fold difference. This does not yet indicate the transport mechanism as affinities are hard to relate to

specificity and transport rates. From the binding models of IDPs to folded proteins it would be of high importance to distinguish which binding type relates to capsid hexamers: whether Nup folds upon interaction (fly-casting mechanism [258]), finds the best conformation to fit in the binding pocket (conformational selection mechanism [259]) or does not require any major energy or time expense for rapid yet specific binding ('archetypal' fuzzy and multivalent complex [11]). One of the established tools that I could implement is to use smFRET. smFRET would allow me to measure any changes in structure and dynamics in the labelled FG-Nup upon binding the structured protein [125, 11]. I can also foresee that the established in this thesis FLIM-FRET approach can also be used to study the interaction of the capsid hexamer/pentamer with the FG-Nups in the central channel. Discovering the exact binding mechanism between the FG-Nups (specifically Nup153, Nup98 and Nup62) and the capsid hexamers/pentamers would not just advance our fundamental understanding of HIV-1s' ability to adapt to hijack the NPCs' transport machinery but also shine light on the transport or uncoating mechanisms taking place at the NPC. This might even lead to potential drug discoveries that can help specifically block the found mechanism of transport, thereby curing AIDS.

A Appendix: supporting data

A.1 Confocal images of FG-Nups used in FLIM-FRET analysis

This Appendix shows supporting data of all representative FLIM-FRET images, segmentation of which resulted in the fluorescence decay curves shown in figures 3.19,3.16,3.17. All figures three channels: Hoechst channel for automated segmentation and the marker of the nucleus in cyan, donor channel labeled with Atto532 in yellow and the acceptor channel labelled with LD650 in magenta.

FIGURE A.1: Representative confocal images of Nup98 38+ FRET mutants. Columns represent imaging channels: cyan is Hoechst staining, yellow is donor (Atto532) labeling and high-resolution TCSPC acquisition channel, magenta is acceptor (LD650) labeling channel. Different representative FRET mutants images are organised in rows according to N_{res}.
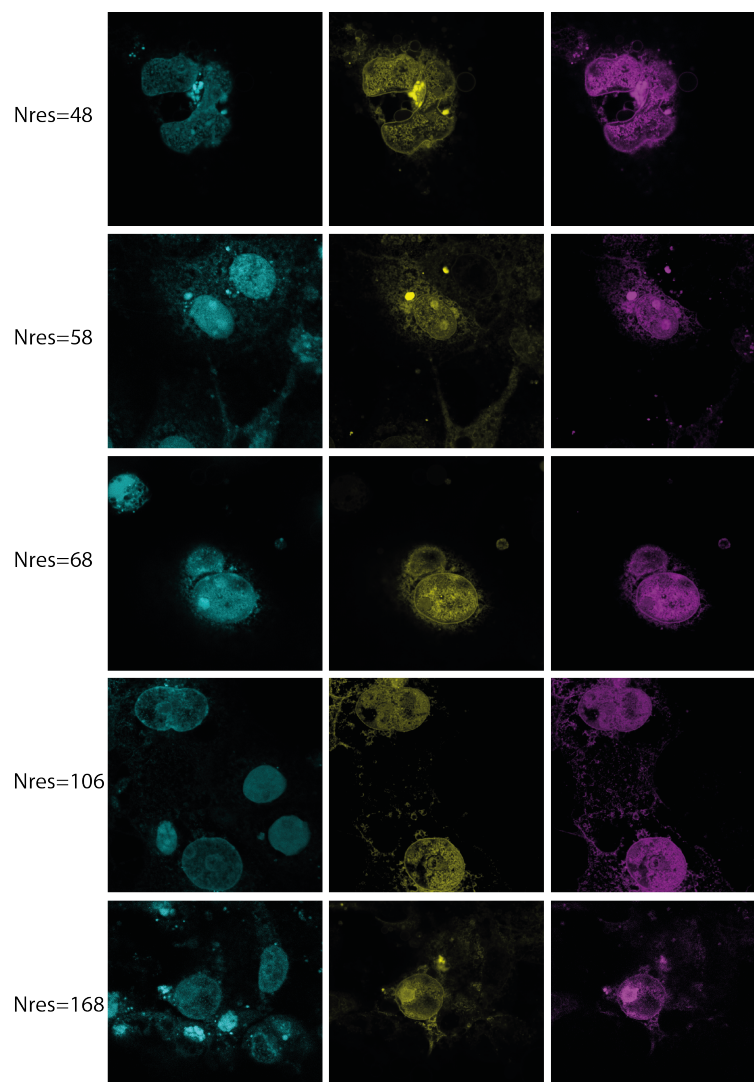
FIGURE A.2: Representative confocal images of Nup62 13+ FRET mutants. Columns represent imaging channels: cyan is Hoechst staining, yellow is donor (Atto532) labeling and high-resolution TCSPC acquisition channel, magenta is acceptor (LD650) labeling channel. Different representative FRET mutants images are organised in rows according to N_{res}.
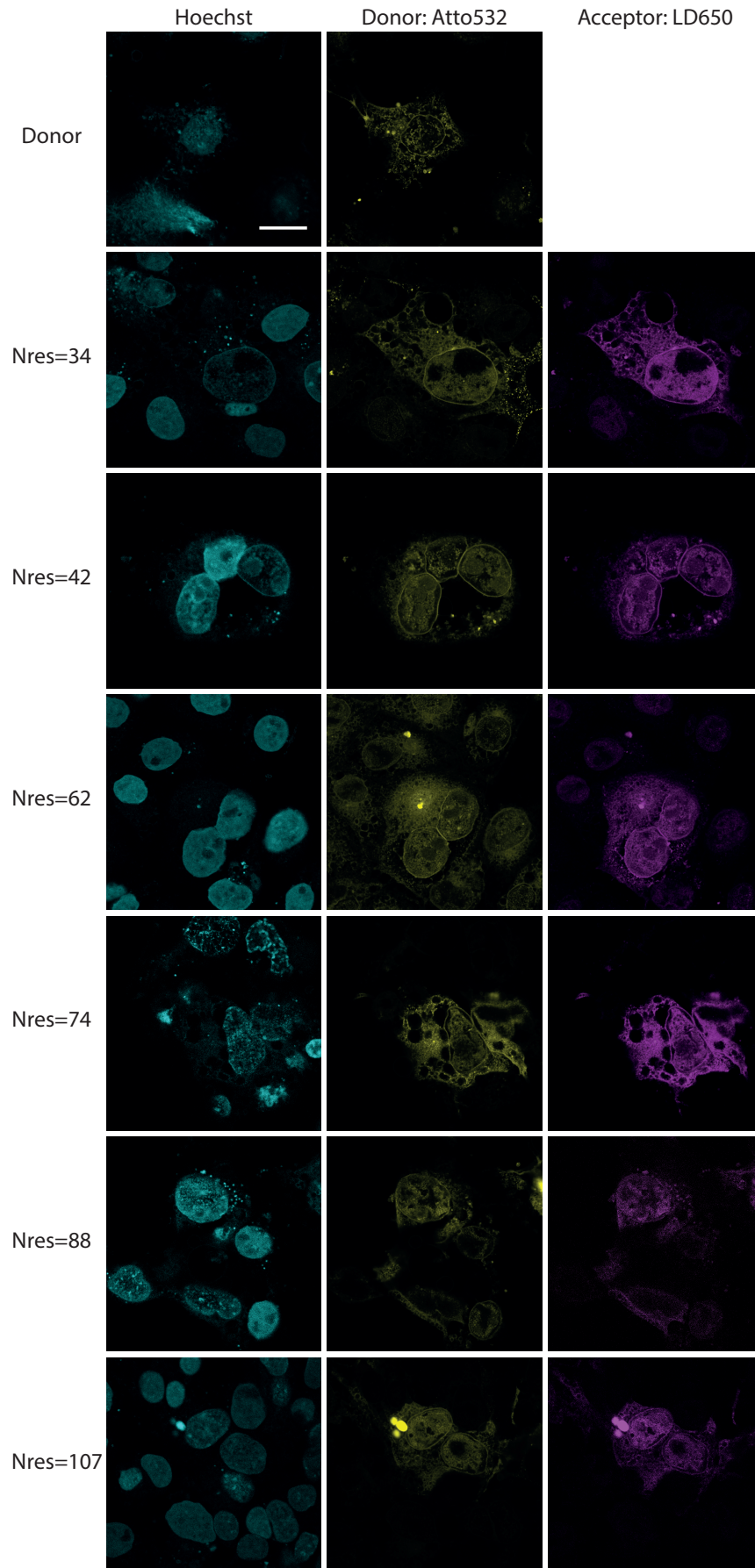
FIGURE A.3: Representative confocal images of Nup62 119+ FRET mutants, part 1. Columns represent imaging channels: cyan is Hoechst staining, yellow is donor (Atto532) labeling and high-resolution TCSPC acquisition channel, magenta is acceptor (LD650) labeling channel. Different representative FRET mutants images are organised in rows according to N_{res}.
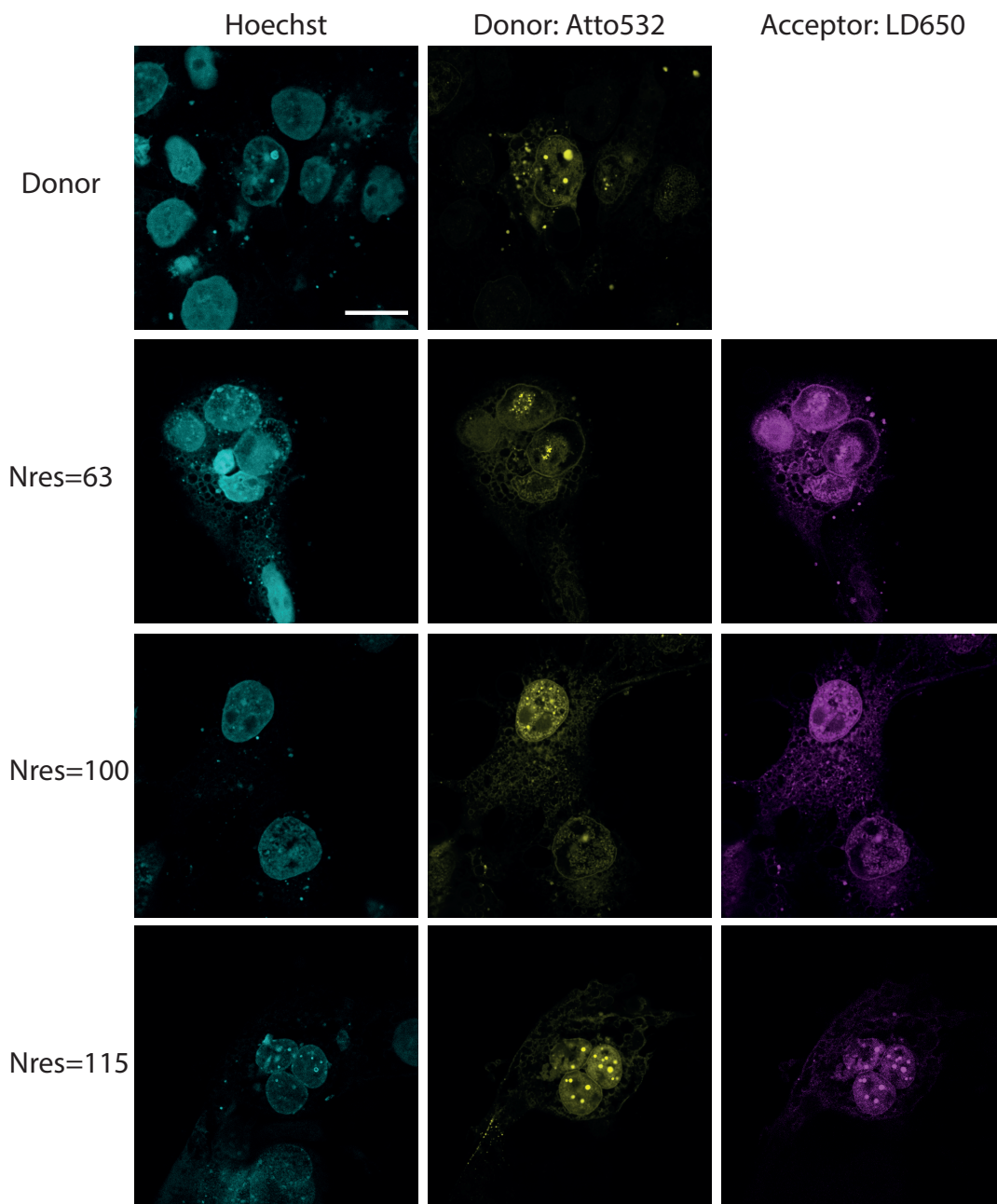
FIGURE A.4: Representative confocal images of Nup62 119+ FRET mutants, part 2. Columns represent imaging channels: cyan is Hoechst staining, yellow is donor (Atto532) labeling and high-resolution TCSPC acquisition channel, magenta is acceptor (LD650) labeling channel. Different representative FRET mutants images are organised in rows according to N_{res}.
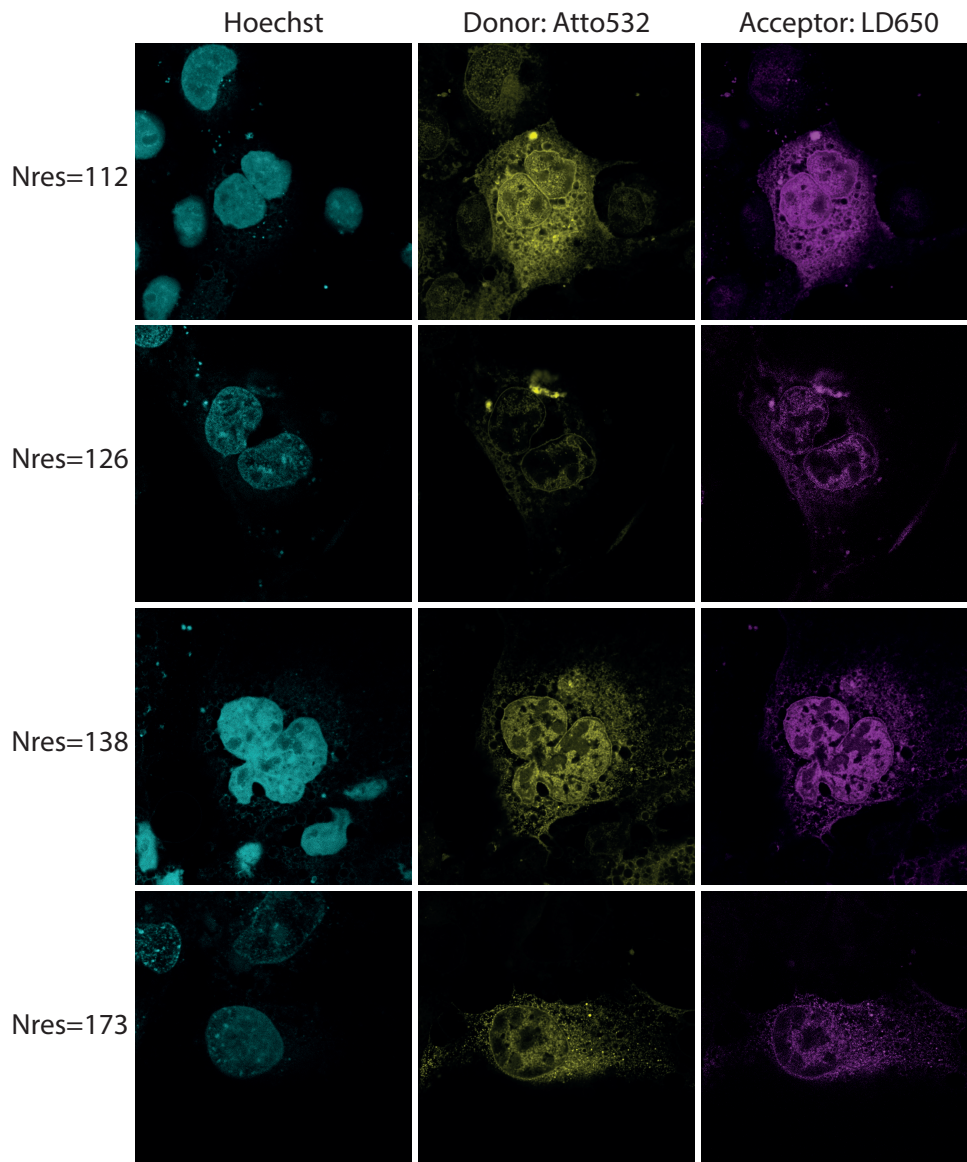
A.2   Materials and methods in HIV-1 in the context of nuclear pore and nucleocytoplasmic transport

HIV-1 capsid protein expression and purification

CANC protein for capsids assembly was expressed in *E.coli* BL21(DE3)AI cells at 37°C in LB (Lysogeny Broth) medium. Expression was induced with 1mM IPTG and 0.02% arabinose at O.D.=0.6-0.8 and cells were harvested after 6-8 hours by centrifugation in a Beckmann centrifuge, rotor JLA 8.100. Amber mutant was expressed the same way with PylRS and tRNA plasmid and adding up to 1mM SCO to the expression culture at O.D.=0.3-0.5.The bacterial pellet was resuspended in lysis buffer (50 mM TRIS, 1M NaCl, 10% w/v glycerol, 1 mM EDTA, 2 mM DTT, pH 8.3) supplemented with protease inhibitor 1mM PMSF, then lysed by lysozyme treatment (room temperature, 30 min) and subsequent sonication. The nucleic acid content was precipitated by 0,2% w/v PEI (1 h, on ice). Insoluble cell debris was pelleted at 20,000g for 45 min in Beckmann centrifuge, rotor JA 25.50. The protein in supernatant was precipitated by ammonium sulfate up to 25% saturation (1 h, 4°C), and pelleted at 15,000 g for 15 min. The pellet was resuspended in binding buffer (50 mM phosphate buffer, 500 mM NaCl, pH 7.5), and centrifuged to remove undissolved material. The protein solution was filtered through 0,45 nm PVDF filter and applied to HisTrap FF 1 ml. The column was developed with 50 mM imidazole in binding buffer. The eluate was applied to Superdex75 column equilibrated with the storage buffer (50 mM Hepes, 500 mM NaCl, 5% glycerol, 10 μM ZnCl2, 1 mM TCEP, pH 7.5). The peak fractions containing CANC were collected, concentration was adjusted to 3 mg/ml, and the aliquoted protein was flash-frozen in liquid nitrogen and stored at -80°C.

HIV-1 capsid assembly and labelling

Capsids were assembled by mixing together proteins (1:2:2 R18L mutant:WT:I91TAG) to a total concentration of 0.3mM with 2mg/ml of labelled DNA in 50 mM Tris-HCl, 1mM EDTA and 1mM TCEP (pH 8.0) buffer up to 250 mM dilution of total NaCl. 25nmols of Cy3 H-tet (Jena Bioscience) dye was added to the reaction for 10 minutes at room temperature. The reaction was then moved to 37°C for 1 hour. Assembled and labelled capsids were spun down at 10,000g for 5 minutes. Precipitate was carefully resuspended in 50 mM Tris-HCl, 250 mM NaCl, 1mM EDTA and 1mM TCEP (pH 8.0), spun down again to remove the rest of non-specifically labelled dye, resuspended and used for further experiments.

HIV-1 capsid hexamer expression and purification

Capsid hexamers were purified according to the protocols by Pornillos *et al.* [231, 260]. HIV-1 CA proteins were expressed by IPTG induction *E.coli* BL21(DE3)AI cells for 6–12 hrs at 25°C. Cell pellets were resuspended in 50 mM Tris, 50 mM NaCl, 1mM PMSF, 200 mM β-mercaptoethanol (βME), pH 8 and lysed by sonication. Insoluble cell debris were pelleted at 45,000g for 45 min in Beckmann centrifuge, rotor JA 25.50.

The protein in supernatant was precipitated by ammonium sulfate up to 25% saturation (1 h, 4°C), and pelleted at 15,000 g for 15 min. The precipitate was resuspended in 50mM Tris, 50mM NaCl, 200mM βME, pH 8. Insoluble fraction was pelleted again and the supernatant was dialised against 25mM 3-(N-morpholino)propanesulfonic acid (MOPS), 100mM βME, pH 6 for at least 4 hours and 1l per 1l of expression. Dialysed solution was applied to HiTrap SP FF column and developed with a linear gradient from 0 to 1 M in dialysis buffer. Peak fractions were pooled, dialysed against 20 mM Tris, 40 mM NaCl, 60 mM βME, pH 8, and applied to a Q-Sepharose column (GE Healthcare). Eluted fractions were pooled and dialysed against 20 mM Tris, 40 mM NaCl, pH 8 buffer to cross-link monomers into hexamers. Success of the cross-linking was varified on the non-reducing SDS-PAGE (sample buffer: 125 mM Tris, pH 6.8, 20% glycerol, 4% SDS, 0.1% bromophenol blue). The final concentration was adjusted to 30 mg/ml, and the aliquoted protein was flash-frozen in liquid nitrogen and stored at -80°C.

Plasmids used:

Plasmid	Internal no.
pET-11a-CANC-WT	1992
pET-11a-CANC(R18L)	1993
pET-11a-CANC(I91TAG)	1994
pET-11a-CA(A14C,E45C,W185A,M186A)	1988
pEvol-PylRS-tRNA	1412

DNA oligo (GGCTAGAAGGATCCATATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGA-GAATTAGATCGACCTATAGTGCAG) used was ordered from Sigma. The labelled oligo with a dye modification at 5' was ordered from iba.

Permeabilised cells assay and imaging

HeLa Kyoto cells were cultured at 37°C in 5% CO2 atmosphere in Dulbecco's modified Eagle's medium with 1 g/ml glucose supplemented with 1% penicillin–streptomycin, 1% L-Glutamine and 10% FBS. The cells were passaged every 2-3 days up to maximum of 15-18 passages. Cells were seeded 1 or 2 days before the experiment at 7000-10000 cells per well in a 8-well Lab-Tek II chambered coverglass (Thermo Scientific Nunc, 155383). Before transport experiments cells were taken out of the incubator, washed in room temperature PBS and then stained for 10 min at room temperature with 20 nM Hoechst 33342 (Sigma, B2261) for nuclear reference label. Cells were then washed once with transport buffer (1XTB: 20 mM Hepes, 110 mM KOAc, 5 mM NaOAc, 2 mM MgOAc, 1 mM EGTA, pH 7.3) and permeabilised with digitonin (40 mg/mL) for 10 min. Cells were then washed 3 times with 1XTB supplied with 5 mg/ml PEG 6000 to avoid osmotic shock. Then the transport mix was added in 1XTB with PEG and 1 mM DTT: 1μM importin-β, 1μM importin-α,4μM RanGDP, 2μM NTF2 and 2mM GTP. Assembled and labelled HIV-1 capsids were resuspended

in the final solution mix (150 μl) and added to the cells. Similarly, they were resuspended in HeLa cytoplasmic extract (ProteinONE, 500μg) and added to the cells. Cells were immediately taken to the microscope and imaged using Leica SP8 STED microscope from the Advanced Light Microscope Facility (ALMF) at EMBL. HCX PL APO 100x/1.4 oil objective (Leica, Germany) was used. Emission was detected from three channels: Hoechst, Cy3 and Cy5. Immunostaining with anti-capsid antibody (rabbit raised, kind gift from Prof Barbara Müller, 1:1000) and labelled with anti-rabbit-Alexa647 (ThermoFisher,A-21246, 1:1000) was performed as described in section2.

Anisotropy measurements

Fluorescence anisotropy was measured using an FG-fragment of Nup153(875-1475aa) with S1391C modified residue for maleimide labelling with Cy3b (GE Healthcare) (labelling reaction is described in section 2). Measurements were done using cuvette in 150uL of PBS, 2mM DTT, 2mM MgAc (pH 7.5) using Jasco fluorescence spectrometer (Jasco FP8500) equipped with FDP-837 automatic polarisers. Unlabled proteins like importin-β (positive control) and CA hexamers were gradually added to the Nup153. First the G-factor was calculated as a ratio of the fluorescence spectrum of the Nup153 in solution I_{HV} and I_{HH}. Here I is the intensity, hh means both excitation and emission polarisers are in horisontal positions and hv means only the excitation polariser is in horizontal position, while the emission polariser is in vertical. The factor G (instrument sensitivity ratio towards vertically and horizontally polarised light) was then used to calculate anisotropy using:

$$r = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}} \tag{A.1}$$

Negative staining EM

Success of capsid assembly was confirmed by imaging the samples with an electron microscope using negative staining. Carbon coated 300 meshes Quantifoil Cu grids were glow-discharged for 10 s in a vacuum chamber. Then, 3mul of sample was adsorbed on a grid for 2 min, blotted with Whatman's filter paper and washed 3 times with buffer, then 3 times with a solution of 2% uranyl acetate. Once the grids were dry, the sample was imaged using a Morgagni 268 microscope (FEI) located in the Electron Microscopy Core Facility (EMCF) at EMBL.

Acknowledgements

I would like to express my gratitude to those people who made this work possible and helped me through these four years to achieve the results summarised in this thesis and to grow as a scientist and as a person.

I am grateful to Edward, who has been a great mentor and advisor and was always available to answer any burning questions or fix any microscope issues. I have learned a lot from many of our fruitful discussions and will carry this experience with me through the years. I would also like to thank my thesis advisory committee members: Dr Martin Beck, Prof Hans-Georg Kräusslich, and Dr Alba Diz-Muñoz for your support, important feedback and insightful advice.

I am very thankful to the whole Lemke lab, in its current and former composition. I sincerely appreciate your help and support. I would like to express my special thanks to Piau Siong, who became my second mentor and my main scientific support in the past few years. Thank you for sharing your expertise and being such a great team-mate! Thanks to Jun for guiding me at the very beginning of my PhD and teaching me joys of cell labelling. Thanks to Giulia for being the best office-mate and a great senior PhD student. I hope you really enjoy London and wish you all the best for your future career. Thanks to Christine for being the most reliable person imaginable, to your immediate responses to any issues and your kind support. I hope your company will thrive for years and will manage to cure cancer. Thanks to Joana for being a wonderful lab manager, and a great help. Thank you for always being so open and so honest. Thank you, Iker, for being an excellent walking-home-through-the-forest-mate and for all your support in the lab. Thanks to Giorgia for being a great running partner. Thanks to Gemma and Ivana for being the great labelling support group. Thanks to Aritra for interesting conversations about fluorescence and scaling laws. Thank you, Miao for joining the project and for your helpful efforts. Thank you Christopher, Paul, Alba, Nataliia, Daniel, Nike, Panos, Catherine, Niccolo,Swati.

I would also like to thank current and former members of the Kräusslich lab and especially Martin. Thank you, Martin, for teaching me everything about the HIV-1 capsid and for always being able to answer all my questions and thank you, Kräusslich lab members for welcoming me in your group for a few days.

I would like to express my gratitude to all the people who make EMBL such a special organisation. Thanks to the Graduate office, Monika, Matija and Carolina for making and a great PhD programme and helping us get through it. Thanks to the all the core facilities who provide great resources, helpful advice and support, especially the Advanced Light Microscopy facility and Marko Lampe in particular, the Mechanical workshop and Helmuth Schaar in particular and the kitchen staff. Thank you, the administrative support. Thank you very much, members of the EDC committee. Thank you for letting me organise many events, for your support and for making EMBL a better and even more welcoming working place. I hope you will continue setting the high standards in equality and diversity in science. Special thanks to Malvika, Laura, Cris and Paul in being the EDC dream-team.

I am extremely lucky to have met many amazing people at EMBL, who also became my good friends. Thank you so much for your support, for all the mornings, afternoons and evening we spent together, for all the coffees, beers and wines we had, for all the trips, walks and runs we did, for all the fun and crazy parties we organised and attended and all the movies we watched together. Thank you Oeyvind, Carina, Luca, Joran, George, Sanji, Renato, Diana, Filipe, Sebastian, Ashna, Areeb. Tusen takk, Merci beaucoup, Grazzi hafna, Aapaka bahut bahut dhanyavaad, Muito obrigado, Vielen Dank!

Thanks a lot to my family for your tremendous support and all the love you transmit from Moscow, Ulm, Berlin and Liverpool!

I do not want to sound too cheesy, however I matured a lot over these four years. I have had grate time and would like to thank my loving husband for always supporting me, being fondue of me, not letting me feel blue or getting feta up with my PhD. Through this experience I hope one day to become a paneer in my field of research and more importantly brie gouda at it. Thank you, for being the most wonderful person you are and the most supportive partner one can dream of. Thank you for brightening every minute of my day and for your brilliant sense of humour. And, of course, a big thank you for proofreading this thesis!

Bibliography

[1]   Martin Beck and Ed Hurt. *The nuclear pore complex: Understanding its function through structural insight*. 2017.

[2]   Lui Ki and Huang Ying. *RanGTPase: A key regulator of nucleocytoplasmic trafficking*. 2009.

[3]   Dirk Görlich, Michael J. Seewald, and Katharina Ribbeck. "Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation". In: *EMBO Journal* (2003).

[4]   Marie-Therese Mackmull et al. "Landscape of nuclear transport receptor cargo specificity". In: *Molecular Systems Biology* (2017).

[5]   Margarida Gama-Carvalho and Maria Carmo-Fonseca. *The rules and roles of nucleocytoplasmic shuttling proteins*. 2001.

[6]   Arturo Vera Rodriguez, Steffen Frey, and Dirk Görlich. "Engineered SUMO/protease system identifies Pdr6 as a bidirectional nuclear transport receptor". In: *The Journal of cell biology* (2019).

[7]   Metin Aksu et al. "Xpo7 is a broad-spectrum exportin and a nuclear import receptor". In: *Journal of Cell Biology* (2018).

[8]   Bruno Catimel et al. "Biophysical Characterization of Interactions Involving Importin-α during Nuclear Import". In: *Journal of Biological Chemistry* (2001).

[9]   Lisa A. Strawn et al. "Minimal nuclear pore complexes define FG repeat domains essential for transport". In: *Nature Cell Biology* (2004).

[10]   Bastian B Hülsmann, Aksana A Labokha, and Dirk Görlich. "The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model." In: *Cell* 150.4 (2012), pp. 738–51.

[11]   Sigrid Milles et al. "Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors." English. In: *Cell* 163.3 (2015), pp. 734–45.

[12]   Alessandro Ori et al. "Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines." In: *Molecular systems biology* 9 (2013), p. 648.

[13]   Bernhard Hampoelz et al. "Structure and Assembly of the Nuclear Pore Complex". In: *Annual Review of Biophysics* (2019).

[14]   J. Kosinski et al. "Molecular architecture of the inner ring scaffold of the human nuclear pore complex". en. In: *Science* 352.6283 (2016), pp. 363–365.

[15] Julia Mahamid et al. "Visualizing the molecular sociology at the HeLa cell nuclear periphery". In: *Science* (2016).

[16] Tal Maimon et al. "The human nuclear pore complex as revealed by cryo-electron tomography". In: *Structure* (2012).

[17] Shyamal Mosalaganti et al. "In situ architecture of the algal nuclear pore complex". In: *Nature Communications* (2018).

[18] Martin Beck and Wolfgang Baumeister. *Cryo-Electron Tomography: Can it Reveal the Molecular Sociology of Cells in Atomic Detail?* 2016.

[19] Katharina Ribbeck and Dirk Görlich. "The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion". In: *EMBO Journal* (2002).

[20] N. Shulga and D. S. Goldfarb. "Binding Dynamics of Structural Nucleoporins Govern Nuclear Pore Complex Permeability and May Mediate Channel Gating". In: *Molecular and Cellular Biology* (2003).

[21] Daniel Stoffler et al. "Calcium-mediated structural changes of native nuclear pore complexes monitored by time-lapse atomic force microscopy". In: *Journal of Molecular Biology* (1999).

[22] Floria Lupu et al. "Nuclear Pore Composition Regulates Neural Stem/Progenitor Cell Differentiation in the Mouse Embryo". In: *Developmental Cell* (2008).

[23] Julien Sellés et al. "Nuclear pore complex plasticity during developmental process as revealed by super-resolution microscopy". In: *Scientific Reports* (2017).

[24] Sasikumar Rajoo et al. "Stoichiometry and compositional plasticity of the yeast nuclear pore complex revealed by quantitative fluorescence microscopy." In: *Proceedings of the National Academy of Sciences of the United States of America* (2018).

[25] Melissa Kane et al. "Nuclear pore heterogeneity influences HIV-1 infection and the antiviral activity of MX2". In: *eLife* (2018).

[26] Hema Chug et al. "Crystal structure of the metazoan Nup62•Nup58•Nup54 nucleoporin complex". In: *Science* (2015).

[27] Chiaki Mizuguchi-Hata et al. "Quantitative regulation of nuclear pore complex proteins by O-GlcNAcylation". In: *Biochimica et Biophysica Acta - Molecular Cell Research* (2013).

[28] Ruchika Sachdev et al. "The C-terminal domain of Nup93 is essential for assembly of the structural backbone of nuclear pore complexes". In: *Molecular Biology of the Cell* (2012).

[29] Daniel H. Lin et al. "Architecture of the symmetric core of the nuclear pore". In: *Science* (2016).

[30] Zhige Wu et al. "Nup62, associated with spindle microtubule rather than spindle matrix, is involved in chromosome alignment and spindle assembly during mitosis". In: *Cell Biology International* (2016).

[31] Chieko Hashizume et al. "Nucleoporin Nup62 maintains centrosome homeostasis". In: *Cell Cycle* (2013).

[32] Kiyomitsu Miyachi et al. "Profile and clinical significance of anti-nuclear envelope antibodies found in patients with primary biliary cirrhosis: A multicenter study". In: *Journal of Autoimmunity* (2003).

[33] I. Han, E. S. Oh, and J. E. Kudlow. "Responsiveness of the state of O-linked N-acetylglucosamine modification of nuclear pore protein p62 to the extracellular glucose concentration". In: *Biochemical Journal* (2000).

[34] Yayoi Kinoshita et al. "Role for NUP62 depletion and PYK2 redistribution in dendritic retraction resulting from chronic stress". In: *Proceedings of the National Academy of Sciences of the United States of America* (2014).

[35] Ana M.C. Faria et al. "The nucleoporin Nup96 is required for proper expression of interferon-regulated proteins and functions". In: *Immunity* (2006).

[36] Xue Qin Chen et al. "Identification of genes encoding putative nucleoporins and transport factors in the fission yeast Schizosaccharomyces pombe: A deletion analysis". In: *Yeast* (2004).

[37] J. S. Rosenblum and G. Blobel. "Autoproteolysis in nucleoporin biogenesis". In: *Proceedings of the National Academy of Sciences* (2002).

[38] J. Enninga, A. Levay, and B. M. A. Fontoura. "Sec13 Shuttles between the Nucleus and the Cytoplasm and Stably Interacts with Nup96 at the Nuclear Pore Complex". In: *Molecular and Cellular Biology* (2003).

[39] Khanh Huy Bui et al. "Integrated structural analysis of the human nuclear pore complex scaffold". In: *Cell* (2013).

[40] E. R. Griffis. "Nup98 Localizes to Both Nuclear and Cytoplasmic Sides of the Nuclear Pore and Binds to Two Distinct Nucleoporin Subcomplexes". In: *Molecular Biology of the Cell* (2003).

[41] Guillaume Chatel et al. "Domain topology of nucleoporin Nup98 within the nuclear pore complex". In: *Journal of Structural Biology* (2012).

[42] Colin E.J. Pritchard et al. "RAE1 is a shuttling mRNA export factor that binds to a GLEBS-like NUP98 motif at the nuclear pore complex through multiple domains". In: *Journal of Cell Biology* (2000).

[43] Daniel H. Lin and André Hoelz. "The Structure of the Nuclear Pore Complex (An Update)". In: *Annual Review of Biochemistry* (2019).

[44] Alec E. Hodel et al. "The three-dimensional structure of the autoproteolytic, nuclear pore-targeting domain of the human nucleoporin Nup98". In: *Molecular Cell* (2002).

[45] M. Oka et al. "The Mobile FG Nucleoporin Nup98 Is a Cofactor for Crm1-dependent Protein Export". In: *Molecular Biology of the Cell* (2010).

[46] Tobias M. Franks and Martin W. Hetzer. *The role of Nup98 in transcription regulation in healthy and diseased cells*. 2013.

[47] E. R. Griffis. "Nup98 Is a Mobile Nucleoporin with Transcription-dependent Dynamics". In: *Molecular Biology of the Cell* (2002).

[48] Stéphanie Morchoisne-Bolhy et al. "Intranuclear dynamics of the Nup107-160 complex". In: *Molecular Biology of the Cell* (2015).

[49] Hermann B.roder Schmidt and Dirk Görlich. "Nup98 FG domains from diverse species spontaneously phase-separate into particles with nuclear pore-like permselectivity". In: *eLife* (2015).

[50] Katharina Ribbeck and Dirk Görlich. "The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion." In: *The EMBO journal* 21.11 (2002), pp. 2664–71.

[51] I. G. Macara. "Transport into and out of the Nucleus". In: *Microbiology and Molecular Biology Reviews* (2001).

[52] Michael P. Rout et al. "Virtual gating and nuclear transport: The hole picture". In: *Trends in Cell Biology* (2003).

[53] Reiner Peters. "Translocation through the nuclear pore complex: Selectivity and speed by reduction-of-dimensionality". In: *Traffic* (2005).

[54] R. Y. H. Lim et al. "Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport". In: *Proceedings of the National Academy of Sciences* (2006).

[55] S. T. Milner. "Polymer brushes". In: *Science* (1991).

[56] Yusuke Sakiyama et al. "Spatiotemporal dynamics of the nuclear pore complex transport barrier resolved by high-speed atomic force microscopy". In: *Nature Nanotechnology* (2016).

[57] Gustav Emilsson et al. "Polymer brushes in solid-state nanopores form an impenetrable entropic barrier for proteins". In: *Nanoscale* (2018).

[58] Steffen Frey and Dirk Görlich. "A Saturated FG-Repeat Hydrogel Can Reproduce the Permeability Properties of Nuclear Pore Complexes". In: *Cell* (2007).

[59] Steffen Frey, Ralf P. Richter, and Dirk Görlich. "FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties". In: *Science* (2006).

[60] Aksana A. Labokha et al. "Systematic analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes". In: *EMBO Journal* (2013).

[61] Sigrid Milles et al. "Facilitated aggregation of FG nucleoporins under molecular crowding conditions". In: *EMBO Reports* (2013).

[62] Samir S. Patel et al. "Natively Unfolded Nucleoporins Gate Protein Diffusion across the Nuclear Pore Complex". In: *Cell* (2007).

[63] Roderick Y.H. Lim and Larisa E. Kapinos. "How to operate a nuclear pore complex by kap-centric control". In: *Nucleus* (2015).

[64] Alan R. Lowe et al. "Importin-β modulates the permeability of the nuclear pore complex in a Ran-dependent manner". In: *eLife* (2015).

[65] Alan R Lowe et al. "Selectivity mechanism of the nuclear pore complex characterized by single cargo tracking." In: *Nature* 467.7315 (2010), pp. 600–3.

[66] Larisa E. Kapinos et al. "Karyopherins regulate nuclear pore complex barrier and transport function". In: *Journal of Cell Biology* (2017).

[67] George J. Stanley et al. "Quantification of Biomolecular Dynamics inside Real and Synthetic Nuclear Pore Complexes using Time-Resolved Atomic Force Microscopy". In: *ACS Nano* (2019).

[68] Chi W. Pak et al. "Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein". In: *Molecular Cell* (2016).

[69] Tyler S. Harmon et al. "Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins". In: *eLife* (2017).

[70] Yuan Lin, Simon L. Currie, and Michael K. Rosen. "Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs". In: *Journal of Biological Chemistry* (2017).

[71] Erik W. Martin and Tanja Mittag. *Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions*. 2018.

[72] V. N. Uversky. "Intrinsically Disordered Proteins". In: *Brenner's Encyclopedia of Genetics: Second Edition*. 2013. ISBN: 9780080961569.

[73] Peter E. Wright and H. Jane Dyson. *Intrinsically disordered proteins in cellular signalling and regulation*. 2015.

[74] H. Broder Schmidt and Dirk Görlich. *Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles*. 2016.

[75] Justin Yamada et al. "A Bimodal Distribution of Two Distinct Categories of Intrinsically Disordered Structures with Separate Functions in FG Nucleoporins". In: *Molecular & Cellular Proteomics* (2010).

[76] Chad Gu et al. "The Role of Cohesiveness in the Permeability of the Spatial Assemblies of FG Nucleoporins". In: *Biophysical Journal* (2019).

[77] Andrei Vovk et al. "Simple biophysics underpins collective conformations of the intrinsically disordered proteins of the nuclear pore complex". In: *eLife* (2016).

[78] Mario Tagliazucchi et al. "Effect of charge, hydrophobicity, and sequence of nucleoporins on the translocation of model particles through the nuclear pore complex". In: *Proceedings of the National Academy of Sciences of the United States of America* (2013).

[79] Kai Huang et al. "Molecular model of the nuclear pore complex reveals a thermoreversible FG-network with distinct territories occupied by different FG motifs". In: *bioRxiv* (2019).

[80] Luke K. Davis et al. "Intrinsically disordered nuclear pore proteins show ideal-polymer morphologies and dynamics". In: *bioRxiv* (2019).

[81] Marina Mikhaylova et al. "Resolving bundled microtubules using anti-tubulin nanobodies". In: *Nature Communications* (2015).

[82] Jonas Ries et al. "A simple, versatile method for GFP-based super-resolution microscopy via nanobodies". In: *Nature Methods* (2012).

[83] Georgyi V. Los et al. "HaloTag: A novel protein labeling technology for cell imaging and protein analysis". In: *ACS Chemical Biology* (2008).

[84] Antje Keppler et al. "A general method for the covalent labeling of fusion proteins with small molecules in vivo". In: *Nature Biotechnology* (2003).

[85] Arnaud Gautier et al. "An Engineered Protein Tag for Multiprotein Labeling in Living Cells". In: *Chemistry and Biology* (2008).

[86] Steffen J. Sahl, Stefan W. Hell, and Stefan Jakobs. *Fluorescence nanoscopy in cell biology*. 2017.

[87] Stephen R. Adams et al. "New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: Synthesis and biological applications". In: *Journal of the American Chemical Society* (2002).

[88] David Virant et al. "A peptide tag-specific nanobody enables high-quality labeling for dSTORM imaging". In: *Nature Communications* (2018).

[89] Howard M. Goodman et al. "Amber suppression: A Nucleotide change in the anticodon of a tyrosine transfer RNA". In: *Nature* (1968).

[90] Jingchun Sun et al. "Relationships among stop codon usage bias, its context, isochores, and gene expression level in various eukaryotes". In: *Journal of Molecular Evolution* (2005).

[91] Wei Wan, Jeffery M. Tharp, and Wenshe R. Liu. *Pyrrolysyl-tRNA synthetase: An ordinary enzyme but an outstanding genetic code expansion tool*. 2014.

[92] Wenshe Liu et al. "Genetic incorporation of unnatural amino acids into proteins in mammalian cells". In: *Nature Methods* (2007).

[93] Tilman Plass et al. "Amino acids for diels-alder reactions in living cells". In: *Angewandte Chemie - International Edition* (2012).

[94] Annika Borrmann et al. "Genetic Encoding of a Bicyclo[6.1.0]nonyne-Charged Amino Acid Enables Fast Cellular Protein Imaging by Metal-Free Ligation". In: *ChemBioChem* (2012).

[95] Ivana Nikić et al. "Minimal tags for rapid dual-color live-cell labeling and super-resolution microscopy." In: *Angewandte Chemie (International ed. in English)* 53.8 (2014), pp. 2245–9.

[96] Ivana Nikic et al. "Minimal tags for rapid dual-color live-cell labeling and super-resolution microscopy". In: *Angewandte Chemie - International Edition* (2014).

[97] Chayasith Uttamapinant et al. "Genetic Code Expansion Enables Live-Cell and Super-Resolution Imaging of Site-Specifically Labeled Cellular Proteins". In: *Journal of the American Chemical Society* (2015).

[98] I. Nikić et al. "Debugging Eukaryotic Genetic Code Expansion for Site-Specific Click-PAINT Super-Resolution Microscopy". In: *Angewandte Chemie - International Edition* 55.52 (2016).

[99] M. G.L. Gustafsson. "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy". In: *Journal of Microscopy* (2000).

[100] Lothar Schermelleh et al. "Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy". In: *Science* (2008).

[101] M. G. L. Gustafsson. "Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution". In: *Proceedings of the National Academy of Sciences* (2005).

[102] S W Hell and J Wichmann. "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy." In: *Optics letters* (1994).

[103] Eric Betzig et al. "Imaging intracellular fluorescent proteins at nanometer resolution". In: *Science* (2006).

[104] Samuel T. Hess, Thanu P.K. Girirajan, and Michael D. Mason. "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy". In: *Biophysical Journal* (2006).

[105] Michael J Rust, Mark Bates, and Xiaowei Zhuang. "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)." In: *Nature methods* (2006).

[106] Mike Heilemann et al. "Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes." In: *Angewandte Chemie (International ed. in English)* 47.33 (2008), pp. 6172–6.

[107] Mike Heilemann et al. "Super-resolution imaging with small organic fluorophores". In: *Angewandte Chemie - International Edition* (2009).

[108] Graham T. Dempsey et al. "Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging". In: *Nature Methods* (2011).

[109] Jonas Fölling et al. "Fluorescence nanoscopy by ground-state depletion and single-molecule return". In: *Nature Methods* (2008).

[110] Mark Bates et al. "Multicolor super-resolution fluorescence imaging via multiparameter fluorophore detection". In: *ChemPhysChem* (2012).

[111] Kim I Mortensen et al. "Optimized localization analysis for single-molecule tracking and super-resolution microscopy." In: *Nature methods* (2010).

[112] Niccolò Banterle et al. "Fourier ring correlation as a resolution criterion for super-resolution microscopy". In: *Journal of Structural Biology* (2013).

[113] Robert P.J. Nieuwenhuizen et al. "Measuring image resolution in optical nanoscopy". In: *Nature Methods* (2013).

[114] Siân Culley et al. "Quantitative mapping and minimization of super-resolution optical imaging artifacts". In: *Nature Methods* (2018).

[115] A. Loschberger et al. "Super-resolution imaging visualizes the eightfold symmetry of gp210 proteins around the nuclear pore complex and resolves the central channel with nanometer resolution". In: *Journal of Cell Science* (2012).

[116] Fabian Göttfert et al. "Coaligned dual-channel STED nanoscopy and molecular diffusion analysis at 20 nm resolution". In: *Biophysical Journal* (2013).

[117] Alessandro Ori et al. "Cell type-specific nuclear pores: A case in point for context-dependent stoichiometry of molecular machines". In: *Molecular Systems Biology* (2013).

[118] Anna Szymborska et al. "Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging." en. In: *Science (New York, N.Y.)* 341.6146 (2013), pp. 655–8.

[119] Jervis V. Thevathasan et al. "Nuclear Pores as Universal Reference Standards for Quantitative Microscopy". In: *Biophysical Journal* (2019).

[120] Alexander von Appen et al. "In situ structural analysis of the human nuclear pore complex." In: *Nature* 526.7571 (2015), pp. 140–3.

[121] Tino Pleiner et al. "Nanobodies: site-specific labeling for super-resolution imaging, rapid epitope-mapping and native protein complex isolation". en. In: *eLife* 4 (2015), e11349.

[122] Evgenia Platonova, Christian M Winterflood, and Helge Ewers. "A simple method for GFP- and RFP-based dual color single-molecule localization microscopy." In: *ACS chemical biology* 10.6 (2015), pp. 1411–6.

[123] Guo Fu et al. "Investigating molecular crowding within nuclear pores using polarization-PALM". In: *eLife* (2017).

[124] Th Forster. "Energiewanderung und Fluoreszenz". In: *Die Naturwissenschaften* (1946).

[125] Benjamin Schuler and William A. Eaton. *Protein folding studied by single-molecule FRET*. 2008.

[126] Stanislav Kalinin et al. "Detection of Structural Dynamics by FRET: A Photon Distribution and Fluorescence Lifetime Analysis of Systems with Multiple States". In: *Journal of Physical Chemistry B* (2010).

[127] Mark C. Overton and Kendall J. Blumer. "G-protein-coupled receptors function as oligomers in vivo". In: *Current Biology* (2000).

[128] R. G. H. Immink et al. "Analysis of MADS box protein-protein interactions in living plant cells". In: *Proceedings of the National Academy of Sciences* (2002).

[129] Atsushi Miyawaki et al. "Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin". In: *Nature* (1997).

[130] Adam Hoppe, Kenneth Christensen, and Joel A. Swanson. "Fluorescence resonance energy transfer-based stoichiometry in living cells". In: *Biophysical Journal* (2002).

[131] Jacco Van Rheenen, Michiel Langeslag, and Kees Jalink. "Correcting Confocal Acquisition to Optimize Imaging of Fluorescence Resonance Energy Transfer by Sensitized Emission". In: *Biophysical Journal* (2004).

[132] T. S. Karpova et al. "Fluorescence resonance energy transfer from cyan to yellow fluorescent protein detected by acceptor photobleaching using confocal microscopy and a single laser". In: *Journal of Microscopy* (2003).

[133] Hongzhan Xu et al. "Evidence for biphasic uncoating during HIV-1 infection from a novel imaging assay". In: *Retrovirology* 10 (2013), p. 1.

[134] Y. Gu et al. "Quantitative fluorescence resonance energy transfer (FRET) measurement with acceptor photobleaching and spectral unmixing". In: *Journal of Microscopy* (2004).

[135] M. E. Dickinson et al. "Multi-spectral imaging and linear unmixing add a whole new dimension to laser scanning fluorescence microscopy". In: *BioTechniques* (2001).

[136] Gerald W. Gordon et al. "Quantitative fluorescence resonance energy transfer measurements using fluorescence microscopy". In: *Biophysical Journal* (1998).

[137] D. Phillips and R. L. Christensen. "Time Correlated Single-Photon Counting (Tcspc) Using Laser Excitation". In: *Instrumentation Science & Technology* (1985).

[138] R. R. Duncan et al. "Multi-dimensional time-correlated single photon counting (TCSPC) fluorescence lifetime imaging microscopy (FLIM) to detect FRET in cells". In: *Journal of Microscopy* (2004).

[139] Wolfgang Becker. *Advanced Time-Correlated Single Photon Counting Techniques.* 2005. ISBN: 978-3-540-26047-9.

[140] Johan Philip and Kjell Carlsson. "Theoretical investigation of the signal-to-noise ratio in fluorescence lifetime imaging". In: *Journal of the Optical Society of America A* (2003).

[141] Michael Wahl and PicoQuant GmbH. *Time-correlated single photon counting.* 2014. ISBN: 0125241402.

[142] Joseph R. Lakowicz. *Principles of fluorescence spectroscopy.* 2006. ISBN: 0387312781.

[143] Joseph R. Lakowicz et al. "Correction for incomplete labeling in the measurement of distance distributions by frequency-domain fluorometry". In: *Analytical Biochemistry* (1991).

[144] Hai-Jui Lin, Petr Herman, and Joseph R. Lakowicz. "Fluorescence lifetime-resolved pH imaging of living cells". In: *Cytometry* (2003).

[145] A. V. Agronskaia, L. Tertoolen, and H. C. Gerritsen. "Fast fluorescence lifetime imaging of calcium in living cells". In: *Journal of Biomedical Optics* (2004).

[146] Hiromi Kurokawa et al. "High resolution imaging of intracellular oxygen concentration by phosphorescence lifetime". In: *Scientific Reports* (2015).

[147] Hideki Itoh et al. "Direct organelle thermometry with fluorescence lifetime imaging microscopy in single myotubes". In: *Chemical Communications* (2016).

[148] Kohki Okabe et al. "Intracellular temperature mapping with a fluorescent polymeric thermometer and fluorescence lifetime imaging microscopy". In: *Nature Communications* (2012).

[149] Marina K. Kuimova et al. "Molecular rotor measures viscosity of live cells via fluorescence lifetime imaging". In: *Journal of the American Chemical Society* (2008).

[150] L. Pytowski et al. "A novel technique for mapping viscosity in discrete subcellular locations with a BODIPY based fluorescent probe". In: *bioRxiv* ().

[151] Carlos Manlio Díaz-García et al. "Quantitative in vivo imaging of neuronal glucose concentrations with a genetically encoded fluorescence lifetime sensor". In: *Journal of Neuroscience Research* (2019).

[152] A. Periasamy and R.M. Clegg. *FLIM Microscopy in Biology and Medicine.* Taylor and Francis Group, CRC Press, 2010.

[153] L. Marcu, P.M.W French, and D.S. Elson. *Fluorescence lifetime spectroscopy and imaging. Principles and Applications in Biomedical Diagnostics.* CRC Press, Taylor and Francis Group, 2015.

[154] Lydia Sauer. "Review of clinical approaches in fluorescence lifetime imaging ophthalmoscopy". In: *Journal of Biomedical Optics* (2018).

[155] Shilpi Rajoria et al. "FLIM-FRET for Cancer Applications". In: *Current Molecular Imaging* (2014).

[156] James McGinty et al. "Wide-field fluorescence lifetime imaging of cancer". In: *Biomedical Optics Express* (2010).

[157] Farzad Fereidouni, Arjen N. Bader, and Hans C. Gerritsen. "Spectral phasor analysis allows rapid and reliable unmixing of fluorescence microscopy spectral images". In: *Optics Express* (2012).

[158] Iwo König et al. "Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells". In: *Nature Methods* (2015).

[159] Daniel Nettels, Armin Hoffmann, and Benjamin Schuler. "Unfolded protein and peptide dynamics investigated with single-molecule FRET and correlation spectroscopy from picoseconds to seconds". In: *Journal of Physical Chemistry B* (2008).

[160] Benjamin Schuler et al. "Application of confocal single-molecule FRET to intrinsically disordered proteins". In: *Methods in Molecular Biology* (2012).

[161] E. Haas et al. "Distribution of end-to-end distances of oligopeptides in solution as estimated by energy transfer." In: *Proceedings of the National Academy of Sciences* (1975).

[162] A. Grinvald, E. Haas, and I. Z. Steinberg. "Evaluation of the Distribution of Distances Between Energy Donors and Acceptors by Fluorescence Decay". In: *Proceedings of the National Academy of Sciences* (2006).

[163] Benjamin Schuler et al. "Single-Molecule FRET Spectroscopy and the Polymer Physics of Unfolded and Intrinsically Disordered Proteins". In: *Annual Review of Biophysics* (2016).

[164] Sean C. Warren et al. "Rapid Global Fitting of Large Fluorescence Lifetime Imaging Microscopy Datasets". In: *PLoS ONE* (2013).

[165] Waldemar Schrimpf et al. "PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data". In: *Biophysical Journal* (2018).

[166] Michael Rubinstein and Ralph H. Colby. *Polymer Physics*. 2003. ISBN: 019852059X.

[167] Yuri E. Nesmelov et al. "Structural kinetics of myosin by transient time-resolved FRET". In: *Proceedings of the National Academy of Sciences* (2011).

[168] Thomas Otavio Peulen, Oleg Opanasyuk, and Claus A.M. Seidel. "Combining Graphical and Analytical Methods with Molecular Simulations to Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately". In: *Journal of Physical Chemistry B* (2017).

[169] Ivan Peran et al. "Unfolded states under folding conditions accommodate sequence-specific conformational preferences with random coil-like dimensions". In: *Proceedings of the National Academy of Sciences* (2019).

[170] Gustavo Fuertes et al. "Decoupling of size and shape fluctuations in heteropolymeric sequences reconciles discrepancies in SAXS vs. FRET measurements". In: *Proceedings of the National Academy of Sciences* (2017).

[171] P. G De Gennes. *Scaling concepts in polymer physics. Cornell university press.* 1979. ISBN: 080141203X.

[172] Paul J. Flory. *Principles of Polymer Chemistry.* 1953. ISBN: 1461422124.

[173] Somendra M. Bhattacharjee, Achille Giacometti, and Amos Maritan. *Flory theory for polymers.* 2013.

[174] Thomas E. Gartner and Arthi Jayaraman. *Modeling and Simulations of Polymers: A Roadmap.* 2019.

[175] H. Hofmann et al. "Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy". In: *Proceedings of the National Academy of Sciences* (2012).

[176] Feng Ding, Ramesh K. Jha, and Nikolay V. Dokholyan. "Scaling behavior and structure of denatured proteins". In: *Structure* (2005).

[177] Liu Hong and Jinzhi Lei. "Scaling law for the radius of gyration of proteins and its dependence on hydrophobicity". In: *Journal of Polymer Science, Part B: Polymer Physics* (2009).

[178] Joshua A. Riback et al. "Commonly used FRET fluorophores promote collapse of an otherwise disordered protein". In: *Proceedings of the National Academy of Sciences of the United States of America* (2019).

[179] André Lampe et al. "Multi-colour direct STORM with red emitting carbocyanines". In: *Biology of the Cell* (2012).

[180] Peter Dedecker et al. "Localizer: fast, accurate, open-source, and modular software package for superresolution microscopy". In: *Journal of Biomedical Optics* (2012).

[181] Shunit Gal-Ben-Ari et al. "PKR: A Kinase to Remember". In: *Frontiers in Molecular Neuroscience* (2019).

[182] Jun Hee Kang. "Precise mapping of disordered regions of nucleoporins in situ". PhD thesis. Ruperto-Carola University of Heidelberg, 2016, pp. 1–113.

[183] Ivana Nikic et al. "Labeling proteins on live mammalian cells using click chemistry". In: *Nature Protocols* (2015).

[184] Jiong Ma et al. "Super-resolution mapping of scaffold nucleoporins in the nuclear pore complex". In: *Journal of Cell Science* (2017).

[185] Susan M. Abmayr et al. "Preparation of Nuclear and Cytoplasmic Extracts from Mammalian Cells". In: *Current Protocols in Molecular Biology* (2006).

[186] Alan J. Herr, John F. Atkins, and Raymond F. Gesteland. "Coupling of Open Reading Frames by Translational Bypassing". In: *Annual Review of Biochemistry* (2000).

[187] Rafael Maldonado and Alan J. Herr. "Efficiency of T4 gene 60 translational bypassing". In: *Journal of Bacteriology* (1998).

[188] D R Finlay et al. "Inhibition of in vitro nuclear transport by a lectin that binds to nuclear pores." In: *The Journal of cell biology* 104.2 (1987), pp. 189–200.

[189] Bin Li and Jennifer J. Kohler. *Glycosylation of the nuclear pore*. 2014.

[190] Lena Voith von Voithenberg et al. "Recognition of the 3 splice site RNA by the U2AF heterodimer involves a dynamic population shift". In: *Proceedings of the National Academy of Sciences* (2016).

[191] Anna Cleta Croce and G. Bottiroli. "Autofluorescence spectroscopy and imaging: A tool for biomedical research and diagnosis". In: *European Journal of Histochemistry* (2014).

[192] Jérémy Surre et al. "Strong increase in the autofluorescence of cells signals struggle for survival". In: *Scientific Reports* (2018).

[193] Kathrin Lang et al. "Genetic encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic diels-alder reactions". In: *Journal of the American Chemical Society* (2012).

[194] Qinsi Zheng et al. *Ultra-stable organic fluorophores for single-molecule research*. 2014.

[195] Malte Köllner and Jürgen Wolfrum. "How many photons are necessary for fluorescence-lifetime measurements?" In: *Chemical Physics Letters* (1992).

[196] P. R. Barber et al. "Multiphoton time-domain fluorescence lifetime imaging microscopy: Practical application to protein-protein interactions using global analysis". In: *Journal of the Royal Society Interface*. 2009.

[197] Tonni Grube Andersen et al. "Improving analytical methods for protein-protein interaction through implementation of chemically inducible dimerization". In: *Scientific Reports* (2016).

[198] J. Chen et al. "Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue." In: *Proceedings of the National Academy of Sciences* (1995).

[199] Barbara K. Müller et al. "Pulsed interleaved excitation". In: *Biophysical Journal* (2005).

[200] Jelle Hendrix et al. "Pulsed interleaved excitation fluctuation imaging". In: *Biophysical Journal* (2013).

[201] Armin Hoffmann et al. "Mapping protein collapse with single-molecule flourescence and kinetic synchrotron radiation circular dichroism spectroscopy". In: *Proceedings of the National Academy of Sciences of the United States of America* (2007).

[202] André Hoelz, Joseph S. Glavy, and Martin Beck. *Toward the atomic structure of the nuclear pore complex: When top down meets bottom up*. 2016.

[203] Simon J. Elsässer et al. "Genetic code expansion in stable cell lines enables encoded chromatin modification". In: *Nature Methods* (2016).

[204] Yayoi Kinoshita et al. "Nuclear distributions of NUP62 and NUP214 suggest architectural diversity and spatial patterning among nuclear pore complexes". In: *PLoS ONE* (2012).

[205] Ralf Jungmann et al. "Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT". In: *Nature Methods* (2014).

[206] Joerg Schnitzbauer et al. "Super-resolution microscopy with DNA-PAINT". In: *Nature Protocols* (2017).

[207] Anne Monette, Nelly Panté, and Andrew J. Mouland. "HIV-1 remodels the nuclear pore complex". In: *Journal of Cell Biology* (2011).

[208] Lorenzo Albertazzi et al. "Quantitative FRET analysis with the E0GFP-mCherry fluorescent protein pair". In: *Photochemistry and Photobiology* (2009).

[209] Björn Hellenkamp et al. "Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study". In: *Nature Methods* (2018).

[210] A. Plochowietz et al. "Stable end-sealed DNA as robust nano-rulers for: In vivo single-molecule fluorescence". In: *Chemical Science* (2016).

[211] Peter J. Verveer, Anthony Squire, and Philippe I.H. Bastiaens. "Global analysis of fluorescence lifetime imaging microscopy data". In: *Biophysical Journal* (2000).

[212] S. Pelet et al. "A fast global fitting algorithm for fluorescence lifetime imaging microscopy based on image segmentation." In: *Biophysical journal* (2004).

[213] Michelle A. Digman et al. "The phasor approach to fluorescence lifetime imaging analysis". In: *Biophysical Journal* (2008).

[214] Jieqiong Lou et al. "Phasor histone FLIM-FRET microscopy quantifies spatiotemporal rearrangement of chromatin architecture during the DNA damage response". In: *Proceedings of the National Academy of Sciences of the United States of America* (2019).

[215] Henryk Szmacinski, Vladimir Toshchakov, and Joseph R. Lakowicz. "Application of phasor plot and autofluorescence correction for study of heterogeneous cell population". In: *Journal of Biomedical Optics* (2014).

[216] Irina V. Gopich and Attila Szabo. "Theory of the energy transfer efficiency and fluorescence lifetime distribution in single-molecule FRET". In: *Proceedings of the National Academy of Sciences of the United States of America* (2012).

[217] Edward P. O'Brien et al. "How accurate are polymer models in the analysis of Förster resonance energy transfer experiments on proteins?" In: *Journal of Chemical Physics* (2009).

[218] Paul J. Flory. "Statistical Mechanics of Dilute Polymer Solutions". In: *J. Chem. Phys.* 17.1347 (1949).

[219] Jonathan B. Grimm et al. "A general method to improve fluorophores for live-cell and single-molecule microscopy". In: *Nature Methods* (2015).

[220] N. Lindlein et al. "High numerical aperture imaging with different polarization patterns". In: *Optics Express* (2007).

[221] G. Paci and E.A. Lemke. "Molecular determinants of large cargo transport into the nucleus". In: *bioRxiv* (2019).

[222] Valerie Le Sage and Andrew J. Mouland. *Viral subversion of the nuclear pore complex*. 2013.

[223] Edward M. Campbell and Thomas J. Hope. "HIV-1 capsid: the multifaceted key player in HIV-1 infection". In: *Nature Reviews Microbiology* 13.8 (2015), pp. 471–483.

[224] Kallesh Danappa Jayappa, Zhujun Ao, and Xiaojian Yao. *The HIV-1 passage from cytoplasm to nucleus: The process involving a complex exchange between the components of HIV-1 and cellular machinery to access nucleus and successful integration*. 2012.

[225] Barbie K. Ganser-Pornillos, Anchi Cheng, and Mark Yeager. "Structure of Full-Length HIV-1 CA: A Model for the Mature Capsid Lattice". In: *Cell* (2007).

[226] Gongpu Zhao et al. "Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics." In: *Nature* 497.7451 (2013), pp. 643–6.

[227] N. Pante. "Nuclear Pore Complex Is Able to Transport Macromolecules with Diameters of 39 nm". In: *Molecular Biology of the Cell* (2002).

[228] Katsiaryna Bichel et al. "HIV-1 capsid undergoes coupled binding and isomerization by the nuclear pore protein NUP358". In: *Retrovirology* (2013).

[229] Amanda J. Price et al. "Host Cofactors and Pharmacologic Ligands Share an Essential Interface in HIV-1 Capsid That Is Lost upon Disassembly". In: *PLoS Pathogens* (2014).

[230] Barbie K. Ganser et al. "Assembly and analysis of conical models for the HIV-1 core". In: *Science* (1999).

[231] Owen Pornillos et al. "X-Ray Structures of the Hexameric Building Block of the HIV Capsid". In: *Cell* (2009).

[232] Daryl A. Bosco et al. "Catalysis of cis/trans isomerization in native HIV-1 capsid by human cyclophilin A". In: *Proceedings of the National Academy of Sciences of the United States of America* (2002).

[233] Y. Li, A. K. Kar, and J. Sodroski. "Target Cell Type-Dependent Modulation of Human Immunodeficiency Virus Type 1 Capsid Disassembly by Cyclophilin A". In: *Journal of Virology* (2009).

[234] Chuang Liu et al. "Cyclophilin A stabilizes the HIV-1 capsid through a novel non-canonical binding site". In: *Nature Communications* (2016).

[235] Cindy Buffone et al. "Nup153 Unlocks the Nuclear Pore Complex for HIV-1 Nuclear Translocation in Nondividing Cells". In: *Journal of Virology* (2018).

[236] Francesca Di Nunzio. *New insights in the role of nucleoporins: A bridge leading to concerted steps from HIV-1 nuclear entry until integration*. 2013.

[237] Francesca Di Nunzio et al. "Nup153 and Nup98 bind the HIV-1 core and contribute to the early steps of HIV-1 replication". In: *Virology* (2013).

[238] Kenneth A Matreyek et al. "Nucleoporin NUP153 phenylalanine-glycine motifs engage a common binding pocket within the HIV-1 capsid protein to mediate lentiviral infectivity." In: *PLoS pathogens* 9.10 (2013), e1003693.

[239] Ke Peng et al. "Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid". In: *eLife* (2014).

[240] Ashwanth C. Francis and Gregory B. Melikyan. *Live-cell imaging of early steps of single HIV-1 infection*. 2018.

[241] Stephen A. Adam, Rachel Sterne Marr, and Larry Gerace. "Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors". In: *Journal of Cell Biology* (1990).

[242] Mary Shannon Moore and Eric D. Schwoebel. "Nuclear Import in Digitonin-Permeabilized Cells". In: *Current Protocols in Cell Biology*. 2001.

[243] Aurelia Cassany and Larry Gerace. "Reconstitution of nuclear import in permeabilized cells". In: *Methods in Molecular Biology* (2008).

[244] Mira Grättinger et al. "In vitro assembly properties of wild-type and cyclophilin-binding defective human immunodeficiency virus capsid proteins in the presence and absence of cyclophilin A". In: *Virology* (1999).

[245] Piau Siong Tan et al. "Two Differential Binding Mechanisms of FG-Nucleoporins and Nuclear Transport Receptors". In: *Cell Reports* (2018).

[246] Iker Valle Aramburu and Edward A. Lemke. *Floppy but not sloppy: Interaction mechanism of FG-nucleoporins and nuclear transport receptors*. 2017.

[247] Akash Bhattacharya et al. "Structural basis of HIV-1 capsid recognition by PF74 and CPSF6". In: ().

[248] P. Pawlica et al. "Functional Evidence for the Involvement of Microtubules and Dynein Motor Complexes in TRIM5 -Mediated Restriction of Retroviruses". In: *Journal of Virology* (2014).

[249] Z. Lukic et al. "HIV-1 Uncoating Is Facilitated by Dynein and Kinesin 1". In: *Journal of Virology* (2014).

[250] Viacheslav Malikov et al. "HIV-1 capsids bind and exploit the kinesin-1 adaptor FEZ1 for inward movement to the nucleus". In: *Nature Communications* (2015).

[251] Adarsh Dharan and Edward M. Campbell. "Role of Microtubules and Microtubule-Associated Proteins in HIV-1 Infection". In: *Journal of Virology* (2018).

[252] Abraham L. Brass et al. "Identification of host proteins required for HIV infection through a functional genomic screen". In: *Science* (2008).

[253] Frauke Christ et al. "Transportin-SR2 Imports HIV into the Nucleus". In: *Current Biology* (2008).

[254] J. C. Valle-Casuso et al. "TNPO3 Is Required for HIV-1 Replication after Nuclear Import but prior to Integration and Binds the HIV-1 Core". In: *Journal of Virology* (2012).

[255] Owen Pornillos, Barbie K. Ganser-Pornillos, and Mark Yeager. "Atomic-level modelling of the HIV capsid". In: *Nature* (2011).

[256] Jaclyn Tetenbaum-Novatt et al. "Nucleocytoplasmic transport: A role for non-specific competition in karyopherin-nucleoporin interactions". In: *Molecular and Cellular Proteomics*. 2012.

[257] Sigrid Milles and Edward A. Lemke. "Mapping multivalency and differential affinities within large intrinsically disordered protein complexes with segmental motion analysis". In: *Angewandte Chemie - International Edition* (2014).

[258] Benjamin A. Shoemaker, John J. Portman, and Peter G. Wolynes. "Speeding molecular recognition by using the folding funnel: The fly-casting mechanism". In: *Proceedings of the National Academy of Sciences of the United States of America* (2000).

[259] Peter Csermely, Robin Palotai, and Ruth Nussinov. "Induced fit, conformational selection and independent dynamic segments: An extended view of binding events". In: *Trends in Biochemical Sciences* (2010).

[260] Owen Pornillos et al. "Disulfide Bond Stabilization of the Hexameric Capsomer of Human Immunodeficiency Virus". In: *Journal of Molecular Biology* (2010).