Characterization of Nuclear Envelope Proteins by Fluorescence Fluctuation Spectroscopy

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Jared Michael Gluek Hennen

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Joachim D. Mueller, Advisor

November 2019

Acknowledgements

The work found in this thesis could not have been completed without the help and support I received over the past 7 years. I would like to thank all of you who have made this work possible.

Dr. Joachim Mueller for the years of patient guidance and advice. Your high standard for research is inspiring and leads to great results. Dr. Yan Chen for our conversations about the direction of my various projects and what their end goals should be. It's easy to get lost in the weeds of research but your concern always kept me directed. Dr. Gant Luxton for showing our lab how much opportunity there was in studying the nuclear envelope.

My fellow lab members: Dr. Kwang-Ho Hur for developing much of the theory necessary to perform this work and the invaluable insights gained from our conversations. Isaac Angert for constantly improving the instrumentation, including fixing a 9 month issue with our brand new ventilation system using only some string, tape, and cardboard, a solution which no engineer can seem to improve on. John Kohler for bringing an expertise in photophysics to study the details of some of the experiments and analysis done here. Siddarth Reddy Karuka for often providing an extra hand or mind whenever needed, leading to helpful discussions on nuclear envelope experiments, career development, or how to best enjoy living in Minnesota.

Finally, my family and friends for the part they played in getting me here and for all of their support. From inspiring my love of science in the first place, to listening to me ramble about some obstacle in my research, to giving me an escape where I don't think about it at all, you all kept me grounded throughout this process.

This thesis is dedicated to my biggest supporter, closest confidant, best friend, and wife Jennifer Smith.

I could not have done this without you.

Having you in my corner means I can accomplish anything.

Abstract

The nuclear evelope (NE) consists of two concentric membranes surrounding the nucleus of eukaryotic cells. The two nuclear membranes are separated by a ~40 nm wide fluid layer, known as the lumen. While the NE is well known as a physical barrier separating the genetic material-containing nucleus from the cytoplasm, the NE has also been recognized as a central hub for cellular signaling. Proteins found in the lumen or at the nuclear membranes are critical to the NE's role in cellular functions, with mutations in these proteins associated with human diseases such as muscular dystrophy and cancer. Despite their significance to human health, the interactions of NE proteins have been limited to *in vitro* studies. This thesis lays the foundation for investigating the interactions and assembly of NE proteins in their native environment by bringing fluorescence fluctuation spectroscopy to the NE of the living cell. After adapting the fluorescence fluctuation technique to the challenging environment of the NE, we succeeded in quantifying the self-oligomerization of proteins within the NE as well as determining their mobility. Our application focuses mainly on constituent proteins of the Linker of Nucleoskeleton and Cytoskeleton (LINC) complex, which are important in nucleocytoplasmic force transduction, with the goal to determine their assembly states spectroscopically. The results of these studies demonstrate the power of the fluorescence fluctuation methods described in this thesis and provide a starting point for extending this work to other NE proteins with the goal to characterize their assembly within their native cellular environment.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACF	autocorrelation function
CC	coiled-coil
CCF	cross-correlation function
DC	dual-color
EGFP	enhanced green fluorescent protein
ER	endoplasmic reticulum
FFS	fluorescence fluctuation spectroscopy
HSP	hetero-species partitioning
INM	inner nuclear membrane
KASH	Klarsich/ANC-1/SYNE homology
LINC	linker of nucleoskeleton and cytoskeleton
mGL	modified Gaussian-Lorentzian
MSQ	mean-segmented Q
NA	numerical aperture
NE	nuclear envelope
NE_D	dorsal nuclear envelope
NE_V	ventral nuclear envelope
NPC	nuclear pore complex
NTD	N-terminal domain
ONM	outer nuclear membrane
OV	observation volume
PNS	perinuclear space

PSF	point spread function
RIPSF	radially integrated point spread function
SC	single-color
SR	spectrin-like repeat
SS	signal sequence
SUN	Sad1/UNC-84
tsMSQ	time-shifted mean-segmented Q
tsQ	time-shifted Q

1. Introduction

Cellular organisms are divided into three domains: Bacteria, Archaea, and Eukarya (1). Both bacteria and archaea consist primarily of simple, single cell organisms while eukaryotes, as members of Eukarya are called, may be either single or multi celled, and include all plants and animals (1). All eukaryotes share a common ancestor in an early archaea which evolved to contain the nucleus and other internal membrane-bound organelles, the fundamental distinguishing feature of eukaryotes (1). As the largest and most readily visible of these organelles, the nucleus (Fig. 1.1) was the first formally discovered in 1710 (2).

Harboring the cell's genetic material, the nucleus is separated from the surrounding cytoplasm by the nuclear envelope (NE) (3). Having only evolved once, the evolutionary benefit of the nucleus is not definitively known (4). Theories as to how the nucleus evolved and what advantage it may have given early eukaryotes can be broadly grouped into two categories. Proponents of an autogenous model of the origin of the nucleus argue that cells may have evolved a full nucleus internally (5). They cite the existence of prokaryotic bacteria containing internal membranes, which can surround the genetic material, as evidence for their model (6). The endosymbiotic origin model argues for a rare event caused by the merging of two distinct cells, with the plasma membranes of each cell forming the NE's unique double membrane structure (4).

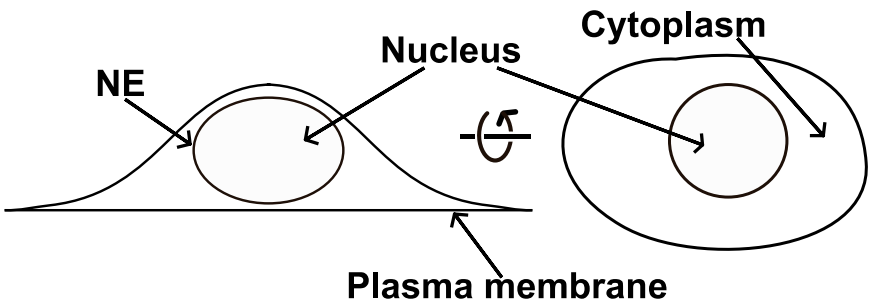

Figure 1.1. Simple illustration of a eukaryotic cell. Side view (left) and top view (right) of a eurkaryotic cell highlighting the nucleus surrounded by the NE. The cytoplasmic region makes up the rest of the cell which is bounded by the plasma membrane.

1.1 The Nuclear Envelope

The NE is composed of two concentric membranes, the inner and outer nuclear membrane (INM and ONM, respectively), separated by a ~40 nm wide space called the lumen or perinuclear space (PNS) (3). This structure presents a barrier to the transmission of information and matter between the cytoplasm and the nucleus, such as the export of RNA into the cytoplasm. To overcome this physical barrier, the NE of eukaryotic cells contain channel-like structures called nuclear pore complexes (NPCs) (Fig. 1.2) (3) whose primary role is to facilitate and regulate the transport of molecules and proteins between the nucleoplasm and cytoplasm (7). Located at regions where the INM and ONM are fused to form a contiguous membrane, the NPC is a large (~120 MDa) protein complex which consists of 34 unique proteins and is ~100 nm in diamater (8). There are a variety of mechanisms for how proteins may translate across the NE via the NPC. Small molecules less than 40 kDa in size have been shown to diffuse passively across the NE, requiring no external energy or chaperone molecules to pass through the NPC (7). Large proteins often require a localization signal which binds to transport proteins, such as importin-β, which facilitate passage of protein complexes up to 39 nm in diameter (9, 10). Although rare, there is evidence of some larger complexes bypassing the NPC altogether by forming vesicles on the nuclear membrane that then pass across the lumen, a strategy employed by the herpesvirus (11, 12).

Beyond acting as a physical barrier between the cytoplasm and nucleoplasm, the NE plays a significant role in many cellular functions. Some of these functions rely on

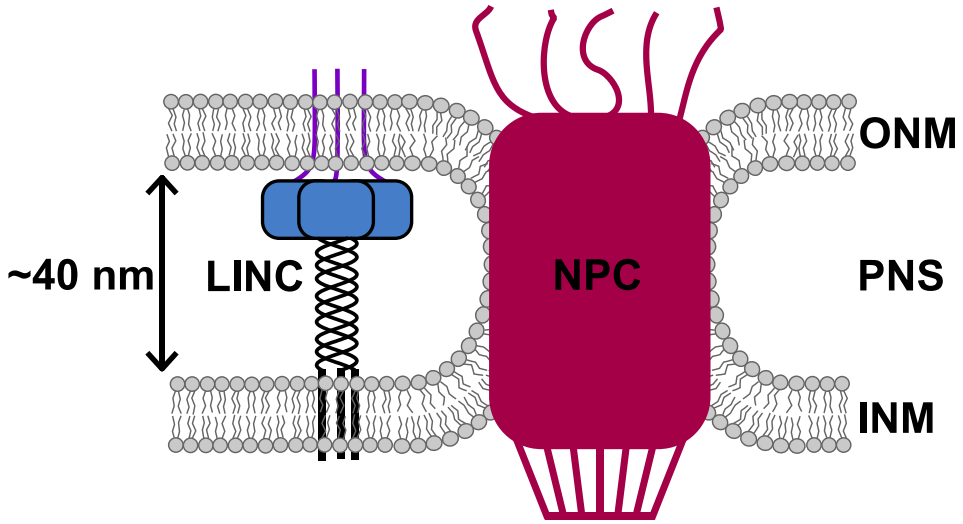

Figure 1.2. Illustration of the NE. The NE consists of the cytoplasm-facing ONM and the nucleoplasm-facing INM which are separated by the ~40 nm wide PNS or lumen. The NPC resides in fenestrations formed by the fusion of the INM and ONM. The Linker of Nucleoskeleton and Cytoskeleton (LINC) complex spans the entire NE, extending into both the cytoplasm and nucleoplasm.

the dynamic nature of the NE within the cell. During cell division, a complicated process of dissolving the NE takes place (13). Proteins found within the NE must be shuttled to the contiguous endoplasmic reticulum (ER) or may be retained through the formation of small NE vesicles (14, 15). The NE and nucleus as a whole can also dramatically change its position within the cell, contrasting with the standard image of a centrally located nucleus (16). This positioning can vary by cell type such as in epithelial cells where the nucleus moves between basal and apical positions depending on the cell cycle, the purpose of which is still unclear (17). It can also depend on the cell's activity such as the rearward nuclear positioning observed in migrating cells which has been connected to the efficiency of cell migration (18). Potential explanations for how nuclear positioning impacts the cell include its integration into the cytoskeleton (19) and its large role in cellular signaling, for example the ability of the nucleus to transmit mechanical forces across the cell as a result of its stiff nature (20).

The nucleus, and in turn the boundary-defining NE, has been shown via micropipette aspiration experiments to be significantly stiffer than the plasma membrane at the periphery of the cell (21). This stiffness is primarily due to the nuclear lamina network, a rigid network composed of interconnected filaments which are connected to the INM via proteins (20). The stiffness of the NE makes the nucleus a significant limiting factor to cell migration through confined spaces, with no migration observed through very small pores (<10% the nucleus cross section) despite protrusion of the cytoplasm through the pores (22). One possible advantage for high nuclear stiffness is the prevention of DNA damage due to compression, as increased apoptosis has been observed in cells after passing through small pores (23). Alterations to nuclear stiffness have a significant impact

on disease and cancer development (24, 25). Specifically, the interplay between reduced NE stiffness and increased cytoskeletal forces has been implicated in higher metastasis potential for cancer cells (26). The stiffness measured for many of these experiments relate to the overall stiffness of the nucleus, not local properties of the individual membranes which make up the NE.

We will see in Chapter 3 that the dynamics of the NE on small, local scales is of particular importance for our studies. The techniques we will use probe regions < 1 μm in diameter and may be affected by local perturbations (27). While the elastic properties of simple membranes, such as the plasma membrane of red blood cells, have been carefully examined (28), this only provides a baseline for understanding local motion of the NE. Importantly, the two membranes of the NE, which are only separated by a thin fluid layer (Fig. 1.2), are expected to be hydrodynamically coupled, potentially leading to coordinated motion of both membrane surfaces (29). This structure may also lead to fluctuations in the spacing between the two membranes, which have been proposed to exist (30). The situation in the NE is further complicated by the presence of protein complexes which directly link the INM and ONM, including the previously mentioned NPCs (Fig. 1.2). These lumen-spanning complexes may act as stents which increase the coupling of the membranes, however their contributions to membrane stiffness is not well understood (31). While perturbations to the NE with subsequent relaxation have been observed through imaging techniques (32) the limitations of the techniques used mean they observe inherently large scale and slow motional modes, occurring over the course of minutes. The majority of the data acquired for this thesis deals with time scales below one minute, thus we would not be significantly affected by these slow changes in the NE. Observations

presented in Chapter 3 uncover the existence of much smaller and faster undulations of the nuclear membranes and their effect on our experiments.

While physical properties of the NE such as its stiffness, permeability, and size play a significant role in the cell, it has become increasingly clear that important chemical and mechanical signaling events are coordinated by proteins within the NE (33, 34). These proteins have been found to be critical to a wide-variety of fundamental cellular processes, including cell cycle progression, DNA repair, gene expression, genome organization, lipid synthesis, and mechanotransduction (35–40). The vital role of the NE within the cell is further demonstrated by recent research showing an association between mutations in genes encoding NE proteins and several human diseases, including DYT1 dystonia, muscular dystrophy, cancer, and many other laminopathies (33, 41). While the behavior of these proteins has been studied extensively in the NE of living cells, quantitative studies of their assembly has so far been limited to *in vitro* investigations. The central goal of this thesis is to directly study the assembly of NE proteins in their native environment by utilizing fluorescence fluctuation spectroscopy (FFS), which has previously proven successful in quantifying protein oligomerization in other regions of living cells (42–44). One group of proteins that we will be using as a model system throughout this thesis, and has been associated with many of the diseases mentioned earlier, are the constituent proteins that form the Linker of Nucleoskeleton and Cytoskeleton (LINC) complex (Fig. 1.2) (45).

1.2 The LINC Complex

The LINC complex spans the lumen, extending into both the cytoplasm and nucleoplasm and coupling the cytoskeleton and nucleoskeleton (35). This is accomplished by interactions between the *Klarsicht*/*ANC-1*/*SYNE* homology (KASH) proteins found at the ONM and the *Sad1*/*UNC-84* (SUN) proteins localized to the INM (35). The N-terminus of KASH proteins reside in the cytoplasm and interact with components of the cytoskeleton while the KASH domain containing C-terminus includes a transmembrane domain and the luminal KASH peptide (35). SUN proteins consist of the SUN domain at the C-terminus which binds to the KASH peptide (35), while their N-terminus interacts with the lamina network, chromatin, and other proteins within the nucleoplasm (46).

By directly coupling the cytoskeleton and nucleoskeleton, the LINC complex can transmit forces between the cytoplasm and nucleoplasm, playing a significant role in vital cellular functions. The LINC complex has been shown to be required for the coupling of the nucleus to the actin network which is necessary for proper nuclear positioning during cell migration (18). Furthermore, SUN proteins have been shown to directly impact telomere positioning which has implications for cell division (47). This interaction with chromosomes has also implicated LINC complexes in altering force-dependent gene expression (48). The association of mutations in SUN and KASH proteins with human diseases further demonstrates the significance of the LINC complex. For example, neuromuscular diseases such as ataxia and muscular dystrophy have been connected to defects in KASH proteins which result in mislocalized nuclei (47).

Despite their importance in both cellular functions and human health, NE proteins such as SUN remain poorly characterized. This is largely due to a lack of tools available to quantify protein interactions within their native environment in living cells. Much of what is known about LINC complex proteins was found through *in vitro* techniques such as x-ray crystallography (49), immunoprecipitation (50), or pull-down assays (51) as well as *in silico* modeling techniques (52). These techniques, while providing important information, are limited by the difference between the environment which the proteins are in during the experiment and their native environment of the NE in living cells.

1.3 Overview of Thesis

The main aim of my thesis work is to address the lack of techniques available for probing proteins within the NE of living cells by examining the suitability of FFS. An overview of FFS and the instrumentation used throughout this thesis is given in Chapter 2. We also discuss previously developed analytical techniques that are used heavily in this thesis. These techniques are then applied to the NE for the first time in Chapter 3. A significant obstacle to this application is the presence of previously unreported nuclear membrane undulations. These undulations create fluctuations in the local volume of the NE, thus introducing additional fluctuations to the fluorescence signal. By accounting for this effect, we show that we can accurately characterize protein interactions in the NE, both in the lumen and at the nuclear membranes.

Having shown that we can characterize the oligomeric state of proteins in the NE, we apply this method to SUN proteins in Chapter 4 and study how these proteins

oligomerize to form the foundation of the LINC complex. We observe differences between the oligomerization in the cytoplasm and in the NE which further highlights the need for methods which can be used in the NE. Chapter 5 seeks to simplify these experiments by providing an improved analysis procedure which removes multiple detector artifacts that previously required correction. We extend our work to dual-color measurements in Chapter 6, allowing us to quantify heterointeractions. Here we observe interactions between the constituent proteins of the LINC complex, KASH and SUN, for the first time in the NE of living cells. Finally, Chapter 7 summarizes the results of the thesis along with proposing future research in the area.

Chapters 3, 4, and 5 are based on work previously published (53–55). Chapter 6 is based on work submitted for publication to Biophysical Journal.

2. Fluorescence Fluctuation Spectroscopy (FFS)

This chapter introduces the background behind the FFS techniques used throughout the thesis. An overview of the concept of FFS is given as well as the basic instrumentation used. Finally, we briefly discuss established analytical procedures which were previously developed and are used in this thesis.

2.1 Concept of FFS

FFS experiments utilize small (~1 fl or less) observation volumes in which excitation of fluorescent molecules occurs (Fig. 2.1A) (56). These volumes can be constructed either through a confocal microscope with single photon excitation or, as is the case for this thesis, two-photon excitation (discussed further in Section 2.2) (56). As fluorescent molecules, such as the fluorescent protein EGFP, pass through the observation volume they emit bursts of photons. This emitted fluorescence signal is stochastic in nature and statistical analysis techniques are needed to extract properties of the sample from the signal. While many analysis methods exist, they all exploit one or more of three key signatures provided by the fluorescence bursts. The frequency of these bursts gives information about the concentration of the sample (57). The duration of the bursts corresponds to the residence or diffusion time of the molecule which can be connected to its diffusion coefficient (57). Finally, the average amplitude of the bursts correspond to the average counts per second per molecule, or brightness (λ), of the sample (58). The methods used to analyze the fluorescent signal and obtain concentration, diffusion time, and brightness are discussed in Section 2.3.

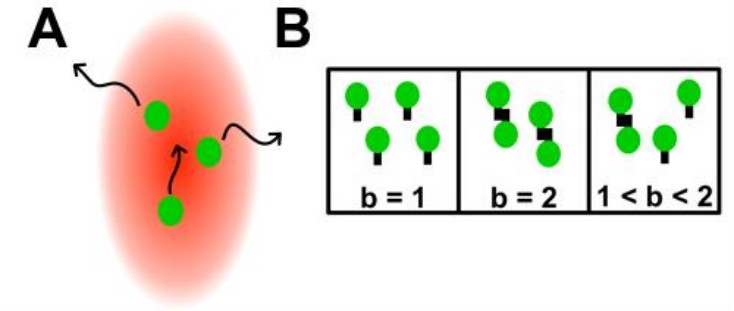

Figure 2.1. Concepts of FFS. A) Fluorescent molecules pass through a small two photon excitation volume. B) *b* reports the average stoichiometry of the sample. Figure adapted from Hennen et al. (55).

2.1.1 FFS brightness

The goal of this thesis is to characterize the dynamics of functional proteins in the NE of living cells. By tagging a protein of interest with a fluorescent label, brightness acts as a direct indicator of that protein's stoichiometry (58). In order to use this relation between brightness and stoichiometry we first determine the brightness of a monomeric calibration standard which we define as $\lambda_{monomer}$. If the fluorescent label is then tagged to a protein which remains monomeric we will measure a brightness of $\lambda = \lambda_{monomer}$. Attaching the fluorescent label to a protein which forms dimers we would expect the measured brightness to double, so that $\lambda = 2\lambda_{monomer}$. This is made explicit through the definition of the normalized brightness (59),

$$b = \frac{\lambda}{\lambda_{monomer}} \quad . \qquad 2.1$$

The analysis methods used to obtain λ, which are described in Section 2.3, measure the ensemble average stoichiometry of the sample. Therefore an entirely monomeric sample will have *b* = 1, while a dimeric sample will have *b* = 2, and a mixture of the two states will have 1 < *b* < 2 (Fig. 2.1B) (59).

2.2 Instrumentation

In order to achieve the small point spread function (PSF) required for FFS experiments we use two-photon excitation. This process requires near simultaneous absorption of two photons which together have enough energy to excite the fluorescent

label (60). Compared to other FFS methods, such as confocal microscopy, two-photon excitation provides two advantages critical for this thesis. Due to the high density of photons required to achieve two-photon excitation, the excitation volume is spatially limited to regions of high flux (60). Thus, unlike with a one-photon confocal setup, our excitation and observation volume are identical, resulting in no out of focus excitation and therefore no out of focus photobleaching. In addition, co-excitation of two different fluorescent labels by two-photon absorption, as discussed further in Chapter 6, is typically achievable at a single excitation wavelength for many common pairs of fluorophores, while one-photon excitation almost always requires two distinct wavelengths. Achieving perfect overlap of two laser beams at the focus of high numerical aperture (NA) objectives is a significant technical challenge (61). Any deviation between the two beams leads to imperfect coexcitation, introducing artifacts to the crosscorrelation signal which are difficult to predict and correct. These experimental challenges are avoided by two-photon excitation at a single wavelength.

The experimental setup used for the majority of the experiments in this thesis is shown in Fig. 2.2. Our experiments were performed on a two-photon microscope with a Zeiss 63x C-Apochromat water immersion objective with NA = 1.2 (42, 62). We used an excitation wavelength of 1000 nm and an average power after the objective of 0.3 - 0.4 mW. Emitted photons were detected by an avalanche photodiode (APD, SPCM-AQ-141, Perkin-Elmer, Dumberry, Quebec), recorded by a Flex04-12D card (correlator.com, Bridgewater, NJ) with a sampling frequency of 20 kHz, and analyzed with programs written in IDL 8.5 (Research Systems, Boulder, CO).

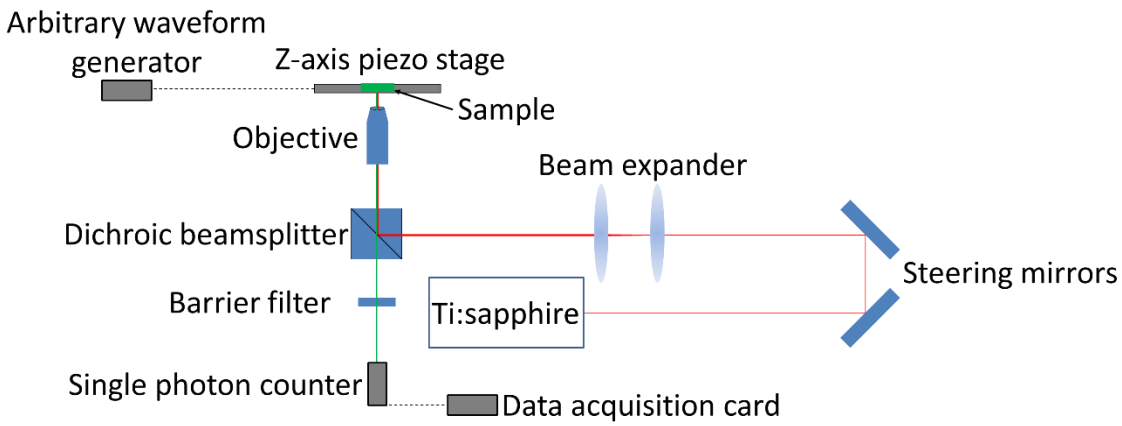

Figure 2.2. FFS instrument diagram. This diagram shows the setup used for single color measurements. Dual color measurements add an additional dichroic mirror and photon counter after the barrier filter (63). Figure taken from Hennen et al. (64).

In addition to the standard FFS measurements outlined in Section 2.1, this thesis also involves measurements where the PSF is scanned axially through a sample which we refer to as z-scans (43). All z-scans were performed using an arbitrary waveform generator (Model No. 33522A, Agilent Technologies, Santa Clara, CA) to move a PZ2000 piezo stage (ASI, Eugene, OR) axially. The driving signal from the arbitrary waveform generator was a linear ramp function with a peak-to-peak amplitude of 1.6 V and a period of 10 seconds. The peak-to-peak voltage corresponded to 24.1 μm of axial travel at a speed of $v_z = 4.82$ μm/s.

2.3 Analysis Procedures

A suite of analytical techniques is used to obtain information from FFS measurements. The next two sections describe established methods which are used throughout this thesis, while Section 2.3.3 describes a recently developed FFS method.

2.3.1 Mandel's Q parameter

Mandel's *Q* parameter (65) is an experimentally determined value commonly used in FFS and throughout this thesis. Defined as (65)

$$Q = \frac{\left\langle \Delta k(t)^2 \right\rangle}{\left\langle k(t) \right\rangle} - 1 \qquad 2.2$$

where *k* are photon counts and $\langle\ \rangle$ represents the average over *t*. *Q* directly relates to the brightness by (66)

$$Q = \frac{\gamma_2 \lambda \mathrm{B}_2\left(\tau_s, \tau_D\right)}{\tau_s} \qquad 2.3$$

where γ_2 is the shape factor which is discussed in more detail in Section 2.3.4, B_2 is the binning function, τ_s is the photon count sampling time, and τ_D is the diffusion time of the fluorescent labels. The binning function accounts for the possibility of a population of fluorescent proteins exiting our PSF during the sampling time which reduces the measured *Q*. In the case where $\tau_S << \tau_D$, B_2 reduces to τ_s^2 and Eq. 2.3 becomes (66)

$$Q = \gamma_2 \lambda \tau_s \ . \qquad 2.4$$

Eqs. 2.2 and 2.3 allow us to obtain λ, which can be combined with a monomeric calibration standard, $\lambda_{monomer}$, in Eq. 2.1 to obtain the normalized brightness or average stoichiometry of the system. The mean fluorescence intensity, $\langle F \rangle$, is connected to the concentration of protein by (66)

$$\langle F \rangle = cV\lambda_{monomer} \ , \qquad 2.5$$

where *c* is the concentration of monomeric subunits and *V* is the volume of the PSF. Plotting the normalized brightness vs. concentration results in a brightness titration curve (59) (Fig. 2.3).

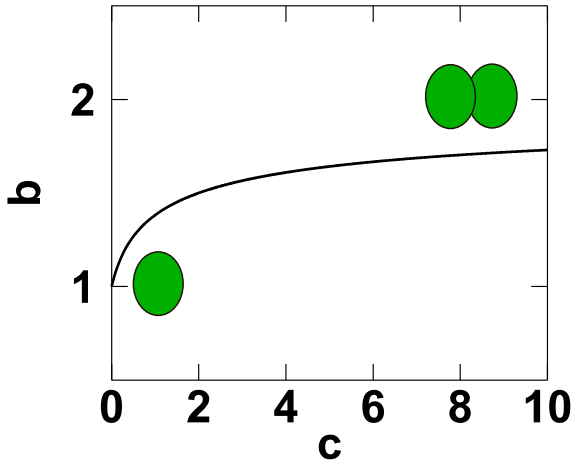

Figure 2.3. Brightness titration curve illustration. This shows the expected behavior for a plot of *b* vs. *c* for a monomer / dimer equilibrium. The proteins (green ovals) are monomeric (single green oval, bottom left), which corresponds to a brightness of *b* = 1, at low concentrations, but form dimers at high concentrations (pair of green ovals, top right) with *b* = 2. Brightness values between 1 and 2 represent mixtures of monomers and dimers. As the concentration increases *b* shifts from 1 to 2 as the equilibrium approaches a pure dimeric state.

A major drawback to the formulation of *Q* shown in Eq. 2.2 is that it assumes ideal photon detectors. It therefore requires corrections to account for artifacts such as deadtime and afterpulsing (67). These corrections can be unwieldy, particularly at highcount rates, and we therefore prefer to use what we refer to as the time-shifted *Q* parameter given by (68)

$$tsQ(\tau_s) = \frac{\langle \Delta k(t) \Delta k(t+\tau_s) \rangle}{\langle k(t) \rangle} \quad . \qquad 2.6$$

Eq. 2.6 calculates the correlation in the photon counts between neighboring photon count bins, with the -1 term from Eq. 2.2 dropped because there is no contribution from shot noise to the correlation term. The timescales of detector artifacts, such as deadtime and afterpulsing, are much shorter than a typical value of τ_s and therefore any spurious correlations will have decayed in that time (68). Thus, this formulation of *Q* effectively removes many artifacts due to non-ideal detectors. These advantages, and the *tsQ* analogue to Eq. 2.3, are discussed in more detail in Chapter 5.

To determine the normalized brightness (Eq. 2.1) from λ, control experiments are performed on a calibration standard such as EGFP in the cytoplasm to determine a calibration brightness (λ_{EGFP}) (44). Next, measurements of the protein of interest are performed to determine λ and thereby calculate normalized brightness, $b = \lambda / \lambda_{EGFP}$, to characterize the average oligomeric state of the EGFP-tagged protein (43). In addition, to obtain rigorous results within cells it has been shown that the data must be split into segments on the order of ~5 s and the results of these segments are averaged together

(27). This approach minimizes the effect of additional slow fluctuations in the recovered intensity.

2.3.2 Fluorescence correlation spectroscopy (FCS) analysis

Another common approach to analyzing FFS data uses the autocorrelation function, $G(\tau)$, (69) which measures the strength of correlation in the FFS signal as a function of the lag time, τ. This is experimentally calculated for τ > 0 using

$$G(\tau) = \frac{\langle \Delta k(t) \Delta k(t+\tau) \rangle}{\langle k \rangle^2}. \qquad 2.7$$

Unlike with *tsQ* (Eq. 2.6), τ is varied to construct an autocorrelation curve. A typical example of such a curve for a measurement of a fluorescent dye, TexasRed, is shown in Fig. 2.4A.

The autocorrelation function can be fitted to model functions describing diffusion for different PSF shapes. The PSF is often approximated as a 3D Gaussian beam which results in a correlation function of (70)

$$G(\tau) = G(0)\left(1+\frac{\tau}{\tau_D}\right)^{-1}\left(1+\frac{\tau}{s^2\tau_D}\right)^{-1/2}, \qquad 2.8$$

where $G(0)$ is the amplitude at τ = 0 and *s* is given by the ratio of the radial and axial beamwaists, $s = z_0/\omega_0$. Because the axial extent of our PSF is much larger than the radial extent ($s > 2$), we find that the approximation for a 2D Gaussian beam,

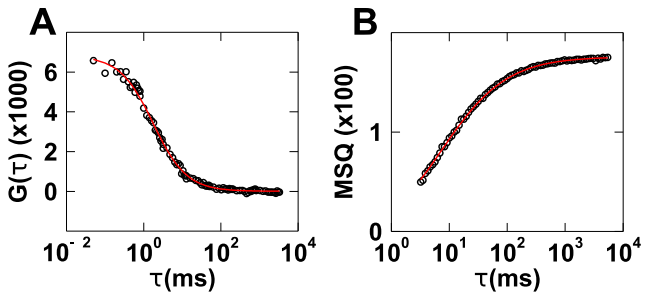

Figure 2.4. FFS analysis of a measurement of TexasRed. A) Autocorrelation (circles) fit to Eq. 2.9 (solid line). B) MSQ (circles) fit to Eq. 2.11 (solid line).

$$G(\tau) = \frac{G(0)}{1+\tau/\tau_D} \ , \qquad 2.9$$

agrees well with experimental data (Fig. 2.4A). In fact, this approximation is the exact form for measurements in very thin samples, such as those obtained in the ~40 nm thick NE. This is because all measurements performed in the NE for this thesis were done with the PSF located near the geometric center of the ventral or dorsal NE of U2OS cells. Due to the large lateral extent of the nucleus in these cells (typically > 10 μm in diameter) (71), the NE is effectively a flat sheet in these regions and can be treated as a 2D sample. More information on this measurement procedure can be found in Hennen et al. (64).

FCS analysis provides the diffusion time τ_D of the sample and is needed to either calculate the brightness using Eq. 2.3 or justify the use of the approximation described by Eq. 2.4. As with calculating *Q*, the photon count data are first segmented before $G(\tau)$ is calculated for measurements within cells.

2.3.3 Mean-segmented Q analysis

Mean-segmented Q (MSQ) analysis is a technique recently developed by Dr. Kwang-Ho Hur (72). Originally created to perform brightness measurements in the presence of non-stationary signals in E. coli, this approach calculates *Q* as a function of segment length, *T*. Given by (72)

$$\mathrm{MSQ}(T) = \frac{1}{n}\sum_{i=1}^{n} \frac{\left\langle \Delta k^2 \right\rangle_i}{\left\langle k \right\rangle_i} - 1 \ , \qquad 2.10$$

where *n* is the number of segments of length *T* in the data set and $\langle \ \rangle_i$ denotes the time-average of the *i*th segment, a typical MSQ curve is constructed by calculating the MSQ value for a range of *T* values, as seen for a simple fluorescent dye, TexasRed (Fig. 2.4B). As with the autocorrelation function, this curve can be fit to a model of simple diffusion through a 2D Gaussian PSF (72)

$$\mathrm{MSQ_D} = Q\left(1 - \frac{\mathrm{B}_2\left(T, \tau_D\right)}{T^2}\right) - \frac{\tau_S}{T} \ , \qquad 2.11$$

allowing us to obtain both *Q* and τ_D. Unlike the *Q* analysis described in Section 2.3.1 and FCS in section 2.3.2, MSQ directly accounts for the segmenting of data. This can remove biases found in the previously described methods, particularly in the presence of long timescale fluctuations as described in more detail in Chapter 3. Combining this with Eq. 2.4 allows us to obtain the brightness and thus the average stoichiometry.

2.3.4 Z-scan intensity profile

When focused on the ~40 nm thick NE, the ~1 μm thick PSF extends significantly into the cytoplasm and nucleoplasm (Fig. 2.5A). As a result, a significant concentration of fluorescent proteins outside the NE would contaminate our fluorescent signal with unwanted fluctuations. In order to ensure the majority of our fluorescent signal comes from the NE we must determine the concentration profile within the cell (73). This is done by performing axial z-scans with our PSF through the cell (Fig. 2.5B) (43). The resulting intensity profile is a convolution of the PSF with the cell's concentration profile (Fig. 2.5C).

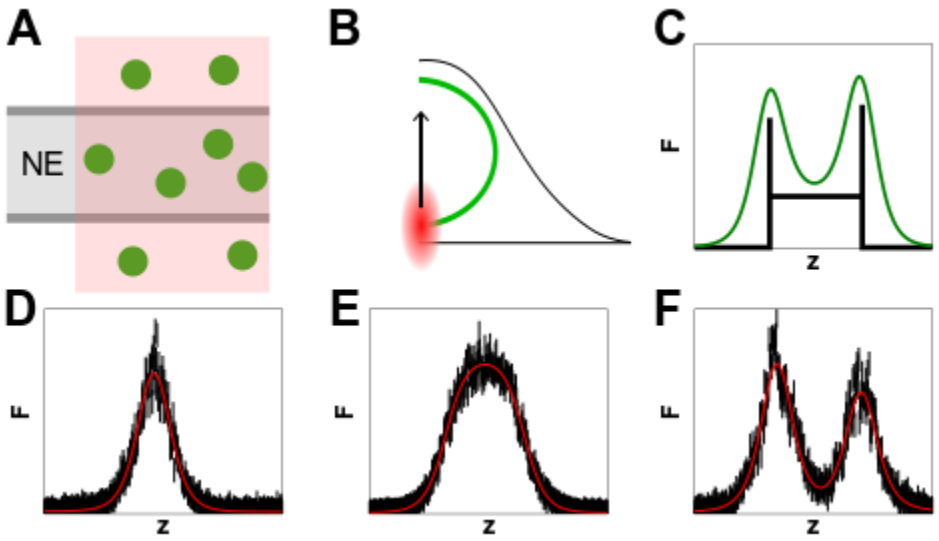

Figure 2.5. Overview of z-scan analysis. A) The PSF (red) extends beyond the NE (grey), exciting fluorescent proteins (green) in the cytoplasm and nucleoplasm. B) Z-scans are performed by scanning the PSF (red) axially through a cell expressing fluorescent proteins. C) The concentration profile (black lines) is convolved with the PSF and results in an intensity profile (green line). D) Intensity profile (black lines) of a z-scan through a thin sample fit to a δ layer (solid red line). E) Intensity profile (black lines) of a z-scan through a thick sample fit to a slab layer (solid red line). F) Intensity profile (black lines) of a z-scan through a cell expressing the NE-localized protein SS-EGFP fit to a δsδ model (solid red line).

To obtain the concentration profile we must perform a deconvolution using a model of our PSF (73). Performing a z-scan through a very thin, δ, layer results in the radially integrated PSF (RIPSF) (Fig. 2.5D) which can be compared to models (43). While PSFs are commonly modeled by either a 3D Gaussian or Gaussian-Lorentzian function, we have found a heuristic model of a modified Gaussian-Lorentzian (mGL) describes our PSF well (43). The RIPSF for the mGL model is described by

$$RIPSF\left(z\right)=\frac{\pi\omega_0^2}{4n}\left(1+\left(z/z_0\right)^2\right)^{1-n\left(1+y\right)} \qquad 2.12$$

where ω_0 is the radial beam waist, z_0 is the axial beam waist, *y* adjusts the shape of the decay, and *n* = 1 for single photon excitation while *n* = 2 for two photon excitation. The parameters z_0 and *y* are determined experimentally by performing z-scans through δ layers while ω_0 can be determined using a sample of known concentration (43).

A z-scan through a thick sample results in a broader, slab, geometry (Fig. 2.5E). The intensity profile of a NE localized protein consists of two thin NE layers with a thick nucleoplasm layer between them, which we refer to as the delta-slab-delta (δsδ) model (Fig. 2.5F) (73). By fitting our z-scans to this model we can determine the fraction of intensity we will receive from outside the NE while performing FFS measurements. Throughout this thesis we limit our measurements to non-NE fluorescence of <10% which allows us to treat our sample as a single thin delta layer (42).

In addition to obtaining a concentration profile, z-scans allow us to obtain the shape factor, γ_2, given by

$$\gamma_2 = \frac{V_{PSF^2}(s)}{V_{PSF}(s)} \qquad 2.13$$

where *V*(*s*) denotes the volume of the PSF embedded in the sample *s* and the subscripts *PSF²* and *PSF* denote the squared and linear PSF. It has been shown that this value can change by ~2x depending on the thickness of the sample which means proper determination of stoichiometry through FFS requires a fitting the RIPSF to a z-scan to determine the thickness (Fig. 2.5E) (43). By combining the techniques laid out here, previous work has been able to accurately characterize protein interactions in the nucleoplasm, cytoplasm, and plasma membrane of living cells (27, 42, 43). This thesis will utilize these methods and extend them to allow for quantification of proteins within the NE.

3. Quantitative Brightness Analysis of Protein Oligomerization in the Nuclear Envelope

As discussed in Chapter 1, the majority of NE proteins are poorly characterized. This chapter explores the potential of FFS for characterizing protein behavior in the NE. FFS techniques had been successfully applied to study protein assembly in the nucleoplasm, cytoplasm, and at the plasma membrane of living cells (42–44). This chapter demonstrates that the NE is a uniquely challenging environment for performing quantitative FFS. Specifically, we identified a slow fluctuation process unique to the NE which affects the fluorescence signal from luminal proteins. This phenomenon confounded standard analysis techniques, requiring the application of the newly developed mean-segmented Q (MSQ) approach. Here we lay the foundation for further studies in the NE and improvements on the technique later in this thesis. This work was done in conjunction with Dr. Kwang-Ho Hur who derived the MSQ theory and its software implementation. My contributions entailed designing and performing the experiments as well as performing data analysis.

3.1 Introduction

Here, we demonstrate that conventional analysis of FFS data taken in the NE leads to perplexing results, prompting the use of the recently described mean-segmented Q (MSQ) analysis method (72). MSQ identified the existence of an additional fluctuation process linked to the subcellular environment of the NE. After characterizing this

fluctuation process, we applied MSQ to investigate the oligomeric state of the luminal domains of the INM Sad1/UNC-84 (SUN) protein SUN2 and the ONM protein nesprin-2, the luminal domains of which interact within the PNS to form the core of the linker of cytoskeleton and nucleoskeleton (LINC) complex (74). This NE-spanning molecular bridge is responsible for the mechanical integration of the nucleus with the rest of the cell, which is required for several fundamental cellular functions including DNA damage repair, meiotic chromosome pairing, nuclear positioning, and the mechano-regulation of gene expression (16, 46, 48). Recently published in vitro studies show that the luminal domain of SUN2 homotrimerizes to form binding sites for the luminal Klarsicht/ANC-1/SYNE homology (KASH) peptide of nesprin-2 between adjacent SUN2 monomers (49). We tested this model by measuring SUN2 oligomerization in the NE and demonstrate that live-cell FFS provides a promising approach for studying NE proteins in their native environment.

3.2 Materials and Methods

3.2.1 Experimental setup

The instrumentation used in this chapter is the same as described in Section 2.2. A few select measurements were performed with a Zeiss 20x Plan-Apochromat immersion-less objective (NA = 0.8).

3.2.2 Sample preparation

All experiments were conducted using transiently transfected U2OS cells obtained from ATCC (Manassas, VA) and maintained in DMEM with 10% fetal bovine serum (Hycolone Laboratories, Logan, UT). Cells were subcultured into 24-well glass bottom plates (In Vitro Scientific, Sunnyvale, CA) prior to transfection. GenJet (SignaGen Laboratories, Rockville, MD) was used to transiently transfect cells 12-24 hours prior to measurement, according to the manufacturer's instructions. The growth medium was replaced with Dulbecco's phosphate-buffered saline containing calcium and magnesium (Biowhittaker, Walkerville, MD) immediately before measuring. Information regarding the construction of the cDNA constructs used in this work is described in the Supplemental Information.

3.2.3 Z-scan analysis

As described in Section 2.3.4, a modified squared Gaussian-Lorentzian (mGL) PSF was used to analyze z-scan FFS data (43). Z-scan calibration was performed as described (43) in order to determine the radial and axial beam-waist (ω_0 and z_0) as well as the axial decay parameter *y*, resulting in values of $\omega_0 = 0.45 \pm 0.05$ μm, $z_0 = 1.0 \pm 0.1$ μm, and $y = 2.4 \pm 0.3$. The z-scan intensity profile $F(z)$ was fit by the δsδ-model (42) to identify the intensity contributions from the ventral NE (NE_V) at height z_V, the dorsal NE (NE_D) at height z_D, and the nuclear layer in between, $F(z) = F_V(z, z_V) + F_N(z, z_V, z_D) + F_D(z, z_D)$. Because the NE is much thinner than the

axial size of the PSF, it is modeled by an infinitesimally thin layer, which we refer to as a δ-layer, while the nucleoplasm is modeled by a slab or s-layer (43). The intensity $F\left(z_V\right)$ represents the signal with the focus of the PSF on the NE_V, while $f_V = F_V\left(z_V, z_V\right)/F\left(z_V\right)$ describes the fraction of the fluorescence intensity from the NE_V at this location. The corresponding intensity fraction f_D from the NE_D is defined analogously.

3.2.4 FFS analysis

The analysis procedure in this chapter follows much of what is discussed in Section 2.3. The autocorrelation function (ACF) of the fluorescence intensity was calculated and fit to a model of two-dimensional diffusion,

$$G\left(\tau\right) = \frac{G\left(0\right)}{1+\tau/\tau_d} \tag{3.1}$$

with fluctuation amplitude $G\left(0\right)$, lag time τ, and diffusion time τ_d. Some ACF data was fit adding an exponential correlation term,

$$G\left(\tau\right) = \frac{A_0}{\left\langle F\right\rangle \Delta t}\exp\left(-\tau/T_0\right) \tag{3.2}$$

with characteristic decay time T_0 and amplitude factor A_0 to Eq. 3.1.

MSQ divides the FFS data into segments of duration *T* and determines the *Q* value averaged over all segments, MSQ(*T*). The model describing MSQ for a single diffusing species is (72)

$$\mathrm{MSQ}_{\mathrm{diff}}\left(T\right)=Q\left(1-\frac{B_2\left(T,\tau_d\right)}{T^2}\right)-\frac{\Delta t}{T}, \qquad 3.3$$

where $B_2\left(T,\tau_d\right)$ is the second-order binning function (66, 75). Fitting the MSQ data to this model determines Mandel's *Q* and the diffusion time τ_d. The MSQ of two diffusing species is given by

$$\mathrm{MSQ}_{diff,2species}\left(T\right)=f_1\,\mathrm{MSQ}_{\mathrm{diff}}\left(Q_1,\tau_{d1},T\right)+f_2\,\mathrm{MSQ}_{\mathrm{diff}}\left(Q_2,\tau_{d2},T\right), \qquad 3.4$$

where f_1 and f_2 are the fractional intensities with $f_2=1-f_1$. The amplitudes of each process, $Q_1 f_1$ and $Q_2 f_2$, of Eq. 3.4 are related to the effective normalized brightness *b* of the MSQ curve by

$$\frac{Q_1 f_1}{Q_2 f_2}=\frac{b_1\left(b-b_2\right)}{b_2\left(b_1-b\right)}, \qquad 3.5$$

where b_i is the normalized brightness of the i^{th} component, and is derived utilizing the relations $b=b_1 f_1+b_2 f_2$ and $Q_1/Q_2=b_1/b_2$.

3.2.5 Experimental protocol

After selecting a fluorescent cell by epifluorescence microscopy, FFS data was acquired by first taking a z-scan through the nucleus followed by focusing the two-photon beam on the NE_V to collect intensity fluctuations for ~60 s. Next, the beam was focused on the NE_D to acquire intensity fluctuations followed by an additional z-scan. Because FFS

analysis assumes a stationary process (76), any fluctuation data with sudden intensity jumps or slow drifts in the intensity were rejected. Comparison of the initial and final z-scan intensity profile served to identify mechanical drift of the stage or cell motion that occurred during the FFS measurement process. Measurements where the initial and final z-scan intensity profile differed were discarded. The z-scan FFS data were analyzed as described above.

3.3 Results and Discussion

3.3.1 Z-scan FFS and conventional analysis of EGFP within the NE

EGFP has been successfully used as a brightness marker of protein oligomerization in the cytoplasm, nucleoplasm, and at the plasma membrane of living cells (42–44). Since the NE was a new environment for brightness studies, we needed to establish the suitability of EGFP as a quantitative brightness marker in this subcellular compartment. We targeted EGFP to the contiguous lumen of the ER and the NE by fusing the signal sequence (SS) of the luminal protein torsinA to the N-terminus of EGFP (77). Expression of SS-EGFP in U2OS cells resulted in its efficient localization to the NE as corroborated by z-scan intensity profiles taken through the nucleus (Fig. 3.1B). A typical z-scan intensity profile from a SS-EGFP-expressing cell displayed two prominent peaks corresponding to fluorescence emanating from the NE_V and NE_D (Fig. 3.1C). $F(z)$ was fit to a three-layer δsδ model consisting of a thick nucleoplasmic layer separating two thin NE layers. The fit demonstrated that the NE layers were the main sources of fluorescence in this SS-EGFP-expressing cell (Fig. 3.1C). The intensity fraction originating at the NE_V

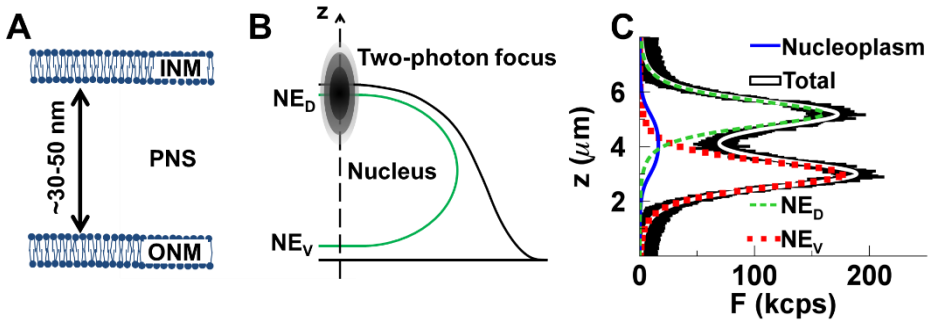

Figure 3.1. Z-scan through the NE. (A) Schematic of the NE with the perinuclear space (PNS) surrounded by an inner and outer nuclear membrane (INM/ONM). (B) Illustration of a z-scan through a cell expressing a fluorescent protein in the NE (green line). (C) Z-scan intensity profile (black line) for SS-EGFP with a fit to a δsδ-model showing the intensity contributions from the NE_V and NE_D as well as the nucleoplasmic contribution.

and NE_D for this intensity profile were 94% and 93%, respectively. For the remainder of this work, only cells with NE intensity fractions > 90% were used for FFS measurements. This criterion significantly simplifies data analysis, as it ensures that non-NE fluorescence contributions are sufficiently small to be ignored (42).

FFS measurements with the two-photon focus at either the NE_V or NE_D of SS-EGFP-expressing cells were used to determine the normalized brightness *b* of SS-EGFP in the NE using conventional brightness analysis (43, 68). We observed an unexpected increase of *b* for SS-EGFP from approximately one to two with number concentration *N* (Fig. 3.2A). Repeating the experiment with a tandem-dimeric $EGFP_2$ (44) fused behind the torsinA SS (SS-$EGFP_2$) resulted in a similar increase of *b* with *N* from two to values exceeding three (Fig. 3.2A). In contrast, brightness measurements of EGFP and $EGFP_2$ in the cytoplasm resulted in stable *b* values of one and two, as expected for a monomeric and dimeric protein (Fig. 3.2B). The *N*-dependent increase in *b* was also observed for SS-tagged mTurquoise (SS-mTurquoise) and enhanced yellow fluorescent protein (SS-EYFP) in the NE, indicating that this behavior was not specific to EGFP (Fig. 3.2C). Moreover, we observed that the decay shape of the ACF for SS-EGFP exhibited a pronounced broadening with *N* (Figs. 3.2D-E). In particular, the single-species diffusion model (Eq. 3.1) agreed well with data measured at low *N* values but no longer accurately described data measured as *N* increased.

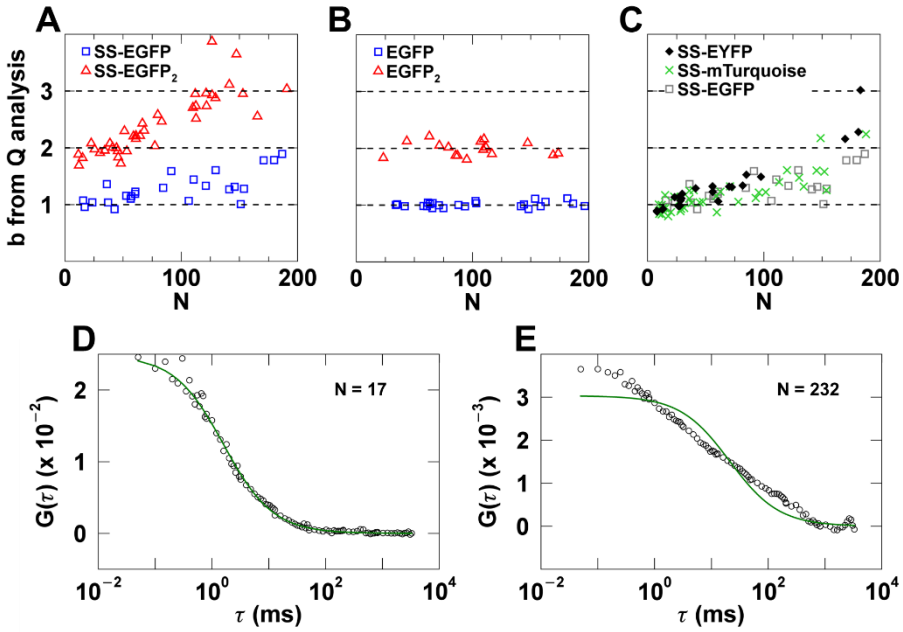

Figure 3.2. Conventional analysis of FFS data generated by fluorescent proteins in the NE or cytoplasm. (A-C) Brightness *b* from *Q* analysis versus number concentration *N*. (A) SS-EGFP and SS-EGFP$_2$ in the NE. (B) EGFP and EGFP$_2$ in the cytoplasm with mean values of b = 1.01 ± 0.05 and b = 1.99 ± 0.1, respectively. (C) SS-EYFP, SS-mTurquoise, and SS-EGFP in the NE. (D-E) ACF of SS-EGFP in the NE at different *N* values. ACFs were fit to a single-species diffusion model (green line).

3.3.2 MSQ analysis of FFS experiments performed in the NE

Since established FFS analysis methods yielded perplexing results, we turned to MSQ, which has proved useful to characterize fluctuations caused by diffusion in the presence of additional slow variations of the fluorescence signal (72). MSQ first divides the FFS intensity trace into segments of period *T* (Fig. 3.3A, top), then calculates the local Q value of each segment, which are averaged to determine MSQ(*T*). This process is repeated for a range of segment times *T* (Fig. 3.3A, bottom). Measurement of cytoplasmic EGFP resulted in an MSQ curve that initially increased with *T* and then plateaued (Fig. 3.3B). The value of MSQ(*T*) reflects the average *Q* taken over *T*, which only accounts for fluctuations with a correlation time less than *T* (72). Thus, the initial increase in the MSQ value with *T* reflects the inclusion of longer time scale dynamics. Once *T* is large enough that fluctuations with the longest correlation time are sampled, the MSQ value plateaus. Fitting the MSQ curve to the single-species diffusion model (Eq. 3.3) resulted in a diffusion time τ_d = 0.57 ± 0.22 ms and *Q* = 0.016 ± 0.001, which corresponded to *b* = 1.12 (Fig. 3.3B). These values were consistent with the results obtained from the ACF (τ_d = 0.68 ± 0.04 ms) and conventional brightness analysis (*b* = 1.05).

Note that our standard algorithms for calculating the ACF and brightness are also performed on segmented data using a predetermined *T* because this approach improves the robustness of FFS analysis of cellular data (27). We typically use a segment time of 6.5 s, which is within the plateau region of the MSQ curve generated for EGFP in the cytoplasm (Fig. 3.3B), and therefore accounts for all correlated fluctuations. Under these

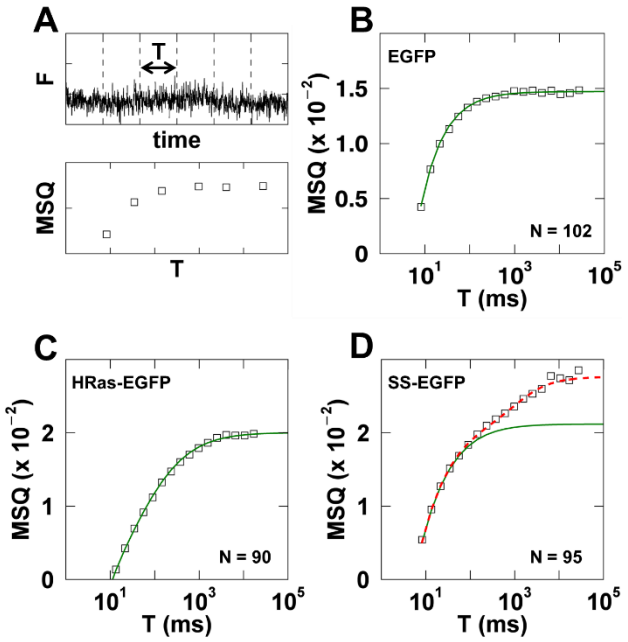

Figure 3.3. MSQ analysis of FFS data taken in the cytoplasm, at the plasma membrane, and in the NE. (A) Illustration of MSQ calculation. Top: Segmentation of FFS data into *T*. Bottom: Plot of MSQ data vs. *T*. (B-C) MSQ (squares) versus segment time with fit to a single diffusion model (solid green line). (B) EGFP in the cytoplasm with fitted $Q = 0.016 \pm 0.001$ (corresponding to $b = 1.05$) and $\tau_d = 0.57 \pm 0.22$ ms. (C) HRas-EGFP at the plasma membrane with fitted $Q = 0.021 \pm 0.003$ (corresponding to $b = 0.99$) and $\tau_d = 16 \pm 1$ ms. (D) MSQ (squares) versus segment time from SS-EGFP in the NE. Dashed red line represents a fit to a diffusion model plus exponential correlation process with fitted $Q = 0.026 \pm 0.004$ (corresponding to $b = 1.13$) and $\tau_d = 1.7 \pm 0.3$ ms. The diffusion component of the fit is shown as a solid green line.

conditions, we expect that MSQ and traditional FFS analysis would lead to similar results, as experimentally verified above.

To determine if MSQ analysis can be used for b measurements in thin cell layers, we next measured the EGFP-tagged peripheral membrane protein HRas (HRas-EGFP) at the plasma membrane (42). The MSQ curve for HRas-EGFP also featured a plateau region, although at larger *T* than observed for EGFP in the cytoplasm (Fig. 3.3B-C). This result reflects the lower mobility of HRas-EGFP, which leads to a longer persistence of correlated fluctuations due to diffusion. A fit to the single-species diffusion model (Eq. 3.3) was sufficient to describe the MSQ of HRas-EGFP (Fig. 3.3C). After accounting for the gamma factor of a thin layer (43), the fitted *Q* value corresponded to *b* = 0.99. This value indicated that HRas-EGFP was monomeric, consistent with our previously reported result based on conventional brightness analysis (42). This agreement was expected, as the segment time of 6.5 s is within the plateau region of the MSQ curve (Fig. 3.3C).

The results described above demonstrate that MSQ reliably identifies the brightness of proteins within thick and thin sample layers. However, the MSQ of SS-EGFP within the NE produced a curve without a clearly identifiable plateau region and could not be described by the single-species diffusion model (solid green line, Fig. 3.3D). The absence of a plateau region indicates that conventional FFS analysis is no longer applicable as some fluctuations persist longer than the segment time. We found empirically that adding an additional process with an MSQ term of

$$\mathrm{MSQ}_{\exp}\left(T,T_0\right)=A_0\left(1-\frac{\mathrm{B}_{2,\exp}\left(T/T_0\right)}{\left(T/T_0\right)^2}\right) \qquad 3.6$$

resulted in good agreement with the MSQ curve obtained for SS-EGFP in the NE (dashed red line, Fig. 3.3D). Eq. 3.6 represents the MSQ of an exponential correlation process (Eq. 3.2) with amplitude A_0, decay time T_0 and binning function

$$\mathrm{B}_{2,\exp}(x) = 2\left(x + \mathrm{e}^{-x} - 1\right). \qquad 3.7$$

Eqs. 3.6 and 3.7 were derived following the procedure described by Hur and Mueller (72).

To test whether the combination of a diffusion and exponential correlation process described the intensity fluctuations of SS-EGFP in the NE, we applied this model to all SS-EGFP data by fitting each experimental MSQ curve to the sum of Eqs. 3.3 and 3.4. The fitted *Q* and τ_d values from the diffusion process stayed approximately constant at low (*N* = 16, *Q* = 0.022 ± 0.001, τ_d = 1.2 ± 0.3 ms) and high (*N* = 91, *Q* = 0.023 ± 0.001 and τ_d = 1.4 ± 0.4 ms) *N* values (Figs. 3.4A and B). Converting all fitted *Q* values to *b* revealed that SS-EGFP is monomeric (*b* = 1.06 ± 0.14) at all measured *N* (Fig. 3.4C). Similarly, we found that the brightness of SS-EGFP$_2$ in the NE was *N*-independent with a mean value of *b* = 1.97 ± 0.19 consistent with a dimeric protein.

However, a strong difference in A_0 was observed when the MSQ data of SS-EGFP in the NE with low and high *N* values were compared (Fig. 3.4A-B). The fitted A_0 of SS-EGFP and SS-EGFP$_2$ increased approximately proportionally with intensity as demonstrated by a linear regression through the origin with a slope of 1.0 x 10^{-4} kHz^{-1} (green line, Fig. 3.4D). The diffusion time τ_d of all fitted values for SS-EGFP in the NE

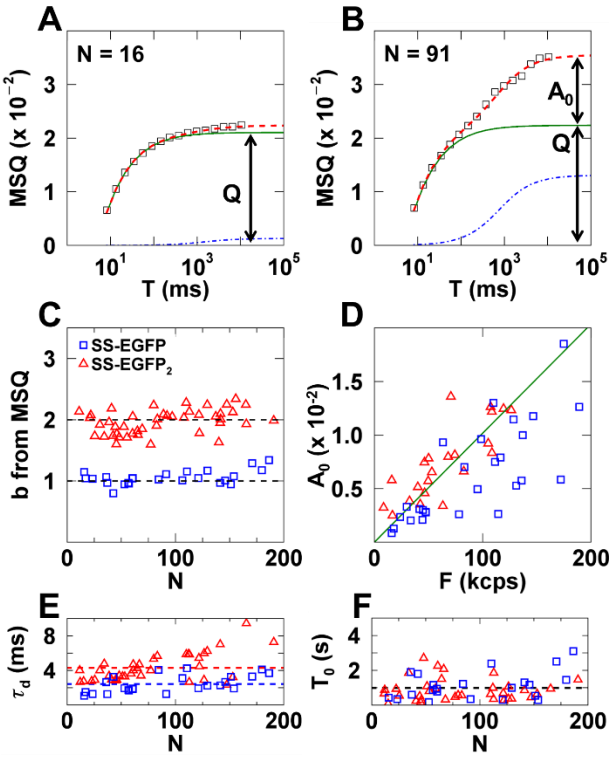

Figure 3.4. MSQ analysis of FFS data from SS-EGFP and SS-EGFP$_2$ in the NE. (A-B) MSQ (squares) versus segment time for SS-EGFP at different *N* values. Dashed red line represents fit to a single diffusion plus exponential correlation model. The diffusion component is shown as a solid green line, while the exponential correlation component is given by the dot-dashed blue line. (C) Plot of *b* vs. *N* for SS-EGFP and SS-EGFP$_2$ calculated from MSQ analysis. (D) Plot of A_0 vs. mean fluorescence intensity with a linear fit (solid green line). (E) Plot of τ_D vs. *N*. Dashed lines indicate the mean value. (F) Plot of T_0 vs. *N*. The mean value is marked by the dashed line. (C-F) Data from SS-EGFP and SS-EGFP$_2$ are plotted as blue squares and red triangles, respectively.

was *N*-independent with a mean and standard deviation of 2.4 ± 1 ms (dashed blue line, Fig. 3.4E). Similarly, the τ_d measured for SS-EGFP$_2$ was also independent of *N* with a mean and standard deviation of 4.3 ± 2 ms (dashed red line, Fig. 3.4E). Finally, the decay time T_0 of the exponential correlation process appeared to be *N*-independent with a mean of 1.7 ± 1.3 s (Fig. 3.4F).

The MSQ analysis of FFS data taken for SS-EGFP and SS-EGFP$_2$ in the NE supported a model where intensity fluctuations are caused by a combination of diffusion and a process with an exponential correlation term. Because we suspected that the unusual shape of the ACF (Fig. 3.2E) was a direct consequence of this additional process, we tested for agreement between MSQ and ACF analysis. Fitting the MSQ curve identified four parameters, Q, τ_d, A_0, and T_0. These parameters served to calculate the predicted ACF curve using the equation

$$G(\tau) = \frac{Q}{\langle F \rangle \Delta t} \frac{1}{1+\tau/\tau_d} + \frac{A_0}{\langle F \rangle \Delta t} \exp\left(-\frac{\tau}{T_0}\right), \tag{3.8}$$

where $Q/(\langle F \rangle \Delta t)$ corresponds to the fluctuation amplitude $G(0)$ of the diffusion process from Eq. 3.1, and the second term represents the exponential correlation. However, this approach failed to reproduce the experimental ACF for FFS data collected for SS-EGFP in the NE (Fig. 3.5). While the MSQ fit was in good agreement with the MSQ data (Fig. 3.5A), the calculated ACF curve deviated significantly from the experimental ACF (dashed red line, Fig. 3.5B). In particular, our model (Eq. 3.8) overestimated the amplitude of the exponential correlation process in the ACF.

We determined that this discrepancy stemmed from the finite segment time (T = 6.5 s) used in calculating the ACF, which was too short to sample all correlated fluctuations of the exponential correlation process. Therefore, we modified the previous equation using the theoretical approach described by Hur and Mueller (72) to include the effect of the finite segment time,

$$G(\tau) = \frac{Q}{\langle F \rangle \Delta t} \frac{1}{1+\frac{\tau}{\tau_d}} + \frac{A_0}{\langle F \rangle \Delta t} \left(e^{-\frac{\tau}{T_0}} - \frac{\mathrm{B}_{2,\exp}(T/T_0) + \mathrm{B}_{2,\exp}(|T-2\tau|/T_0) - 2\,\mathrm{B}_{2,\exp}(\tau/T_0)}{2\left(\frac{T-\tau}{T_0}\right)^2} \right). \quad (3.9)$$

The second term reflects the estimator bias for an exponential correlation e^{-t/T_0}, which leads to binning functions (75) as described previously (72). The application of Eq. 3.9 rather than Eq. 3.8 to predict the ACF led to a good agreement between data and model (solid red line, Fig. 3.5B).

3.3.3 Origin of the exponential correlation process

To test whether the exponential correlation process is caused by intensity fluctuations of individual SS-EGFP proteins or by a collective process, we measured the ACF for SS-EGFP in the NE using first a water immersion objective with high NA and then an immersion-less objective with lower NA (Fig. 3.6). The ACFs were fit to Eq. 3.9 (dashed

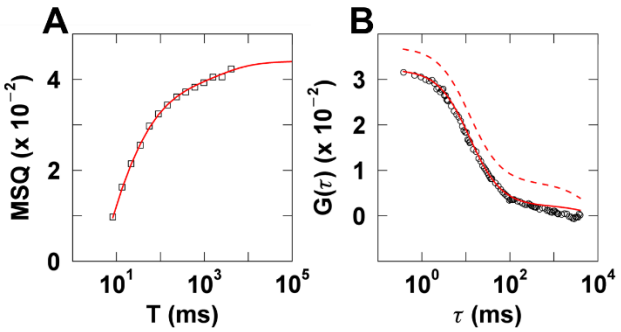

Figure 3.5. Comparison between MSQ and ACF for SS-EGFP in the NE. (A) MSQ (squares) vs. *T* for SS-EGFP with a fit to the diffusion plus exponential correlation model (red line). (B) Experimental ACF (circles) calculated with a segment time *T* = 6.5 s. Predicted ACF curves based on fitted MSQ parameters using Eq. 3.8 (dashed red line) and Eq. 3.9 (solid red line).

red line, Fig. 3.6) with the diffusion term (solid green line) and the exponential correlation term (dot-dashed blue line) shown for comparison. These fits provide insight into the effect of NA on two sets of parameters, the temporal parameters τ_d and T_0 and the amplitude parameters *Q* and A_0. The diffusion time τ_d for SS-EGFP in the NE increased from 1.5 ms for the high to 4.6 ms for the low NA objective, reflecting the larger radial beam waist of the latter. On the other hand, T_0 was reduced by a factor of ~2 (T_{0water} = 1.6 s, T_{0air} = 0.7 s). This ~7 fold difference between the change seen in τ_d and T_0 implies the new process is not diffusive.

The fitted *Q* value changed by a factor of ~30 following the switch from the high (Q_{water} = 0.02) to the low (Q_{air} = 0.0007) NA objective. Since *Q* represents a single molecule property, which is independent of concentration or volume, the decrease reflects the change in excitation intensity and collection efficiency between both objectives. If A_0 also represents a molecular property, then we would expect to observe a corresponding 30-fold reduction in amplitude upon switching objectives. However, A_0 only changed by a factor of 2 (A_{0water} = 0.0025, A_{0air} = 0.0012), which indicates that the new process is not related to intensity fluctuations of individual molecules. Collective phenomena, on the other hand, are expected to lead to intensity fluctuations that scale with the observed volume, which provides a potential explanation why the amplitude A_0 reduced much less than 30-fold.

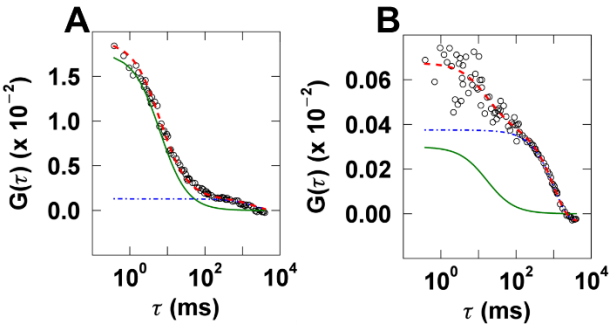

Figure 3.6. ACF of SS-EGFP in the NE taken with a high and a low NA objective. ACF taken with a high NA water immersion objective (A) and a low NA immersion-less objective (B). Both experimental ACF curves (circles) were fit to Eq. 3.9 (dashed red line). The fitted diffusion (solid green line) and exponential (dot-dashed blue line) component are shown for comparison.

Thus, we investigated whether the exponential correlation process is consistent with the presence of local volume fluctuations of the NE. Fluctuations in the relative position of the INM and ONM would change the distance *h* separating these membranes, altering the luminal volume enclosed between both membranes (Fig. 3.7A). Fluctuations in *h*, volume, and fluorescence intensity all would have the same relative standard deviation,

$$\sqrt{\langle \Delta h^2 \rangle} \Big/ \langle h \rangle = \sqrt{\langle \Delta V^2 \rangle} \Big/ \langle V \rangle = \sqrt{\langle \Delta F^2 \rangle} \Big/ \langle F \rangle = c \ . \qquad (3.10)$$

We further anticipate that the fluctuations in *h* are a property intrinsic to the NE and therefore independent of the expressed protein concentration, which implies that the variance $\langle \Delta h^2 \rangle$ is approximately constant. As a consequence, the relative standard deviation $\sqrt{\langle \Delta h^2 \rangle} \Big/ \langle h \rangle$ of Eq. 3.10 has to be equal to a constant value *c*. Furthermore, the A_0 value of the volume fluctuations is determined by $A_0 = \langle \Delta F^2 \rangle \Big/ \langle F \rangle \Delta t$, which can be rewritten using Eq. 3.10,

$$A_0 = c^2 \Delta t \langle F \rangle . \qquad (3.11)$$

This equation predicts that A_0 is directly proportional to the fluorescence intensity $\langle F \rangle$ as was experimentally observed (Fig. 3.4D). Eq. 3.11 was applied to determine that *c* = 0.06 from the fitted slope of Fig. 3.4D and a sampling time Δt = 50 µs. Since *c* represents the relative standard deviation of $\sqrt{\langle \Delta h^2 \rangle} \Big/ \langle h \rangle$, knowing the average distance $\langle h \rangle$

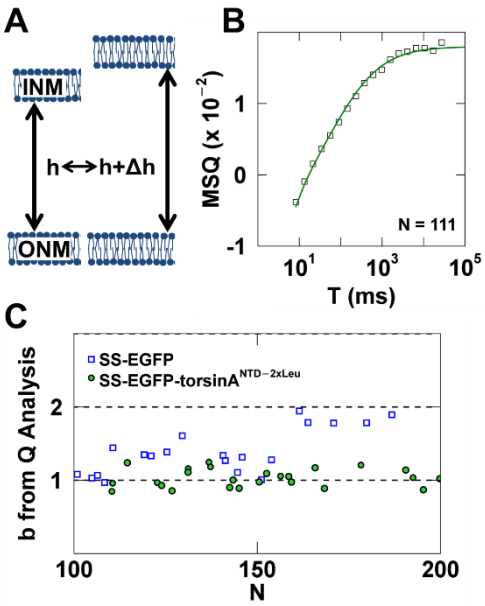

Figure 3.7. Distance fluctuations between the INM and ONM. (A) Illustration of fluctuations in the distance *h* separating the two nuclear membranes. (B) MSQ curve (squares) for SS-EGFP-torsinA$^{NTD-2xLeu}$ in the NE fit to a single-species diffusion model (green line). (C) Plot of *b* from conventional brightness analysis vs. *N* for SS-EGFP and SS-EGFP-torsinA$^{NTD-2xLeu}$ in the NE.

specifies the fluctuation amplitude $\sqrt{\langle \Delta h^2 \rangle}$. Assuming an average distance of 40 nm between the INM and the ONM, we predicted that subtle variations (~2 nm) in the axial distance between these membranes would be sufficient to give rise to the observed exponential correlation process. The equilibrium out-of-plane motion of a biomembrane in a viscous environment is described by an exponential decay of each spatial mode with its own characteristic time (78). Thus, the observation of an exponential correlation process indicates that the range of modes that are observed in our studies are very restricted. This is plausible since NPCs embedded in the NE at high density provide local restraints for the motion of the membranes.

We further predicted that FFS measurements of nuclear membrane-associated proteins would not be sensitive to these volume fluctuations, as axial motion by a few nanometers within the PSF is experimentally undetectable. To test this prediction, SS-EGFP was fused with a previously characterized transmembrane domain generated by two leucine substitutions within the hydrophobic membrane-associating N-terminal domain (NTD) of torsinA (SS-EGFP-torsinA$^{NTD\text{-}2xLeu}$) (79). The MSQ of FFS data collected for SS-EGFP-torsinA$^{NTD\text{-}2xLeu}$ in the NE with N = 111 resulted in a curve that plateaued at large T. The data were described by the single-species diffusion model with Q = 0.018 and τ_d = 26 ms, which is consistent with the absence of the exponential correlation process observed for SS-EGFP in the NE (Fig. 3.7B). Repeated measurements in the NE of SS-EGFP-torsinA$^{NTD\text{-}2xLeu}$ expressing cells resulted in a mean value of τ_D = 40 ± 20 ms, which was significantly slower than the value obtained for SS-EGFP in the same sub-cellular compartment (τ_d =2.4 ± 1.0 ms), as expected for a membrane-bound protein.

Because the plateau region for the MSQ curve obtained for SS-EGFP-torsinA$^{NTD-2xLeu}$ in the NE includes T = 6.5 s, we were able to perform conventional brightness analysis. The b of SS-EGFP-torsinA$^{NTD-2xLeu}$ in the NE was monomeric (b = 1.03 ± 0.13) and N-independent (Fig. 3.7C), as opposed to the increase in b observed for SS-EGFP using the same conventional analysis (Fig. 3.2A). This difference further supports the proposed model of nanometer-sized fluctuations in the axial distance between the INM and ONM.

3.3.4 Quantifying the oligomerization of physiologically relevant NE proteins.

The results described above establish the suitability of EGFP as a quantitative brightness marker in the NE and the framework necessary for interpreting FFS data collected in this subcellular compartment. To test the applicability of this approach for the study of NE proteins, we investigated the oligomerization of nesprin-2 and SUN2. Because EGFP-tagged nesprin-2 and SUN2 constructs were previously shown to be immobile in the NE (80), we selected protein domains that diffuse within the PNS for our initial FFS experiments. Specifically, we measured the oligomerization of the SS-EGFP-tagged KASH peptide of nesprin-2 (SS-EGFP-KASH2) and luminal domain of SUN2 (SS-EGFP-SUN2$^{261-731}$) in the NE.

Since the MSQ curves of SS-EGFP-KASH2 measured in the NE lacked a plateau region, the data was fit to the diffusion and exponential correlation model described above (Fig. 3.8A). The fitted b values were N-independent with an average brightness of b = 1.05 ± 0.1 (Fig. 3.8B), indicating that SS-EGFP-KASH2 was monomeric in the NE.

The diffusion time for SS-EGFP-KASH2 (τ_d = 4.0 ± 1.4 ms) in the NE was similar to the diffusion time observed for SS-EGFP$_2$ (τ_d = 4.3 ± 1.6 ms) in the same subcellular compartment (Fig. 3.4E), suggesting that SS-EGFP-KASH2 diffuses through the PNS.

In contrast, FFS measurements taken in the NE of SS-EGFP-SUN2$^{261\text{-}731}$-expressing cells revealed features in the MSQ curve not shared by any of the other constructs previously measured in the NE. This is illustrated by the significant differences in the results of MSQ analysis for FFS data collected for SS-EGFP and SS-EGFP-SUN2$^{261\text{-}731}$ in the NE at *N* = 14 (Fig. 3.8C). We deliberately chose FFS data with a low *N* to ensure that the influence of volume fluctuations on MSQ data is negligible. As expected, the MSQ curve for SS-EGFP was well characterized by the single-species diffusion model, confirming the absence of detectable PNS volume fluctuations, with a plateau value corresponding to a monomeric brightness. The maximum value of the MSQ curve of SS-EGFP-SUN2$^{261\text{-}731}$ significantly exceeded the maximum for SS-EGFP. Because the influence of volume fluctuations is negligible at this low concentration, the higher MSQ value corresponds to an increased brightness, indicating the presence of SS-EGFP-SUN2$^{261\text{-}731}$ oligomerization at *N* = 14. A single diffusion species was unable to reproduce the shape of the MSQ curve for SS-EGFP-SUN2$^{261\text{-}731}$. Increasing the number of diffusing species to two (Eq. 3.4) was sufficient to describe the experimental curve (green line, Fig. 3.8C). The fit identified diffusion times of τ_{d1} = 8.5 ± 0.2 ms and τ_{d2} = 150 ± 6 ms. Comparison of these values with the diffusion time of the membrane-bound SS-EGFP-torsinA$^{\text{NTD-2xLeu}}$ (Fig. 3.7B) revealed that τ_{d1} was too fast to

represent a membrane-associated protein, while τ_{d2} was very slow and likely represents membrane-associated SS-EGFP-SUN2$^{261\text{-}731}$.

Fitting the experimental MSQ for SS-EGFP-SUN2$^{261\text{-}731}$ to Eq. 3.4 determined the total amplitude $Q = f_1 Q_1 + f_2 Q_2$, which was converted into a brightness *b*. This *b* increased in an *N*-dependent manner from approximately 1 to 3 (Fig. 3.8D), which is consistent with a model of trimerization of SS-EGFP-SUN2$^{261\text{-}731}$. The fit also identified the diffusion times of the two components. Plotting τ_d vs. *N* for SS-EGFP-SUN2$^{261\text{-}731}$ revealed that the fast and slow diffusion times were approximately *N*-independent (Fig. 3.8E). The larger scatter in the fast diffusion times is probably caused by the lower MSQ amplitude of this component compared to the slow process.

Protein trimerization involves a concentration-dependent mixture of monomers, dimers, and trimers. Since a two-species fit was sufficient to describe the experimental MSQ, the proportions of these oligomeric states cannot be determined. In fact, the two-species fit only determines the amplitudes $f_1 Q_1$ and $f_2 Q_2$, which are insufficient to determine the individual *b* values associated with Q_1 and Q_2. Interestingly, the ratio

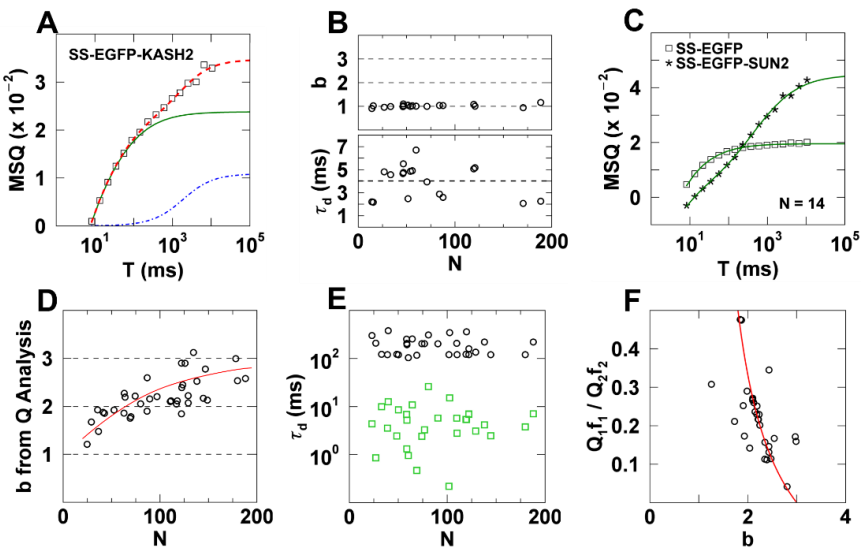

Figure 3.8. Measuring the oligomerization of SS-EGFP-KASH2 and SS-EGFP-SUN2$^{261\text{-}731}$ in the NE. (A) MSQ data for SS-EGFP-KASH2 fit to a diffusion plus exponential correlation model (dashed red line) with separated diffusion (dot-dashed blue line) and exponential correlation (solid green line) components. (B) Plot of *b* vs. *N* (top) and τ_d vs *N* (bottom) for SS-EGFP-KASH2. (C) MSQ data for SS-EGFP (squares) and SS-EGFP-SUN2$^{261\text{-}731}$ (stars) at *N* = 14 with SS-EGFP fit to a single-species diffusion model (green line) and SS-EGFP-SUN2$^{261\text{-}731}$ fit to a two-species diffusion model (green line). (D) Plot of *b* from conventional *Q* analysis vs. *N* for SS-EGFP-SUN2$^{261\text{-}731}$ with a red line added to guide the eye. (E) Plot of τ_d vs. *N* for SS-EGFP-SUN2$^{261\text{-}731}$ showing the fast (green squares) and slow (black circles) components with an average of 5.8 and 190 ms, respectively. (F) Plot of the ratio f_1Q_1/f_2Q_2 of the fast to slow diffusing species vs. *b* of SS-EGFP-SUN2$^{261\text{-}731}$. Red line represents relation for monomer-trimer model with Q_1 and Q_2 corresponding to the monomer and trimer.

f_1Q_1/f_2Q_2 , which indicates the relative contribution of the fast and slow population to the MSQ curve, decreased with increasing *b* (Fig. 3.8F). This suggests that the amplitude of the fast diffusing species decreases relative to the slow diffusing species as the average oligomerization of SS-EGFP-SUN2$^{261\text{-}731}$ increases.

Eq. 3.5 establishes a relation between f_1Q_1/f_2Q_2 and *b*. A simple model assuming a monomer / trimer transition (b_1 and b_2) with the monomer and trimer representing the fast and slow components, respectively, agreed with the data (Fig. 3.8F, red line). This implies that the trimeric form of SS-EGFP-SUN2$^{261\text{-}731}$ associates with the nuclear membrane and thus diffuses slowly. Since SUN2 oligomerization is required for KASH-binding (49), it is possible that this membrane association is mediated by interaction of SS-EGFP-SUN2$^{261\text{-}731}$ with the KASH peptides of endogenous nesprin proteins present at the ONM. Relatedly, the increase of the fast diffusing component as *b* decreases suggests that monomeric SS-EGFP-SUN2$^{261\text{-}731}$ diffuses through the PNS.

The MSQ analysis of SS-EGFP-SUN2$^{261\text{-}731}$ was performed without accounting for the exponential correlation process. This is justified, because at high *N*, where the volume fluctuations at the NE are most easy to observe, the brightness is close to three. Under these conditions SS-EGFP-SUN2$^{261\text{-}731}$ is associated with the membrane, and therefore the fluorescence of the sample is not affected by the distance fluctuations of the nuclear membranes.

3.4 Conclusions

This chapter describes the unique challenges for FFS and brightness analysis when quantifying protein oligomerization in the NE. We discovered the existence of a novel intensity fluctuation process in the NE, which renders conventional brightness analysis of FFS data in this subcellular compartment unreliable. Furthermore, we provided evidence to suggest that nanometer fluctuations in the separation of the INM and ONM are the cause of this process. The experimental data revealed that an exponential correlation term was sufficient to model the PNS volume fluctuations with an amplitude that was approximately proportional to fluorescence intensity. The incorporation of this exponential correlation term into MSQ analysis enabled the calculation of the oligomeric state and diffusion time for EGFP-tagged luminal proteins. We further recognized that the correlation time of the PNS volume fluctuations was sufficiently slow to lead to discrepancies between ACF and MSQ analysis, which originated from the fixed segment period used for calculating the ACF. To correct for these inconsistencies, we introduced a correlation term (Eq. 3.9) that accounts for the finite segment time, and thereby allows for future analysis of FFS data taken in the NE not only by MSQ but also by ACF analysis. Finally, we demonstrated the ability of MSQ analysis to identify the oligomeric state of EGFP-tagged NE protein constructs encoding the KASH peptide of nesprin-2 or the luminal domain of SUN2. The brightness of the luminal SUN2 construct increased from a monomer to approximately a trimer, which agrees with results obtained *in vitro* (49).

In summary, this work lays the foundation for future quantitative characterization of protein oligomerization in the NE by FFS and brightness analysis. Future efforts are

needed to fully characterize the nature of the distance fluctuations between the nuclear membranes and their influence on FFS experiments. While the presence of these fluctuations complicates the interpretation of FFS data collected in the NE, they also provide a means for distinguishing between NE proteins that are freely diffusing through the PNS and proteins that are associated with the nuclear membranes.

4. Fluorescence Fluctuation Spectroscopy Reveals Differential SUN Protein Oligomerization in Living Cells

Having established that the MSQ technique accurately reports stoichiometry in the NE, we apply it to study the oligomerization of SUN proteins. In this chapter, we show that we can determine binding curves of biologically relevant NE proteins in their native environment. Comparing results from measurements in the cytoplasm to those performed in the NE shows the importance of performing these measurements in the proper space. Dr. Cosmo A. Saunders created the DNA constructs as well as performed immunofluorescence imaging on fixed cells. My contributions entailed designing and performing all FFS experiments as well as analysis of the FFS data.

4.1 Introduction

Mammals encode two major SUN proteins, SUN1 and SUN2, which are widely expressed in somatic cells (81). Consistent with their high level of sequence similarity (i.e. mouse SUN1 and SUN2 share 65% identify), SUN1 and SUN2 perform redundant functions during the DNA damage response (37), radial neuronal migration in the developing mouse cerebral cortex and hippocampus (82), as well as synaptic nuclear anchorage in mouse skeletal muscle (83). These redundancies may be due to the ability of both SUN1 and SUN2 to interact promiscuously with the KASH peptide of several KASH proteins including nesprin-1, -2, and -3 (84). Nevertheless, examples of SUN protein function specificity also exist. For example, SUN1 is differentially required for meiotic

chromosome pairing (45, 85) and NPC insertion and distribution (86, 87). However, the molecular mechanisms underlying the redundant and specific functions of SUN1 and SUN2 remain unclear.

Recent *in vitro* studies reveal that SUN2 trimerizes due to the presence of a coiled-coil (CC) containing helical region within its luminal domain (49, 51, 88). SUN2 oligomerization is essential for KASH-binding, which is further stabilized by an intermolecular disulfide bond formed between conserved cysteine residues in the SUN domain and KASH peptide (49). Despite these important mechanistic advances, the *in vivo* relevance of SUN protein trimerization remains unclear due to the lack of suitable methods for measuring protein assembly states within the NE. Here, we sought to address this deficiency by extending the application of FFS (89) to quantify protein-protein interactions in the NE in living cells.

4.2 Materials and Methods

4.2.1 Antibodies

Anti-EGFP mouse monoclonal antibody MAB3580 was purchased from Sigma-Aldrich (Cambridge, MA) and was used at a dilution of 1:200 for immunofluorescence. Anti-SUN1 (ab74758) and- SUN2 (ab87036) were used at a 1:200 dilution for validating the shRNA-mediated depletion of SUN1 or SUN2 by immunofluorescence. Secondary antibodies were from two different sources. From Jackson ImmunoResearch Laboratories, Inc. (Bar Harbor, ME), we purchased goat anti-mouse secondary antibodies conjugated to Alexa Fluor 488 or rhodamine. From

ThermoFischer Scientific (Waltham, MA), we purchased goat anti-rabbit secondary antibodies conjugated to Dylight 488 or 561. All secondary antibodies were used at a 1:200 dilution for immunofluorescence.

4.2.2 Reagents

DAPI was purchased from ThermoFisher Scientific. Restriction enzymes were either purchased from New England Biolabs (NEB, Ipswich, MA) or Promega (Madison, WI). Phusion DNA polymerase, T4 DNA ligase, and T4 PolyNucleotide Kinase (PNK) were also purchased from NEB. All other chemicals were from Sigma-Aldrich (St. Louis, MI) unless otherwise specified. Wizard SV Gel and PCR Clean-Up System was from Promega. GeneJet Plasmid Midiprep Kit was from ThermoFisher Scientific (Waltham, MA).

4.2.3 Cell culture

U2OS cells obtained from the ATCC (Manassas, VA) were cultured using standard sterile technique in DMEM medium with 10% fetal bovine serum from Hyclone Laboratories (Logan, UT).

4.2.4 DNA constructs

NC (TR30015), SUN1 (TF300647B / FI302582), and SUN2 (TF300646A/FI302577) HuSH-29 shRNA constructs in pRFP-C-RS were purchased from

OriGene Technologies, Inc. (Rockville, MD). The SS-EGFP and SS-$EGFP_2$ constructs were generated as follows using a previously described human SS-EGFP-torsinA construct (77). EGFP from SS-EGFP-torsinA was PCR amplified using the primers SS-EGFP-F and SS-EGFP-R (Table 4.1), which contain 5' *Nhe*I and *Eco*RI cut sites, respectively. The PCR product was purified and digested alongside SS-EGFP-torsinA with *Nhe*I and *Eco*RI. Following gel purification, the digested PCR product and plasmid were ligated together to create SS-EGFP. To generate SS-$EGFP_2$, EGFP was amplified from SS-EGFP-torsinA using the primers SS-$EGFP_2$-F and SS-$EGFP_2$-R (Table 4.1), which contains 5' *Bsr*GI and *Apa*I cut sites, respectively. In addition, SS-$EGFP_2$-F encodes a 10 amino acid linker (GHGTGSTGSG) following the *Bsr*GI site, while SS-EGFP2-R encodes a mutated *Bsr*GI site that disrupts the 3' *Bsr*GI present in EGFP. The resulting PCR product was then purified and digested beside SS-EGFP with *Bsr*GI and *Apa*I. The digested PCR product and plasmid were purified and ligated to make SS-$EGFP_2$.

Previously described EGFP-tagged full-length mouse SUN1 and SUN2 constructs (18) were used as templates for the generation of the SS-EGFP-tagged luminal SUN1 and SUN2 constructs. To create SS-EGFP-$SUN2^{261\text{-}731}$, the sequence encoding amino acids 261-731 was PCR amplified from EGFP-SUN2 using the SS-EGFP-$SUN2^{261\text{-}731}$-F and SS-EGFP-$SUN2^{261\text{-}731}$-R primer pair (Table 4.1), which contain 5' *Bsr*GI and *Eco*RI cut sites, respectively. In addition, SS-EGFP-$SUN2^{261\text{-}731}$-F encodes a 10 amino acid linker (GHGTGSTGSG) following the *Bsr*GI site. The PCR product was purified and digested alongside SS-EGFP with *Bsr*GI and *Eco*RI. Following gel purification, the digested PCR product and plasmid were ligated together to create SS-EGFP-$SUN2^{261\text{-}731}$. SS-EGFP-$SUN2^{520\text{-}731}$ and SS-EGFP-$SUN2^{595\text{-}731}$ were both generated via Kinase, Ligase,

Primer Name	**DNA Sequence**	**5' RE Site**
SS-EGFP-F	GTGGCTAGCGTGAGCAAGGGCGAGGAG	*Nhe*I
SS-EGFP-R	GACTGACTGAATTCCTACTTGTACAGCTCGTCCATG	*Eco*RI
SS-EGFP$_2$-F	GCTGTACAAG**GGGCACGGGACCGGGTCTACAGGGAGCGGG**AGCGTGAGCAAGGGCGAGG	*Bsr*GI
SS-EGFP$_2$-R	**AACGGGCCCGGCTGCCAATCATGACTG**TTACTTATACAGCTCGTCCATGCCG	*Apa*I
SS-EGFP-SUN1$^{457-913}$-F	GCTGTACAAG**GGGCACGGGACCGGGTCTACAGGGAGCGGG**AGGGTGGACGATTCCAAGG	*Bsr*GI
SS-EGFP-SUN1$^{457-913}$-R	GAATTCCTACTGGATGGGCTCTCCG	*Eco*RI
SS-EGFP-SUN1$^{702-913}$-F	ACATCCGAGGCTATTGTGTC	-
SS-EGFP-SUN1$^{777-913}$-F	TGGTACTTCTCACAGTCACC	-
SS-EGFP-SUN2$^{261-731}$-F	GCTGTACAAG**GGGCACGGGACCGGGTCTACAGGGAGCGGG**TCCTGGTGGGCAGCAAAAG	*Bsr*GI
SS-EGFP-SUN2$^{261-731}$-R	TTTTGAATTCCTAGTGGGCAGGCTCTC	*Eco*RI
SS-EGFP-SUN2$^{520-731}$-F	TTGGTGAGCCGCCGC	-
SS-EGFP-SUN2$^{595-731}$-F	TGGTACCACTCCCAGTCAC	-
SS-EGFP-SUN$^{\Delta}$-R	CCCGCTCCCTGTAGACC	-
SS$^{\Delta}$-F	GTGGCTAGCGTGAGCAAGGG	-
SS$^{\Delta}$-R	CATGGATCCGAGCTCGGTACC	-

Table 4.1. Primers used to generate the constructs used in this chapter. The F or R in the primer name refers to forward or reverse, respectively. Restriction enzyme cut sites are underlined. The sequence encoding the linker is bolded.

*Dpn*I treatment where 2µL PCR product was treated with T4 ligase, T4 PNK, and *Dpn*I in T4 ligase buffer in a 20µL reaction for 20 minutes at room temperature. The forward primers used to create SS-EGFP-SUN2$^{520\text{-}731}$ and SS-EGFP-SUN2$^{595\text{-}731}$ were SS-EGFP-SUN2$^{520\text{-}731}$-F and SS-EGFP-SUN2$^{595\text{-}731}$-F, respectively (Table 4.1). The same reverse primer, SS-EGFP-SUN2$^{\Delta}$-R (Table 4.1), was used for both SS-EGFP-SUN2$^{520\text{-}731}$ and SS-EGFP-SUN2$^{595\text{-}731}$.

To create SS-EGFP-SUN1$^{457\text{-}913}$, the sequence encoding amino acids 457-913 was PCR amplified from EGFP-SUN1 using the primers SS-EGFP-SUN1$^{457\text{-}913}$-F and SS-EGFP-SUN2$^{457\text{-}913}$-R (Table 4.1), which contain 5' *Bsr*GI and *Eco*RI cut sites, respectively. SS-EGFP-SUN1$^{457\text{-}913}$-F also encodes a 10 amino acid linker (GHGTGSTGSG) following the *Bsr*GI site. The PCR product was purified and digested beside SS-EGFP with *Bsr*GI and *Eco*RI. Following gel purification, the digested PCR product and plasmid were ligated together to create SS-EGFP-SUN1$^{457\text{-}913}$. SS-EGFP-SUN1$^{702\text{-}913}$ and SS-EGFP-SUN1$^{777\text{-}913}$ were generated via the Kinase, Ligase, *Dpn*I method as described above. The forward primers used to create SS-EGFP-SUN1$^{702\text{-}913}$ and SS-EGFP-SUN1$^{777\text{-}913}$ were SS-EGFP-SUN1$^{702\text{-}931}$-F and SS-EGFP-SUN2$^{777\text{-}913}$-F, respectively (Table 4.1). The same reverse primer, SS-EGFP-SUN$^{\Delta}$-R (Table 4.1), was used for both SS-EGFP-SUN1$^{702\text{-}913}$ and SS-EGFP-SUN1$^{777\text{-}913}$.

The cytoplasmic EGFP-tagged SUN1 and SUN2 constructs were generated via Kinase, Ligase, *Dpn*I reactions as follows. EGFP-SUN2$^{260\text{-}731}$ and EGFP-SUN1$^{457\text{-}913}$ were made first using the primers SS$^{\Delta}$-F and SS$^{\Delta}$-R (Table 4.1). Kinase, Ligase, *Dpn*I treatments were used to make EGFP-SUN2$^{520\text{-}731}$ and EGFP-SUN2$^{595\text{-}731}$. The forward

primers used to create EGFP-SUN2$^{520-731}$ and EGFP-SUN2$^{595-731}$ from EGFP-SUN2$^{260-731}$ were SS-EGFP-SUN2$^{520-731}$-F and SS-EGFP-SUN2$^{595-731}$-F, respectively (Table 4.1). The same reverse primer, SS-EGFP-SUN2^{deletion}-R (Table 4.1), was used for both EGFP-SUN2$^{520-731}$ and EGFP-SUN2$^{595-731}$. EGFP-SUN1$^{777-913}$ was generated via Kinase, Ligase, *Dpn*I treatment from EGFP-SUN1$^{457-913}$ using the primers SS-EGFP-SUN1$^{777-913}$-F and SS-EGFP-SUN1$^{777-913}$-R (Table 4.1).

4.2.5 Transfections

Transient transfections of cDNA and shRNA constructs were performed using GenJet from SignaGen Laboratories (Rockville, MD) or Lipofectamine LTX from Invitrogen (Carlsbad, CA) according to the manufacturers' instructions. Measurements were performed 24 or 48 hours after transfection for FFS experiments in the absence or presence of shRNA, respectively. Immediately before measurement, the growth medium was replaced with Dulbecco's phosphate-buffered saline (PBS) with calcium and magnesium from Biowhittaker (Walkerville, MD). Brightness measurements in the presence of shRNA were only performed on cells expressing turboRFP.

4.2.6 Fixed- and live-cell epifluorescence microscopy

All fixed-cell imaging was performed on an Eclipse Ni-E microscope driven by NIS-Elements software using a 40X /1.30 NA Plan Fluor oil immersion objective lens (Nikon Instruments Inc., Melville, NY), a SOLA solid state white-light excitation subsystem (Lumencor), and a CoolSNAP ES2 CCD camera (Photometrics). A custom DAPI filter set

(#49028 with exciter: ET395/25x, dichroic: T425LPXR, and emitter: ET460/50m, Chroma Technology, Bellows Falls, VT) for the SOLA light source was used. EGFP (C-FL EGFP Zero Shift, #96362, Nikon Instruments Inc.), and Texas Red (C-FL Texas Red Zero Shift, #96365, Nikon Instruments Inc.) filter sets were also used.

Twelve hours prior to their transfection, U2OS cells were grown in 24-well plates with #1.5 glass coverslip bottoms or 35 mm dishes with #1.5 glass coverslip bottoms from In Vitro Scientific (Sunnyvale, CA). Cells were then washed twice with live imaging media composed of HBSS (GIBCO®, Invitrogen) containing essential and nonessential MEM amino acids (Invitrogen), 2.5 g/L glucose, 2 mM glutamine, 1 mM sodium pyruvate, 20 mM HEPES (pH 7.4)), and transferred to a 37° C Okolab full-enclosure incubator (Ottaviano, Italy) with temperature control attached to an Intelligent Imaging Innovations (3I, Denver, CO) Marianas 200 Microscopy Workstation built on a Zeiss AxioObserver Z.1 stand (Jena, Germany) and driven by SlideBook 6.0 from 3I. All live cell epifluorescence images were acquired with a Zeiss Plan-Apochromat 63x/1.4NA oil objective, a Sutter DG4 light source (Novato, CA), and a Photometrics CoolSnap HQ2 CCD (Tucson, AZ). A BrightLine Sedat filter set, optimized for DAPI, FITC, TRITC, and CY5 from Semrock (Rochester, NY) was used.

4.2.7 FFS measurements

The instrumental setup for most of the experiments in this chapter was described in Section 2.2. When performing measurements in the presence of shRNA a dichroic mirror with a center wavelength of 515 nm (515DCLPXR, Chroma Technology, Bellows

Falls, VT) was used to split the emission light of EGFP and turboRFP into a 'green' and 'red' detection channel and an additional short pass filter centered at 512 nm (FF01-512/SP, Semrock, Rochester, NY) was added to the 'green' channel to remove any reflected light from turboRFP. Additional information on the instrumentation and analysis techniques used in this chapter can be found in Sections 2.2 and 2.3, respectively.

The number concentration *N* represents the average number of labeled protein monomers in the observation volume. Because the brightness of an individual EGFP protein is given by I_{EGFP}, the time-averaged fluorescence intensity of a measurement is proportional to the number concentration, $\langle F \rangle = \lambda_{EGFP} N$. We experimentally calculate *N* by dividing the average intensity $\langle F \rangle$ by I_{EGFP} (44). This procedure is valid for FFS experiments in the cytoplasm as well as at the NE. The observation volume V_O is given by the overlap between the two-photon PSF and the fluorescent sample. Converting the number concentration into a molar concentration is achieved by $c = N/(V_O N_A)$, where N_A is Avogadro's number. The volume $V_O^{(C)}$ of a cytoplasmic FFS experiment is measured using a previously published procedure (43). In contrast, the volume $V_O^{(NE)}$ and therefore the labeled protein concentration of FFS experiments at the NE cannot be determined experimentally. However, molar concentrations in the NE can be estimated as described below.

4.2.8 Brightness modeling

A monomer/dimer/trimer equilibrium reaction was used to model *b* as a function of *N*. The number of molecules of monomers, dimers, and trimers (N_1, N_2, N_3) was determined by the reactions $A + A \rightleftharpoons A_2$ and $A + A_2 \rightleftharpoons A_3$ with the dissociation coefficients K_{MD} and K_{DT}, respectively. By definition, the normalized brightness of an *n*-mer is given by *b* = *n*. The brightness of this mixture of species is $b = \sum_{i=1}^{3} b_i^2 N_i \Big/ \sum_{i=1}^{3} b_i N_i$. The total number of monomeric proteins in the observation volume is $A_0 = A + 2A_2 + 3A_3$. The same analysis was performed for other binding equilibrium models, such as the monomer/trimer reaction $3A \rightleftharpoons A_3$ with a dissociation coefficient defined by $K_{MT}^2 = [A]^3 / [A_3]$. A detailed description of the modeling is found in the Supplemental Materials. Fitting of the experimental data to *b* binding models was accomplished using bootstrapping (90). Confidence intervals of the estimated parameters were also determined from the bootstrap algorithm.

The observation volume $V_O^{(NE)}$ of NE measurements is small compared to the observation volume $V_O^{(C)}$ of cytoplasmic FFS experiments, which is reflected in the measured *N*. To facilitate the comparison of *b* changes with concentration of a protein in both compartments, it is useful to translate between the measured *N* in both environments. This is achieved by the molar concentration defined by *c* = $N^{(C)} \Big/ \left(V_O^{(C)} N_A \right)$ =

$N^{(NE)} \Big/ \left(V_O^{(NE)} N_A \right)$, which demonstrates that the values of *N* in the cytoplasm and the NE are proportional to one another, $N^{(C)} = N^{(NE)} V_O^{(C)} \Big/ V_O^{(NE)}$. While the observation volume at the NE cannot be measured, it can be modeled as the product of the cross-sectional area of the PSF and the thickness *h* of the NE layer, $V_O^{(NE)} = \frac{\rho w_0^2}{4} h$. Given previously published measurements of NE thickness (91), we assume that the NE has an average thickness *h* = 40 nm, which leads to a volume of 6.9 10^{-3} fl. This value is 34-fold smaller than the observation volume in the cytoplasm. Thus, the multiplication of *N* in the NE by 34 compensates for the difference in observation volume and determines the equivalent cytoplasmic *N*. The observation volume in the cytoplasm with a fully embedded PSF was determined to be 0.23 fl. These values served to calculate molar dissociation coefficients, which were quoted to one significant digit to reflect the uncertainties of the estimate.

4.2.9 Immunofluorescence

Cells grown on #1.5 coverslips were fixed in -20° C methanol as previously described (92). Coverslips were mounted on slides using Fluoromount purchased from Thermo Fischer Scientific.

4.3 Results and Discussion

4.3.1 Quantifying NE protein-protein interactions in living cells

FFS characterizes fluctuating fluorescence signals generated by fluorescently labeled proteins passing through an optical observation volume of a confocal or two-photon microscope (89). Subsequent brightness analysis of these fluorescence fluctuations provides quantitative information about the stoichiometry of the labeled proteins (44). We recently combined FFS with z-scans for quantifying protein-protein interactions at the plasma membrane and in thin cell sections (42, 43). Given the ~30-50 nm thickness of the PNS, we explored the use of FFS and z-scans as a method to quantify protein-protein interactions within the NE. For simplicity, we refer to the INM, ONM, and PNS collectively as the NE for the remainder of this work.

A z-scan consists of an axial scan of the two-photon spot through a cell expressing a fluorescently labeled protein (Fig. 4.1A) (43). The resulting axial intensity profile or 'z-scan' characterizes the subcellular distribution of the labeled proteins. Thus, a labeled NE protein produces a z-scan with two peaks along the trajectory, which correspond to signals generated within the ventral and dorsal NE. To quantify the fluorescence contributions from the NE and the nucleoplasm, z-scans were analyzed as previously described (42, 73). FFS data was collected with the two-photon spot repositioned at either NE.

We initially tested the feasibility of FFS and brightness analysis within the NE by measuring the normalized brightness (*b*) of EGFP in this subcellular compartment. EGFP was targeted to the ER lumen/PNS by adding the signal sequence (SS) from a luminal

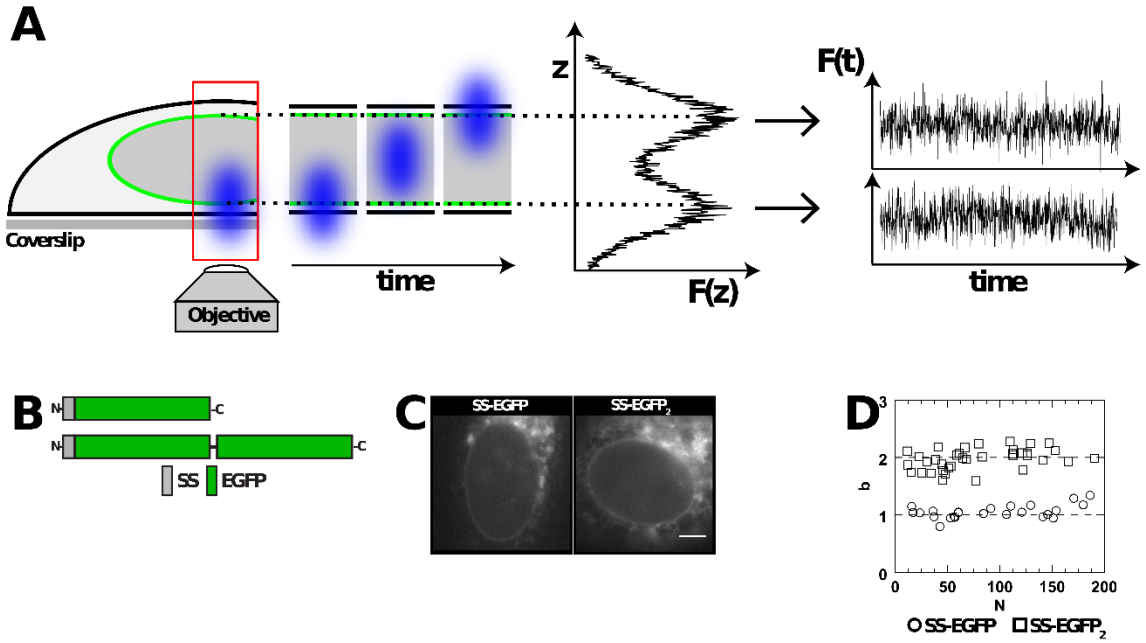

Figure 4.1: FFS and brightness analysis in the NE. A) Identification of the dorsal (NE_D) and ventral (NE_V) NEs in a cell expressing EGFP-tagged NE proteins by z-scan FFS. Fluorescence intensity fluctuations are measured at either NE. B) Constructs used in this figure. C) Representative epifluorescence images of U2OS cells expressing the indicated constructs. Scale bar: 5 μm. D) Brightness analysis of the cells described in C. Each data point represents the average *b* measured in a single cell.

protein, torsinA (77) (Fig. 4.1B). The *b* indicates the average oligomeric state of an EGFP-tagged protein (93), i.e. a monomer and a dimer corresponds to a *b* of one and two, respectively. To confirm that brightness accurately reports stoichiometry within the NE, we measured the *b* of a dimeric EGFP construct (SS-EGFP$_2$) (44). Following expression, both constructs localized to the ER/NE (Fig. 4.1C). Z-scans and FFS data for these constructs were collected from expressing cells, followed by the calculation of *b* and number concentration (*N*), which is the average number of EGFPs within the observation volume. Further details on *b* and *N* can be found in Sections 2.3 and 4.2.7.

Independent of *N*, the mean and standard deviation of *b* from SS-EGFP- or SS-EGFP$_2$-expressing cells was 1.06 ± 0.14 and 1.98 ± 0.18, respectively (Fig. 4.1D). These results are within the experimental uncertainty consistent with SS-EGFP being monomeric and SS-EGFP$_2$ dimeric within the NE. This work establishes FFS and brightness analysis as a powerful method for probing the *in vivo* biochemical and biophysical behavior of NE proteins within their native cellular environment.

4.3.2 SUN2 oligomerization in the NE

Next, we sought to detect *in vivo* SUN2 trimerization using FFS and brightness analysis. Since quantitative brightness analysis of FFS data requires labeled proteins to be mobile (76) and full-length EGFP-tagged SUN2 was shown to be highly immobile by fluorescence recovery after photobleaching (FRAP) (80), we generated and measured the oligomerization of a SS-EGFP-tagged construct that encodes the entire luminal domain of SUN2 (SS-EGFP-SUN2$^{261-731}$) (Fig. 4.2A). We also generated SS-EGFP-SUN2$^{520-731}$ and

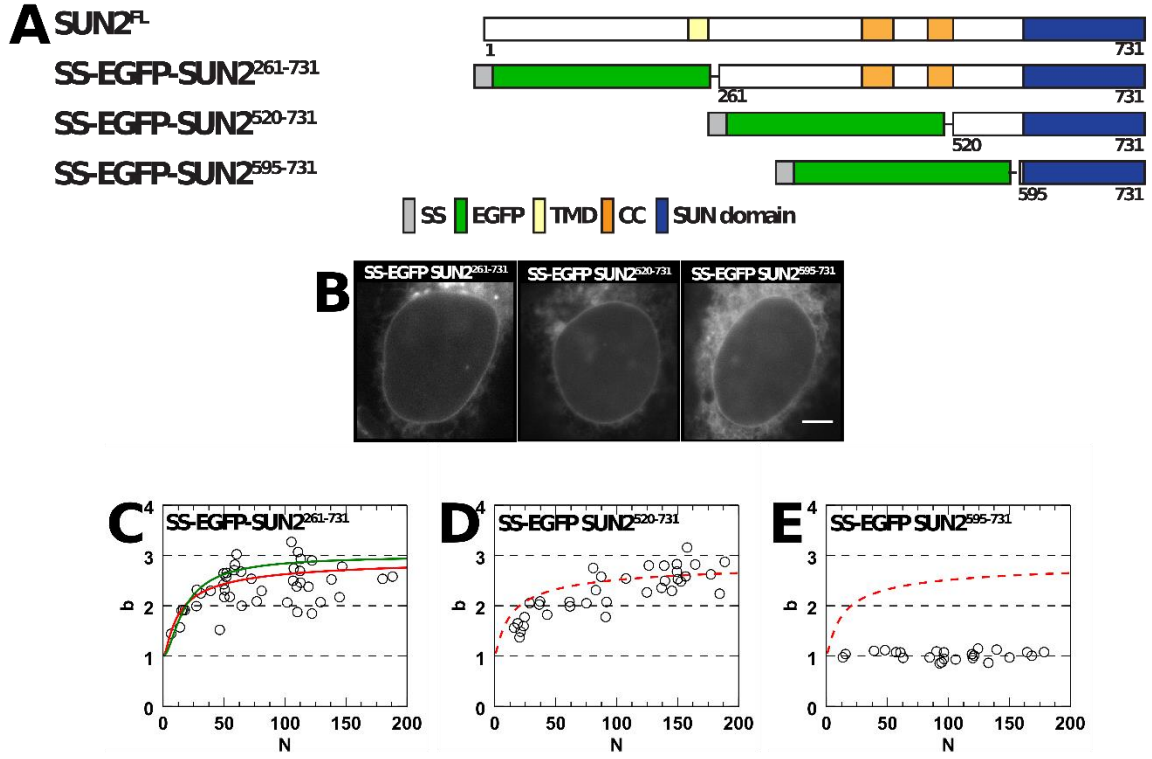

Figure 4.2: SUN2 oligomerization in the NE. A) Constructs used in this figure. B) Representative epifluorescence images of U2OS cells expressing the indicated constructs. Scale bar: 5 μm. C-E) Plots of *b* vs. *N* for the indicated constructs. The data in C were fit to a monomer/dimer/trimer binding model (solid red line), which is shown in D and E (dashed red line), with K_{MD} = 4100 (1000 μM) $^{+5900}_{-4000}$, K_{DT} = 0.06 (0.01 μM) $^{+3}_{-0.04}$, and a monomer/trimer binding model (solid green line) with K_{MT} = 26 (6 μM) ± 7.

SS-EGFP-SUN2$^{595-731}$, which encode a non-canonical CC through the SUN domain and the SUN domain alone, respectively (Fig. 4.2A). The absence of a transmembrane domain liberates these constructs from membrane-induced constraint, enabling diffusion throughout the ER lumen/PNS (Fig. 4.2B). Nevertheless, it should be noted that the transmembrane domain itself could influence SUN protein oligomerization. This possibility should be addressed in future studies.

Plotting *b* vs. *N* revealed a concentration-dependent increase in *b* with values mainly between 2 and 3 for $N > 50$ (Fig. 4.2C). These data show that SS-EGFP-SUN2$^{261-731}$ exhibits an average oligomerization state that is between a dimer and a trimer. While the *b* binding curve has not yet achieved saturation within the experimentally accessible concentration range, the data are approaching the next integer *b* of 3, suggesting a limiting trimeric stoichiometry as supported by fitting the data to a monomer/dimer/trimer binding model (Fig. 4.2C). The fitted monomer/dimer ($K_{MD} = 4100$) and dimer/trimer ($K_{DT} = 0.06$) dissociation coefficients, in units of *N* for the monomer/dimer and dimer/trimer reaction, indicate that dimers are a minority species since $K_{MD} > K_{DT}$. Thus, the *b* binding curve can also be modeled by a monomer/trimer equilibrium (Fig. 4.2C). Molar values for the dissociation coefficients were estimated as described in Section 4.2.8 and are quoted in the figure caption.

The *b* of SS-EGFP-SUN2$^{520-731}$ increased with *N* and approached 3 at high *N* values (Fig. 4.2D). While these data are on average slightly below the binding curve for $N < 100$, the differences are small. Unlike either SS-EGFP-SUN2$^{261-731}$ or SS-EGFP-SUN2$^{520-731}$, the *b* values obtained for SS-EGFP-SUN2$^{595-731}$ remain close to one and do not increase with concentration (Fig. 4.2E). These results demonstrate SUN2

trimerization within the NE of living cells and that the SUN domain is not sufficient for this oligomerization, consistent with previously reported *in vitro* studies (49, 88).

While SUN2$^{261-731}$ contains both types of CC, SUN2$^{520-731}$ only possesses the non-canonical CC. Despite this difference, both constructs displayed similar *N*-dependent oligomerization within the NE (Figs. 4.2C-D) suggesting that the non-canonical CC may be sufficient for SUN2 trimerization in agreement with the use of a similar human SUN2 construct to solve the crystal structure of SUN2 trimers (49).

4.3.3 SUN1 oligomerization in the NE

To provide insight into the conservation of SUN protein trimerization, we next investigated the oligomerization of SUN1 within the NE. Like SUN2, EGFP-tagged full-length SUN1 was also shown by FRAP to be immobile (80). Consequently, we generated an SS-EGFP-tagged construct encoding the entire luminal domain of SUN1 (SS-EGFP-SUN1$^{457-931}$) (Fig. 4.3A). Due to the lack of structural information available for SUN1, we limited our analysis to this construct, a construct analogous to SS-EGFP-SUN2$^{520-731}$ (SS-EGFP-SUN1$^{702-913}$), and another that encodes the SUN1 SUN domain (SS-EGFP-SUN1$^{777-913}$) (Fig. 4.3A). All three constructs localized to the peripheral ER and NE (Fig. 4.3B).

The *b* of SS-EGFP-SUN1$^{457-913}$ increased linearly over the entire range of *N* (Fig. 4.3C), which prohibited the fitting of this data to a binding curve and estimating the stoichiometry of this construct. Nevertheless, the highest *b* value measured for SS-EGFP-SUN1$^{457-913}$ was ~5, indicating the presence of higher-order oligomeric states

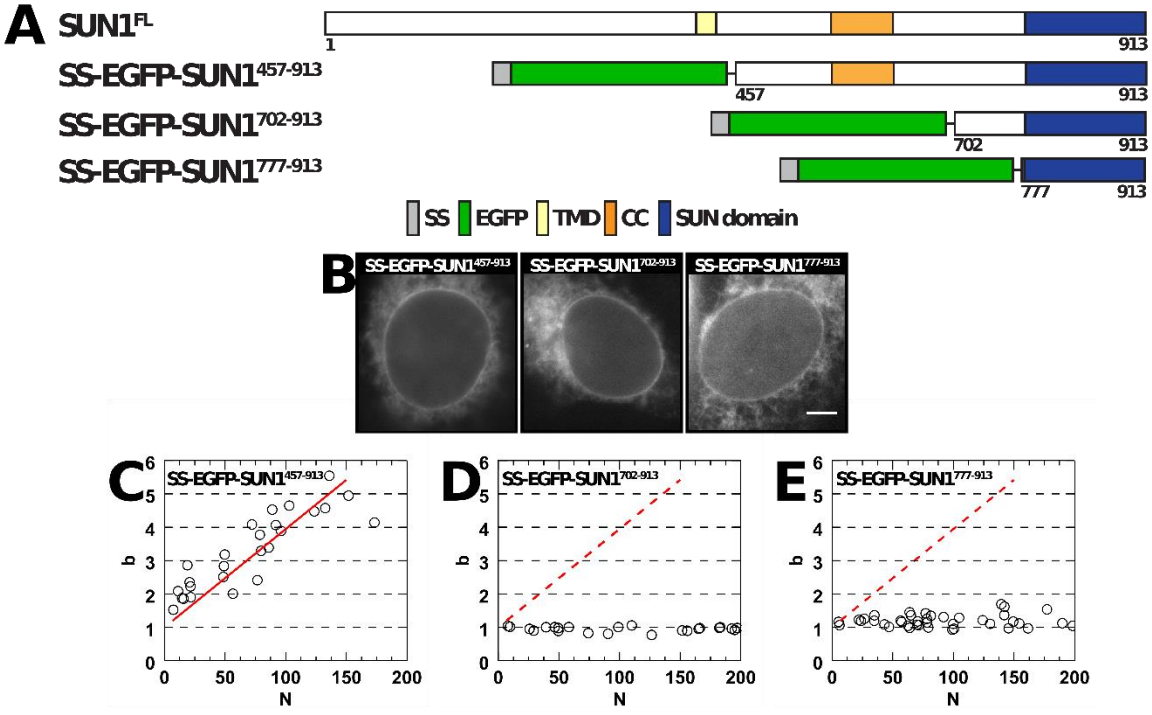

Figure 4.3: SUN1 oligomerization in the NE. A) Constructs used in this figure. B) Representative epifluorescence images of U2OS cells expressing the indicated constructs. Scale bar: 5 μm. C-E) Plots of *b* vs. *N* for the indicated constructs. The data in C were fit to a linear regression (solid red line), which is shown in D and E (dashed red line).

than those observed for SS-EGFP-SUN2$^{261-731}$. In contrast to what was observed for SS-EGFP-SUN2$^{520-731}$, the *b* of SS-EGFP-SUN1$^{702-913}$ did not appreciably increase above one over the range of measured *N* (Fig. 4.3D). Finally, a lack of oligomerization similar to what was observed for SS-EGFP-SUN2$^{595-731}$, was also reflected by the *b* values obtained for the SUN domain encoding SS-EGFP-SUN1$^{777-913}$ construct (Fig. 4.3E). These results suggest that trimers are not the limiting assembly state for all SUN proteins, which is in agreement with a previous report of the existence of immobile macromolecular assemblies of SUN1 within the NE composed of dimers and tetramers (86).

4.3.4 SUN protein oligomerization in the cytoplasm

Given the requirement of SUN2 trimerization for KASH-binding (49), SUN protein oligomerization may represent an important target for the regulation of LINC complex assembly. In fact, a recent report shows that the two canonical CCs of SUN2 display distinct oligomeric states, the modulation of which regulates the ability of SUN2 to interact with KASH peptides (51). However, the mechanisms responsible for regulating SUN2 oligomerization within the NE remain unknown. As an initial step towards defining these mechanisms, we quantified the oligomerization of constructs encoding EGFP-tagged luminal domains of SUN1 or SUN2 in the heterologous subcellular environment of the cytoplasm.

Cytoplasmic expression of these constructs was achieved by removing the SS (Fig. 4.4A-B). FFS experiments were performed in the cytoplasm of these cells as previously described (43, 94). The *b* of EGFP-SUN2$^{261-731}$ increased with *N* and appeared

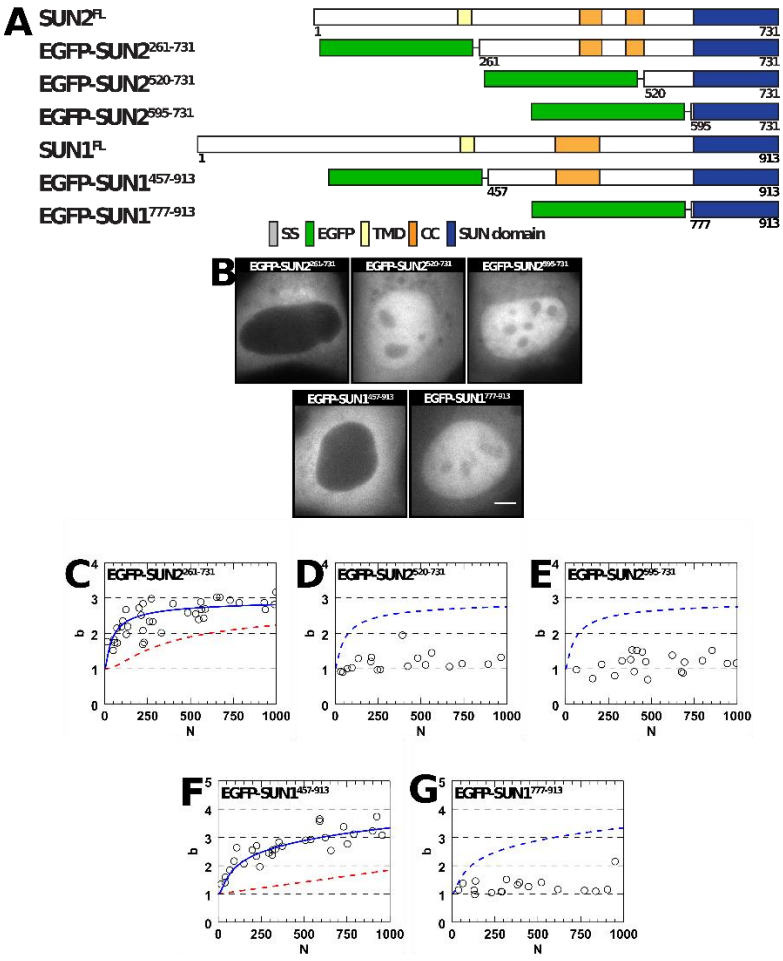

Figure 4.4: SUN1 and SUN2 oligomerization in the cytoplasm. A) Constructs used in this figure. B) Representative epifluorescence images of U2OS cells expressing the indicated constructs. Scale bar: 5 μm. C-G) Plots of the *b* vs. *N* for the indicated constructs. The data in C were fit to a trimeric binding model (solid blue line), which is shown in D and E (dashed blue line) with K_{MD} = 8000 (60 μM) ± 4000 and K_{DT} =0.3 (0.002 μM) ± 0.2. The data in F were fit to a monomer/trimer/hexamer binding model (solid blue line) with K_{MT} = 100 (0.7 μM) ± 60 and a trimer-hexamer dissociation coefficient K_{TH} = 1500 (10 μM) ± 400, which is then shown in G (dashed blue line). Estimated binding curves (dashed red lines) for the data obtained in the NE for the indicated constructs are presented in C and F by converting *N* from the NE to its cytoplasmic value.

to reach a limiting value of 3 at N > 600 (Fig. 4.4C), which agrees with the estimated stoichiometry of the analogous construct in the NE. A fit of the cytosolic b data measured for cytoplasmic EGFP-SUN2$^{261\text{-}731}$ to a monomer/dimer/trimer binding curve is shown with K_{MD} = 8000 and K_{DT} = 0.3 (Fig. 4.4C). Because $K_{DT} > K_{MD}$, the dimer population is negligible and a monomer/trimer equilibrium is sufficient to describe the b binding curve for EGFP-SUN2$^{261\text{-}731}$ (Fig. 4.S1A).

Only the cytoplasmic observation volume is experimentally obtainable (94). Thus, we computed an estimated NE observation volume to express N values in the NE as an approximate equivalent cytoplasmic N by accounting for the volume difference (see Section 4.2.8). Applying this procedure converts the trimeric NE b binding curve (Fig. 4.2C) to a predicted b binding curve in the cytoplasm (Fig. 4.4C). A comparison of the trimeric NE binding curve and the predicted cytoplasmic binding curve demonstrates that the observed b increase with N in the cytoplasm is significantly more pronounced than what was observed in the NE (see figure legends for more details). In contrast, both EGFP-SUN2$^{520\text{-}731}$ and EGFP-SUN2$^{595\text{-}731}$ appeared to be monomeric as their b values remained near one (Figs. 4.4D-E). These results reveal that the oligomerization of EGFP-tagged SUN2$^{261\text{-}731}$ and SUN2$^{520\text{-}731}$ is sensitive to as-of-yet unidentified environmental factors.

In addition, differences in the behavior of the SUN1$^{457\text{-}913}$ constructs in the cytoplasm relative to the NE were also observed. While the b of SS-EGFP-SUN1$^{457\text{-}913}$ and EGFP-SUN1$^{457\text{-}913}$ rises with increasing N, the b increase of EGFP-SUN1$^{457\text{-}913}$ within the cytoplasm slows at higher concentrations unlike the linear increase with N we observed in the NE (Fig. 4.4F, see figure legend for more details). However, the b of

EGFP-tagged SUN1$^{457-913}$ rises much faster in the cytoplasm as compared to the predicted *b* curve, which was converted from the NE *b* binding reaction (Fig. 4.3C).

Unlike EGFP-SUN2$^{261-731}$, EGFP-SUN1$^{457-913}$ does not reach the same *b* levels in the cytoplasm as in the NE. Instead, the *b* data slightly exceed 3 in the cytoplasm at high *N* values. A fit of the *b* values to a monomer/dimer/trimer binding model reveals that the tail of the binding curve ($N > 500$) is not reproduced by the model (Fig. 4.S1B). This misfit implies the need for at least one stoichiometric state in excess of a trimer for EGFP-SUN1$^{457-913}$ in the cytoplasm. While a monomer/trimer/hexamer binding model was sufficient to describe the experimental data, a monomer/dimer/tetramer binding model could also reproduce the data (Fig. 4.S1C), consistent with a previous report of the existence of SUN1 dimers and tetramers (86). Currently, our data cannot distinguish between these different binding models. Finally, the *b* of EGFP-SUN1$^{777-913}$ remains close to one over the measured concentration range (Fig. 4.4G), in agreement with the behavior observed for SS-EGFP-SUN1$^{777-913}$ in the NE. Since it did not display appreciable oligomerization within the NE, we did not measure the oligomerization of EGFP-tagged SUN1$^{702-913}$ in the cytoplasm.

A more pronounced *N*-dependent *b* increase for EGFP-tagged SUN2$^{261-731}$ and SUN1$^{457-913}$ was observed in the cytoplasm than in the NE (Figs. 4.4C and F). A potential explanation for these results could be the presence of unlabeled endogenous SUN proteins within the NE that compete with these labeled SUN constructs, leading to a reduction in *b*. This competition leads to an apparent reduction in the measured binding affinity (59), which potentially explains the shift of the binding curve of EGFP-tagged SUN2$^{261-731}$ and SUN1$^{457-913}$ to lower *N*. Unlike the binding affinity, the saturating value of

b remains unchanged and is approached after the exogenously expressed EGFP-tagged protein concentration exceeds the endogenous concentration as well as the K_D value, thereby identifying the stoichiometry of EGFP-tagged protein complexes (59).

To directly verify that the reported brightness values for the SS-EGFP-tagged luminal domains of SUN1 and SUN2 (Figs. 4.2C and 4.3C) are not lowered by the endogenous population, we performed additional measurements of both SS-EGFP-SUN2$^{261\text{-}731}$ and SS-EGFP-SUN1$^{457\text{-}913}$ in U2OS cells expressing a short hairpin RNA (shRNA), which efficiently depleted either endogenous SUN1 or SUN2, or a non-coding (NC) control shRNA (Fig. 4.S2A). SS-EGFP-SUN2$^{261\text{-}731}$ was previously observed to approach the saturating value of $b = 3$ for values of $100 < N < 200$ (Fig. 4.2C). Therefore, we selected SUN2 shRNA-expressing cells with concentrations of SS-EGFP-SUN2$^{261\text{-}731}$ in this range, measured *b,* and then calculated the median and quartile values (Fig. 4.S2B). Comparing *b* in the absence of shRNA to *b* in the presence of either NC or SUN2 shRNA resulted in p-values of 0.88 and 0.85, respectively. These results suggest that the *b* measured for SS-EGFP-SUN2$^{261\text{-}731}$ is unaffected by the absence or presence of endogenous SUN2, and thus the reported *b* values accurately reflect the average stoichiometry of SS-EGFP-SUN2$^{261\text{-}731}$ at concentrations of $100 < N < 200$.

Similarly, we compared *b* measurements in the presence and absence of NC or SUN1-depleting shRNA (Fig. 4.S3A) over the range of $100 < N < 200$. As with SS-EGFP-SUN2$^{261\text{-}731}$ we measured *b* and determined the median and quartiles. The *b* measurements performed on SS-EGFP-SUN1$^{457\text{-}913}$ in the presence of either NC or SUN1 shRNA resulted in p-values of 0.78 and 0.61, respectively, indicating that the measured

brightness is unaffected by the absence or presence of endogenous SUN1 (Fig. 4.S3B). Taken together with our observations of SS-EGFP-SUN2$^{261-731}$, these results demonstrate that the presence of endogenously expressed unlabeled SUN proteins has a negligible impact on our reported brightness at the higher concentrations we measured. Furthermore, the loss of SUN2$^{520-731}$ oligomerization in the cytoplasmic environment as compared to the NE (Figs. 4.2D and 4.4D) cannot be caused by endogenous competition in the NE, since such competition can only lower the brightness in the NE environment.

Thus, our data imply the existence of unidentified regulators of SUN protein oligomerization within the NE, which may be chemical in nature. For instance, the contiguous ER lumen and PNS have a high calcium concentration and an oxidizing environment that favors the formation of disulfide bonds (95). Since the conserved cation loop in the SUN domain of SUN2 is required for KASH-binding, which also requires SUN2 trimerization (49), we anticipate that SUN protein oligomerization may be sensitive to changes in the concentration of calcium within the PNS. In addition, SUN1 oligomerization involves interchain disulfide bonds, which leads us to speculate that SUN protein oligomerization may be influenced by the redox potential of the ER lumen/PNS (95). Alternatively, luminal proteins such as the AAA+ ATPase torsinA, which was recently shown to localize to and be required for the assembly of transmembrane actin-associated (TAN) lines in migrating fibroblasts (92), may structurally regulate SUN protein oligomerization.

4.3.5 Models of SUN1 and SUN2 oligomerization

We propose that SUN2 monomers are in equilibrium with SUN2 trimers in the NE, with no evidence for a significantly populated dimeric state (Fig. 4.5A). Currently, we cannot distinguish between two models of SUN1 oligomerization that are not mutually exclusive (Figs. 4.5B'-B"). In the first, SUN1 oligomerizes via a monomer/dimer/tetramer reaction (Fig. 4.5B'). In the second, SUN1 oligomerizes via a monomer/trimer/hexamer reaction (Fig. 4.5B"). Both reactions would ultimately lead to the assembly of higher-order SUN1 oligomers through progressive oligomerization (i.e. monomer to trimer to hexamer to n-mer). The second model is favored based on recently published computational modeling results from the Mofrad laboratory, which suggest that, like SUN2, SUN1 is capable of forming stable homo-trimers (31). Unlike SUN2 homo-trimers, however, they found that SUN1 homo-trimers were able to form lateral complexes via the association of their SUN domains. This model is consistent with our observation of *b* values in excess of trimers for the case of SS-EGFP-SUN1$^{457\text{-}913}$. Future work will be needed to carefully explore the modeling predictions put forward by the Mofrad laboratory.

The ability of SUN1 to form higher-order oligomers than SUN2 may be related to its ability to form rings around NE-associated meiotic telomeres (96) and to localize to NPCs (97). These SUN1-specific localizations may explain the differential requirement for SUN1 during meiotic chromosome pairing and DNA double stranded break repair (85, 96, 98) as well as NPC insertion and distribution (86, 87). Moreover, SUN2 trimerization may be required for actin-dependent nuclear movement as demonstrated by the specific recruitment of SUN2 to TAN lines in migrating fibroblasts (18). Future efforts aimed at

further understanding the mechanisms of the differential oligomerization of SUN1 and SUN2 will provide important insights into how LINC complex-dependent mechanotransduction and nuclear-cytoplasmic communication.

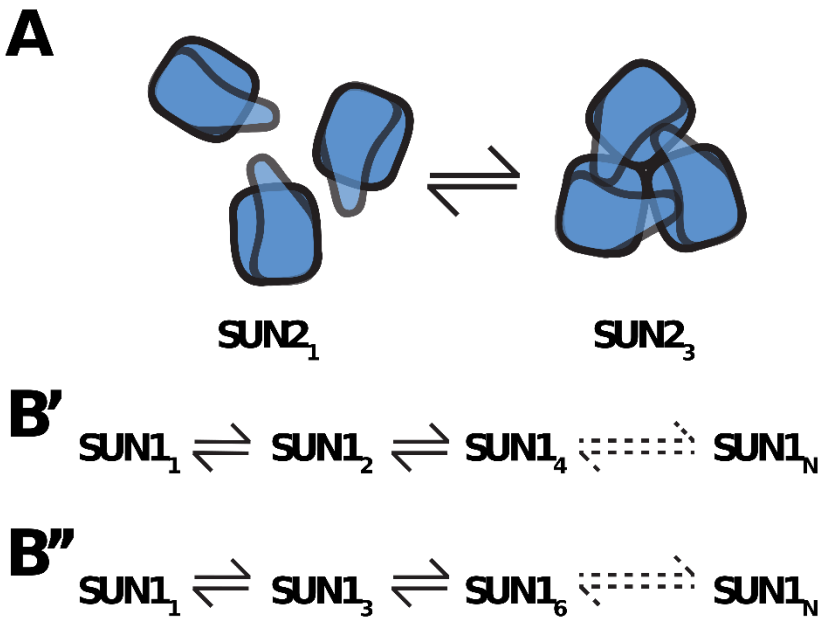

Figure 4.5: Models of SUN1 and SUN2 oligomerization. A) Working model of SUN2 oligomerization. B' and B'') Working models of SUN1 oligomerization with the NE, which lead to the assembly of higher-order SUN1 oligomers ($SUN1_N$).

4.4 Supplemental Figures

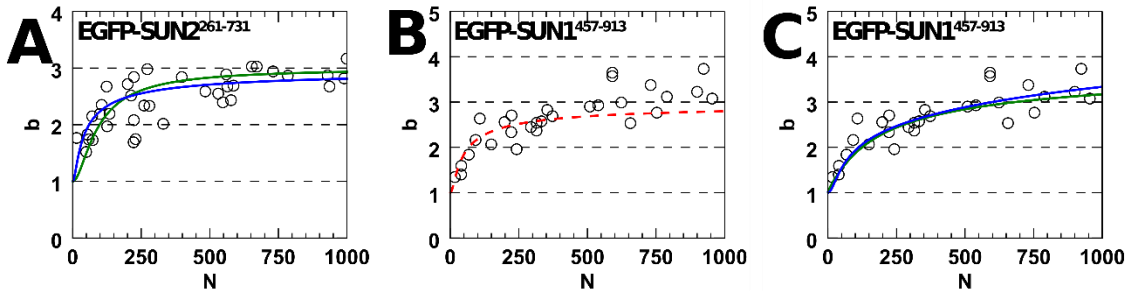

Figure 4.S1: Cytoplasmic EGFP-SUN2$^{261-731}$ and EGFP-SUN1$^{457-913}$ *b* data fit to different models. (A) The data from Fig. 4.4C were fit to the monomer/dimer/trimer binding model (blue line) from the same figure and a monomer/trimer binding model (green line) with K_{MT} = 130 (1 μM) ± 9. **(B)** The data in Fig. 4.4F were fit to a monomer/dimer/trimer binding model (dashed red line). **(C)** Data from B were fit to the monomer/trimer/hexamer binding model (blue line) from Fig. 4.4F and a monomer/dimer/tetramer binding model (green line) with K_{MD} = 110 (0.8 μM) ± 30 and K_{DT} = 90 (0.7 μM) ± 30.

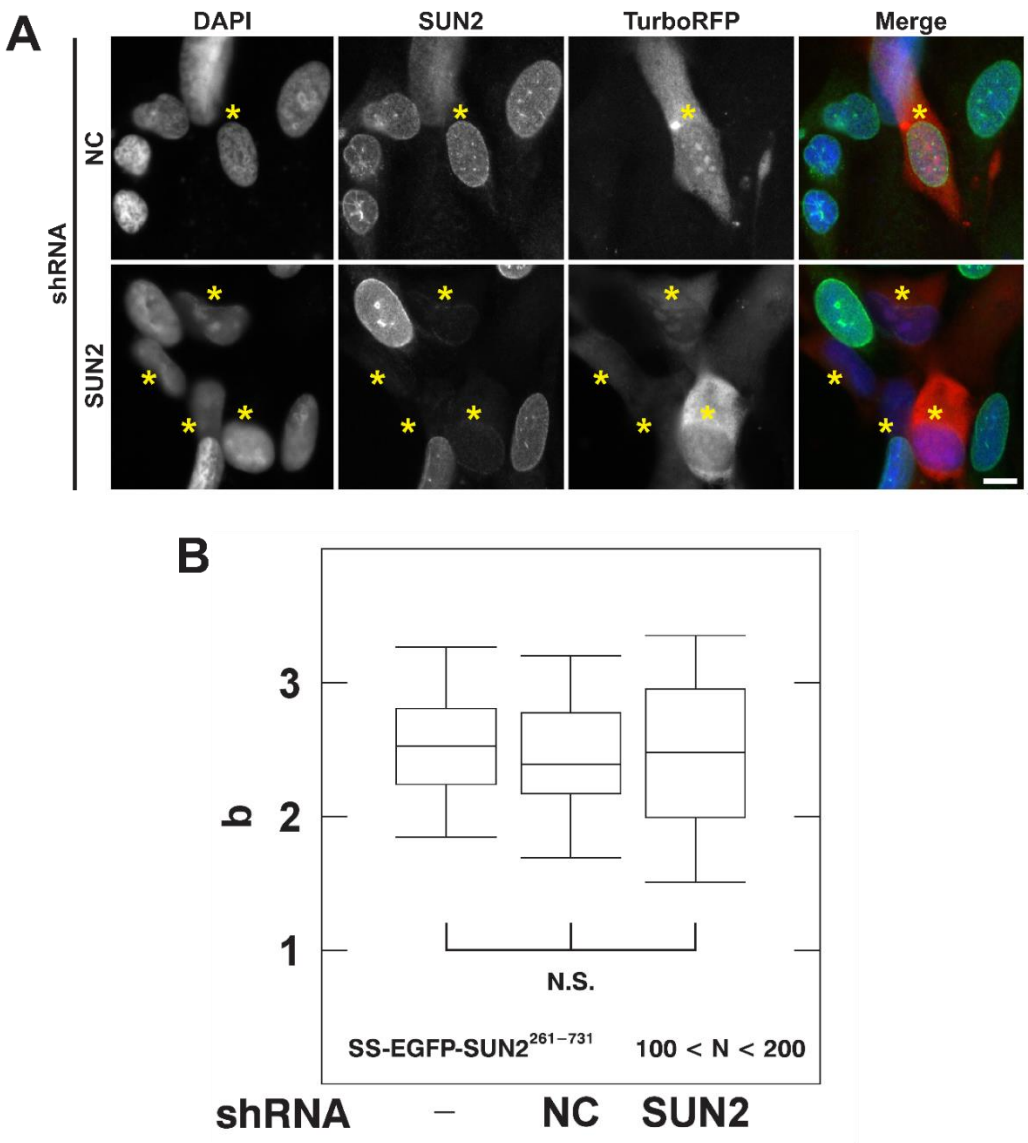

Figure 4.S2. SS-EGFP-SUN2$^{261-731}$ oligomerization in cells depleted of endogenous SUN2. (A) Representative epifluorescence images of U2OS cells expressing the indicated shRNA constructs. Yellow asterisks: shRNA-expressing cells. Scale bar: 10 μm. **(B)** Box and whisker plot of *b* for SS-EGFP-SUN2$^{261-731}$ with $100 < N < 200$ in the absence of shRNA (-) as well as the presence of noncoding (NC) or SUN2-depleting shRNA.

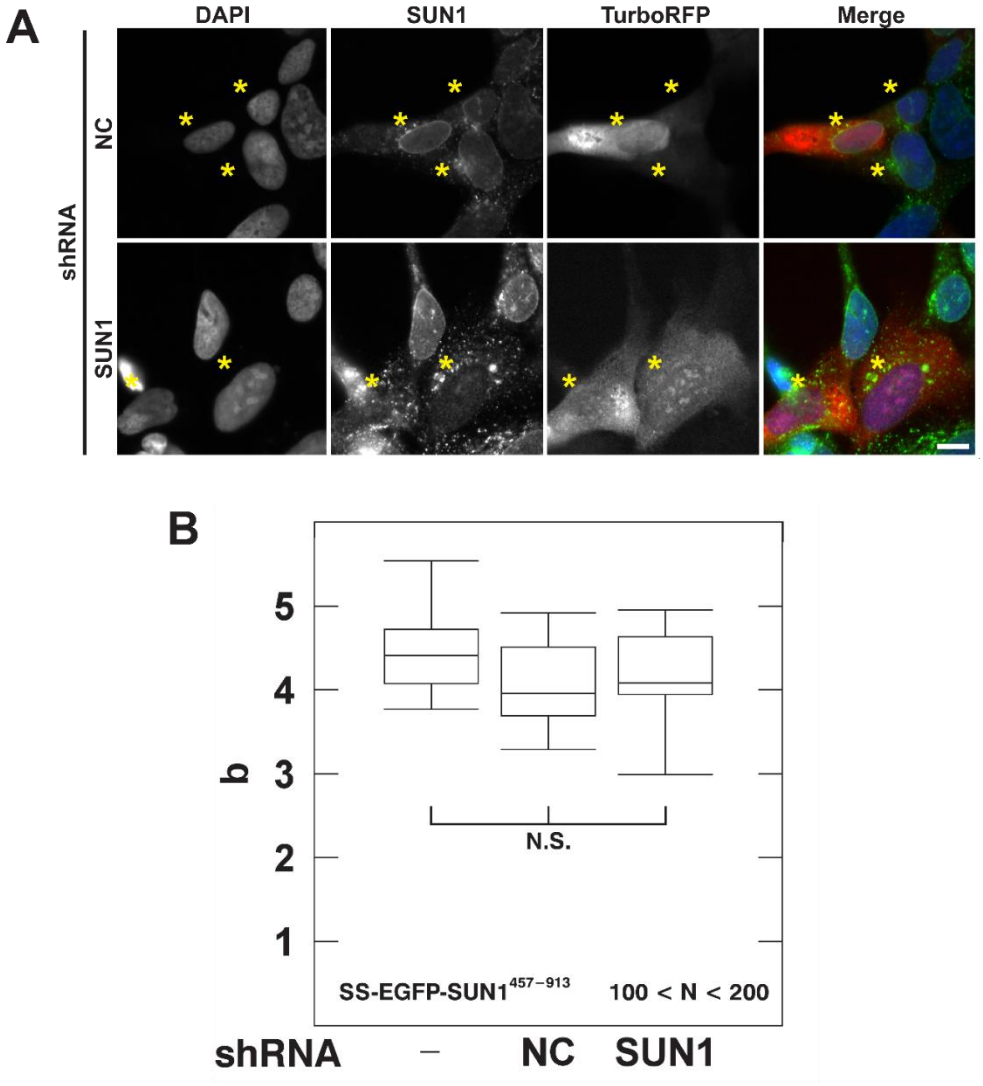

Figure 4.S3. SS-EGFP-SUN1$^{457-913}$ oligomerization in cells depleted of endogenous SUN1. (A) Representative epifluorescence images of U2OS cells expressing the indicated constructs. Yellow asterisks: shRNA-expressing cells. Scale bar: 10 μm. **(B)** Box and whisker plot of *b* for SS-EGFP-SUN1$^{457-913}$ with 100 < *N* < 200 in the absence of shRNA (-) as well as the presence of noncoding (NC) shRNA or SUN1-depleting shRNA.

5. Protein Oligomerization and Mobility within the Nuclear Envelope Evaluated by the Time-Shifted Mean-Segmented Q Factor

The previously discussed MSQ analysis has fairly significant drawbacks that were not discussed in detail. Specifically, dead-time and afterpulsing effects must be accounted for in the analysis. In this chapter we introduce the time-shifted MSQ (tsMSQ) method which overcomes many of these issues. The theory behind tsMSQ analysis and the relevant equations were derived by Dr. Kwang-Ho Hur who also performed measurements and data analysis for cytoplasmic EGFP. Measurements of SUN3 in the NE were performed by Siddarth Reddy Karuka. My contributions entailed developing decorrelated tsMSQ analysis, designing and performing all other experiments, and analyzing the resulting FFS data.

5.1 Introduction

As Chapter 3 showed, the double membrane system of the NE proved to be a challenging environment for traditional FFS techniques which prompted us to use the recently developed mean-segmented Q (MSQ) method for analysis (72). Nanometer-sized undulations of the INM and ONM introduce fluctuations in the local thickness of the NE that are superimposed on the fluorescence intensity fluctuations caused by NE proteins passing through the observation volume (OV) of the microscope. MSQ proved to be essential for separating these different noise sources and

characterizing the oligomeric state of proteins in the NE (53). Unfortunately, the application of MSQ is not straightforward as it requires compensation for non-ideal detector effects to remove significant biases from the collected data (67). While methods to account for these effects exist (67, 68), they necessitate the difficult task of properly calibrating each detector. Moreover, the correction procedure becomes unreliable once the amplitude of the bias correction approaches the amplitude of the signal. In addition, the original formulation of MSQ lacks error analysis (72), which is a significant shortcoming that prevents the statistical testing of models and the determination of uncertainties in fit parameters. This chapter addresses these challenges by introducing time-shifted MSQ (tsMSQ) together with a simple procedure for proper error analysis. Experimental verification of the theory demonstrates that tsMSQ is inherently robust with respect to non-ideal detector effects, which significantly simplifies the analysis of FFS data obtained within the NE, and is suitable for model analysis with error estimates.

To demonstrate the utility of tsMSQ for analyzing FFS data collected within the NE, we used it to quantify the assembly and dynamics of the EGFP-tagged luminal domain of the Sad1/UNC-84 (SUN) protein SUN2, which is an INM protein and a key component of the linker of nucleoskeleton and cytoskeleton (LINC) complex (74). The decision to initially analyze the luminal domain of SUN2 provided an important test case for tsMSQ, since we previously reported that this protein homo-trimerizes within the NE as determined by MSQ analysis of FFS data (54). We then used tsMSQ to determine the assembly state of the luminal domain of the germline-restricted SUN protein, SUN3 (99), which is currently unknown. Taken together, these results establish tsMSQ as a simple, yet powerful method

for analyzing FFS data taken within the NE, which eliminates the need for an in-depth knowledge of detector effects.

5.2 Material and Methods

5.2.1 Experimental setup

The instrumentation used in this chapter is the same as described in Section 2.2.

5.2.2 Measurement procedure

Calibration measurements were performed on cells transiently transfected with EGFP. Transfected cells were identified using brief epifluorescence illumination. The selected cell was centered and the two-photon laser spot was focused into the cytoplasm of the cell. Fluorescence intensity fluctuation data was acquired for ~60 s followed by z-scans. These data were used to determine the molecular brightness, which is also referred to simply as brightness, λ_{EGFP} as previously described (43, 68). In cells expressing NE localized proteins, z-scans were used to identify cells with NE intensity fractions >90% for FFS measurements (53). The point spread function (PSF) was focused on the ventral NE and fluorescence intensity fluctuations data were acquired for ~60 s to ~300 s; the same procedure was repeated at the dorsal NE. A detailed description of the measurement protocol is found in Hennen et al.(64). The FFS data was analyzed as described in the results section of this chapter.

5.2.3 Sample preparation

Experiments were performed using transiently transfected U2OS cells (ATCC, Manassas, VA), which were maintained in DMEM with 10% FBS (Hyclone Laboratories, Logan, UT) before being sub-cultured into 24-well glass bottom plates (In Vitro Scientific, Sunnyvale, CA) 24 hours prior to transfection. Transfections were performed using GenJet (SignaGen Laboratories, Rockville, MD) 12-24 hours prior to measurement according to the protocol provided by the manufacturer. Growth medium was replaced immediately before measuring with DPBS containing calcium and magnesium (BioWhittaker, Walkerville, MD).

5.2.4 DNA constructs

The SS-EGFP and SS-EGFP-SUN2$^{261-731}$ constructs were described previously (53). The SS-EGFP-SUN3$^{30-320}$construct was generated as follows. First, the cDNA encoding full length (FL) SUN3 was PCR amplified using the SUN3FL-F and SUN3FL-R primer pair (Table 5.1) from a prep of cDNA isolated from the testes of post-embryonic day 21 male mice, which was a gift from Drs. Vivian Bardwell and David Zarkower (University of Minnesota, Minneapolis, MN). The PCR product was purified and digested beside pEGFP-C1 with *Eco*RI and *Xho*I. Following gel purification, the digested PCR product and plasmid were ligated together to create EGFP-SUN3^{FL}. To create SS-EGFP-SUN3$^{30–320}$, the sequence encoding amino acids 30-320 was PCR-amplified from EGFP-SUN3^{FL} using the primers SS-EGFP-SUN3$^{30-320}$-F and SS-EGFP-SUN3$^{30-320}$-R (Table 5.1). SS-EGFP-SUN3$^{30-320}$-F also encodes a 10-amino

acid linker (GHGTGSTGSG) following the *Bsr*GI site. The resulting PCR product was purified and digested beside SS-EGFP with *Bsr*GI and *Xho*I. Following gel purification, the digested PCR product and plasmid were ligated together to create SS-EGFP-SUN3$^{30–320}$. The cytoplasmic EGFP-tagged SUN3$^{30-320}$ construct was generated via a T4 polynucleotide kinase (PNK), T4 DNA ligase, *Dpn*I reaction after EGFP-SUN3$^{30-320}$ was PCR amplified using the primers SSΔ-F and SSΔ-R (Table 5.1). All of the constructs used in this work were sequence-validated by the University of Minnesota Genomics Center. Phusion DNA polymerase, T4 DNA ligase, and T4 PNK were purchased from New England Biolabs (NEB, Ipswich, MA). Restriction enzymes were either purchased from NEB or Promega (Madison, WI). Wizard SV Gel and PCR Clean-Up System was purchased from Promega. GeneJet Plasmid Midiprep Kit was purchased from ThermoFisher Scientific (Waltham, MA). All other chemicals were purchased from Sigma-Aldrich (St. Louis, MI) unless otherwise specified.

Primer Name	**DNA Sequence**	**5' RE Site**
SUN3^{FL}-F	TTTTCTCGAGATGTTAACTCGATCATGGAAGATTATCC	*Xho*I
SUN3^{FL}-R	AAAAAGAATTCCTAAGTGTAATCACTGGGGATGCCG	*Eco*RI
SS-EGFP-SUN3$^{30-320}$-F	GCTGTACAA**GGGGCACGGGACCGGG**TCTACAGGGAGCGGGAAAGAAACAGAGTTTCCTCA	*Bsr*GI
SS-EGFP-SUN3$^{30-320}$-R	AAAACTCGAGCTAAGTGTAATCACTGGGGATGC	*Xho*I
SS$^{\Delta}$-F	GTGGCTAGCGTGAGCAAGGG	-
SS$^{\Delta}$-R	CATGGATCCGAGCTCGGTACC	-

Table 5.1: Primers used to generate the constructs used in this chapter. The F or R in the primer name refers to forward or reverse, respectively. Restriction enzyme (RE) cut sites are underlined. The sequence encoding the linker is bolded.

5.3 Background

We recently showed that conventional FFS analysis of NE proteins can introduce severe distortions in the recovered brightness values and confound data interpretation as a result of NE membrane undulations (53). Analysis methods which ignore the temporal information of fluctuations, such as cumulant analysis, PCH, and FIDA are unable to distinguish the diffusion and membrane undulation processes and therefore are unsuited to recover brightness in the NE (100). Methods that include temporal information, such as FCS, TIFCA, and FIMDA (100), do not account for the finite data segment length incorporated in the analysis algorithm, which has the potential to introduce biases when applied to slow processes as was observed recently due to the NE membrane undulations (53). To remedy this issue, we introduced MSQ analysis to obtain brightness values that accurately reflect the oligomeric state of FP-tagged proteins within the NE (72). The MSQ curve is determined by dividing the recorded photon counts into segments with period T (Fig. 5.1A). Q is calculated for each segment and subsequently averaged to yield the data point MSQ(T). By repeating this process for different values of T, the MSQ curve is constructed (Fig. 5.1B). Fitting of the MSQ curve is used to recover both the brightness and the diffusion time of the labeled protein. Using this method we demonstrated that the luminal domain of SUN2 undergoes a monomer / trimer transition, whereas the luminal domain of the related protein SUN1 forms higher-order oligomers within the NE (54).

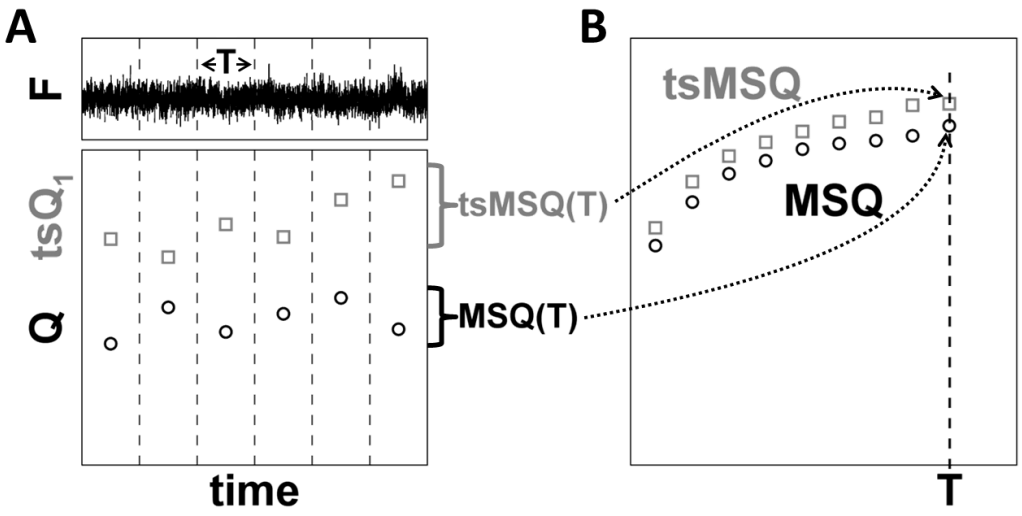

Figure 5.1. Construction of MSQ and tsMSQ curves. A) The recorded fluorescence intensity signal is divided into segments of period *T* (top panel). *Q* or tsQ_1 is calculated for each of these segments (bottom panel). B) The Q-values for the segment time *T* are converted by an algorithm into an MSQ(*T*) value. Similarly, the tsQ_1 values are converted into a tsMSQ(*T*) value. Repeating this procedure for a range of segment times identifies the experimental MSQ or tsMSQ curve.

5.4 Theory

This section describes derivations of functions that are important for tsMSQ. Interpretation and usage of the derived equations is found in Section 5.5.

5.4.1 Sampling time dependence of the Q-parameter

The Q-parameter of FFS data sampled with a time resolution T_S is calculated by (65)

$$Q(T_S) = \frac{\langle \delta k_i^{\,2} \rangle}{\langle k_i \rangle} - 1 \qquad 5.1$$

The raw data k_i represents the photon counts detected at time $i \times T_S$, while $\langle\ \rangle$ symbolizes the population mean. The deviation or fluctuation of k_i from the mean $\langle k_i \rangle$ is given by $\delta k_i \equiv k_i - \langle k_i \rangle$. For an arbitrary sampling time, the Q-parameter of a single diffusing species with brightness λ and diffusion time τ_D is determined by (68, 75)

$$Q(T_S) = Q_0 \frac{B_2(T_S, \tau_D)}{T_S^2}\ , \qquad 5.2$$

where $B_2(T_S, \tau_D)/T_S^2$ accounts for the decrease in $Q(T_S)$ relative to the asymptotic Q-factor Q_0 and, in the case of a two-dimensional Gaussian OV, is given by

$B_2\left(T_S,\tau_D\right)=2\tau_D\left(\left(\tau_D+T_S\right)\ln\left(1+T_S/\tau_D\right)-T_S\right)$ (66). This decrease in $Q\left(T_S\right)$ is significant when $T_S \geq \tau_D$, which is referred to as undersampling (101). In the case of oversampling ($T_S \ll \tau_D$), $B_2\left(T_S,\tau_D\right)/T_S^2$ converges to 1 and the Q-value $Q\left(T_S\right)$ reduces to the limiting value of Q_0, which is directly related to λ by Eq. 5.2. Since FFS data obtained in cells are typically oversampled, Eq. 2.4 is usually quoted in the literature. In this chapter we use Eq. 5.2, which is correct for all sampling times, as it facilitates comparison with the time-shifted FFS theory described in Section 5.4.3.

5.4.2 Eliminating the shot noise term from MSQ

The MSQ value of FFS data sampled with a time resolution T_S is calculated by dividing the data into segments of duration T and determining the expectation value of the Q-estimator $\hat{Q}_T$ across all data segments (Fig. 5.1) as previously described (72). Here, we introduce an updated definition of MSQ by adding the term T_S/T,

$$\mathrm{MSQ}\left(T\right)\equiv\left\langle\hat{Q}_T\left(T_S\right)\right\rangle+\frac{T_S}{T}\ . \qquad 5.3$$

This redefinition removes the bias caused by the shot-noise of the detector, which provides no relevant information about the sample. Moreover, it simplifies the expression of the MSQ curve to

$$\mathrm{MSQ}(T) = Q_0 \frac{B_2(T_S, \tau_D)}{T_S^2} - Q_0 \frac{B_2(T, \tau_D)}{T^2}. \qquad 5.4$$

The first term represents the sampling-time correction of the asymptotic Q-parameter. The second term characterizes the remaining degree of correlation within a segment of length *T* and depends on the second-order binning function $B_2(T, \tau_D)$ for FP-tagged proteins with diffusion time τ_D (66, 75). For simplicity, we omit the parameter τ_D in the derivations below. The MSQ curves with and without the shot-noise bias term are provided in Sup. Fig. 5.S1.

5.4.3. tsMSQ for a diffusing species

tsMSQ and MSQ are conceptually very similar algorithms that are applied to the same FFS data sampled with frequency $1/T_S$ (Fig. 5.1). However, while MSQ calculates Mandel's Q-parameter using Eq. 5.1, tsMSQ is based on a generalized form of the Q-factor (68),

$$tsQ_1 \equiv \frac{\langle \delta k_i \delta k_{i+1} \rangle}{\langle k_i \rangle}, \qquad 5.5$$

which we refer to as time-shifted Q-value. In the oversampling limit, where $T_S \ll \tau_D$, the time-shifted covariance is given by $\langle \delta k_i \delta k_{i+1} \rangle = \gamma_2 \lambda^2 T_S^2 N G_2(T_S) = Q_0 \langle k_i \rangle G_2(T_S)$ (75). $G_2(T_S)$ is the second-order normalized autocorrelation function of a single diffusing

species and accounts for the correlation between successive data points. Thus, the oversampled, time-shifted Q-factor is written as

$$tsQ_1 = Q_0 G_2\left(T_S\right) \ . \tag{5.6}$$

A generalization of Eq. 5.6 valid for all sampling times,

$$tsQ_1 = Q_0 \frac{tsB_2\left(T_S\right)}{T_S^2} \ , \tag{5.7}$$

is derived in Section 5.4.4. The factor $tsB_2\left(T_S\right)/T_S^2$ describes the reduction of tsQ_1 from Q_0 due to diffusion, where tsB_2 represents the time-shifted binning function of second order defined in Section 5.4.4.

While Eq. 5.7 specifies the population value, experimental data are always finite and require a statistical estimator. Because estimators frequently introduce biases, their expectation value must be critically assessed. We begin by assuming a segment with *M* data points, which corresponds to a segment time period of $T = M \times T_S$, and define the following estimator of tsQ_1,

$$tsQ_1\left(T\right) \equiv \frac{\overline{\Delta k_i \Delta k_{i+1}}}{\overline{k_i}} \ , \tag{5.8}$$

which we express as a function of *T* to emphasize the estimator's dependence on the segment time. $\overline{k_i} = \sum_{i=1}^{M} k_i \big/ M$ is the estimator of the mean and

$$\overline{\Delta k_i \Delta k_{i+1}} = \frac{1}{M-1} \sum_{i=1}^{M-1} \left(k_i - \overline{k}_L\right)\left(k_{i+1} - \overline{k}_R\right) \tag{5.9}$$

is the estimator of the time-shifted covariance $\langle \delta k_i \delta k_{i+1} \rangle$ with $\bar{k}_L = \frac{1}{M-1}\sum_{j=1}^{M-1} k_j$ and $\bar{k}_R = \frac{1}{M-1}\sum_{j=1}^{M-1} k_{j+1}$.

To evaluate the expectation value of the estimator of tsQ_1 we rewrite the fluctuation estimators in terms of δk_i,

$$\begin{aligned} \Delta k_i &= k_i - \bar{k}_L = \delta k_i - \frac{1}{M-1}\sum_{j=1}^{M-1} \delta k_j \quad , \\ \Delta k_{i+1} &= k_{i+1} - \bar{k}_R = \delta k_{i+1} - \frac{1}{M-1}\sum_{j=1}^{M-1} \delta k_{j+1} \quad . \end{aligned} \tag{5.10}$$

Inserting these expressions into Eq. 5.8 results in

$$tsQ_1(T) = \left(\frac{1}{M-1}\sum_{i=1}^{M-1} \delta k_i \delta k_{i+1} - \frac{1}{(M-1)^2}\sum_{i=1}^{M-1}\sum_{j=1}^{M-1} \delta k_i \delta k_{j+1} \right) \Bigg/ \frac{1}{M}\sum_{i=1}^{M} k_i \, . \tag{5.11}$$

Since the ensemble average of a ratio is equal to the ratio of the respective ensemble averages, we obtain

$$\left\langle tsQ_1(T) \right\rangle = \frac{\langle \delta k_i \delta k_{i+1} \rangle}{\langle k_i \rangle} - \frac{1}{(M-1)^2}\sum_{i=1}^{M-1}\sum_{j=1}^{M-1} \frac{\langle \delta k_i \delta k_{j+1} \rangle}{\langle k_i \rangle} , \tag{5.12}$$

where the first term is equal to the ideal tsQ_1-factor given by Eq. 5.5 and the second term reflects the bias introduced by using the estimator of the mean instead of the population mean. Because the second term becomes negligible as *M* increases, the estimator is asymptotically unbiased. Evaluating Eq. 5.12 (see Section 5.4.4) leads to

$$\left\langle tsQ_1(T)\right\rangle = tsQ_1 - \frac{(T-2T_S)T_S}{(T-T_S)^2} - Q_0\frac{tsC_2(T)}{(T-T_S)^2}. \quad 5.13$$

For convenience we defined a new function,

$$tsC_2(T) = \left(B_2(T) + B_2(T-2T_S) - 2B_2(T_S)\right)/2\ , \quad 5.14$$

in the third term of Eq. 5.13, which describes the influence of the correlation of the data due to diffusion. The shot noise for $i = j+1$ gives rise to the second term. Finally, to eliminate the shot noise term we define tsMSQ by

$$tsMSQ(T) \equiv \left\langle tsQ_1\right\rangle + \frac{(T-2T_S)T_S}{(T-T_S)^2}, \quad 5.15$$

which leads to an analytical expression for tsMSQ of a diffusing species,

$$tsMSQ(T) = Q_0\frac{tsB_2(T_S)}{T_S^2} - Q_0\frac{tsC_2(T)}{(T-T_S)^2}. \quad 5.16$$

The above equation applies to a single species, but a straightforward generalization of tsMSQ to account for a mixture of *S* species is provided by,

$$tsMSQ(T) = \sum_{i=1}^{S} f_i\, tsMSQ_i(T)\ , \quad 5.17$$

where $tsMSQ_i$ is the time-shifted MSQ of the *i*-th species and f_i is its intensity fraction.

5.4.4 Derivation of tsQ$_1$

This section derives an analytical expression for tsQ$_1$ given by Eq. 5.7 based on concepts introduced in earlier work (66, 75). To model the mean $\langle k_i \rangle$ and time-shifted covariance $\langle \delta k_i \delta k_{i+1} \rangle$ of photon counts, consider a system of N_{tot} non-interacting fluorescent molecules diffusing in the sample volume V. The fluorescent intensity for a single molecule depends on the position of the molecule relative to the PSF during the sampling time T_S. The integrated intensity W_i of a single molecule with trajectory $\vec{r}'(t')$ over the time interval from iT_S to $(i+1)T_S$ is (75)

$$W_i = \int_{iT_S}^{(i+1)T_S} \lambda PSF\left(\vec{r}'(t')\right) dt' . \qquad 5.18$$

The molecule is equally likely to occupy any location in the sample volume, which is expressed by the probability density $P(\vec{r}') = 1/V$. This leads to the expectation value $\langle W \rangle^{(1)} = \lambda T_S V_{OV}/V$, where $V_{OV} = \int_V PSF(\vec{r}) d\vec{r}$ is the volume of the overlap between the sample and the PSF, or the OV (75). The superscript (1) denotes that the expectation value refers to a single molecule. Because independent molecules contribute equally to the signal, the mean photon count is

$$\langle k_i \rangle = N_{tot} \langle W \rangle^{(1)} = \lambda T_S N , \qquad 5.19$$

where $N = N_{tot} V_{OV}/V$ is the mean number of molecules within the OV.

The time-shifted covariance $\langle \delta k_i \delta k_{i+1} \rangle$ for independent fluorescent molecules is related to the correlation $\langle W_i W_{i+1} \rangle^{(1)}$ of the time-integrated intensity of a single molecule by (68),

$$\langle \delta k_i \delta k_{i+1} \rangle = N_{tot} \langle W_i W_{i+1} \rangle^{(1)} \ . \qquad 5.20$$

By defining $\vec{r}_i' \equiv \vec{r}'(t_i')$ to simplify notation, the expectation value of the time-shifted correlation $\langle W_i W_{i+1} \rangle^{(1)}$ is given by (68)

$$\langle W_i W_{i+1} \rangle^{(1)} = \lambda^2 \int_{(i+1)T_S}^{(i+2)T_S} \int_{iT_S}^{(i+1)T_S} \langle PSF(\vec{r}_1') PSF(\vec{r}_2') \rangle^{(1)} dt_1' dt_2' \qquad 5.21$$

with probability density $P(\vec{r}_1', \vec{r}_2') = \mathrm{Pro}(\vec{r}_1' | \vec{r}_2') P(\vec{r}_1')$. The propagator Pro is given by

$$\mathrm{Pro}(\vec{r}_1' | \vec{r}_2') = \frac{1}{(4\pi D |t_2' - t_1'|)^{d/2}} Exp\left(- \frac{\|\vec{r}_2' - \vec{r}_1'\|^2}{4D|t_2' - t_1'|} \right) , \qquad 5.22$$

where d is the spatial dimension and *D* is the diffusion coefficient (75). For a stationary signal the correlation only depends on the time difference $t_2' - t_1'$. Since $t_2' > t_1'$,

$$\langle PSF(\vec{r}_1') PSF(\vec{r}_2') \rangle^{(1)} = \gamma_2 \frac{V_{OV}}{V} G_2(t_2' - t_1') \ , \qquad 5.23$$

where G_2 is the second-order normalized correlation function (75),

$$G_2(t_2' - t_1') = \int_V \int_V PSF(\vec{r}_2') \mathrm{Pro}(\vec{r}_2' | \vec{r}_1') PSF(\vec{r}_1') d\vec{r}_1' d\vec{r}_2' \Big/ \int_V PSF^2(\vec{r}') d\vec{r}' \ . \qquad 5.24$$

With these definitions Eq. 5.21 is rewritten as

$$\begin{aligned}\left\langle W_i W_{i+1}\right\rangle^{(1)} &= \gamma_2\lambda^2\frac{V_{OV}}{V}\int_{T_S}^{2T_S}\int_0^{T_S} G_2(t_2'-t_1')dt_1'dt_2' \\ &= \gamma_2\lambda^2\frac{V_{OV}}{V}\int_0^{T_S}\int_0^{T_S} G_2(T_S+t_2-t_1)dt_1dt_2\end{aligned} \quad . \qquad 5.25$$

We define the time-shifted binning function tsB_2 as

$$tsB_2\left(T_S\right) \equiv \int_0^{T_S}\int_0^{T_S} G_2(T_S+t_2-t_1)dt_1dt_2 \ , \qquad 5.26$$

which describes the under-sampling corrected correlation due to time-shifting by T_S. This equation is simplified by the variable substitution $\tau = t_2 - t_1$, resulting in

$$tsB_2\left(T_S\right) = \int_{-T_S}^{T_S}\left(T_S - |\tau|\right)G_2(T_S+\tau)d\tau \ , \qquad 5.27$$

which is suitable for deriving the expression of $tsB_2\left(T_S\right)$ for specific correlation functions G_2.

For the three-dimensional Gaussian (3DG) PSF, $tsB_2\left(T_S\right)$ is given by

$$\begin{aligned}tsB_2(T_S)/\tau_D^2 &= -2r^2 + 4r\sqrt{r^2+t_S} - 2r\sqrt{r^2+2t_S} \\ &+ \frac{r}{s}\left\{2\left(1+t_S\right)\ln\frac{\sqrt{r^2+t_S}+s}{\sqrt{r^2+t_S}-s} + 2t_S\ln\frac{\sqrt{r^2+2t_S}-s}{\sqrt{r^2+2t_S}+s} + \ln\frac{\left(s-r\right)\left(s-\sqrt{r^2+2t_S}\right)}{\left(s+r\right)\left(s+\sqrt{r^2+2t_S}\right)}\right\},\end{aligned} \qquad 5.28$$

where $r^2 = z_0^2/w_0^2$, $s = \sqrt{r^2-1}$, and $t_S = T_S/\tau_D$. For the two-dimensional Gaussian (2DG) PSF, the formula for $tsB_2\left(T_S\right)$ is (68)

$$tsB_2\left(T_S\right)/\tau_D^2 = \left(1+2t_S\right)\ln\left(1+2t_S\right) - 2\left(1+t_S\right)\ln\left(1+t_S\right). \qquad 5.29$$

Note that tsB_2 is related to the regular binning function,

$$tsB_2\left(T_S\right)=\frac{B_2\left(2T_S\right)-2B_2\left(T_S\right)}{2}, \qquad 5.30$$

as can be verified by direct substitution of the integral expression defining B_2 (75). Utilizing Eqs. 5.20, 5.25, and 5.26 yields a formula for the time-shifted covariance of the photon counts,

$$\left\langle \delta k_i \delta k_{i+1} \right\rangle = Q_0 \left\langle k_i \right\rangle \times \frac{tsB_2\left(T_S\right)}{T_S^2} \ . \qquad 5.31$$

Finally, by combining Eqs. 5.5 and 5.31 we arrive at the general expression for the time-shifted Q-parameter of Eq. 5.7.

5.4.5 Estimator mean of tsMSQ

In general, the covariance of photon counts k_i and k_j with $i<j$ is expressed by (75)

$$\left\langle \delta k_i \delta k_j \right\rangle = \gamma_2 \lambda^2 N \times \frac{B_2\left((j-i)T_S+T_S\right)+B_2\left((j-i)T_S-T_S\right)-2B_2\left((j-i)T_S\right)}{2} . \qquad 5.32$$

The double sum $\sum_{i=1}^{M-1}\sum_{j=1}^{M-1}\left\langle \delta k_i \delta k_{j+1} \right\rangle$ is broken into three parts. First, for $i=j+1$ there are *M*-2 terms, so we obtain

$$\sum_{i=2}^{M-2}\left\langle \delta k_i^{\,2}\right\rangle = \left(M-2\right)\left(\gamma_2\lambda^2 NB_2\left(T_S\right)+\lambda T_S N\right)\ , \tag{5.33}$$

where we used $\left\langle \delta k_i^2\right\rangle - \left\langle k_i\right\rangle = \gamma_2\lambda^2 NB_2\left(T_s\right)$ and $\left\langle k_i\right\rangle = \lambda N T_S$ (66, 75). Second, for $i > j+1$, the substitution variable $k = i-j-1$ runs from 1 to *M*-3, and for each *k* there are *M*-*k*-2 terms. Summation of these terms leads to

$$\sum_{i=3}^{M-1}\sum_{j+1=2}^{i>j+1}\left\langle \delta k_{j+1}\delta k_i\right\rangle = \gamma_2\lambda^2 N\,\frac{B_2\left(\left(M-2\right)T_S\right)-\left(M-2\right)B_2\left(T_S\right)}{2}. \tag{5.34}$$

Finally, for $i < j+1$, $k = j+1-i$ runs from 1 to *M* – 1, and for each *k* there are *M* – *k* terms. The partial sum evaluates to

$$\sum_{j+1=2}^{M}\sum_{i=1}^{i<j+1}\left\langle \delta k_i\delta k_{j+1}\right\rangle = \gamma_2\lambda^2 N\,\frac{B_2\left(MT_S\right)-MB_2\left(T_S\right)}{2}\ . \tag{5.35}$$

The double sum is determined by adding the partial sums for each of the three conditions,

$$\sum_{i=1}^{M-1}\sum_{j=1}^{M-1}\left\langle \delta k_i\delta k_{j+1}\right\rangle = \left(M-2\right)\lambda N T_S + \gamma_2\lambda^2 NtsC_2\left(MT_S\right), \tag{5.36}$$

where the first term comes from the shot-noise contributions in Eq. 5.33, while the second term describes the effect of correlations between data points within a segment. The mean of the estimator $\left\langle tsQ_1\left(T\right)\right\rangle$ is determined by inserting Eq. 5.36 into Eq. 5.12 and using the relation $T = MT_S$.

5.4.6 tsMSQ for an exponential correlation process

An exponential correlation $G_{2,\exp}(\tau) = e^{-|\tau|/T_0}$ with T_0 as the characteristic time is needed to account for the additional intensity fluctuation process experienced by proteins residing within the lumen of the NE. This correlation process is caused by undulations in the gap size separating the INM and ONM (53). The oversampled, time-shifted Q-factor, is given in accordance with Eq. 5.6 by

$$tsQ_{1,\exp} = A_0 G_{2,\exp}(T_S) \ , \tag{5.37}$$

where we used A_0 instead of Q_0 for the amplitude to emphasize the difference in the physical origin of the fluctuation process. Thus, following the same steps as presented in Sections 5.4.3 and 5.4.4,

$$tsQ_{1,\exp} = A_0 \frac{tsB_{2,\exp}(T_S)}{T_S^2} \tag{5.38}$$

is the general form of the time-shifted Q-factor for an exponential correlation process that is valid for all sampling times with $tsB_{2,\exp}$ denoting the time-shifted binning function of second order for an exponential correlation, which is derived below.

The undulation process is slow enough to introduce estimator bias into MSQ (53). Here we derive the corresponding expression for the time-shifted estimator for the exponential correlation process,

$$\left\langle tsQ_1(T) \right\rangle_{\exp} = \frac{\left\langle \delta k_i \delta k_{i+1} \right\rangle_{\exp}}{\left\langle k_i \right\rangle} - \frac{1}{(M-1)^2} \sum_{i=1}^{M-1} \sum_{j=1}^{M-1} \frac{\left\langle \delta k_i \delta k_{j+1} \right\rangle_{\exp}}{\left\langle k_i \right\rangle} \ . \tag{5.39}$$

The time-shifted covariance $\left\langle \delta k_i \delta k_j \right\rangle_{\exp}$ for $j > i$ is equal to $A_0 \langle k \rangle G_{2,\exp}\left(\left(j-i\right)T_S\right)$ in the oversampling limit and to $A_0 \langle k \rangle tsB_{2,\exp}\left(\left(j-i\right)T_S\right)$ in general. The time-shifted binning function for the exponential correlation is according to Eq. 5.26

$$tsB_{2,\exp}\left(\left(j-i\right)T_S\right) = \int_{jT_S}^{(j+1)T_S} \int_{iT_S}^{(i+1)T_S} e^{-|t_2'-t_1'|/T_0} dt_1' dt_2' \ , \tag{5.40}$$

which reduces to the analytical expression

$$tsB_{2,\exp}\left(\left(j-i\right)T_S\right) = T_0^2 \left(e^{-(j-i-1)T_S/2T_0} - e^{-(j-i+1)T_S/2T_0}\right)^2 . \tag{5.41}$$

For the special case of $j-i=1$ we obtain $tsB_{2,\exp}\left(T_S\right) = T_0^2\left(1-e^{-T_S/T_0}\right)^2$.

The evaluation of the double sum in Eq. 5.39 closely follows the steps in Section 5.4.4 and yields a bias term due to the exponential correlation of

$$\sum_{i=1}^{M-1}\sum_{j=1}^{M-1}\left\langle \delta k_i \delta k_{j+1} \right\rangle_{\exp} = \left(M-2\right)\langle k \rangle + A_0 \langle k \rangle \frac{tsC_{2,\exp}\left(T\right)}{\left(T-T_S\right)^2} \ , \tag{5.42}$$

where $tsC_{2,\exp}\left(T\right) = \left(B_{2,\exp}\left(T\right) + B_{2,\exp}\left(T-2T_S\right) - 2B_{2,\exp}\left(T_S\right)\right)\Big/2$. The binning function for the exponential correlation is given by (53)

$$B_{2,\exp}\left(T\right) = 2T_0^2\left(-1+\frac{T}{T_0}+e^{-T/T_0}\right) . \tag{5.43}$$

These results provide an analytical expression for $\left\langle tsQ_1\left(T\right)\right\rangle_{\exp}$ of Eq. 5.39, which is converted by Eq. 5.15 into an expression for tsMSQ of an exponential correlation process,

$$tsMSQ_{\exp}(T) = A_0\left(\frac{tsB_{2,\exp}(T_S)}{T_S^2} - \frac{tsC_{2,\exp}(T)}{(T-T_S)^2}\right). \qquad 5.44$$

The FFS signal of diffusing proteins in the lumen of the NE is composed of two independent sources of correlated fluctuations: the correlations due to diffusion and an exponential correlation process caused by membrane undulations. Since both processes are present, the time-shifted covariance for oversampling is given by sum of the individual contributions, $\langle \delta k_i \delta k_{i+1} \rangle_{d+\exp} = Q_0 \langle k_i \rangle G_{2,d}(T_S) + A_0 \langle k_i \rangle G_{2,\exp}(T_S)$, where we used the subscript *d* to identify diffusion. Following the same derivation steps described above we arrive at an analytical expression for tsMSQ,

$$\begin{aligned} tsMSQ_{d+\exp}(T) = Q_0&\left(\frac{tsB_{2,d}(T_S)}{T_S^2} - \frac{tsC_{2,d}(T)}{(T-T_S)^2}\right) \\ &+ A_0\left(\frac{tsB_{2,\exp}(T_S)}{T_S^2} - \frac{tsC_{2,\exp}(T)}{(T-T_S)^2}\right), \end{aligned} \qquad 5.45$$

where the first and second term account for the diffusion and the exponential correlation process, respectively.

5.4.7 Dead-time and afterpulsing effects on MSQ and tsMSQ

For the sake of simplicity, we treat the case of oversampled data with negligible estimator bias ($T \to \infty$). In this limit, the ideal MSQ and tsMSQ approach the asymptotic Q-value, $\mathrm{MSQ} = Q_0$ and $\mathrm{tsMSQ} = Q_0$. Here we consider the first-order or leading correction term due to dead-time and afterpulsing. Quantities biased by afterpulsing and

dead-time are denoted by an asterisk and a prime, respectively. Further details on dead-time and afterpulsing effects on FFS can be found elsewhere (67, 68, 102).

Afterpulsing leads to an increased number of observed counts $\langle k_i \rangle^* = \langle k_i \rangle (1+P)$, where *P* is the probability of the detector to generate a spurious count following a real event (67, 68, 102). The second-order factorial cumulant, $\langle \delta k_i^2 \rangle - \langle k_i \rangle$, is changed to $\left(\langle \delta k_i^2 \rangle - \langle k_i \rangle \right)^* = \left(\langle \delta k_i^2 \rangle - \langle k_i \rangle \right)(1+P)^2 + 2P\langle k_i \rangle$ in the presence of afterpulsing (102), where the factor $2P\langle k \rangle$ is due to the shot noise. Thus, the afterpulsing affected MSQ is given by

$$\mathrm{MSQ}^* = Q_0^* = \frac{\left(\langle \delta k_i^2 \rangle - \langle k_i \rangle \right)^*}{\langle k_i \rangle^*} = Q_0 (1+P) + \frac{2P}{1+P} \approx Q_0 + 2P, \qquad 5.46$$

where we used $P \ll 1$ (*P* is typically on the order of 0.01). In the case of tsMSQ, the correlation of afterpulsing between the two consecutive counts is negligible because the characteristic time of afterpulsing is a few microseconds which is much smaller than the sampling time T_S used in FFS (68). Thus, we treat the consecutive counts k_i and k_{i+1} as independent events with the same *P*, which leads to $\langle \delta k_i \delta k_{i+1} \rangle^* = \langle \delta k_i \delta k_{i+1} \rangle (1+P)^2$ (68, 102), and an afterpulsing affected tsMSQ of

$$\mathrm{tsMSQ}^* = tsQ_1 (1+P) = Q_0 (1+P) \ . \qquad 5.47$$

Dead-time causes the detector to miss photon counts for the dead-time period $\tau_\dagger$ after a detection event (67, 68, 102). The dead-time affected MSQ amplitude relevant for FFS experiments in cells is given by (67, 68)

$$MSQ' = Q_0' \approx Q_0 - 2\delta\langle k_i \rangle \ , \tag{5.48}$$

with the parameter $\delta = \tau_\dagger / T_S$.

We now consider the effect of dead-time on tsMSQ. Two consecutive photon count measurements can be treated as independent experiments since the characteristic dead-time is much smaller than the sampling time. Therefore, the time-shifted covariance in the presence of dead-time is (68)

$$\langle \delta k_i \delta k_{i+1} \rangle' = \gamma_2 \lambda^2 N T_S^2 - \delta\left(4\lambda T_S N \times \gamma_2 \lambda^2 N T_S^2 + 2\gamma_3 \lambda^3 N T_S^3\right) \ , \tag{5.49}$$

which is approximated by

$$\langle \delta k_i \delta k_{i+1} \rangle' = \gamma_2 \lambda^2 N T_S^2 - 4\delta\lambda T_S N \times \gamma_2 \lambda^2 N T_S^2 \ . \tag{5.50}$$

In addition, for a typical FFS experiment the effect of dead-time on the mean count is negligible, $\langle k_i \rangle' \approx \langle k_i \rangle$ (68). Thus, the dead-time affected tsMSQ is given by

$$tsMSQ' \approx tsQ_1 - 4\delta\gamma_2 \lambda^2 N T_S^2 \approx Q_0(1 - 4\delta\langle k_i \rangle) . \tag{5.51}$$

Finally, to account for the combined effect of dead-time and afterpulsing, the correction terms for each effect are summed (67). For MSQ and tsMSQ we obtain

$$MSQ^{*\prime} \approx Q_0 - 2\delta\langle k_i \rangle + 2P \qquad 5.52$$

and

$$tsMSQ^{*\prime} \approx Q_0 \left(1 - 4\delta\langle k \rangle + P\right) \qquad 5.53$$

to first order in δ and *P*. The difference between Eqs. 5.52 and 5.53 is striking as illustrated in Sup. Fig. 5.S2 using typical dead-time and afterpulsing values for our detectors. Note that the predicted influence of the two non-ideal detector effects on tsMSQ is negligible, while its influence on MSQ is significant.

5.5 Results

5.5.1 MSQ in the presence of dead-time and afterpulsing

Our original definition of MSQ includes a term due to shot noise (72), which alters the amplitude of the curve (Fig. 5.S1), but carries no information about the sample. Since this term is identical for all MSQ curves, we introduce here an alternative definition of MSQ given by Eq. 5.3 that removes the shot noise term and simplifies the direct comparison between MSQ curves from different samples.

However, even with this improved definition the application of MSQ to cellular FFS data has to be done cautiously as illustrated by a simple control experiment. MSQ curves were calculated from FFS experiments performed in the cytoplasm of U2OS cells expressing varying levels of EGFP (Fig. 5.2A). Theory (Eq. 5.4) dictates that the MSQ curves only depend on the brightness and diffusion time of EGFP. Thus, the experimental MSQ curves are expected to be independent of the EGFP expression level. However, this

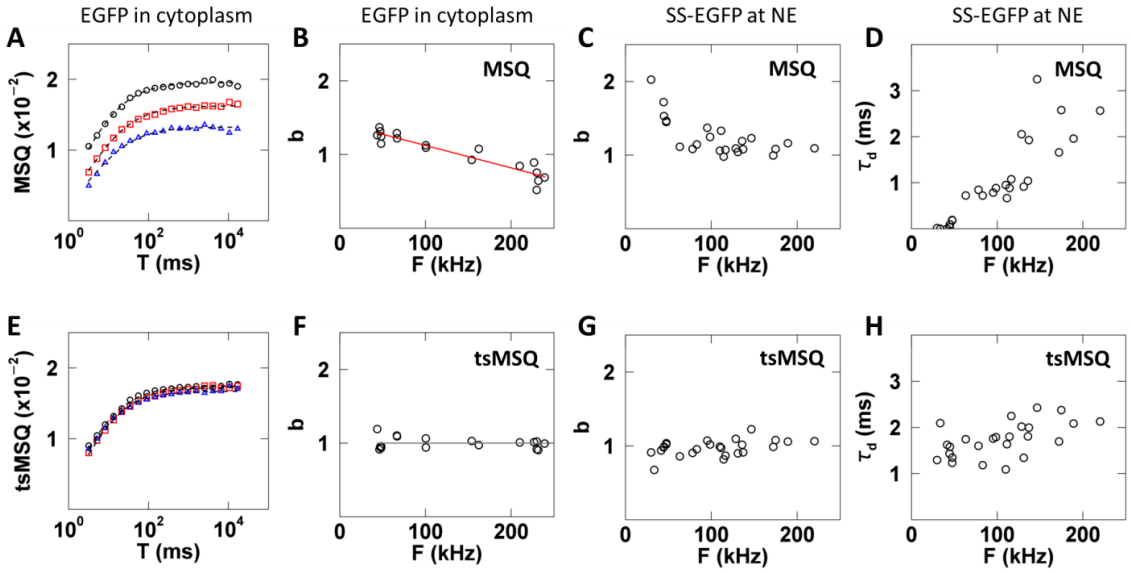

Figure 5.2. Results of MSQ and tsMSQ analysis of FFS data collected for EGFP and SS-EGFP within the cytoplasm and NE, respectively. A – D) Results based on MSQ analysis include biases. E – H) Results based on reanalysis by tsMSQ removes bias. A) MSQ curves from EGFP expressing cells with low (black circles), medium (red squares), and high (blue triangles) intensities with fits (dashed lines). B) Biased *b* from MSQ vs. intensity for EGFP expressing cells (n = 17) with a linear fit (red line) representing first-order non-ideal detector effects (Eq. 5.52). C) Biased brightness from MSQ vs. intensity for SS-EGFP expressing cells (n = 13). D) Biased diffusion time from MSQ for SS-EGFP vs. intensity. E) tsMSQ curves from EGFP expressing cells with low (black circles), medium (red squares) and high (blue triangles) intensities with fits (dashed lines). F) *b* from tsMSQ for EGFP vs. intensity and the average brightness (grey line). G) Brightness from tsMSQ for SS-EGFP vs. intensity. H) Diffusion time from tsMSQ for SS-EGFP vs. intensity.

prediction is not supported by the data, as they differ in both amplitude and shape (Fig. 5.2A). The brightness value determined from fits to the MSQ curve taken from ~15 cells shows a strong dependence on the fluorescence intensity (Fig. 5.2B). The explanation for this discrepancy between experiment and theory is found in non-ideal detector effects, i.e. afterpulsing and dead-time (67, 68). Because the determination of *Q* is biased by non-ideal detector effects (68), the MSQ curve is affected as well (Eq. 5.52).

The presence of an exponential correlation process due to membrane undulations at the NE (53) complicates the situation further as it changes the bias. This effect is demonstrated by FFS data with EGFP targeted to the lumen of the NE by a signal sequence (SS-EGFP), which was analyzed by MSQ fitting without accounting for non-ideal detector effects. Since EGFP is a monomer throughout the concentration range measured, a constant brightness of $b = 1$ was expected. However, we observed an unexpected increase in brightness at low intensities (Fig. 5.2C), erroneously implying the presence of EGFP dimers ($b = 2$) at low intensities. Unlike our observation in the cytoplasm, the biased brightness in the NE is not fully captured by the simple model of Eq. 5.52 (Figs. 5.2B and C). This deviation from the model is most likely due to the presence of an additional fit term to account for the NE membrane undulations. At low intensities the effect of the membrane undulations on the fluorescence signal is minimal and the diffusion and extra fit term combine to result in an increased *b* and low τ_D (Figs. 5.2C and D). As the intensity increases, the diffusion and undulation process grow more distinct, allowing the fitter to resolve both processes which leads to *b* values in the NE that approach the behavior observed in the cytoplasm (Sup. Fig. 5.S3). We also expected that

the diffusion time of SS-EGFP in the lumen would be independent of concentration, yet we observed an apparent increase in the diffusion time with intensity (Fig. 5.5D), which is caused by subtle shape changes in the MSQ curve as a result of non-ideal detector effects.

5.5.2 tsMSQ in the presence of dead-time and afterpulsing

The potential for erroneous interpretations, as shown in Section 5.5.1, emphasizes the importance of including non-ideal detector effects in the MSQ analysis of cellular FFS data. While algorithms for modeling the influence of detector afterpulsing and dead-time exist (68), they are cumbersome, require detailed calibration measurements, and become unreliable once the correction amplitude approaches the amplitude of the signal. Furthermore, proper use of these algorithms is nontrivial and represents a significant barrier for quantifying protein-protein association and mobility within the NE by FFS.

In Section 5.4.3-5.4.4, we derived a modified and improved form of MSQ that overcomes these complications by using the time-shifted Q-value (tsQ_1) instead of Mandel's *Q* factor. It has been previously noted that tsQ_1 is far less susceptible to non-ideal detector effects than *Q* (68), making it a superior choice for application in cells. The tsQ_1 value is calculated by splitting the photon count record into segments of length *T* and then applying Eq. 5.8 to each segment (Fig. 5.1A). Next, the average is calculated and the time-shifted MSQ (tsMSQ) value is calculated from Eq. 5.15. This is repeated for a range of segment times *T* (Fig. 5.1B).

The same FFS data that revealed biased behavior with MSQ were reevaluated using tsMSQ. Unlike the MSQ curves (Fig. 5.2A), the tsMSQ curves determined from cells expressing varying levels of EGFP now coincide (Fig. 5.2E). The brightness recovered by fitting tsMSQ curves taken from cells expressing a range of EGFP levels was independent of fluorescence intensity (Fig. 5.2F). tsMSQ analysis of FFS data collected for SS-EGFP within the NE included an exponential correlation process (Eq. 5.45). This analysis demonstrated that both brightness and diffusion time are independent of intensity (Figs. 5.2G and H). No bias was detected in any of the results obtained by tsMSQ, which agrees with the theoretical prediction described in Section 5.4.7 and demonstrates that tsMSQ is an effective tool for removing bias due to non-ideal detector effects.

5.5.3 Determining goodness-of-fit for tsMSQ

Because the tsMSQ curve is constructed by repeatedly resegmenting the same data set (Fig. 5.3A), the individual data points of the curve are not statistically independent but self-correlated, which confounds traditional goodness-of-fit tests. This problem is immediately evident when fitting a typical tsMSQ curve for data from cytoplasmic EGFP. The computed residuals are significantly smaller in magnitude than expected (Fig. 5.3B) and result in an abnormally low χ_ν^2 of 0.04. These unusual values arise from the self-correlation of the tsMSQ curve, which invalidates the χ^2 test. However, goodness-of-fit tests are crucial for cellular applications of tsMSQ in order to accept and reject fit models as well as to identify uncertainties in the fitted parameters.

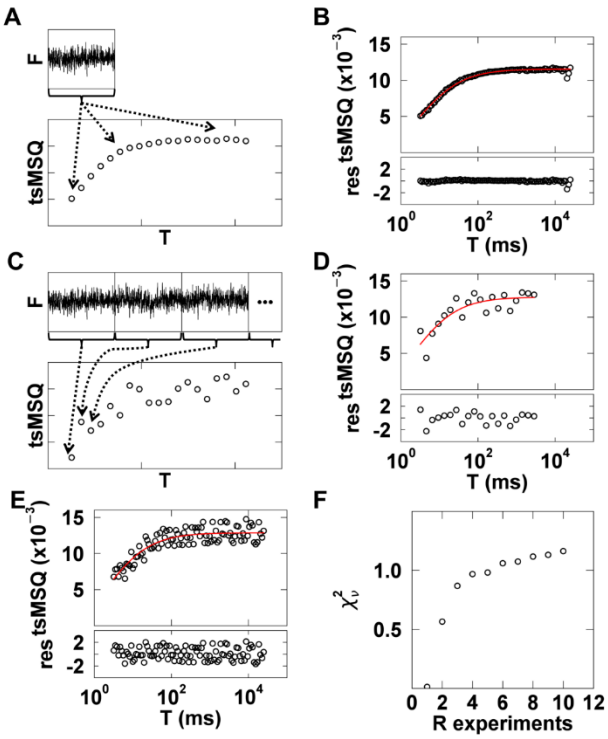

Figure 5.3. χ^2 curve fitting of tsMSQ data. A) Illustration of tsMSQ curve construction using the same experimental data set to determine each tsMSQ point along the curve. B) Experimental tsMSQ curve from cytoplasmic EGFP with fit to Eq. 5.16 (red line) and residuals (bottom panel). C) Illustration of the construction of a decorrelated tsMSQ curve, where a long photon count record is split into separate experiments (top) and each point on the tsMSQ curve is calculated from a unique experiment (bottom). D) Experimental decorrelated tsMSQ curve for cytoplasmic EGFP, constructed as described in the previous panel, with fit (red) and residuals. E) Experimental decorrelated tsMSQ curve for cytoplasmic EGFP, constructed by randomly selecting from 10 experiments for each data point, with fit (red) and residuals. F) The dependence of χ^2_ν on the number of experiments used to construct the decorrelated tsMSQ. Four cells expressing cytoplasmic EGFP were used and the average over the four cells was calculated (black circles).

The most straightforward way to avoid self-correlation is to calculate each data point in the tsMSQ curve from an independently measured data set. This is achieved by performing a longer FFS measurement and dividing the data record into distinct data sets, each of which represents an individual experiment (Fig. 5.3C). The tsMSQ curve can now be constructed with each data point calculated using a unique experimental realization. To do this, we assign each tsMSQ data point to a different experiment, eliminating any self-correlations (Fig. 5.3C). A tsMSQ curve for cytoplasmic EGFP constructed in this way shows an increase in scatter between neighboring data points, indicating the absence of self-correlation (Fig. 5.3D). Moreover, a fit of the curve to a model of a single diffusing species (Eq. 5.16) resulted in residuals that were distributed as expected with a χ_ν^2 value of 0.95 (Fig. 5.3D).

In practice, we have found that not every data point has to originate from a unique experimental realization. For example, we assigned each point of the tsMSQ curve randomly to one out of ten data sets. Because each tsMSQ curve typically has more than 20 data points, some of the points are calculated from the same experiment. Nevertheless, a fit of the tsMSQ curve determined from R = 10 experiments to the single species model resulted in a χ_ν^2 of 1.2 with reasonable residuals (Fig. 5.3E). By changing the random assignment for each tsMSQ data point, we constructed 100 tsMSQ curves from R unique experiments, and the curves were fitted to a single species model. We repeated this process for FFS data taken from four cells expressing EGFP. The χ_ν^2 values of all fits were averaged to identify the trend of χ_ν^2 as a function of the number R of experiments

(Fig. 5.3F). We observed that the averaged χ_ν^2 plateaus for $R \geq 5$ with a reduced chi-square value close to one. For $R < 5$, the averaged χ_ν^2 drops steeply with a value near zero for $R = 1$, which confirms the presence of self-correlations in tsMSQ, resulting in a reduction of χ_ν^2 as observed in Fig. 5.3B. This result provides important guidelines for the experimental construction of tsMSQ curves suitable for goodness-of-fit testing. Since plateauing is observed for $R \geq 5$, the tsMSQ curve should be constructed by randomly selecting data points from five or more experiments to ensure sufficient decorrelation for meaningful error analysis. We have found that 30 to 60 seconds of data is sufficient for a single FFS experiment performed in cells. Thus, the data acquisition time should be at least five times longer, and we suggest a total measurement time between 3 and 5 minutes.

5.5.4 Applying tsMSQ to the luminal domain of SUN2

We applied decorrelated tsMSQ to SS-EGFP-SUN2$^{261\text{-}731}$ to provide a point of comparison with our recent FFS studies of the same protein using the original MSQ, which included corrections for dead-time and afterpulsing (53, 54). FFS data were taken in U2OS cells expressing SS-EGFP-SUN2$^{261\text{-}731}$ as previously described (54). No corrections for non-ideal detector effects were applied to the tsMSQ data in this study. The decorrelated tsMSQ curves were in good agreement with our previously proposed model of two diffusing species (Eq. 5.17), as exemplified by the data presented in Fig. 5.4A. The uncertainty in tsMSQ was determined from the experimental variance of tsQ_1 over all

segments. We obtained reasonable residuals and a χ^2 value of 1.0 for the fit to a two-species model (Fig. 5.4A), which supports the chosen fit model. The fitted brightness values of SS-EGFP-SUN2$^{261\text{-}731}$ collected from a large number of measured cells followed the same trend as previously reported (54). The *b* increased with increasing *N* before reaching a saturating value in the vicinity of $b = 3$ (Fig. 5.4B). This data was fit to a monomer / trimer transition model as previously described with a dissociation coefficient $K = 70 \pm 40$, in agreement with our previously reported value (54) (Fig. 5.4B red line). The fit of the tsMSQ curves identified a fast and a slow species with average diffusion times of ~10 and 300 ms, respectively. The diffusion times remain approximately constant as a function of the protein concentration (Fig. 5.4C), which is consistent with our previous results (53). We hypothesized that these two distinct species represent a fast population of freely diffusing luminal proteins and a slow population associated with the nuclear membrane. While we cannot directly identify the oligomeric state of these two populations, tsMSQ provides information about their relative contributions to the total *Q* value using Eqs. 5.16 and 5.17. Specifically, the ratio of the relative amplitudes $f_i Q_i$ identified by the tsMSQ is instructive. High values of the ratio $f_1 Q_1 / f_2 Q_2$ indicate that the fast species (subscript 1) has a larger amplitude and low values indicate the slow species (subscript 2) dominates. Plotting this ratio vs. the total brightness shows that the fast species dominates at low brightness values while the slow species dominates at high brightness values. This observation is consistent with a model of fast diffusing monomers and slow diffusing trimers (Fig. 5.4D) and is consistent with our previous results (53). Indeed, a

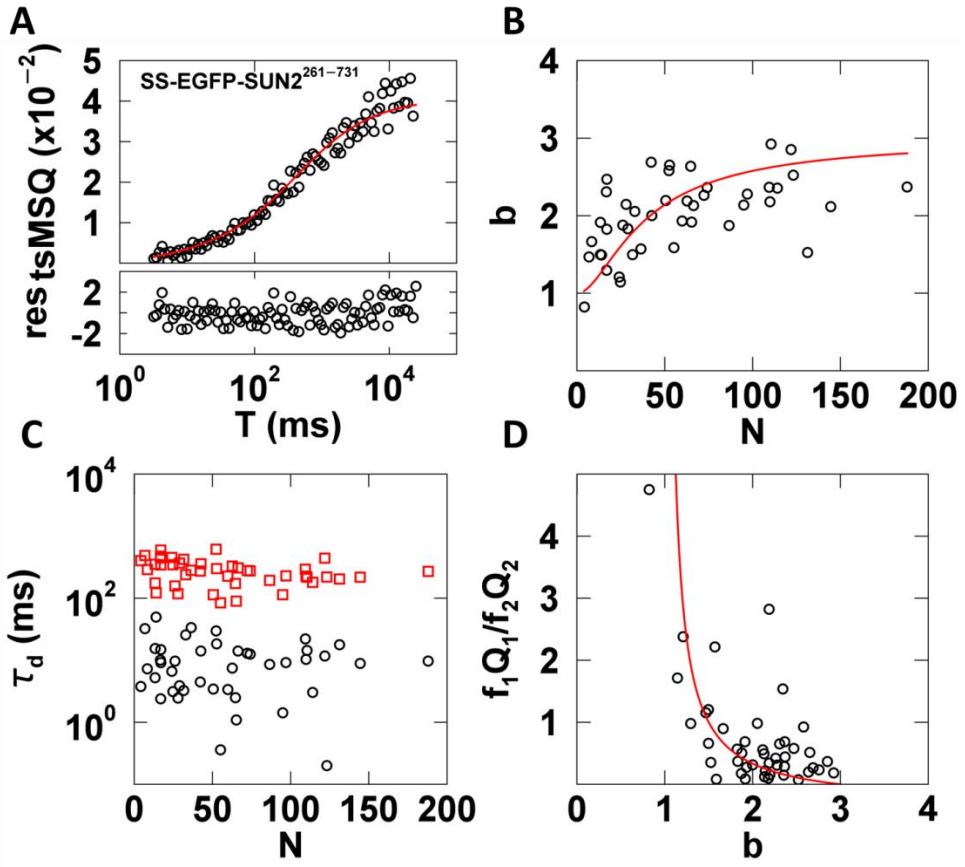

Figure 5.4. tsMSQ analysis of SS-EGFP-SUN2$^{261-731}$ within the NE. A) Decorrelated tsMSQ curve with a fit to a two species diffusion model (Eq. 5.17) with residuals calculated from experimental uncertainty. B) Plot of *b* vs. *N* of SS-EGFP-SUN2$^{261-731}$ in the NE (*n* = 23 cells) with a fit to a monomer / trimer binding model (red curve). C) Diffusion times from tsMSQ fits identify a fast (black circles) and a slow (red squares) diffusing species. D) Plot of relative amplitude of the fast species to the slow species vs. *b* with a model of a transition from fast monomers to slow trimers (red line).

monomer / trimer transition model is in good agreement with the FFS data (Fig. 5.4D). This analysis is completely analogous to the earlier analysis performed with MSQ (53) and confirms that tsMSQ provides the same information content as MSQ. All the results obtained with tsMSQ without corrections for detector artifacts are in agreement with our previous study.

5.5.5 Application of tsMSQ to the luminal domain of SUN3

In addition to SUN1 and SUN2, mammals express three testes-specific SUN proteins: SUN3, SUN4, and SUN5 (103). Based on sequence homology to SUN1 and SUN2, these SUN proteins are also thought to be able to form homo-trimers within the NE (104). To test this hypothesis experimentally in living cells, we applied the same measurement and analysis protocol used for SS-EGFP-SUN2$^{261-731}$ to the SS-EGFP-tagged luminal domain of SUN3 (SS-EGFP-SUN3$^{30-320}$) expressed within the NE. We decided to focus on the SUN3 luminal domain, since full-length NE proteins are frequently too immobile for the type of FFS experiments described here (76). This same strategy was previously used to characterize the oligomeric states of SUN1 and SUN2 in the NE (54). We found that the brightness of SS-EGFP-SUN3$^{30-320}$ increased with increasing N to a brightness of at least 3 without a decrease in slope (Fig. 5.5A). A binding curve with a limiting stoichiometry of 3 would exhibit a clear decrease in slope for $b > 2$. Thus, the lack of an observed decrease in slope in Fig. 5.5A indicates that like SUN1, the oligomerization of SUN3 may not be limited to a trimer (54).

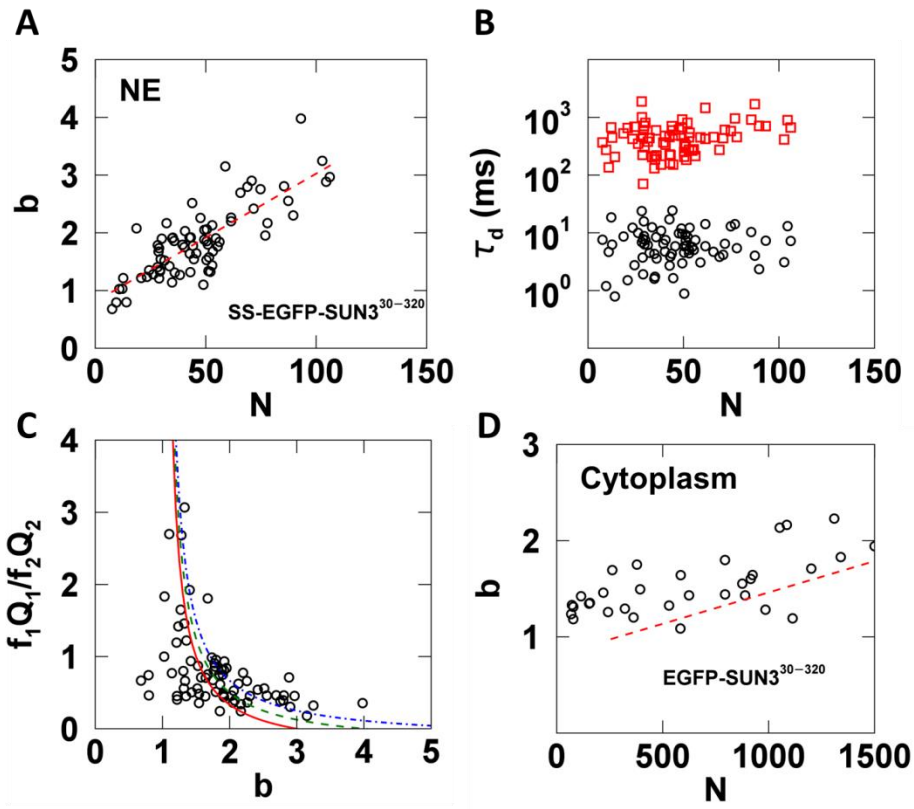

Figure 5.5. tsMSQ analysis of EGFP tagged SUN3$^{30-320}$ within the NE and cytoplasm. A-C) Results from fitting tsMSQ data from SS-EGFP-SUN3$^{30-320}$ within the NE to a two species diffusion model. A) Plot of *b* vs. *N* for SS-EGFP-SUN3$^{30-320}$ measured within the NE (*n* = 41 cells) together with a linear fit to data (red dashed line). B) Diffusion times from two species fits of MSQ curves showing both a fast (black circles) and slow (red squares) component. C) Relative amplitude of the fast component to the slow component vs. brightness. The lines represent a monomer / trimer (solid red), monomer / tetramer (dashed green), and monomer / hexamer (dashed-dotted blue) transition. D) Plot of *b* vs. *N* for EGFP-SUN3$^{30-320}$ within the cytoplasm (*n* = 32 cells) with best fit line from NE (red dashed line) after converting *N* from the NE to its equivalent cytoplasmic value.

The tsMSQ curves for SS-EGFP-SUN3$^{30-320}$ revealed the presence of a fast and a slow species with average diffusion times of ~7 and 500 ms, respectively (Fig. 5.5B). These values are similar to the diffusion times observed for SS-EGFP-SUN2$^{261-731}$ (Fig. 5.4C). The ratio f_1Q_1/f_2Q_2 of the relative amplitudes of the fast and slow species behaved similar to what was observed for SS-EGFP-SUN2$^{261-731}$. Specifically, the fast species dominates at low *b* values, while the slow species dominates at high *b* values (Fig. 5.5C). However, unlike SS-EGFP-SUN2$^{261-731}$, the data cannot be modeled as a simple transition from a fast diffusing monomer to a slow diffusing trimer (Fig. 5.5C, solid red line). While models of a fast diffusing monomer assembling into a slow diffusing tetramer or hexamer (Fig. 5.5C, dashed green or dashed-dotted blue lines respectively) were unable to describe all of the data, these models approached the experimental data at high brightness values, in agreement with our earlier inference that SS-EGFP-SUN3$^{30-320}$ oligomeric states in excess of a trimer exist.

Finally, we performed brightness measurements on the same construct without a signal sequence (EGFP-SUN3$^{30-320}$). This construct is found in the cytoplasm and was measured by FFS as previously described (43). We observed an increase of *b* with *N*, which indicates the presence of oligomerization in the cytoplasm (Fig. 5.5D). Because the OV in the cytoplasm and at the NE are different, a direct comparison of *N* is not meaningful. However, by assuming a NE thickness of 40 nm, the OV in the NE can be estimated and compared to that of the cytoplasm. We have used this approach in the past to convert the occupation number from the NE to the cytoplasm for direct comparison of brightness binding curves (54). Using this method we plotted the best fit line from the NE

data (Fig. 5.5A) after conversion together with the brightness data obtained in the cytoplasm (Fig. 5.5D). While the brightness in the cytoplasm appears to be slightly higher than in the NE at equivalent protein concentrations (Fig. 5.5D, red dashed line), converting *N* in the NE to its cytoplasmic equivalent is only an approximation based on the assumption of a mean thickness of 40 nm for the NE. Increasing this value by only 10 nm would shift the red dashed line resulting in significant overlap with the cytoplasmic data. Because the thickness of the NE is not precisely known, we have to conclude that given current experimental uncertainties no significant difference between the cytoplasmic and NE brightness at equivalent concentrations was found. This conclusion implies that the binding affinity of the SUN3 luminal domain is approximately the same in the NE and the cytoplasmic environment of U2OS cells. In contrast, the binding affinity of the luminal domains of SUN1 and SUN2 was significantly higher in the cytoplasm than in the NE (54). We previously ruled out competition from endogenous SUN1 and SUN2 with their respective EGFP-tagged counterparts as the cause for this change between the behavior in the NE and cytoplasm. Thus, the results of our earlier study suggest the existence of potential regulators in the NE that affect SUN1 and SUN2 oligomerization. Unlike the luminal domains of SUN1 and SUN2, we found that the binding affinities of the SUN3 luminal domain for itself in the cytoplasmic and NE environments were approximately similar. We can rule out the effect of competition from endogenous SUN3 in our experiments, as the expression of SUN3 is limited to the testes in mice (99). While SUN3 can associate with another SUN protein, SUN4 (105), it too is expressed solely within the mammalian male germline (106, 107). It is currently unknown whether or not SUN3 is capable of interacting with SUN1 or SUN2; however, our results described above do not

support the existence of these interactions within the NE of U2OS cells. Therefore, we propose that either the potential regulators of SUN3 oligomerization are only expressed within the NE of male germline cells or that the homo-oligomerization of SUN3 is not subject to regulation.

5.6 Discussion

While the theory underlying tsMSQ presented in this chapter is complex, the application of tsMSQ is relatively straightforward and can be broken down into a few easy steps. FFS data is collected for ~5 minutes in the NE as described in detail in Hennen et al. (64). This photon count record is divided into five data sets, each representing an independent measurement. For a given segment time *T* of the tsMSQ curve, one of these five data sets is chosen at random, segmented into intervals of time *T*, and Eq. 5.8 is used to calculate tsQ_1 on each segment. The average of tsQ_1 over all segments is changed into the tsMSQ value by applying Eq. 5.15 as shown in Fig. 5.4. This procedure is repeated for a range of segment times to construct the decorrelated tsMSQ curve as required for χ^2 curve fitting. Fitting of decorrelated tsMSQ curves to model functions that describe diffusing molecules in the absence (Eqs. 5.16, 5.17) as well as in the presence of an exponential correlation process (Eq. 5.45) should be evaluated using well-established goodness-of-fit criteria to accept or reject each model. We have successfully applied these models to soluble luminal proteins (Eq. 5.45), simple membrane bound proteins (Eq. 5.16), and proteins which transition from luminal to membrane-associated proteins (Eq. 5.17) (53). Evaluation of fit models should be performed on a representative sample

of cells covering the range of expression levels to be measured in order to ensure that a model accurately describes the behavior of a given protein. Upon determining the proper model, fits are used to obtain values for the asymptotic Q-factor, Q_0, and diffusion time, τ_d. Using $b = Q_0 / (\gamma_2 \lambda_{EGFP})$ and $N = \langle F \rangle / \lambda_{EGFP}$, plots of *b* vs. *N* can be constructed and analyzed to determine the extent of oligomerization, if any, as in Figs. 5.4B and 5.5A. Further information may be obtained from the tsMSQ fit results, as described here and in Hennen et al (53).

The membrane undulations at the NE pose a special challenge for conventional FFS analysis methods. To illustrate the problem let us divide commonly used point FFS techniques into two groups. The first group, which includes PCH and moment analysis, exploits the amplitude of fluctuations, but ignores their temporal correlation (100). Because temporal information is discarded, it is impossible to differentiate fluorescence fluctuations caused by the NE membrane undulations from those caused by molecular diffusion, which precludes identification of molecular brightness. The second group, which includes FCS and TIFCA, utilizes temporal correlations (100). However, these analysis methods use algorithms that segment the data using a predetermined length, which significantly improves the robustness of FFS analysis of cellular data (44). We found that the finite and fixed segment length leads to biased results in the presence of slow processes such as the NE membrane undulations (53). This was a key observation that prompted us to use MSQ instead of FCS for the analysis of fluorescence fluctuation data from NE proteins. MSQ and tsMSQ overcome these challenges by providing dynamic information while accounting for and visualizing the effect of data segment length on the analysis.

The application of MSQ to the NE of living cells has proven to be a powerful tool, revealing insights into the dynamics and oligomerization of proteins within the NE as well as the NE itself (53, 54). tsMSQ produces the same results as the original MSQ method and is considerably easier to apply, as it avoids the background work of characterizing detectors. Incorrect application of non-ideal detector corrections is a potential source of error in interpreting FFS results, which is avoided by using tsMSQ. Because of these advantages, we recommend the use of tsMSQ over MSQ. Furthermore, the procedure for producing decorrelated tsMSQ curves provides a strong foundation for future investigations of NE proteins by FFS. These developments serve to strengthen the results and simplify the analysis of FFS data obtained within the NE of living cells.

While it is possible to apply tsMSQ in other cellular compartments, the NE of mammalian cells is the only environment where we have found it to be particularly advantageous over more established techniques (53). The standard analysis techniques have proven to be successful when applied to measurements performed in the nucleoplasm, cytoplasm, and at the plasma membrane (42–44). It is in the presence of a slow fluctuation process not caused by the motion of single molecules, as is the case for the NE membrane undulations, where tsMSQ becomes necessary.

Our observation that both SS-EGFP-SUN2$^{261\text{-}731}$ and SS-EGFP-SUN3$^{30\text{-}320}$ exist as two distinct diffusing species suggests the presence of a luminal and a membrane-associated population. The data obtained for SS-EGFP-SUN2$^{261\text{-}731}$ can be modeled as a simple transition of free monomers to membrane-associated trimers (Fig. 5.6A). While the data from SS-EGFP-SUN3$^{30\text{-}320}$ do not follow a monomer / *n*-mer

transition model, it is clear that there is a transition from low to high oligomeric states as the membrane-associated population increases. These results may indicate the transition of the SUN3 luminal domain from soluble monomers to membrane-associated oligomers with different assembly states, with a strong indication of the presence of oligomers larger than a trimer (Fig. 5.6B).

While the developments described in this chapter represent a significant improvement to the application of FFS within the NE, there is substantial need for further advances. For example, the current use of truncated luminal domains of nuclear membrane proteins is not optimal, as they might not accurately represent the oligomerization behavior and dynamics of the FL proteins. Future developments in combining tsMSQ with an imaging-based approach (108–110) would potentially allow for measurement of relatively immobile proteins, including FL SUN proteins. In addition, SUN proteins only represent one part of the proposed LINC complexes with nesprins being a necessary binding partner. Although we performed FFS on the soluble KASH peptide of nesprin-2 within the perinuclear space (53), it is the interaction of the KASH peptides of nesprins with the SUN domains of SUN proteins which is required for forming functional LINC complexes (81, 111). Investigating the SUN-KASH interaction and its regulation within the NE of living cells requires the use of two differently colored FPs to label each protein species in order to identify their association by dual-color FFS (89). In addition, quantifying the interaction between different SUN proteins via dual-color FFS is of considerable interest, as it will enable testing of previously proposed models of SUN protein hetero-oligomerization (86, 112). Thus, the development of dual-color tsMSQ will be essential for addressing these questions in future studies.

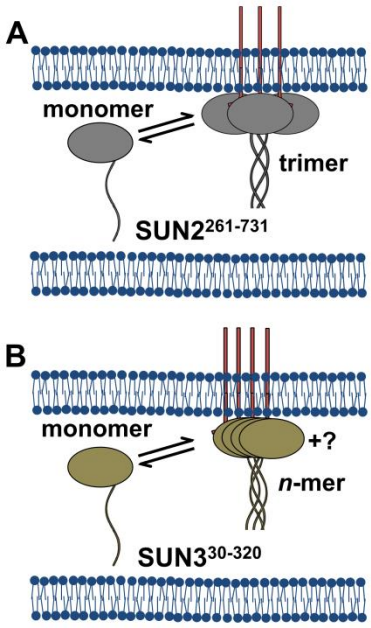

Figure 5.6. Working models for the observed behavior of the luminal domains of SUN2 and SUN3. A) SS-EGFP-SUN2$^{261-731}$ (grey) exists as either freely diffusing, luminal monomers or membrane-associated trimers, potentially due to interactions with endogenous nesprins (red lines) at the ONM. B) SS-EGFP-SUN3$^{30-320}$ (tan) exists as either freely diffusing, luminal monomers or membrane-associated oligomers. The size of these membrane associated oligomers has yet to be determined.

Chapter 6. Identifying Hetero-Protein Complexes in the Nuclear Envelope

This chapter extends the previous work to dual-color measurements, thereby allowing us to quantify hetero-protein interactions. The theory behind dual-color tsMSQ was developed by Dr. Kwang-Ho Hur. My contributions entailed designing and performing the experiments and performing data analysis.

6.1 Introduction

Fluorescence fluctuation spectroscopy (FFS) refers to a collection of related biophysical techniques that exploit the stochastic intensity of fluorescently labeled biomolecules passing through a small observation volume (OV) created by confocal or two-photon microscopy (113). The primary parameters accessible by FFS are the concentration, mobility, and oligomeric state of the labeled biomolecule (113). While the original analysis of FFS results was based on the autocorrelation function (ACF), many other methods have been introduced over the years, each with its own strengths and weaknesses (58, 66, 108, 114). An important advance in FFS was the introduction of dual-color (DC) FFS for identifying interactions between two species of biomolecules labeled with spectrally distinct fluorophores (115). In DC FFS, the emission of the fluorophores is separated by color into two detection channels. Heterotypic interactions between the two species lead to synchronized temporal fluctuations in both channels, which are recognized by the cross-correlation function (CCF) of the two detected signals. In addition to CCF,

other analysis techniques have been introduced for quantifying hetero-species interactions from FFS experiments (63, 101).

Because FFS is an equilibrium technique that passively observes fluctuations, it is well suited for applications in live cells (44). Cellular proteins can be conveniently labeled by genetic tagging with one of the many available fluorescent proteins (116). Brightness, which characterizes the intrinsic fluorescence intensity of a molecule, is an important FFS parameter because it contains information about the stoichiometry of fluorescently-tagged protein complexes (117). For example, the brightness of monomeric proteins tagged with EGFP will increase upon their association into homo-oligomers, as each protein complex contains several fluorescent labels. This concept has been generalized to include differently colored fluorophores to characterize hetero-protein complexes (63). While FFS brightness analysis has been successfully used to quantify protein-protein interactions within the cytoplasm, nucleoplasm, and at the plasma membrane (42–44), its extension to the nuclear envelope (NE) has proven challenging (53).

The NE consists of an inner and outer nuclear membrane (INM and ONM, respectively) separated by a ~40 nm thick fluid layer, known as the lumen or perinuclear space. Although the NE has been identified as a critical signaling node of the cell (118), a mechanistic understanding of how these hetero-protein complexes assemble remains limited due to the lack of quantitative biophysical techniques suitable for use in the NE. To address this challenge, we explored the use of single-color (SC) FFS for characterizing homo-protein association in the NE of living cells (53, 54). We found that slow undulations of the nuclear membranes give rise to local volume fluctuations that are not properly accounted for by conventional FFS methods (53). We overcame this obstacle by

introducing mean-segmented Q (MSQ) analysis as well as time-shifted MSQ (tsMSQ), a significantly improved version of MSQ (53, 55, 72). We demonstrated that these techniques successfully identify the mobility and homo-oligomeric state of NE proteins (53).

This study extends FFS analysis of hetero-protein complexes to the NE by introducing DC tsMSQ, a generalized form of regular tsMSQ. The DC tsMSQ framework includes hetero-species partitioning (HSP) (63) and effectively eliminates complications due to spectral crosstalk in the characterization of heterotypic protein interactions as verified by control experiments. Our study demonstrates that local volume fluctuations of the NE are a significant challenge for conventional DC FFS analysis, which prompted the development of DC tsMSQ. We first verified the foundation of DC tsMSQ using pairs of interacting and non-interacting proteins measured in the cytoplasm and in the NE. In addition, control experiments using both luminal and nuclear membrane-associated proteins were conducted to illustrate the influence of the nuclear membrane undulations on DC tsMSQ analysis.

To demonstrate the power of DC tsMSQ, we applied it towards studying the assembly of the linker of nucleoskeleton and cytoskeleton (LINC) complex in the NE of living cells. This NE-spanning molecular bridge mediates mechanical force transmission into the nucleoplasm and is required for several fundamental cellular processes including cell division, DNA damage repair, meiotic chromosome pairing, mechano-regulation of gene expression, and nuclear positioning (81)**.** The LINC complex is formed by a direct transluminal heterotypic interaction between the cytoskeletal-binding ONM *Klarsicht*/*ANC-1*/*SYNE* homology (KASH) proteins and the nuclear lamina-binding INM *Sad1*/*UNC-84*

(SUN) proteins (119). Previous *in vitro* biochemical and structural studies revealed that the luminal domain of SUN2 homo-trimerizes and that the luminal domain of nesprin-2 binds in the grooves formed at the interface of two adjacent interacting SUN2 monomers (49, 120). Thus, the homo-trimerization of the SUN2 luminal domain enables the recruitment of 3 nesprin-2 luminal domains resulting in the assembly of a SUN2-nesprin-2 hetero-hexamer. We recently succeeded in directly measuring the homo-trimerization of the EGFP-tagged luminal domain of SUN2 in the NE of living cells by SC FFS (54). Here we utilize DC tsMSQ to directly observe the heterotypic interactions formed between the EGFP-tagged SUN2 luminal domain and an mCherry-tagged construct that encodes the last three spectrin-like repeats (SRs), the transmembrane domain, and the luminal KASH peptide of nesprin-2 in the NE. This work establishes the theoretical and practical framework necessary for future quantitative studies of LINC complex assembly within the NE of living cells as well as the characterization of additional heterotypic interactions between proteins within this relatively unexplored subcellular environment.

6.2 Material and Methods

6.2.1 Experimental setup

The experimental setup used in this chapter is nearly identical to the one described in Section 2.2. A dichroic mirror centered at 580 nm (FF580-FDi01; Semrock, Rochester, NY) was used to split the emission path into two channels. An additional 84 nm wide bandpass filter centered at 510 nm (FF01-510/84; Semrock) was placed before the green channel to remove any reflected fluorescence from mCherry (63).

6.2.2 Measurement procedure

EGFP (G) calibration measurements were performed in the cytoplasm of EGFP-expressing U2OS cells as previously described (43, 44) to obtain its brightness in both the green and red channels ($\lambda_{g,G}$ and $\lambda_{r,G}$, respectively). Additional calibration measurements were performed in the cytoplasm of U2OS cells expressing EGFP-RARLBD-mCherry (63) to obtain the brightness $\lambda_{r,Ch}$ of mCherry (Ch) in the red channel, which accounts for the previously described two-state brightness of this fluorescent protein (63). These calibration values were then converted to *Q* values for the NE using $Q = \gamma_2 \lambda T_S$ where γ_2 is the shape factor for a 2D Gaussian point spread function and T_S is the sampling time (43, 68). Measurements in the NE were performed as previously described (53, 64) by first using epifluorescence to identify cells expressing the relevant constructs. FFS data were then acquired by taking z-scans through the nucleus which were analyzed as previously described (53, 73). Next, the two-photon beam was focused on the ventral NE followed by the dorsal NE and ~60 seconds of intensity fluctuation data were obtained at each location. These data were analyzed as described in Section 6.3 in order to obtain the normalized HSP brightness vector $\mathbf{b} = \left(b_g, b_r\right)$ (63, 121). The normalized brightness values were corrected for two-state brightness and fluorescence resonance energy transfer (FRET) as previously described (63, 122). The average number of molecules in the observation volume was determined by $N_g = \left\langle F_g \right\rangle / \lambda_{g,G}$ and

$N_r = \left(\langle F_r \rangle - f_{ct} * \langle F_g \rangle \right) / \lambda_{r,Ch}$, where f_{ct} is the spectral crosstalk of EGFP given by $Q_{r,G} / Q_{g,G}$ (63).

6.2.3 Sample preparation

The sample preparation used in this chapter is the same as described in Section 3.2.2.

6.2.4 Reagents

Restriction enzymes (REs) were either purchased from New England Biolabs (NEB, Ipswich, MA) or Promega (Madison, WI). Calf Intestinal Phosphatase, Phusion DNA polymerase, T4 DNA ligase, and T4 polynucleotide kinase were also purchased from NEB. All other chemicals were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise specified. The Wizard SV Gel and PCR Clean-Up System were purchased from Promega while the GeneJet Plasmid Midiprep Kit was purchased from ThermoFisher Scientific (Waltham, MA).

6.2.5 DNA constructs

The generation of SS-EGFP, SS-EGFP-torsinA$^{NTD\text{-}2xLeu}$, SS-EGFP-SUN2$^{595\text{-}731}$, and SS-EGFP-SUN2$^{261\text{-}731}$ constructs were described previously (53, 54). The generation of SS-mCherry-KDEL, the SUN domain-linked SS-EGFP-SUN2$^{595\text{-}731}$-mCherry (SS-

EGFP-SL-mCherry), and mCherry-SR-KASH2 constructs are described in Section 6.6.6 of the Supplemental Materials.

6.3 Theory

6.3.1 Dual-channel tsMSQ$_{\sigma\rho}$

The single-channel tsMSQ was generalized to dual-channel tsMSQ$_{\sigma\rho}$ (Section 6.6.1 of the Supplemental Materials) with the subscripts σ and ρ specifying the detection channel. A single diffusing species is described by

$$\mathrm{tsMSQ}_{\sigma\rho,\mathrm{D}}\left(T;Q_\rho,\tau_D\right)=Q_\rho\,\mathrm{tsF}_\mathrm{D}\left(T;\tau_D\right)\ , \tag{6.1}$$

while the exponential correlation process caused by the local volume fluctuations of the NE is given by

$$\mathrm{tsMSQ}_{\sigma\rho,\mathrm{E}}\left(T;A_\rho,\tau_0\right)=A_\rho\,\mathrm{tsF}_\mathrm{E}\left(T;\tau_0\right)\ , \tag{6.2}$$

as derived in Section 6.6.2 of the Supplemental Materials. A sample consisting of a single diffusing species in the lumen of the NE experiences local volume fluctuations and is described by the addition of Eqs. 6.1 and 6.2 (53),

$$\mathrm{tsMSQ}_{\sigma\rho,\mathrm{DE}}\left(T\right)=Q_\rho\,\mathrm{tsF}_\mathrm{D}\left(T;\tau_D\right)+A_\rho\,\mathrm{tsF}_\mathrm{E}(T;\tau_0)\ . \tag{6.3}$$

The time-shifted correlation functions tsF_D and tsF_E are defined in terms of second-order binning functions $\mathrm{B}_{2,\mathrm{D}}\left(T\right)$ and $\mathrm{B}_{2,\mathrm{E}}\left(T\right)$ (53, 55, 66, 75). Explicit formulas are found in Section 6.6.3 of the Supplemental Materials. Since Eqs. 6.1 and 6.2 must reproduce the

single channel case for $\sigma = \rho$, the amplitudes Q_ρ and A_ρ are given by $Q_\rho = \gamma_2 \lambda_\rho T_S$ and $A_\rho = c^2 \lambda_\rho T_S N$, respectively (53, 55). The brightness of the fluorescent molecule in the ρ channel is λ_ρ, the number of molecules in the observation volume is N, and the shape factor of the observation volume is γ_2. The factor c is determined by the fluctuations in the gap distance h separating the INM and ONM, $c = \sqrt{\langle \delta h^2 \rangle} / \langle h \rangle$ (53). The diffusion time and the characteristic time of the volume fluctuations are given by τ_D and τ_0, respectively.

The two detection channels used for DC FFS are labeled as green (g) and red (r). We define the DC tsMSQ function as

$$\mathbf{tsMSQ}(T) = \begin{bmatrix} \mathrm{tsMSQ_{gg}}(T) \\ \mathrm{tsMSQ_{gr}}(T) \end{bmatrix}. \qquad 6.4$$

We refer to $\mathrm{tsMSQ_{gg}}$ as the autocorrelation tsMSQ of the green detection channel and $\mathrm{tsMSQ_{gr}}$ as the cross-correlation tsMSQ of the green and red channel. Thus, DC tsMSQ of a diffusing species is described by

$$\mathbf{tsMSQ_D}(T) = \bar{\mathbf{Q}}\,\mathrm{tsF_D}(T;\tau_D) \qquad 6.5$$

and in the presence of volume fluctuations the exponential correlation process

$$\mathbf{tsMSQ_E}(T) = \mathbf{A}\,\mathrm{tsF_E}(T,\tau_0) \qquad 6.6$$

is added, resulting in

$$\mathbf{tsMSQ_{DE}}(T) = \mathbf{tsMSQ_{D}}(T) + \mathbf{tsMSQ_{E}}(T) \ . \qquad 6.7$$

DC tsMSQ reduces to the same functional form as SC tsMSQ but with vector amplitudes $\bar{\mathbf{Q}} = (\bar{Q}_g, \bar{Q}_r)$ and $\mathbf{A} = (A_g, A_r)$. Finally, for a fit to a model with *S* diffusing species the **tsMSQ$_D$** term is replaced by

$$\mathbf{tsMSQ_{D}}(T) = \sum_{i=1}^{S} f_i \overline{\mathbf{Q}_i}\, \mathrm{tsF_D}(T; \tau_{D,i}) \ , \qquad 6.8$$

where $\overline{\mathbf{Q}_i}$ and $\tau_{D,i}$ are the Q-vector and diffusion time of the *i*th species, respectively. The fractional intensity f_i of the *i*th species is defined by the ratio of the green-channel intensity of the *i*th species to the total intensity $\langle F_g \rangle = \sum_{j=1}^{S} \langle F_{g,j} \rangle$ of the green channel, $f_i = \langle F_{g,i} \rangle / \langle F_g \rangle$, as described in the Section 6.6.4 of the Supplemental Materials. The total Q-vector of the sample is defined by

$$\overline{\mathbf{Q}} = \sum_{i=1}^{S} f_i \overline{\mathbf{Q}_i} \ . \qquad 6.9$$

6.3.2 DC tsMSQ and HSP

We previously demonstrated that the complications in interpreting the results of FFS experiments in the presence of spectral crosstalk are avoided by HSP analysis (63). HSP requires that there is no spectral leakage of the red-emitting mCherry into the green emission channel, which is accomplished by choosing appropriate filters (63). Thus, the

two fluorescent proteins EGFP (G) and mCherry (Ch) are characterized by their respective Q-vectors $\mathbf{Q}_G = \left(Q_{g,G}, Q_{r,G}\right)$ and $\mathbf{Q}_{ch} = \left(0, Q_{r,Ch}\right)$. The Q-vector (Eq. 6.9) determined by DC tsMSQ analysis may be expressed as a linear combination of the Q-vectors of EGFP and mCherry,

$$\bar{\mathbf{Q}} = b_g \mathbf{Q}_G + b_r \mathbf{Q}_{Ch} \ , \qquad 6.10$$

where the coefficients b_g and b_r represent the normalized brightness associated with the corresponding fluorescent proteins (59, 63). The tuple $\left(b_g, b_r\right)$ characterizes the brightness vector of the hetero-species present in the sample. The hetero-species comprises all EGFP-labeled proteins and hetero-protein complexes carrying both EGFP and mCherry. Monomeric or oligomeric complexes that only contain the mCherry label are partitioned out by HSP. For example, a non-interacting monomeric EGFP-labeled species is described by a HSP brightness vector of (1, 0), while a hetero-dimer containing one EGFP and one mCherry label are described by a HSP brightness vector of (1, 1). A HSP brightness vector of (1, *y*) describes an EGFP-labeled protein that on average is associated with *y* mCherry-labeled proteins (63). A graphical representation of these examples is provided in Fig. 6.S1. The HSP brightness vector $\left(b_g, b_r\right)$ is affected by fluorescent labels with dark states and multiple brightness states as well as by FRET, which bias the interpretation (63, 122). These effects were accounted and corrected for as previously described (63, 122). A derivation of HSP for DC tsMSQ is found in Section 6.6.5 of the Supplemental Materials.

6.4 Results

Initial experiments were performed on a non-interacting pair of proteins located in the lumen of the NE (Fig. 6.1A). Specifically, we used SS-EGFP and SS-mCherry-KDEL, which have been found to be monomeric proteins within the NE by SC FFS (Fig. 6.S2) (53). EGFP and mCherry were chosen as labels because this pair has been characterized extensively by DC FFS and have a wide separation in their emission color (63, 116). The signal sequence (SS) of the luminal ATPase torsinA, which is cleaved after protein expression, ensures the presence of EGFP in the lumen of the NE (123). The C-terminus of SS-mCherry was additionally fused to the endoplasmic reticulum retention signal KDEL to ensure its efficient targeting to the lumen. DC FFS data were collected at the NE of U2OS cells co-expressing SS-EGFP and SS-mCherry-KDEL. The ACFs of the fluorescence collected by the green and red detection channels were calculated (G_{gg} and G_{rr}, respectively) as well as the cross-correlation (CCF) G_{gr} of both channels. We observed a positive CCF amplitude (Fig. 6.2), which was not unexpected for a non-interacting pair of fluorescent proteins, since spectral crosstalk leads to a positive CCF component (124). This crosstalk-induced CCF function is predicted by (124)

$$\mathrm{CCF}_{\mathrm{ct}}(\tau) = G_g(\tau)\, f_{\mathrm{ct}} \frac{\langle F_g \rangle}{\langle F_r \rangle} \qquad 6.11$$

with $\langle F_g \rangle$ and $\langle F_r \rangle$ representing the mean fluorescence intensities of the green and red channel, respectively. The crosstalk in our setup is caused by the detection of the long wavelength emission of EGFP in the red detection channel (125), which is characterized

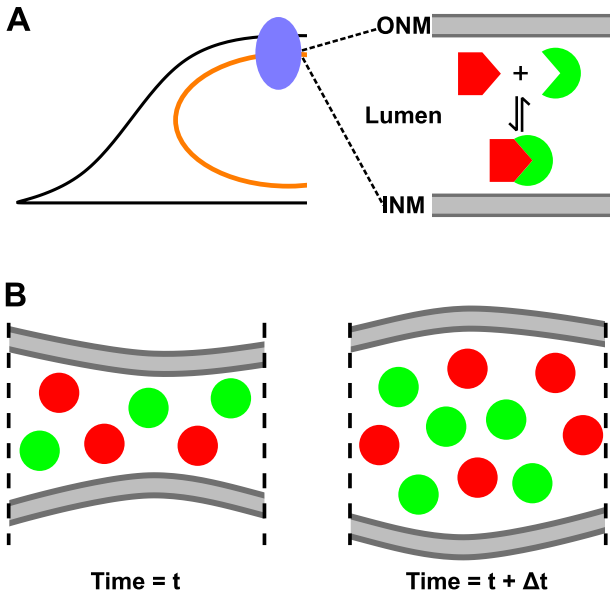

Figure 6.1. FFS at the NE. A) Illustration of a cell expressing potentially interacting green and red fluorescently labeled NE proteins with the two-photon excitation volume (blue oval) focused at its NE (orange circle), which consists of the INM and ONM separated by a ~40 nm thick lumen. B) Illustration of the time-dependent local volume fluctuations caused by nuclear membrane undulations, which give rise to coupled intensity variations of the non-interacting green and red fluorescently labeled proteins.

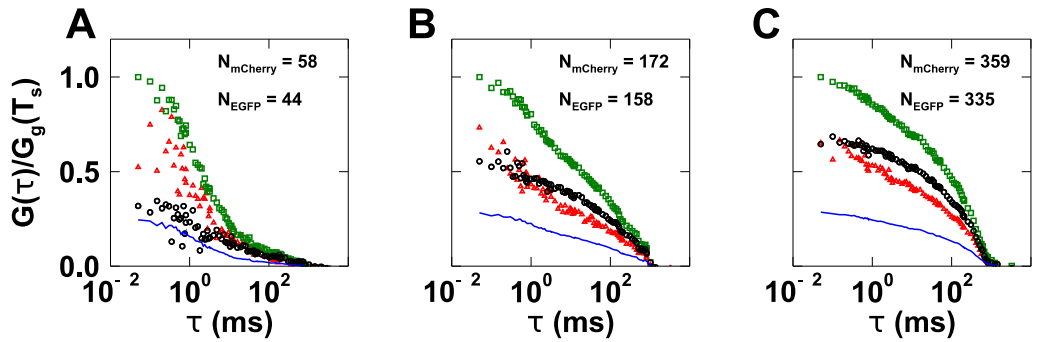

Figure 6.2. ACF and CCF analysis of DC FFS data collected in the NE of cells expressing SS-EGFP and SS-mCherry-KDEL at low (A), medium (B), and high (C) levels of protein expression. The number of proteins in the observation volume is given by N_{EGFP} and $N_{mCherry}$. The red and green channel ACFs are shown with red triangles and green squares, respectively. The CCF is graphed as black circles and the calculated CCF from spectral crosstalk (CCF_{ct}) is shown by the blue line. All correlation curves are divided by the amplitude of the ACF of the green channel, $G_{gg}(T_S)$, with a lag time equal to the sampling time T_S.

by the crosstalk intensity fraction f_{ct}. The observed CCF has to significantly exceed this baseline for a positive identification of heterotypic protein interactions (124). The computed $\mathrm{CCF}_{ct}(\tau)$ (solid blue lines, Fig. 6.2) is not significantly different from the observed CCF for the data taken at low expression levels (Fig. 6.2A), but is significantly below the experimental CCF at medium (Fig. 6.2B) and high expression levels (Fig. 6.2C). This observation is counterintuitive, as it suggests the onset of interactions between EGFP and mCherry-KDEL at higher concentrations. We further noticed a significant change in the shape of the ACF of both channels as well as the CCF with increasing concentration (Fig. 6.2), which was unexpected because the diffusion process is independent of concentration. These changes in the ACF are specific to the NE environment (53) and are not observed in the cytoplasm, as confirmed by a control experiment performed in the cytoplasm of cells expressing EGFP and mCherry. Specifically, measurements performed in these cells showed a concentration-independent shape of the ACF as well as a strong overlap between the computed $\mathrm{CCF}_{ct}(\tau)$ and the CCF over a wide range of expression levels (Fig. 6.S3), indicating the absence of significant interactions between EGFP and mCherry within the cytoplasm.

The unusual behavior of the experimental ACF of FFS data collected within the NE was traced to undulations of the nuclear membranes (53). These membrane undulations introduce local volume fluctuations that modulate the fluorescence intensity signals received from SS-EGFP and SS-mCherry-KDEL (Fig. 6.1B). This additional fluctuation process not only affects the ACF, but also the CCF as the volume fluctuations lead to concomitant intensity variations in both the green and red channels. We previously

demonstrated that the slow membrane undulations lead to biases in ACF analysis (53). To overcome this challenge we replaced ACF with MSQ analysis and more recently with the improved tsMSQ algorithm (53, 55). Both methods have proven to be robust for measuring homotypic protein interactions within the NE of living cells (54, 64). To extend crosscorrelation analysis to DC FFS data obtained in the NE, we generalize single-color to dual-color tsMSQ. Two subscripts are added to distinguish the different tsMSQ functions. The SC tsMSQ from the green channel is identified by tsMSQ_{gg}, while the crosscorrelation tsMSQ between the green and the red channel is described by tsMSQ_{gr} (Fig. 6.3). The autocorrelation tsMSQ_{gg} is calculated from the fluorescence intensity of a single channel using the standard tsMSQ algorithm (Fig. 6.3A). The crosscorrelation tsMSQ_{gr} follows the same procedure but utilizes the intensity traces of both channels (Fig. 6.3B).

We reanalyzed the FFS data of SS-EGFP and SS-mCherry-KDEL taken in the NE (Fig. 6.2) using the experimental tsMSQ_{gg} (green squares) and tsMSQ_{gr} (red circles) curves (Figs. 6.4A and 6.S4A-B). For two non-interacting proteins undergoing simple diffusion in the lumen, the crosscorrelation tsMSQ_{gr} should be entirely determined by the crosstalk f_{ct} from EGFP into the red channel and the membrane undulations of the NE, which contribute a correlation amplitude proportional to the average intensity (53). Thus, the tsMSQ_{gr} curve for two non-interacting luminal proteins is predicted to be

$$\text{tsMSQ}_{\text{gr}}^{(-)}(T) = f_{\text{ct}}\,\text{tsMSQ}_{gg,\text{D}}\left(T;\tau_D,Q_g\right)+\frac{\langle F_r\rangle}{\langle F_g\rangle}\,\text{tsMSQ}_{gg,\text{E}}\left(T;\tau_0,A_g\right), \qquad 6.12$$

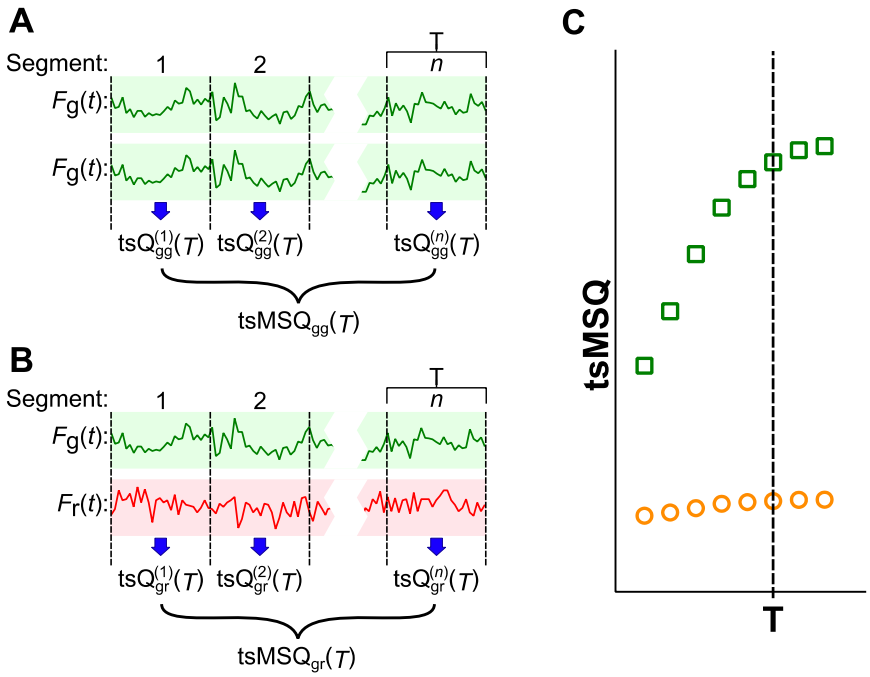

Figure 6.3. Conceptual illustration of the DC tsMSQ algorithm. A) The fluorescence intensity trace of the green channel, F_g, is divided into *n* segments of period *T*. The time-shifted Q value of segment *i*, $\mathrm{tsQ}_{\mathrm{gg}}^{(i)}(T)$, is calculated from F_g and a copy of F_g according to Eq. 6.S3. Averaging over the time-shifted Q values determines the $\mathrm{tsMSQ}_{\mathrm{gg}}$ for segment time *T*. B) The algorithm for calculating the crosscorrelation $\mathrm{tsMSQ}_{\mathrm{gr}}$ follows the same procedure as described in panel A, but replaces the copy of F_g with the intensity of the red channel, F_r. C) The autocorrelation $\mathrm{tsMSQ}_{\mathrm{gg}}$ (green squares) and crosscorrelation $\mathrm{tsMSQ}_{\mathrm{gr}}$ (orange circles) are constructed by repeating the procedures shown in panels A and B for a range of segment times *T*.

where $\mathrm{tsMSQ}_{gg,\mathrm{D}}\left(T;\tau_D,Q_g\right)$ and $\mathrm{tsMSQ}_{gg,\mathrm{E}}\left(T;\tau_0,A_g\right)$ are model functions (Eqs. 6.1 and 6.2) that describe the diffusion and exponential components of $\mathrm{tsMSQ}_{gg,}$, respectively, while $\left\langle F_i \right\rangle$ represents the mean intensity in the i^{th} channel. The green-channel tsMSQ_{gg} was fit to a single species diffusion model with an exponential correlation term (Eq. 6.3) to account for nuclear membrane undulations, enabling the calculation of the predicted crosscorrelation $\mathrm{tsMSQ}_{\mathrm{gr}}^{(-)}$ in the absence of interactions from Eq. 6.12 (blue line, Figs. 6.4A and 6.S4A-B). The predicted $\mathrm{tsMSQ}_{\mathrm{gr}}^{(-)}$ and the experimental tsMSQ_{gr} closely overlap, demonstrating the absence of protein interactions between luminal EGFP and mCherry-KDEL. These results demonstrate that tsMSQ_{gg} accurately predicts the cross-correlation $\mathrm{tsMSQ}_{\mathrm{gr}}^{(-)}$ curve of a non-interacting pair of proteins within the NE.

This initial validation of the theory prompted us to analyze FFS data with DC tsMSQ, which simultaneously describes the green-channel autocorrelation and the cross-correlation tsMSQ (Eq. 6.4), **tsMSQ** = [tsMSQ_{gg}, tsMSQ_{gr}]. We fit both tsMSQ curves to Eq. 6.7, which describes a single diffusion species in the presence of volume fluctuations. The fit and data are in very good agreement (Figs. 6.4B and 6.S4C-D) with reduced chi-squared values close to one. The analysis was performed on n = 16 cells to collect fit parameters over a range of expression values. The amplitude $\mathbf{A}=\left(A_g,A_r\right)$ is expected to increase linearly with the fluorescence intensity, which agrees with the data (Fig. 6.S5A). The characteristic time of the volume fluctuations is independent of concentration

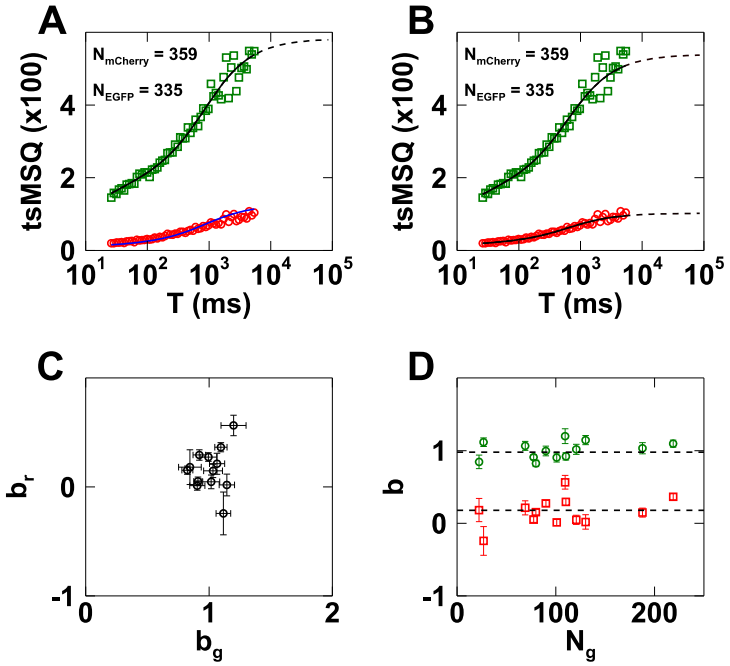

Figure 6.4. DC tsMSQ analysis of measurements performed in the NE of cells (*n* = 16) co-expressing SS-EGFP and SS-mCherry-KDEL. A) tsMSQ curves (symbols) with predicted $\mathrm{tsMSQ}_{\mathrm{gr}}^{(-)}$ curve (blue line) derived from a fit to $\mathrm{tsMSQ}_{\mathrm{gg}}$ (black line). The $\mathrm{tsMSQ}_{\mathrm{gg}}$ and $\mathrm{tsMSQ}_{\mathrm{gr}}$ curves are shown with green squares and red circles, respectively. B) tsMSQ curves (symbols) with fit to Eq. 6.7 (black). C) b_r vs. b_g. D) b_g (green circles) and b_r (red squares) vs. N_g with means (dashed lines) and standard deviations of b_g = 1.0 ± 0.2 and b_r = 0.16 ± 0.13.

with a mean time of 0.3 s, closely mirroring results obtained in previous work (Fig. 6.S6A) (53).

The fitted Q-vector $\bar{\mathbf{Q}} = \left(\bar{Q}_g, \bar{Q}_r\right)$ was converted into the normalized brightness $\mathbf{b} = \left(b_g, b_r\right)$ (Fig. 6.4C), as described in Section 6.3. The values of the brightness plot scatter around $\mathbf{b} = \left(1, 0\right)$, which is consistent for the SS-EGFP species containing one EGFP label and zero mCherry labels. These $\mathbf{b}$ values showed no dependence on the concentration of SS-EGFP (Fig. 6.4D). As discussed in the Theory section, DC tsMSQ was designed to identify HSP parameters. Thus, any purely red-emitting species is filtered out by HSP. As a consequence, we expected the non-interacting SS-mCherry-KDEL species to be invisible to the analysis, which is confirmed by the brightness plot. Consequently, the recovered diffusion times apply to the SS-EGFP species (Fig. 6.S7A). Their values are concentration independent with a mean value of 1.7 ms, which is consistent with previous results (53, 55).

As an additional control, we reanalyzed the cytoplasmic EGFP and mCherry data (Fig. 6.S3) using DC tsMSQ with a fit to a single diffusing species (Eq. 6.5). The data and fit closely match (Fig. 6.S8A-C), with the recovered HSP brightness values $\mathbf{b} = \left(b_g, b_r\right)$ centered near (1, 0) in a brightness plot, which is consistent with the EGFP species (Fig. 6.S8D-E). The fitted diffusion time is concentration independent with a mean value of 0.8 ms (Fig. 6.S7B), which agrees with previously published experiments (27).

After establishing that DC tsMSQ properly identifies non-interacting proteins within the NE and the cytoplasm, we next examined the tandem hetero-dimer SS-EGFP-SL-

mCherry. This construct carries both EGFP and mCherry separated by a linker (SL) and mimics a strongly interacting pair of proteins forming a hetero-dimeric complex. As expected, DC FFS measurements of SS-EGFP-SL-mCherry in the NE produced a cross-correlation $tsMSQ_{gr}$ curve that significantly exceeded the predicted baseline $tsMSQ_{gr}^{(-)}$ for non-interacting proteins (Fig. 6.5A). The DC tsMSQ curves were readily modeled by a fit to Eq. 6.7 (Fig. 6.5A) with reduced chi-squared values of 1.06. The HSP brightness values were centered around (1, 1), reflecting the presence of complexes with an average composition of one EGFP and one mCherry label as expected for the hetero-dimer (Fig. 6.5B). The diffusion time of the hetero-dimer identified by fitting is concentration independent (Fig. 6.S7C). The $\mathbf{A}$ and τ_0 values obtained for the volume fluctuations agree with the expected behavior (Figs. 6.S5B, 6.S6B). These results demonstrate that DC tsMSQ can be used to accurately identify the presence of hetero-dimeric protein complexes in the NE of living cells.

While the fluorescence intensity of luminal proteins is affected by volume fluctuations in the NE (Fig. 6.1B), the fluorescence intensity of membrane-associated proteins is not affected by these volume fluctuations as demonstrated in previous work using SC FFS (53). To test this difference between luminal and membrane-bound proteins in the context of DC tsMSQ, we performed DC FFS experiments in the NE of cells expressing the membrane-bound SS-EGFP-torsin$A^{NTD\text{-}2xLeu}$ and the luminal SS-mCherry-KDEL. Since torsin$A^{NTD\text{-}2xLeu}$ is a transmembrane domain (79), its presence ensures that EGFP is anchored to the nuclear membrane.

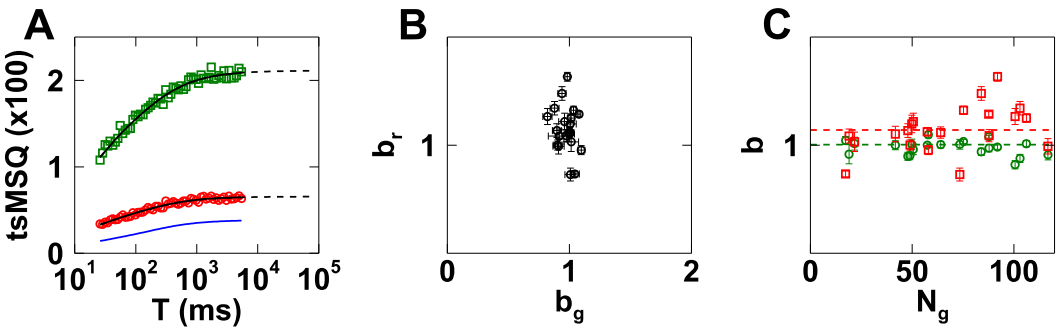

Figure 6.5. DC tsMSQ analysis of measurements obtained in the NE of cells (*n* = 24) expressing SS-EGFP-SL-mCherry. A) tsMSQ curves (symbols) with fits (black lines) and predicted $\mathrm{tsMSQ}_{gr}^{(-)}$ curve (blue line). The tsMSQ_{gg} and tsMSQ_{gr} curves are shown with green squares and red circles, respectively. B) b_r vs. b_g. C) b_g (green circles) and b_r (red squares) vs. N_g with means (dashed lines) and standard deviations of b_g = 0.96 ± 0.06 and b_r = 1.1 ± 0.3.

While the fluorescence signal generated by the luminal SS-mCherry-KDEL includes volume fluctuations, the fluorescence signal generated from the membrane-bound SS-EGFP-torsinA$^{NTD\text{-}2xLeu}$ does not. Since DC tsMSQ filters out any purely red fluorescing species, the data are expected to only contain the membrane-bound green fluorescing species. Consequently, we fit the DC tsMSQ curves to a model containing only a diffusing species and no exponential process (Eq. 6.5), which agreed well with the data (χ_ν^2 = 1.07). The fit recovered HSP brightness values $\mathbf{b}$ that are clustered around (1, 0) (Fig. 6.6B) and show no concentration dependence (Fig. 6.6C), which is consistent with SS-EGFP-torsinA$^{NTD\text{-}2xLeu}$ being monomeric. The diffusion times for SS-EGFP-torsinA$^{NTD\text{-}2xLeu}$ (Fig. 6.S7D) have a mean and standard deviation of 16 ± 4 ms, which is in agreement with previously published results for this construct (53).

We next used DC tsMSQ to study the assembly of LINC complexes composed of SUN2 and nesprin-2 *in vivo* by investigating the ability of the SUN2 luminal domain to form a heterotypic interaction with the luminal domain of nesprin-2. Based on previously published *in vitro* biochemical and structural studies, it is expected that a homo-trimer of SUN2 interact with three nesprin-2 KASH peptides to form a hetero-hexamer (49, 120). Moreover, SUN2 homo-trimerization was shown to be critical for KASH-binding (49, 51). To begin to test this model of LINC complex assembly in the NE of living cells we first performed measurements on cells co-expressing mCherry-SR-KASH2 with SS-EGFP-SUN2$^{595\text{-}731}$. The SUN domain (SUN2$^{595\text{-}731}$) contains the KASH binding sites and has previously been shown to remain monomeric using SC tsMSQ (54). Thus, the SUN

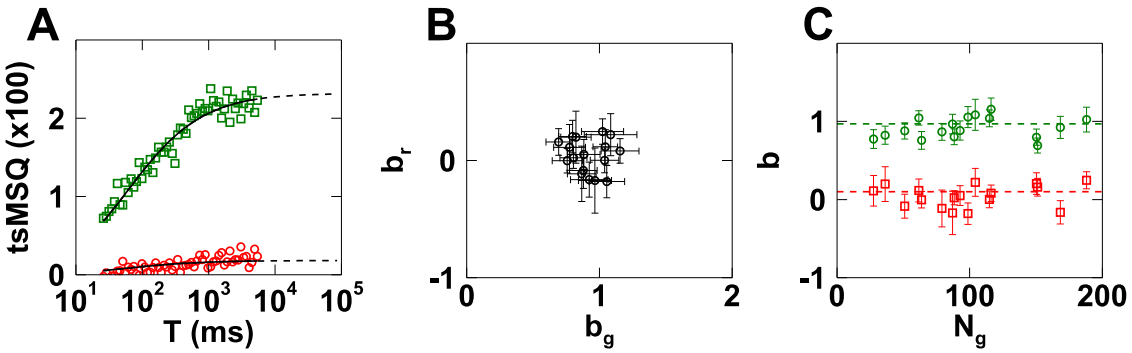

Figure 6.6. DC tsMSQ analysis of measurements performed in the NE of cells (*n* = 20) co-expressing SS-EGFP-torsinA$^{NTD\text{-}2xLeu}$ and SS-mCherry-KDEL. A) tsMSQ curves (symbols) with fits (black lines). The $tsMSQ_{gg}$ and $tsMSQ_{gr}$ curves are shown with green squares and red circles, respectively. B) b_r vs. b_g. C) b_g (green circles) and b_r (red squares) vs. N_g with means (dashed lines) and standard deviations of b_g = 0.97 ± 0.18 and b_r = 0.08 ± 0.12.

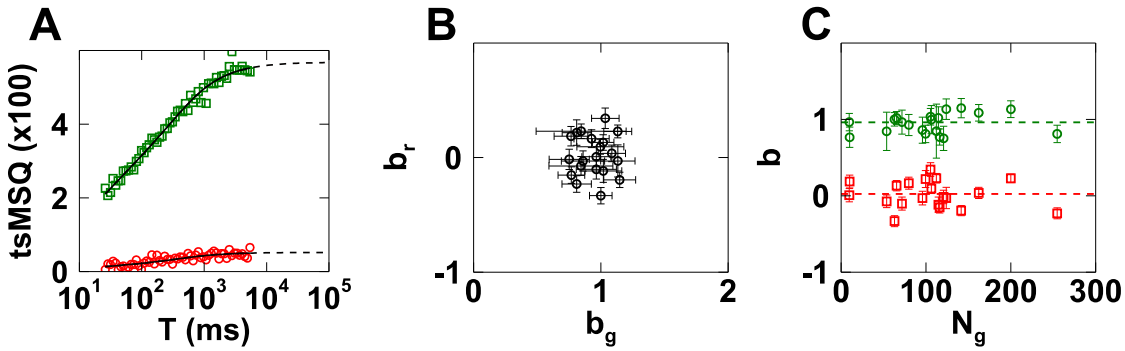

Figure 6.7. DC tsMSQ analysis of measurements performed in the NE of cells (*n* = 20) co-expressing SS-EGFP-SUN2$^{595\text{-}731}$ and mCherry-SR-KASH2. A) tsMSQ curves (symbols) with fits (black lines). The $tsMSQ_{gg}$ and $tsMSQ_{gr}$ curves are shown with green squares and red circles, respectively. B) b_r vs. b_g. C) b_g (green circles) and b_r (red squares) vs. N_g with means (dashed lines) and standard deviations of $b_g = 0.96 \pm 0.13$ and $b_r = 0.0 \pm 0.2$.

domain permits us to directly test for the presence of monomer-monomer interactions with nesprin-2.

Fitting the DC tsMSQ curves to Eq. 6.7 agreed well with the data (χ_ν^2 = 0.88) (Fig. 6.7A). A plot of b_r vs. b_g revealed that the data were clustered around (1, 0) indicating a non-interacting monomeric EGFP-labeled protein species (Fig. 6.7B). Neither b_g nor b_r showed any concentration dependence and both have means consistent with SS-EGFP-SUN2$^{595\text{-}731}$ being unable to interact with mCherry-SR-KASH2 (Fig. 6.7C). In addition, the values of τ_D, $\mathbf{A}$, and τ_0 all agreed with the expectation of a non-interacting SS-EGFP-SUN2$^{595\text{-}731}$ monomer (Figs. 6.S7E, 6.S5C, 6.S6C). Taken together, these results support the model in which the luminal domain of nesprin-2 is unable to interact with monomers of the SUN2 luminal domain.

Unlike SS-EGFP-SUN2$^{595\text{-}731}$, the SUN2 luminal domain-encoding SS-EGFP-SUN2$^{261\text{-}731}$ homo-trimerizes in the NE of living cells as determined by SC FFS (55). Thus, we expected that measurements performed in the NE of cells co-expressing SS-EGFP-SUN2$^{261\text{-}731}$ and mCherry-SR-KASH2 would identify heterotypic interactions between these constructs. Proper analysis by DC tsMSQ of data obtained in cells expressing both proteins required the inclusion of two diffusing species (Eq. 6.8) to describe the data (Fig. 6.8A). The two diffusion times have means of 1.7 and 180 ms (Fig. 6.S7F) which are consistent with a previous SC FFS study of SS-EGFP-SUN2$^{261\text{-}731}$ (55). The $tsMSQ_{gr}$ amplitude is significantly higher than the non-interacting prediction (Fig. 6.8A), indicating the presence of hetero-protein association. Plotting b_r vs. b_g from n = 53 cells revealed that both b_g and b_r increase together (Fig. 6.8B), implying that an increase in the average

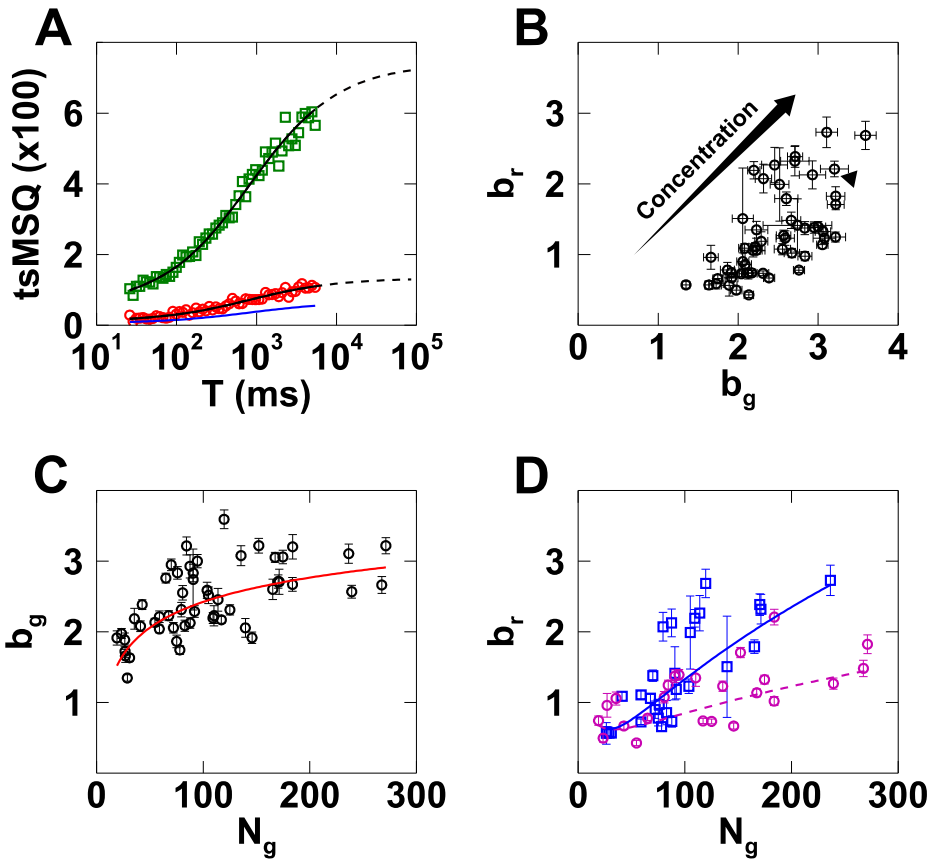

Figure 6.8. DC tsMSQ analysis of measurements performed in the NE of cells (*n* = 53) co-expressing SS-EGFP-SUN2$^{261\text{-}731}$ and mCherry-SR-KASH2. A) tsMSQ curves (symbols) with fits (black lines) with χ^2_ν = 0.88 and predicted $\mathrm{tsMSQ}^{(-)}_{\mathrm{gr}}$ curve (blue line). The $\mathrm{tsMSQ}_{\mathrm{gg}}$ and $\mathrm{tsMSQ}_{\mathrm{gr}}$ curves are shown with green squares and red circles, respectively. B) b_r vs. b_g are correlated and increase with concentration as indicated. C) b_g vs. N_g with a line provided to guide the eye. D) b_r vs. N_g separated by N_r:N_g ratios greater (blue squares) or less (magenta circles) than 4:1, with lines provided to guide the eye.

oligomeric state of SS-EGFP-SUN2$^{261-731}$ is associated with an increase in the number of bound mCherry-SR-KASH2. This result supports the model that nesprin-2 interacts with homo-trimers of SUN2 within the NE. We further notice that $b_r \leq b_g$, indicating that at most one mCherry-SR-KASH2 can associate with each SS-EGFP-SUN2$^{261-731}$ protein within the homo-trimer.

These data show that b_g ranges between 1 and 3, signifying a limiting homo-trimeric state for the SS-EGFP-SUN2$^{261-731}$ in the NE. The concentration dependence of the oligomerization of SS-EGFP-SUN2$^{261-731}$ is visualized by a plot of b_g vs. N_g (Fig. 6.8C). The brightness increases with N_g, approaching a homo-trimeric brightness state, which is identical to the behavior observed for SS-EGFP-SUN2$^{261-731}$ in the absence of mCherry-SR-KASH2 (54). We expect to observe an increase in b_r with N_g, because the luminal domain of nesprin-2 is predicted to bind to homo-trimers of the SUN2 luminal domain, which are populated at higher N_g (Fig. 6.8C). However, the amount of bound mCherry-SR-KASH2 also depends on its concentration, which varies significantly from cell to cell. Thus, the variability in the expression ratio of mCherry-SR-KASH2 to SS-EGFP-SUN2$^{261-731}$ is responsible for the large scatter in the observed b_r values (Fig. 6.8D). To visualize the dependence of b_r on the expression ratio of these two constructs, the data were separated into sets with an N_r:N_g above and below 4:1 (Fig. 6.8D). Data with a N_r:N_g ratio above 4:1 exhibit a strong increase in b_r up to values approaching 3, while the lower ratio data only show a modest increase in b_r with N_g. This reflects the relative reduction in mCherry-SR-KASH2 concentration between both data sets. Notably, the highest brightness values measured for b_g and b_r approach 3 (Figs. 6.8B-D), which suggests the

formation of a SUN2-nesprin-2 hetero-hexamer *in vivo* as predicted by *in vitro* models (49, 120).

6.5 Conclusions

While CCF analysis has been widely used to identify heterotypic protein interactions from FFS measurements performed in cells (126), the presence of nuclear membrane undulations is a significant barrier for its application to FFS measurements performed in the NE. The local volume changes caused by nuclear membrane undulations lead to coupled changes in the emission intensity of both the red and green fluorescent labels (Fig. 6.1B), which result in a spurious crosscorrelation in the CCF, even when accounting for spectral crosstalk. The DC tsMSQ theory developed in this paper provides a comprehensive method that incorporates both spectral crosstalk and volume fluctuations, thus overcoming the issues present in CCF analysis of FFS data measured in the NE. We experimentally verified DC tsMSQ using model systems representing non-interacting protein pairs as well as hetero-dimeric protein complexes and recovered diffusion times and HSP brightness values. The label stoichiometry determined by HSP brightness matched the expected values of the control samples and reliably distinguished interacting from non-interacting proteins. Furthermore, the parameters of the volume fluctuation process agreed with our previously published SC tsMSQ results (53), thus providing additional support for the DC tsMSQ model.

Next, we used DC tsMSQ to investigate the assembly mechanism of LINC complexes, which are formed by the heterotypic interaction of the luminal domains of SUN

and KASH proteins (74). Previously published *in vitro* biochemical and structural studies revealed that a homo-trimer of the SUN2 luminal domain interacts with three nesprin-2 luminal domains (49, 120). Deep KASH peptide-binding grooves are formed at the interface of two adjacent SUN domains upon SUN2 luminal domain homo-trimerization; therefore, SUN2 homo-trimerization is believed to be a necessary precursor to the assembly of a SUN2-nesprin-2 hetero-hexamer (49, 120).

To test this model *in vivo,* we looked for interactions between the luminal domain of nesprin-2 (mCherry-SR-KASH2) and two SS-EGFP-tagged SUN2 luminal domain constructs. We previously showed that the EGFP-tagged luminal domain of SUN2 (SS-EGFP-SUN2$^{261-731}$) forms homo-trimers in the NE, while the SUN domain of SUN2 (SS-EGFP-SUN2$^{595-731}$) remains monomeric (54). As expected, we were unable to detect a heterotypic interaction between SS-EGFP-SUN2$^{595-731}$ and mCherry-SR-KASH2 using DC tsMSQ. However, DC tsMSQ did detect a concentration-dependent heterotypic interaction between SS-EGFP-SUN2$^{261-731}$ and mCherry-SR-KASH2 in the NE. The limiting HSP brightness values measured for these constructs suggest the formation of a SUN2-nesprin-2 hetero-hexamer *in vivo*, which agrees with the *in vitro* results described above (104).

These initial results provide a promising starting point for future quantitative studies of the mechanisms underlying the *in vivo* assembly of functional LINC complexes and their regulation. Combining quantitative modeling of the DC tsMSQ brightness data with targeted mutations that perturb the known SUN2-nesprin-1/2 binding sites should provide a comprehensive approach for investigating the assembly of SUN2-containing LINC complexes. In addition, DC tsMSQ offers a tool to investigate the formation of LINC

complexes composed of lesser studied SUN proteins, such as SUN1, which forms higher-order oligomers than SUN2 (54). While this paper has focused on the application of DC tsMSQ for characterizing the heterotypic interactions of LINC complex proteins, the development of DC tsMSQ offers a promising and general platform for future studies of hetero-protein complex formation in the NE of living cells.

6.6 Supplemental Materials

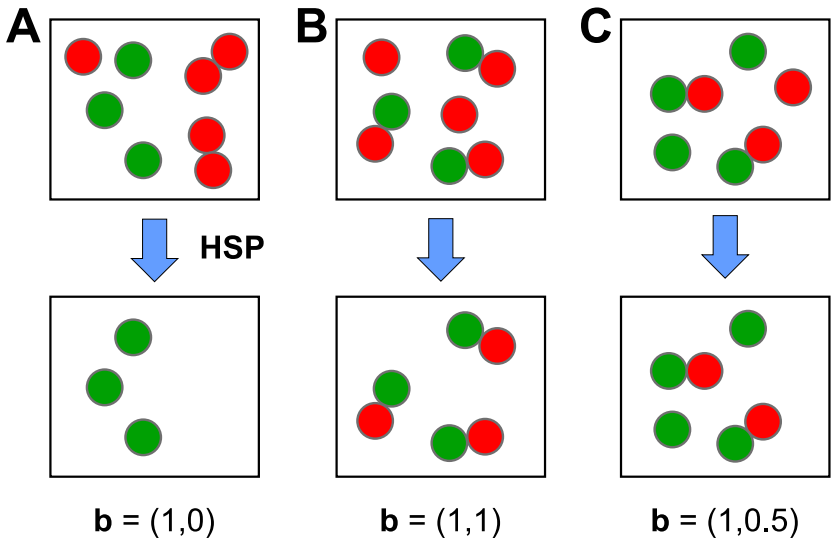

Figure 6.S1. Illustration of HSP analysis. Green and red circles represent an EGFP-labeled protein (G) and an mCherry-labeled protein (R), respectively. A) A sample containing monomeric G, monomeric R and dimeric R_2 is filtered by HSP analysis, which removes R and R_2. The HSP heterospecies is represented by monomeric G, which corresponds to a normalized brightness vector of (1, 0). B) HSP of a sample consisting of heterodimeric GR and monomeric R removes R, which leads to a heterospecies consisting of GR with a corresponding brightness vector of (1, 1). C) A mixture of G, R, and GR is filtered by HSP analysis leading to a heterospecies consisting of a 1:1 mixture of G and GR. This mixture is represented by an average oligomeric state of 1 for G with half an R on average associated with each G. This composition leads to a HSP brightness vector of (1, 0.5).

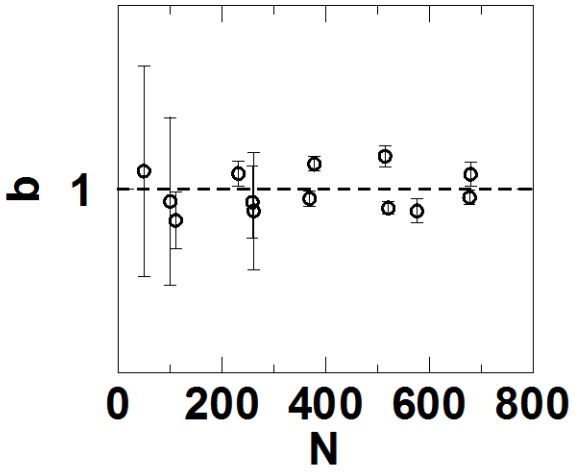

Figure 6.S2. *b* vs. *N* for SS-mCherry-KDEL. Results of SC tsMSQ analysis for measurements performed in the NE of cells (n = 13) expressing SS-mCherry-KDEL. Mean (dashed line) and standard deviation are $b = 0.99 \pm 0.14$.

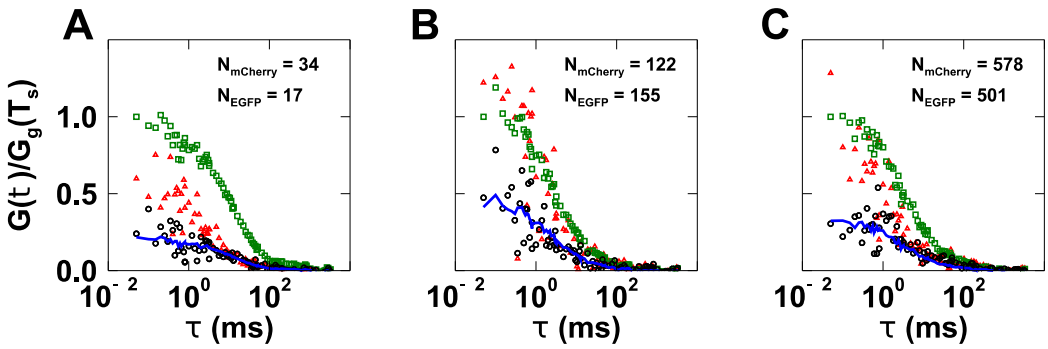

Figure 6.S3. Results of ACF and CCF analysis for measurements performed in the cytoplasm of cells expressing EGFP and mCherry in the cytoplasm at low (A), medium (B), and high (C) levels of protein expression. N_{EGFP} and $N_{mCherry}$ give the number of EGFP and mCherry proteins present in the observation volume, respectively. The red and green channel ACF are shown with red triangles and green squares, respectively. The CCF is graphed as black circles and the calculated CCF from spectral crosstalk (CCF_{ct}) is shown by the blue line. All correlation curves are divided by the amplitude of the ACF $G_{GG}(T_S)$ of the green channel with lag time equal to the sampling time T_S.

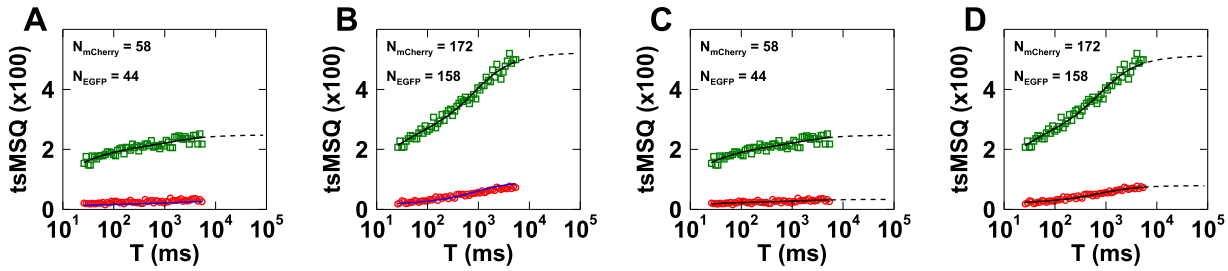

Figure 6.S4. DC tsMSQ curves of measurements performed in the NE of cells co-expressing SS-EGFP and SS-mCherry-KDEL. A-B) tsMSQ curves (symbols) for low (A) and medium (B) protein expression. The predicted $\mathrm{tsMSQ}_{gr}^{(-)}$ curve (blue line) is derived from a fit to tsMSQ_{gg} (black line). The tsMSQ_{gg} and tsMSQ_{gr} curves are shown with green squares and red circles, respectively. C-D) tsMSQ curves (symbols) for low (C) and medium (D) protein expression with fit to Eq. 6.3 (black lines).

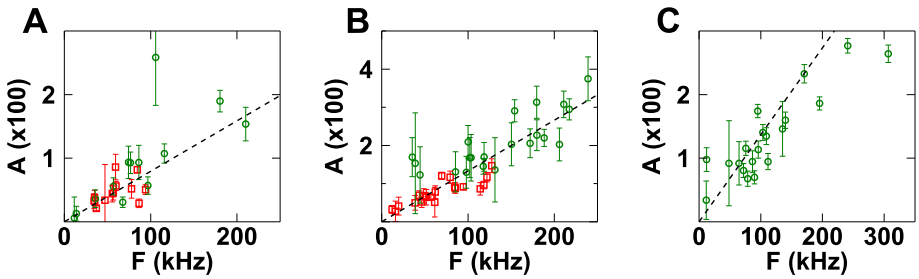

Figure 6.S5. $\mathbf{A}$ vs. *F* from DC tsMSQ analysis of measurements performed in the NE of cells. Plots of $\mathbf{A}$ vs. *F* for different protein pairs along with fitted slope (dashed line). A) $\mathbf{A}$ vs. *F* for SS-EGFP and SS-mCherry-KDEL with slope of 7.9×10^{-5} kHz^{-1}. B) $\mathbf{A}$ vs. *F* for SS-EGFP-SL-mCherry with slope of 1.3×10^{-4} kHz^{-1}. C) $\mathbf{A}$ vs. *F* for SS-EGFP-SUN2$^{595\text{-}731}$ and mCherry-SR-KASH2 with slope of 1.4×10^{-4} kHz^{-1}.

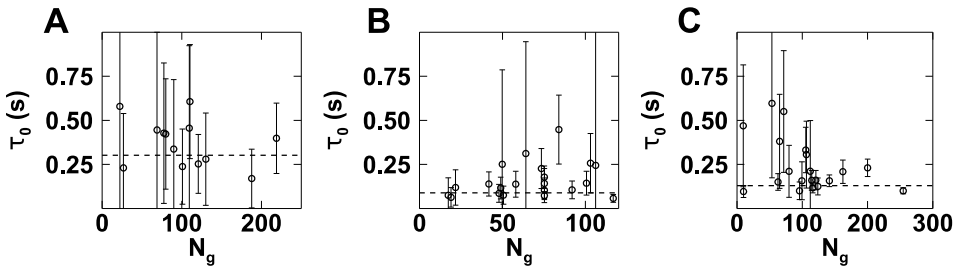

Figure 6.S6. τ_0 vs. N_g from DC tsMSQ analysis of measurements performed in the NE of cells. Plots of τ_0 vs. N_g for different protein pairs and their mean (dashed line). A) τ_0 vs. N_g of SS-EGFP and SS-mCherry-KDEL with mean and standard deviation of τ_0 = 0.30 ± 0.13 s. B) τ_0 vs. N_g of SS-EGFP-SL-mCherry with mean and standard deviation of τ_0 = 0.09 ± 0.05 s. C) τ_0 vs. N_g of SS-EGFP-SUN2$^{595\text{-}731}$ and mCherry-SR-KASH2 with mean and standard deviation of τ_0 = 0.13 ± 0.06 s.

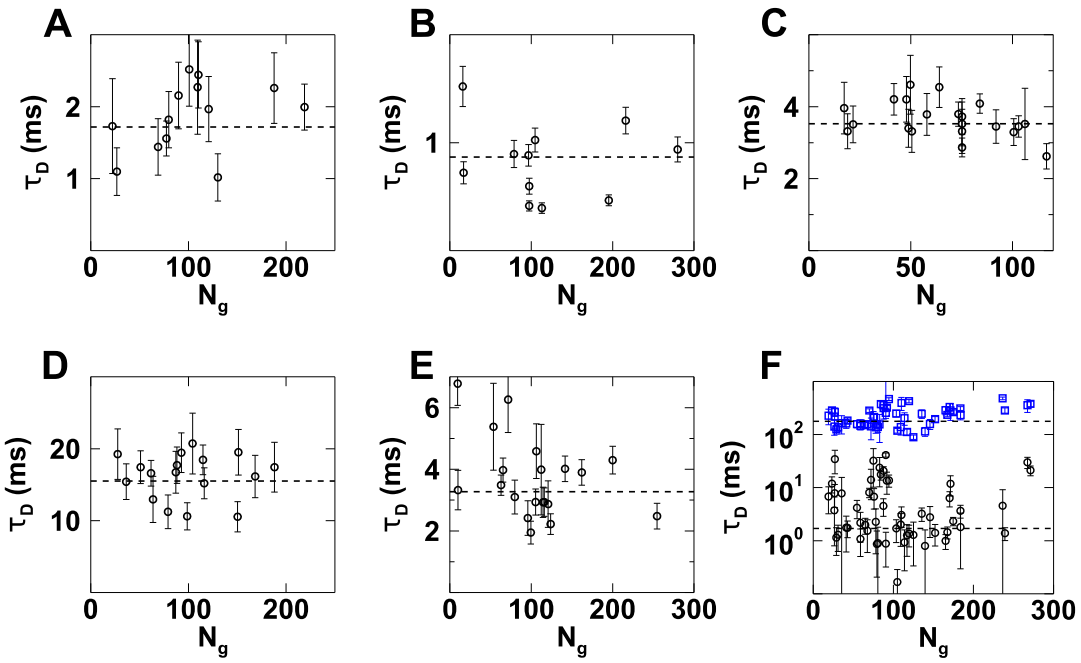

Figure 6.S7. τ_D vs. N_g from DC tsMSQ analysis of measurements performed in cells. Plots of τ_D vs. N_g for different protein pairs measured in the NE unless otherwise stated, along with their mean (dashed line). A) τ_D vs. N_g of SS-EGFP and SS-mCherry-KDEL with mean and standard deviation of τ_D = 1.7 ± 0.5 ms. B) τ_D vs. N_g of EGFP and mCherry measured in the cytoplasm with mean and standard deviation of τ_D = 0.8 ± 0.3 ms. C) τ_D vs. N_g of SS-EGFP-SL-mCherry with mean and standard deviation of τ_D = 3.5 ± 0.6 ms. D) τ_D vs. N_g of SS-EGFP-NTD$^{\text{torsinA-2xLeu}}$ and mCherry-KDEL with mean and standard deviation of τ_D = 16 ± 4 ms. E) τ_D vs. N_g of SS-EGFP-SUN2$^{595\text{-}731}$ and mCherry-SR-KASH2 with mean and standard deviation of τ_D = 3.3 ± 1.0 ms. F) τ_D vs. N_g of SS-EGFP-SUN2$^{261\text{-}731}$ and mCherry-SR-KASH2 $\tau_{D,1}$ (black circles) and $\tau_{D,2}$ (blue squares) vs. N_g with means and standard deviations of $\tau_{D,1}$ = 1.7 ± 1.1 ms and $\tau_{D,2}$ = 180 ± 150 ms.

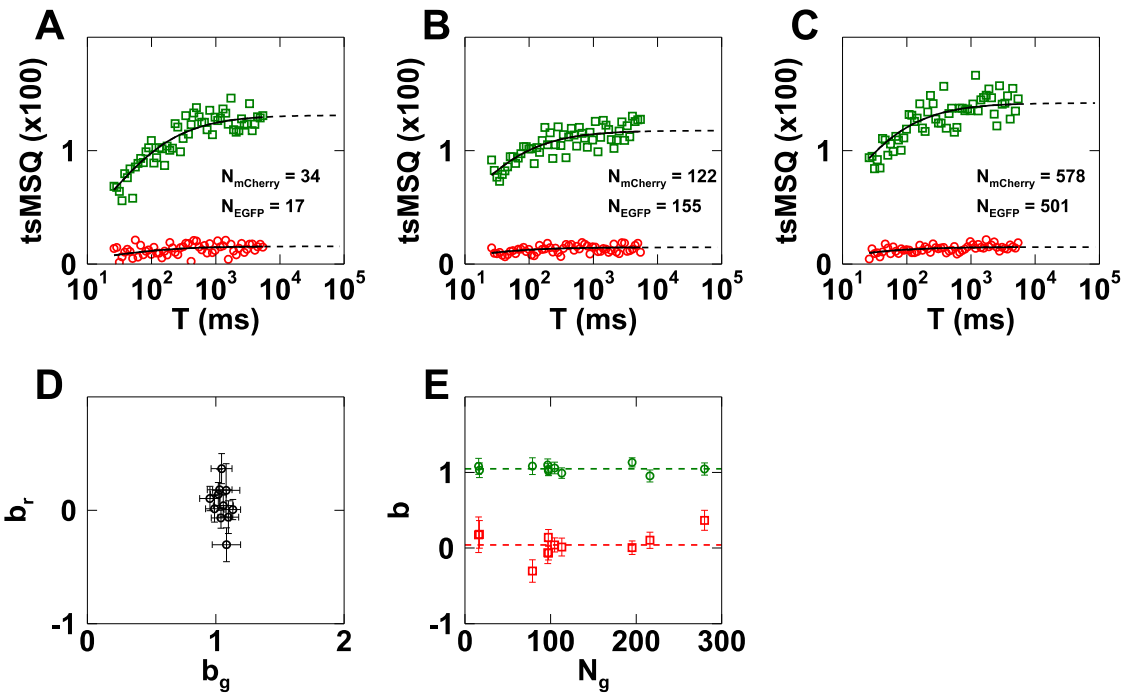

Figure 6.S8. DC tsMSQ analysis of measurements performed in the cytoplasm of cells (*n* = 11) co-expressing EGFP and mCherry. A-C) tsMSQ curves (symbols) for low (A), medium (B), and high (C) protein expression with fit to Eq. 6.1 (black). The $tsMSQ_{gg}$ and $tsMSQ_{gr}$ curves are shown with green squares and red circles, respectively. E) b_r vs. b_g. D) b_g (green circles) and b_r (red squares) vs. N_g with means (dashed lines) and standard deviation of b_g = 1.05 ± 0.06. and b_r = 0.04 ± 0.13.

6.6.1 Definition of tsMSQ$_{\sigma\rho}$

The recorded data $k_{\sigma,i}$ represent the photon counts detected at time $i \cdot T_S$ in channel σ, where $\sigma = g$ and $\sigma = r$ refer to the green and the red detection channels, respectively, and T_S represents the sampling time. We generalize the time-shifted Q-factor (tsQ) of a single channel (68) to two channels by

$$\mathrm{tsQ}_{\sigma\rho} \equiv \frac{\left\langle \delta k_{\sigma,i}\, \delta k_{\rho,i+1} \right\rangle}{\left\langle k_{\sigma,i} \right\rangle} \qquad \text{6.S1}$$

with the deviation of the photon counts from the mean $\left\langle k_{\sigma,i} \right\rangle$ defined by $\delta k_{\sigma,i} \equiv k_{\sigma,i} - \left\langle k_{\sigma,i} \right\rangle$ and $\langle\ \rangle$ symbolizing the population average. The tsMSQ algorithm divides the photon count record into data segments of period $T = MT_S$, with each segment containing M data points, before calculating $\mathrm{tsQ}_{\sigma\rho}$ for each segment. Because the M segment is finite, population averages are only estimated. We define the estimator of $\mathrm{tsQ}_{\sigma\rho}$ from a straightforward generalization of the single channel case (72),

$$\mathrm{tsQ}_{\sigma\rho}(T) \equiv \frac{\overline{\Delta k_{\sigma,i} \Delta k_{\rho,i+1}}}{\bar{k}_{\sigma,i}}\ , \qquad \text{6.S2}$$

where $\bar{k}_{\sigma,i} = M^{-1} \sum_{i=1}^{M} k_{\sigma,i}$ is the estimator of the data segment mean and $\overline{\Delta k_{\sigma,i} \Delta k_{\rho,i+1}} = (M-1)^{-1} \sum_{i=1}^{M-1} \left(k_{\sigma,i} - \bar{k}_{\sigma,L}\right)\left(k_{\rho,i+1} - \bar{k}_{\rho,R}\right)$ is the covariance estimator with $\bar{k}_{\sigma,L} = (M-1)^{-1} \sum_{j=1}^{M-1} k_{\sigma,j}$ and $\bar{k}_{\sigma,R} = (M-1)^{-1} \sum_{j=1}^{M-1} k_{\sigma,j+1}$.

We define the dual-channel tsMSQ in a manner that is analogous to the single-channel tsMSQ with the average of the $\mathrm{tsQ}_{\sigma\rho}$ -estimator and a correction term to account for estimator bias due to shot noise (55),

$$\mathrm{tsMSQ}_{\sigma\rho}(T) \equiv \left\langle \mathrm{tsQ}_{\sigma\rho}(T) \right\rangle + \delta_{\sigma\rho} \frac{(T-2T_S)T_S}{(T-T_S)^2} \quad . \qquad \text{6.S3}$$

The correction term includes a Kronecker delta $\delta_{\sigma\rho}$, because the estimator bias only occurs for $\sigma = \rho$ as shown in Section 6.6.2 below.

6.6.2 tsMSQ$_{\sigma\rho}$ of a single diffusing species and of NE volume fluctuations

The population mean of $\mathrm{tsQ}_{\sigma\rho}$ is determined by inserting the definitions for $\overline{\Delta k_{\sigma,i} \Delta k_{\rho,i+1}}$ and $\bar{k}_{\sigma,i}$ into Eq. 6.S2 and taking the ensemble average,

$$\left\langle \mathrm{tsQ}_{\sigma\rho}(T) \right\rangle = \frac{\left\langle \delta k_{\sigma,i} \delta k_{\rho,i+1} \right\rangle}{\left\langle k_{\sigma,i} \right\rangle} - \frac{1}{(M-1)^2} \frac{\sum_{i=1}^{M-1} \sum_{j=1}^{M-1} \left\langle \delta k_{\sigma,i} \delta k_{\rho,j+1} \right\rangle}{\left\langle k_{\sigma,i} \right\rangle} . \qquad \text{6.S4}$$

A detailed derivation of Eq. 6.S4 for the single-color case ($\sigma = \rho$) was previously described (55). It is straightforward to verify Eq. 6.S4 by repeating this derivation with the subscripts σ and ρ to account for the general case of two detection channels.

We evaluated Eq. 6.S4 for two different fluctuation sources: a single diffusing species and an exponential correlation process due to volume fluctuations. To simplify

notation we use the symbol X with X = D denoting the diffusion process, while X = E refers to the exponential correlation process. Previous work on the single channel case ($\sigma = \rho$) derived a compact expression for both processes (55),

$$\left\langle \mathrm{tsQ}_{\sigma\sigma}\left(T\right)\right\rangle = X_{\sigma}\mathrm{tsF}_{\mathrm{X}}\left(T\right) - \frac{\left(T-2T_S\right)T_S}{\left(T-T_S\right)^2} \ , \qquad \text{6.S5}$$

where X_{σ} is the amplitude and $\mathrm{tsF}_{\mathrm{X}}\left(T\right)$ is the normalized tsMSQ function (55),

$$\mathrm{tsF}_{x}\left(T\right) = \frac{\mathrm{B}_{2,x}\left(2T_S\right) - 2\,\mathrm{B}_{2,x}\left(T_S\right)}{2T_S^2} - \frac{\mathrm{B}_{2,x}\left(T\right) + \mathrm{B}_{2,x}\left(T-2T_S\right) - 2\,\mathrm{B}_{2,x}\left(T_S\right)}{2\left(T-T_S\right)^2} \ , \qquad \text{6.S6}$$

with the second-order binning function $\mathrm{B}_{2,X}\left(T\right)$ (66, 75). For the case of two different channels ($\sigma \neq \rho$), Eq. 6.S4 reduces to

$$\left\langle \mathrm{tsQ}_{\sigma\rho}\left(T\right)\right\rangle = X_{\rho}\mathrm{tsF}_{\mathrm{X}}\left(T\right) \ , \qquad \text{6.S7}$$

which will be shown below. Based on the definition of tsMSQ (Eq. 6.S3), Eqs. 6.S5 and 6.S7 inserted into Eq. 6.S3 provide a general expression for dual-channel tsMSQ,

$$\mathrm{tsMSQ}_{\sigma\rho}\left(T\right) = X_{\rho}\,\mathrm{tsF}_{\mathrm{X}}\left(T\right) \ . \qquad \text{6.S8}$$

Deduction of Eq. 6.S7 follows the same steps as described in the literature for the special case of $\sigma = \rho$ (55). Since repeating this derivation for the case of two different channels ($\sigma \neq \rho$) is relatively straightforward, we only outline the main steps of the derivation and indicate where differences arise. The first term of Eq. 6.S4 is by definition the time-shifted Q-factor $\mathrm{tsQ}_{\sigma\rho}$ (Eq. 6.S2). In the case of oversampling (i.e. T_S is much

smaller than the characteristic correlation time of the fastest process), $\mathrm{tsQ}_{\sigma\rho}$ is related to the normalized correlation function $G_2(\tau)$ with amplitude G_0,

$$\mathrm{tsQ}_{\sigma\rho} = \frac{\left\langle \delta k_{\sigma,i} \delta k_{\rho,i+1} \right\rangle_X}{\left\langle k_{\sigma,i} \right\rangle \left\langle k_{\rho,i} \right\rangle} \left\langle k_{\rho,i} \right\rangle = G_0 \left\langle k_{\rho,i} \right\rangle G_{2,X}(T_S) = X_\rho G_{2,X}(T_S). \qquad \text{6.S9}$$

We further proved for $\sigma = \rho$ that for the general case of arbitrary sampling time T_S the above equation leads to (55)

$$\mathrm{tsQ}_{\sigma\rho} = X_\rho \frac{\mathrm{tsB}_{2,X}(T_S)}{T_S^2}, \qquad \text{6.S10}$$

which also holds for $\sigma \neq \rho$. The time-shifted binning function $\mathrm{tsB}_{2,X}(T)$ of Eq. 6.S10 is defined in terms of the regular binning function $\mathrm{B}_{2,X}(T)$ (55, 66, 75)

$$\mathrm{tsB}_{2,X}(T) \equiv \frac{\mathrm{B}_{2,X}(2T) - 2\,\mathrm{B}_{2,X}(T)}{2}. \qquad \text{6.S11}$$

Similarly, the single-channel covariance $\left\langle \delta k_{\sigma,i} \delta k_{\sigma,j} \right\rangle$ described in (75) was generalized to the dual-channel case,

$$\left\langle \delta k_{\sigma,i} \delta k_{\rho,j} \right\rangle_X = X_\rho \left\langle k_{\sigma,i} \right\rangle \frac{B_{2,X}((j-i)T_S + T_S) + B_{2,X}((j-i)T_S - T_S) - 2B_{2,X}((j-i)T_S)}{2T_S^2}$$

. 6.S12

To evaluate the second term of Eq. 6.S4 for $\sigma \neq \rho$ the double sum is broken into three parts, $i = j+1$, $i > j+1$, and $i < j+1$. Utilizing Eq. 6.S12 and following the steps described in (55), we obtained expressions for the partial sums,

$$\sum_{i=2}^{M-2} \left\langle \delta k_{\sigma,i}\, \delta k_{\rho,i} \right\rangle_X = (M-2)\, X_\rho \left\langle k_{\sigma,i} \right\rangle \mathrm{B}_{2,X}(T_S) / T_S^2 \ , \qquad \text{6.S13}$$

$$\sum_{i=3}^{M-1} \sum_{j+1=2}^{i>j+1} \left\langle \delta k_{\sigma,i} \delta k_{\rho,j+1} \right\rangle_X = X_\rho \left\langle k_{\sigma,i} \right\rangle \frac{\mathrm{B}_{2,X}\left((M-2)T_S\right) - (M-2)\mathrm{B}_{2,X}(T_S)}{2T_S^2} \ , \qquad \text{6.S14}$$

and

$$\sum_{j+1=2}^{M} \sum_{i=1}^{i<j+1} \left\langle \delta k_{\sigma,i} \delta k_{\rho,j+1} \right\rangle_X = X_\rho \left\langle k_{\sigma,i} \right\rangle \frac{\mathrm{B}_{2,X}(MT_S) - M\,\mathrm{B}_{2,X}(T_S)}{2T_S^2} \qquad \text{6.S15}$$

Note that unlike for the case $\sigma = \rho$ (55), there is no shot noise contribution to Eq. 6.S13 for $\sigma \neq \rho$. The expression for the double,

$$\sum_{i=1}^{M-1} \sum_{j=1}^{M-1} \left\langle \delta k_{\sigma,i} \delta k_{\rho,j+1} \right\rangle_X = X_\rho \left\langle k_{\sigma,i} \right\rangle \frac{\mathrm{B}_{2,X}(T) + \mathrm{B}_{2,X}(T-2T_S) - 2\,\mathrm{B}_{2,X}(T_S)}{2T_S^2} \ , \qquad \text{6.S16}$$

was obtained by adding the partial sums and using the relation $T = MT_S$. Finally, inserting Eqs. 6.S10 and 6.S16 as well as $\left\langle k_{\sigma,i} \right\rangle = \lambda_\sigma N T_S$ into Eq. 6.S4 and using the definition of $\mathrm{tsF}_x(T)$ (Eq. 6.S6) yields Eq. 6.S7 for $\sigma \neq \rho$.

6.6.3 Binning function for an exponential correlation and diffusion process

The binning function for the exponential correlation process is defined by (53)

$$\mathrm{B}_{2,\mathrm{E}}\left(T;\tau_0\right)=2\tau_0^2\left(-1+\frac{T}{\tau_0}+e^{-T/\tau_0}\right), \qquad \text{6.S17}$$

where τ_0 is the characteristic time of the correlation process. The binning function for a single diffusing species depends on the shape of the observation volume. For the case of a two-dimensional Gaussian, the function is given by (66)

$$\mathrm{B}_{2,\mathrm{D}}\left(T;\tau_D\right)=2\tau_D^2\left(\left(1+t\right)\ln\left(1+t\right)-t\right), \qquad \text{6.S18}$$

where $t=T/\tau_D$ and τ_D is the diffusion time. For a three-dimensional Gaussian observation volume, the binning function is written as

$$\mathrm{B}_{2,\mathrm{D}}\left(T;\tau_D\right)=4\tau_D^2\left(r^2-r\sqrt{r^2+t}-\frac{r}{s}\left(1+t\right)\times\ln\left(\frac{\left(r-s\right)\left(s+\sqrt{r^2+t}\right)}{\sqrt{1+t}}\right)\right) \qquad \text{6.S19}$$

with $r^2=z_0^2/w_0^2$ and $s=\sqrt{r^2-1}$, where w_0 and z_0 are the radial and axial beam waist, respectively.

6.6.4 HSP

The HSP method requires that $Q_{g,Ch}=0$ (63), which is achieved by choosing appropriate fluorescence emission filters. Consequently, the two fluorescent proteins

EGFP (G) and mCherry (Ch) are characterized by the Q-vectors $\mathbf{Q}_G = \left(Q_{g,G}, Q_{r,G}\right)$ and $\mathbf{Q}_{ch} = \left(0, Q_{r,Ch}\right)$, respectively. Protein complexes formed by G- and Ch- labeled proteins are denoted as G_mR_n, where m and n are the stoichiometry coefficients for EGFP and mCherry. The Q-vector of G_mR_n is given by vector addition of the individual Q-vectors of the labels and yields $\mathbf{Q}\left[\mathrm{G}_m\mathrm{R}_n\right] = m\mathbf{Q}_G + n\mathbf{Q}_{Ch}$ (63). Corrections to this simple relationship have been described to include effects due to the presence of FRET and fluorescent protein dark states (63, 122). While these corrections lead to cumbersome equations, they have no influence on the final result, and are consequently omitted for convenience.

Interacting proteins in cells typically exist as a mixture of assembly states. These mixtures cannot be resolved experimentally once there are more than two assembly states (66) and the measured Q-values represent an averaged property over all assembly states (68). Let us consider a mixture of S species with the *i*th species $\mathrm{G}_{\mathrm{m}[i]}\mathrm{R}_{n[i]}$ containing $m[i]$ EGFP and $n[i]$ mCherry labels with the Q-vector $\mathbf{Q}[i] = m[i]\mathbf{Q}_G + n[i]\mathbf{Q}_R$ and fluorescence intensity $F_\sigma[i]$ in the σ-channel. The corresponding fractional fluorescence intensity is $f_\sigma[i] = F_\sigma[i]/F_{0,\sigma}$, where $F_{0,\sigma}$ is the total intensity of the σ-channel. The brightness mixture of S species is characterized by four averaged or apparent Q values, $\bar{Q}_{\sigma\rho} = \sum_{i=1}^{S} f_\sigma[i]Q_\rho[i]$. The Q-values $\bar{Q}_{g\rho}$ ($\sigma = g$) define the HSP Q-vector $\bar{\mathbf{Q}} = \left(\bar{Q}_g, \bar{Q}_r\right)$ $= \sum_{i=1}^{S} f[i]\mathbf{Q}[i]$ that represents the hetero-species where for convenience we define $f[i] = f_g[i]$. Using the properties of the *i*th species defined above, the HSP Q-vector is rewritten as

$$\bar{\mathbf{Q}} = \langle m \rangle \mathbf{Q}_G + \langle n \rangle \mathbf{Q}_{Ch} \qquad \text{6.S20}$$

where we defined $\langle m \rangle = \sum_{i=1}^{S} f[i] m[i]$ and $\langle n \rangle = \sum_{i=1}^{S} f[i] n[i]$. The scale factor $\langle m \rangle$ represents the fluorescence intensity averaged oligomeric state of the EGFP-labeled protein G of the sample, while $\langle n \rangle$ specifies the intensity-averaged oligomeric state of the mCherry-labeled protein R associated with G. Comparison with Eq. 6.10 demonstrates that the HSP brightness vector $\mathbf{b} = (b_g, b_r) = (\langle m \rangle, \langle n \rangle)$ provides a measure of the average hetero-oligomeric state of the sample.

6.6.5 tsMSQ with HSP analysis

To demonstrate the relation between tsMSQ and the HSP Q-vector of the preceding Section 6.6.4, we initially assume that the diffusion time of all species is approximately identical. For this case, the DC tsMSQ for the diffusion species reduces to a single apparent species,

$$\mathbf{tsMSQ_D}(T) = \sum_{i=1}^{S} f[i] \mathbf{Q}[i] \, \mathrm{tsF_D}(T; \tau_D) = \bar{\mathbf{Q}} \, \mathrm{tsF_D}(T; \tau_D) \, , \qquad \text{6.S21}$$

with an amplitude $\bar{\mathbf{Q}}$ equal to the HSP Q-vector. We now consider the case where the diffusion times $\tau_D[i]$ of the different species are no longer the same. Such a mixture of diffusion times is typically well approximated by the superposition of two single-species $\mathbf{tsMSQ_D}$ curves with distinct diffusion times given by

$$\begin{aligned}\mathbf{tsMSQ_D}\left(T\right) &= \sum_{i=1}^{S} f[i]\mathbf{Q}[i]\,\mathrm{tsF_D}\left(T;\tau_D[i]\right) \\ &\approx f_1\bar{\mathbf{Q}}_1\,\mathrm{tsF}_D\left(T;\tau_{D,1}\right)+ f_2\bar{\mathbf{Q}}_2\,\mathrm{tsF}_D\left(T;\tau_{D,2}\right)\end{aligned} \quad . \qquad 6.S22$$

The HSP Q-vector is determined by adding up the amplitudes of the two components,

$$\bar{\mathbf{Q}} = f_a\bar{\mathbf{Q}}_a + f_b\bar{\mathbf{Q}}_b \ , \qquad 6.S23$$

or generalized to S species by adding up all individual components (Eq. 6.9).

6.6.6 DNA constructs

The SS-mCherry-KDEL construct was generated by first replacing the cDNA encoding EGFP with cDNA encoding mCherry in the previously described SS-EGFP-torsinA^{WT} construct (77, 92). To do this, we first PCR amplified mCherry from pmCherry-N1 using the primers mCherry-F and mCherry-R (Table 6.S1). After purification, the PCR product was digested alongside SS-EGFP-torsinA^{WT} with the restriction enzyme (RE) *Nhe*I. The digested SS-EGFP-torsinA^{WT} construct was treated with Calf Intestinal Phosphatase was then treated with gel purification after which both it and the digested PCR product were gel purified and subsequently ligated together to create SS-mCherry-torsinA^{WT}. Next, SS-mCherry-KDEL was PCR amplified from SS-mCherry-torsinA^{WT} using the primers SS-mCherry-KDEL-F and SS-mCherry-KDEL-R (Table 6.S1). After purification, the PCR product was digested alongside pcDNA3.1+ with the REs *Bam*HI and *Eco*RI, gel purified, and then ligated resulting in the formation of SS-mCherry-KDEL.

To create the SS-EGFP-SUN2$^{595-731}$-mCherry construct, we first needed to generate a construct where the cDNA encoding EGFP was removed from SS-EGFP-SUN$^{261-731}$ and the cDNA encoding mCherry was subsequently fused to the 3' end of the SUN2^{LD}. This construct, SS-SUN2$^{261-731}$-mCherry, was generated by first deleting EGFP from SS-EGFP-SUN2$^{261-731}$ via PCR using the primers SS-EGFP-SUN2$^{261-731}$-ΔEGFP-F and SS-EGFP-SUN2$^{261-731}$-ΔEGFP-R (Table 6.S1). The resulting PCR product was then purified and ligated together via T4 polynucleotide kinase, ligase, *Dpn*I treatment to create SS-SUN2$^{261-731}$. The cDNA encoding SS-SUN2$^{261-731}$ was then PCR amplified using the primers SS-SUN2$^{261-731}$-F and SS-SUN2$^{261-731}$-R (Table 6.S1), after which it was gel purified and digested alongside pmCherry-N1 with *Hind*III and *Sal*I. The digested SS-SUN2$^{261-731}$ PCR product and pmCherry-N1 were then gel purified and ligated together to generate SS-SUN2$^{261-731}$-mCherry. Next, we digested SS-EGFP-SUN2$^{595-731}$ with the REs *Eco*NI and *Not*I, which cut SS-EGFP-SUN2$^{595-731}$ at sequences within the 3' end of the SUN2 luminal domain and mCherry from SS-SUN2$^{261-731}$-mCherry. Following gel purification, the digested vector and insert DNA fragments were ligated together to generate SS-EGFP-SUN2$^{595-731}$-mCherry.

Finally, the mCherry-SR-KASH2 construct was generated by PCR amplifying the cDNA encoding SR-KASH2 from the previously described mRFP1-SR-KASH construct (18) using the primers SR-KASH2-F and SR-KASH2-R. The purified PCR product was then digested alongside pmCherry-C1 with the REs *Kpn*I and *Xho*I. After gel purifying the digested PCR product and pmCherry-C1, they were ligated together to create the mCherry-SR-KASH2 construct.

Primer Name	DNA Sequence (5' → 3')	5' RE Site
mCherry-F	GTGGCTAGCGTGAGCAAGGGCGAGGAG	*Nhe*I
mCherry-R	CACGCTAGCCTTGTACAGCTCGTCCATGCC	*Nhe*I
SS-mCherry-KDEL-F	GCGCGCGGATCCATGAAGCTGGGCC	*Bam*HI
SS-mCherry-KDEL-R	CCCCGGGAATTCTTAAAGTTCATCCT	*Eco*RI
SS-EGFP-SUN2$^{261-731}$-ΔEGFP-F	TCCTGGTGGGCAGCAAAAG	-
SS-EGFP-SUN2$^{261-731}$-ΔEGFP-R	GCTAGCCACCGCCTG	-
SS-SUN2$^{261-731}$-F	AAAAAAAAGCTTATGAAGCTGGGCCGGG	*Hind*III
SS-SUN2$^{261-731}$-R	TTTTGTCGACTTGTGGGCAGGCTCTCCG	*Sal*I
SR-KASH2-F	AAAACTCGAGAAAGCTCCCAGCCGAGAC	*Xho*I
SR-KASH2-R	TTTGGTACCCTAGGTGGGAGGTGGCC	*Kpn*I

Table 6.S1: Primers used to generate the constructs used in this chapter. The F or R in the primer name refers to forward or reverse, respectively. RE cut sites are underlined. The sequence encoding the linker is bolded.

Chapter 7. Summary and Future Directions

Proteins within the NE play a critical role in a multitude of cellular functions (34) and are implicated in many human diseases (127). While previous work to quantify NE protein assembly was relegated to *in vitro* studies, which requires removal of the proteins from their native environment, this thesis lays the groundwork for using FFS to study NE protein complex assembly within the NE of living cells. The ability to quantitatively assess the oligomerization of NE proteins in their native environment represents a significant advance towards studying the mechanistic behavior of these proteins and the human diseases they are associated with.

The NE proved to be a challenging environment for performing FFS measurements. The presence of nuclear membrane undulations, which affect the local volume of the lumen, presented the first obstacle for extending FFS into the NE. Traditional FFS analysis techniques failed to properly account for the slow volume fluctuations, which led to systematic errors in determining the monomeric brightness with a calibration standard. Consequently, the correct interpretation of oligomerization from brightness data is compromised. To overcome this challenge we turned to the MSQ analysis technique which was previously developed to account for non-stationary signals undergoing slow decay (72). Through MSQ analysis we were able to identify and account for the slow fluctuation process associated with the local volume changes and recovered the expected behavior of the calibration standard. Further tests on model proteins, both membrane associated and luminal, supported the hypothesis that we were observing undulations from the nuclear membranes.

While we were able to identify an approximate timescale for these fluctuations and a mean amplitude, the mechanisms and properties behind nuclear membrane undulations should be further explored. By combining lateral scanning of the PSF with spatial-temporal correlation analysis, future work may gain insights into how, or whether, these undulations are structured spatially across the membrane. In particular, we may speculate about the existence of decreased fluctuation amplitude at points where protein complexes, such as the nuclear pore complex, directly couple both membranes together. Measurements of luminal proteins are currently being performed at varying temperatures to test whether the membrane undulations are mainly thermally driven. In addition, future studies may involve applying drugs to the cells to determine whether the membrane undulations are caused by active cellular processes.

The questions of whether these nuclear membrane undulations are present in other cells lines and if the properties we observed are universal have not been fully studied. Preliminary measurements in other human cell lines indicate the existence of the same slow process with similar fluctuation magnitude, but the number of cell lines tested has been fairly limited. For example, some cell lines contain shorter SUN proteins which may affect the spacing of the lumen (128). Finally, while the proportionality factor that relates to a ratio of the fluctuation amplitude and the mean height is fairly stable, some day to day variation has been observed. The cause of this variation is currently unknown, however there is evidence of cell cycle dependence on overall NE kinetics which requires further exploration (32).

The original formulation of MSQ analysis required unwieldy corrections due to detector artifacts, limiting its applicability for live cell measurements. The development of

the tsMSQ algorithm removed this disadvantage and simplified the analysis process. In conjunction, we developed a method for performing goodness-of-fit tests on the tsMSQ curves. We previously relied on finding a model with the fewest parameters that had no significant systematic residuals. By removing the self-correlations inherent to the tsMSQ curve, we obtained proper chi-squared statistics, thus providing statistical backing to our choice of fit models for each of the measured proteins. This method has the disadvantage of requiring a five-fold, or higher, increase in measurement time which can be difficult to obtain in living cells due to cellular motion or other slow changes in the cell environment. Further progress, not discussed in this thesis, was made in the form of a tsMSQ algorithm that relied on bootstrapping statistics in order to generate uncorrelated results from a single data set. Using the bootstrapping method a single measurement was sufficient to recover chi-squared statistics as well as determine uncertainties in the fit parameters. This method is developed and discussed in more detail in Hennen et al. (129).

The utility of the work laid out in this thesis was demonstrated through measurements of constituent proteins of the LINC complex. In particular, SUN2 had the benefit of being extensively studied through *in vitro* techniques and we were able to observe the formation of SUN2 trimers, predicted by *in vitro* results, for the first time in their native cellular environment. After demonstrating agreement between previous *in vitro* studies and our FFS experiments for SUN2, we measured the much less studied SUN1 protein and observed an unexpected oligomerization exceeding a trimer. This was a significant finding because, while an *in vitro* study speculated about the possibility of SUN1 tetramers (86), the similarities in the sequences of SUN1 and SUN2 had lead to the implicit assumption that SUN1 would be trimeric as well (130). Significantly, our observation of

higher order SUN1 oligomers inspired a computational modeling study that predicted the existence of interactions between pairs of SUN1 trimers, leading to the formation of a hexameric complex (52).The importance of performing measurements within the NE was also highlighted by the observation that a truncated SUN2 protein construct which formed trimers within the NE remained monomeric in the cytoplasm.

While SUN proteins are integral to the LINC complex, they must interact with nesprins in order to complete LINC complex assembly. We explored this critical step in LINC complex formation by extending tsMSQ to a dual-color technique. Measurements of SUN2 and the KASH domain of nesprin-2 identified the formation of heterocomplexes. While we were unable to observe a saturation in the binding curves and therefore could not identify the limiting stoichiometry, our data were consistent with the trimer-trimer model for this pair of proteins. This pilot study of the dual-color tsMSQ method is encouraging and suggests follow-up studies of LINC complex formation by dual-color tsMSQ would be valuable. In particular, studying the interactions with SUN1 and KASH may give further information on how it differs from SUN2, building off our single-color FFS study which identified differences in their limiting stoichiometry. This method also allows us to study the assembly of LINC complexes in the presence of mutated forms of SUN and KASH proteins that are associated with human disease which may potentially provide insight on the mechanisms behind these diseases. Futhermore, interactions between LINC complexes and the AAA+ ATPase torsinA have been observed *in vitro.* Dual-color tsMSQ may be used to identify precisely which components of the LINC complex torsinA may interact with while also studying the mechanisms through truncated or mutated protein constructs.

While this thesis lays much of the groundwork for studying proteins in the NE through FFS, there are still open questions in addition to those laid out earlier in this chapter. A significant obstacle for our technique has been the mobility of proteins. Proteins which are largely immobile over the course of the measurement experience photobleaching and provide no diffusion-based fluctuations as they do not pass through the PSF. Future work may utilize scanning FFS to overcome this limitation, essentially replacing the motion of the proteins through the PSF with motion of the PSF across the proteins. This may allow us to study full length SUN proteins, which have been shown to have very low mobility but may contain interactions not seen in the truncated proteins we have measured thus far. While further work is required to solve these open problems, the techniques presented in this thesis provide the first tool box for studying protein assembly within the NE of living cells.

References

1. Woese, C.R., O. Kandler, and M.L. Wheelis. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. U. S. A. 87: 4576–4579.
2. Pederson, T. 2011. The nucleus introduced. Cold Spring Harb. Perspect. Biol. 3.
3. Watson, M.L. 1955. THE NUCLEAR ENVELOPE. J. Biophys. Biochem. Cytol. 1: 257–270.
4. Martin, W.F., S. Garg, and V. Zimorski. 2015. Endosymbiotic theories for eukaryote origin. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 370: 20140330.
5. Fuerst, J.A. 2005. Intracellular compartmentation in planctomycetes. Annu. Rev. Microbiol. 59: 299–328.
6. Fuerst, J.A., and E. Sagulenko. 2011. Beyond the bacterium: planctomycetes challenge our concepts of microbial structure and function. Nat. Rev. Microbiol. 9: 403–413.
7. Wente, S.R., and M.P. Rout. 2010. The nuclear pore complex and nuclear transport. Cold Spring Harb. Perspect. Biol. 2: a000562.
8. Lin, D.H., T. Stuwe, S. Schilbach, E.J. Rundlet, T. Perriches, G. Mobbs, Y. Fan, K. Thierbach, F.M. Huber, L.N. Collins, A.M. Davenport, Y.E. Jeon, and A. Hoelz. 2016. Architecture of the symmetric core of the nuclear pore. Science. 352: aaf1015.

9. Panté, N., and M. Kann. 2002. Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. Mol. Biol. Cell. 13: 425–434.

10. Chook, Y.M., and K.E. Süel. 2011. Nuclear import by karyopherin-βs: recognition and inhibition. Biochim. Biophys. Acta. 1813: 1593–1606.

11. Speese, S.D., J. Ashley, V. Jokhi, J. Nunnari, R. Barria, Y. Li, B. Ataman, A. Koon, Y.-T. Chang, Q. Li, M.J. Moore, and V. Budnik. 2012. Nuclear envelope budding enables large ribonucleoprotein particle export during synaptic Wnt signaling. Cell. 149: 832–846.

12. Mettenleiter, T.C., F. Müller, H. Granzow, and B.G. Klupp. 2013. The way out: what we know and do not know about herpesvirus nuclear egress. Cell. Microbiol. 15: 170–178.

13. Margalit, A., S. Vlcek, Y. Gruenbaum, and R. Foisner. 2005. Breaking and making of the nuclear envelope. J. Cell. Biochem. 95: 454–465.

14. Hetzer, M.W. 2010. The Nuclear Envelope. Cold Spring Harb. Perspect. Biol. 2.

15. Prunuske, A.J., and K.S. Ullman. 2006. The nuclear envelope: form and reformation. Curr. Opin. Cell Biol. 18: 108–116.

16. Gundersen, G.G., and H.J. Worman. 2013. Nuclear positioning. Cell. 152: 1376–1389.

17. Lee, H.O., and C. Norden. 2013. Mechanisms controlling arrangements and movements of nuclei in pseudostratified epithelia. Trends Cell Biol. 23: 141–150.

18. Luxton, G.W.G., E.R. Gomes, E.S. Folker, E. Vintinner, and G.G. Gundersen. 2010. Linear arrays of nuclear envelope proteins harness retrograde actin flow for nuclear movement. Science. 329: 956–959.

19. Maniotis, A.J., C.S. Chen, and D.E. Ingber. 1997. Demonstration of mechanical connections between integrins, cytoskeletal filaments, and nucleoplasm that stabilize nuclear structure. Proc. Natl. Acad. Sci. U. S. A. 94: 849–854.

20. Gruenbaum, Y., A. Margalit, R.D. Goldman, D.K. Shumaker, and K.L. Wilson. 2005. The nuclear lamina comes of age. Nat. Rev. Mol. Cell Biol. 6: 21–31.

21. Dahl, K.N., S.M. Kahn, K.L. Wilson, and D.E. Discher. 2004. The nuclear envelope lamina network has elasticity and a compressibility limit suggestive of a molecular shock absorber. J. Cell Sci. 117: 4779–4786.

22. Wolf, K., M. Te Lindert, M. Krause, S. Alexander, J. Te Riet, A.L. Willis, R.M. Hoffman, C.G. Figdor, S.J. Weiss, and P. Friedl. 2013. Physical limits of cell migration: control by ECM space and nuclear deformation and tuning by proteolysis and traction force. J. Cell Biol. 201: 1069–1084.

23. Harada, T., J. Swift, J. Irianto, J.-W. Shin, K.R. Spinler, A. Athirasala, R. Diegmiller, P.C.D.P. Dingal, I.L. Ivanovska, and D.E. Discher. 2014. Nuclear lamin stiffness is a barrier to 3D migration, but softness can limit survival. J. Cell Biol. 204: 669.

24. Dahl, K.N., P. Scaffidi, M.F. Islam, A.G. Yodh, K.L. Wilson, and T. Misteli. 2006. Distinct structural and mechanical properties of the nuclear lamina in Hutchinson-Gilford progeria syndrome. Proc. Natl. Acad. Sci. U. S. A. 103: 10271–10276.

25. Friedl, P., K. Wolf, and J. Lammerding. 2011. Nuclear mechanics during cell migration. Curr. Opin. Cell Biol. 23: 55–64.

26. Isermann, P., and J. Lammerding. 2013. Nuclear mechanics and mechanotransduction in health and disease. Curr. Biol. CB. 23: R1113-1121.

27. Chen, Y., J.D. Müller, Q. Ruan, and E. Gratton. 2002. Molecular brightness characterization of EGFP in vivo by fluorescence fluctuation spectroscopy. Biophys. J. 82: 133–144.

28. Waugh, R., and E.A. Evans. 1979. Thermoelasticity of red blood cell membrane. Biophys. J. 26: 115–131.

29. Qu, D., G. Brotons, V. Bosio, A. Fery, T. Salditt, D. Langevin, and R. von Klitzing. 2007. Interactions across liquid thin films. Colloids Surf. Physicochem. Eng. Asp. 303: 97–109.

30. Kats, E.I., V.V. Lebedev, and S.V. Malinin. 1998. Dynamic fluctuation phenomena in double membrane films. J. Exp. Theor. Phys. 86: 1149–1155.

31. Jahed, Z., H. Shams, and M.R.K. Mofrad. 2015. A Disulfide Bond Is Required for the Transmission of Forces through SUN-KASH Complexes. Biophys. J. 109: 501–509.

32. Chu, F.-Y., S.C. Haley, and A. Zidovska. 2017. On the origin of shape fluctuations of the cell nucleus. Proc. Natl. Acad. Sci. : 201702226.

33. Dauer, W.T., and H.J. Worman. 2009. The nuclear envelope as a signaling node in development and disease. Dev. Cell. 17: 626–638.

34. Wilson, K.L., and J.M. Berk. 2010. The nuclear envelope at a glance. J. Cell Sci. 123: 1973–1978.

35. Crisp, M., Q. Liu, K. Roux, J.B. Rattner, C. Shanahan, B. Burke, P.D. Stahl, and D. Hodzic. 2006. Coupling of the nucleus and cytoplasm: role of the LINC complex. J. Cell Biol. 172: 41–53.

36. Liu, J., T. Rolef Ben-Shahar, D. Riemer, M. Treinin, P. Spann, K. Weber, A. Fire, and Y. Gruenbaum. 2000. Essential roles for Caenorhabditis elegans lamin gene in nuclear organization, cell cycle progression, and spatial organization of nuclear pore complexes. Mol. Biol. Cell. 11: 3937–3947.

37. Lei, K., X. Zhu, R. Xu, C. Shao, T. Xu, Y. Zhuang, and M. Han. 2012. Inner nuclear envelope proteins SUN1 and SUN2 play a prominent role in the DNA damage response. Curr. Biol. CB. 22: 1609–1615.

38. Schmit, F., M. Korenjak, M. Mannefeld, K. Schmitt, C. Franke, B. von Eyss, S. Gagrica, F. Hänel, A. Brehm, and S. Gaubatz. 2007. LINC, a human complex that is related to pRB-containing complexes in invertebrates regulates the expression of G2/M genes. Cell Cycle Georget. Tex. 6: 1903–1913.

39. Mekhail, K., and D. Moazed. 2010. The nuclear envelope in genome organization, expression and stability. Nat. Rev. Mol. Cell Biol. 11: 317–328.

40. Barbosa, A.D., H. Sembongi, W.-M. Su, S. Abreu, F. Reggiori, G.M. Carman, S. Siniossoglou, and R.G. Parton. 2015. Lipid partitioning at the nuclear envelope controls membrane biogenesis. Mol. Biol. Cell. 26: 3641–3657.

41. Worman, H.J., and W.T. Dauer. 2014. The nuclear envelope: an intriguing focal point for neurogenetic disease. Neurother. J. Am. Soc. Exp. Neurother. 11: 764–772.

42. Smith, E.M., P.J. Macdonald, Y. Chen, and J.D. Mueller. 2014. Quantifying Protein-Protein Interactions of Peripheral Membrane Proteins by Fluorescence Brightness Analysis. Biophys. J. 107: 66–75.

43. Macdonald, P.J., Y. Chen, X. Wang, Y. Chen, and J.D. Mueller. 2010. Brightness Analysis by Z-Scan Fluorescence Fluctuation Spectroscopy for the Study of Protein Interactions within Living Cells. Biophys. J. 99: 979–988.

44. Chen, Y., L.-N. Wei, and J.D. Müller. 2003. Probing protein oligomerization in living cells with fluorescence fluctuation spectroscopy. Proc. Natl. Acad. Sci. U. S. A. 100: 15492–15497.

45. Horn, H.F. 2014. LINC complex proteins in development and disease. Curr. Top. Dev. Biol. 109: 287–321.

46. Chang, W., H.J. Worman, and G.G. Gundersen. 2015. Accessorizing and anchoring the LINC complex for multifunctionality. J. Cell Biol. 208: 11–22.

47. Méjat, A., and T. Misteli. 2010. LINC complexes in health and disease. Nucleus. 1: 40–52.

48. Wang, N., J.D. Tytell, and D.E. Ingber. 2009. Mechanotransduction at a distance: mechanically coupling the extracellular matrix with the nucleus. Nat. Rev. Mol. Cell Biol. 10: 75–82.

49. Sosa, B.A., A. Rothballer, U. Kutay, and T.U. Schwartz. 2012. LINC complexes form by binding of three KASH peptides to domain interfaces of trimeric SUN proteins. Cell. 149: 1035–1047.

50. Haque, F., D. Mazzeo, J.T. Patel, D.T. Smallwood, J.A. Ellis, C.M. Shanahan, and S. Shackleton. 2010. Mammalian SUN protein interaction networks at the inner nuclear membrane and their role in laminopathy disease processes. J. Biol. Chem. 285: 3487–3498.

51. Nie, S., H. Ke, F. Gao, J. Ren, M. Wang, L. Huo, W. Gong, and W. Feng. 2016. Coiled-Coil Domains of SUN Proteins as Intrinsic Dynamic Regulators. Struct. Lond. Engl. 1993. 24: 80–91.

52. Jahed, Z., D. Fadavi, U.T. Vu, E. Asgari, G.W.G. Luxton, and M.R.K. Mofrad. 2018. Molecular Insights into the Mechanisms of SUN1 Oligomerization in the Nuclear Envelope. Biophys. J. 114: 1190–1203.

53. Hennen, J., K.-H. Hur, C.A. Saunders, G.W.G. Luxton, and J.D. Mueller. 2017. Quantitative Brightness Analysis of Protein Oligomerization in the Nuclear Envelope. Biophys. J. 113: 138–147.

54. Hennen, J., C.A. Saunders, J.D. Mueller, and G.W.G. Luxton. 2018. Fluorescence Fluctuation Spectroscopy Reveals Differential SUN Protein Oligomerization In Living Cells. Mol. Biol. Cell. .

55. Hennen, J., K.-H. Hur, S.R. Karuka, G.W.G. Luxton, and J.D. Mueller. 2019. Protein oligomerization and mobility within the nuclear envelope evaluated by the time-shifted mean-segmented Q factor. Methods San Diego Calif. 157: 28–41.

56. Berland, K.M., P.T. So, and E. Gratton. 1995. Two-photon fluorescence correlation spectroscopy: method and application to the intracellular environment. Biophys. J. 68: 694–701.

57. Maiti, S., U. Haupts, and W.W. Webb. 1997. Fluorescence correlation spectroscopy: diagnostics for sparse molecules. Proc. Natl. Acad. Sci. U. S. A. 94: 11753–11757.

58. Chen, Y., J.D. Müller, P.T. So, and E. Gratton. 1999. The photon counting histogram in fluorescence fluctuation spectroscopy. Biophys. J. 77: 553–567.

59. Chen, Y., and J.D. Müller. 2007. Determining the stoichiometry of protein heterocomplexes in living cells with fluorescence fluctuation spectroscopy. Proc. Natl. Acad. Sci. U. S. A. 104: 3147–3152.

60. Denk, W., J.H. Strickler, and W.W. Webb. 1990. Two-photon laser scanning fluorescence microscopy. Science. 248: 73–76.

61. Heinze, K.G., A. Koltermann, and P. Schwille. 2000. Simultaneous two-photon excitation of distinct labels for dual-color fluorescence crosscorrelation analysis. Proc. Natl. Acad. Sci. U. S. A. 97: 10377–10382.

62. Smith, E.M., J. Hennen, Y. Chen, and J.D. Mueller. 2015. In situ quantification of protein binding to the plasma membrane. Biophys. J. 108: 2648–2657.

63. Wu, B., Y. Chen, and J.D. Müller. 2010. Heterospecies partition analysis reveals binding curve and stoichiometry of protein interactions in living cells. Proc. Natl. Acad. Sci. U. S. A. 107: 4117–4122.

64. Hennen, J., I. Angert, K.-H. Hur, G.W. Gant Luxton, and J.D. Mueller. 2018. Investigating LINC Complex Protein Homo-oligomerization in the Nuclear Envelopes of Living Cells Using Fluorescence Fluctuation Spectroscopy. In: Gundersen GG, HJ Worman, editors. The LINC Complex: Methods and Protocols. New York, NY: Springer New York. pp. 121–135.

65. Mandel, L. 1979. Sub-Poissonian photon statistics in resonance fluorescence. Opt. Lett. 4: 205–207.

66. Müller, J.D. 2004. Cumulant Analysis in Fluorescence Fluctuation Spectroscopy. Biophys. J. 86: 3981–3992.

67. Hillesheim, L.N., and J.D. Müller. 2003. The photon counting histogram in fluorescence fluctuation spectroscopy with non-ideal photodetectors. Biophys. J. 85: 1948–1958.

68. Sanchez-Andres, A., Y. Chen, and J.D. Müller. 2005. Molecular brightness determined from a generalized form of Mandel's Q-parameter. Biophys. J. 89: 3531–3547.

69. Müller, J.D., Y. Chen, and E. Gratton. 2003. Fluorescence correlation spectroscopy. Methods Enzymol. 361: 69–92.

70. Qian, H. 1990. On the statistics of fluorescence correlation spectroscopy. Biophys. Chem. 38: 49–57.

71. Wang, Y.-H., A. Hariharan, G. Bastianello, Y. Toyama, G.V. Shivashankar, M. Foiani, and M.P. Sheetz. 2017. DNA damage causes rapid accumulation of phosphoinositides for ATR signaling. Nat. Commun. 8: 2118.

72. Hur, K.-H., and J.D. Mueller. 2015. Quantitative Brightness Analysis of Fluorescence Intensity Fluctuations in E. Coli. PLoS ONE. 10.

73. Smith, E.M., J. Hennen, Y. Chen, and J.D. Mueller. 2015. Z-scan fluorescence profile deconvolution of cytosolic and membrane-associated protein populations. Anal. Biochem. 480: 11–20.

74. Crisp, M., Q. Liu, K. Roux, J.B. Rattner, C. Shanahan, B. Burke, P.D. Stahl, and D. Hodzic. 2006. Coupling of the nucleus and cytoplasm: role of the LINC complex. J. Cell Biol. 172: 41–53.

75. Wu, B., and J.D. Müller. 2005. Time-Integrated Fluorescence Cumulant Analysis in Fluorescence Fluctuation Spectroscopy. Biophys. J. 89: 2721–2735.

76. Hur, K.-H., P.J. Macdonald, S. Berk, C.I. Angert, Y. Chen, and J.D. Mueller. 2014. Quantitative measurement of brightness from living cells in the presence of photodepletion. PloS One. 9: e97440.

77. Goodchild, R.E., and W.T. Dauer. 2004. Mislocalization to the nuclear envelope: an effect of the dystonia-causing torsinA mutation. Proc. Natl. Acad. Sci. U. S. A. 101: 847–852.

78. Seifert, U. 1997. Configurations of fluid membranes and vesicles. Adv. Phys. 46: 13–137.

79. Vander Heyden, A.B., T.V. Naismith, E.L. Snapp, and P.I. Hanson. 2011. Static retention of the lumenal monotopic membrane protein torsinA in the endoplasmic reticulum. EMBO J. 30: 3217–3231.

80. Ostlund, C., E.S. Folker, J.C. Choi, E.R. Gomes, G.G. Gundersen, and H.J. Worman. 2009. Dynamics and molecular interactions of linker of nucleoskeleton and cytoskeleton (LINC) complex proteins. J. Cell Sci. 122: 4099–4108.

81. Meinke, P., and E.C. Schirmer. 2015. LINC'ing form and function at the nuclear envelope. FEBS Lett. 589: 2514–2521.

82. Zhang, X., K. Lei, X. Yuan, X. Wu, Y. Zhuang, T. Xu, R. Xu, and M. Han. 2009. SUN1/2 and Syne/Nesprin-1/2 complexes connect centrosome to the nucleus during neurogenesis and neuronal migration in mice. Neuron. 64: 173–187.

83. Lei, K., X. Zhang, X. Ding, X. Guo, M. Chen, B. Zhu, T. Xu, Y. Zhuang, R. Xu, and M. Han. 2009. SUN1 and SUN2 play critical but partially redundant roles in anchoring nuclei in skeletal muscle cells in mice. Proc. Natl. Acad. Sci. U. S. A. 106: 10207–10212.

84. Stewart-Hutchinson, P.J., C.M. Hale, D. Wirtz, and D. Hodzic. 2008. Structural requirements for the assembly of LINC complexes and their function in cellular mechanical stiffness. Exp. Cell Res. 314: 1892–1905.

85. Ding, X., R. Xu, J. Yu, T. Xu, Y. Zhuang, and M. Han. 2007. SUN1 is required for telomere attachment to nuclear envelope and gametogenesis in mice. Dev. Cell. 12: 863–872.

86. Lu, W., J. Gotzmann, L. Sironi, V.-M. Jaeger, M. Schneider, Y. Lüke, M. Uhlén, C.A.-K. Szigyarto, A. Brachner, J. Ellenberg, R. Foisner, A.A. Noegel, and I. Karakesisoglou. 2008. Sun1 forms immobile macromolecular assemblies at the nuclear envelope. Biochim. Biophys. Acta. 1783: 2415–2426.

87. Talamas, J.A., and M.W. Hetzer. 2011. POM121 and Sun1 play a role in early steps of interphase NPC assembly. J. Cell Biol. 194: 27–37.

88. Zhou, Z., X. Du, Z. Cai, X. Song, H. Zhang, T. Mizuno, E. Suzuki, M.R. Yee, A. Berezov, R. Murali, S.-L. Wu, B.L. Karger, M.I. Greene, and Q. Wang. 2012. Structure of Sad1-UNC84 homology (SUN) domain defines features of molecular bridge in nuclear envelope. J. Biol. Chem. 287: 5317–5326.

89. Slaughter, B.D., and R. Li. 2010. Toward Quantitative "In Vivo Biochemistry" with Fluorescence Fluctuation Spectroscopy. Mol. Biol. Cell. 21: 4306–4311.

90. Efron, B. 1979. Bootstrap Methods: Another Look at the Jackknife. Ann. Stat. 7: 1–26.

91. Franke, W.W., U. Scheer, G. Krohne, and E.D. Jarasch. 1981. The nuclear envelope and the architecture of the nuclear periphery. J. Cell Biol. 91: 39s–50s.

92. Saunders, C.A., N.J. Harris, P.T. Willey, B.M. Woolums, Y. Wang, A.J. McQuown, A. Schoenhofen, H.J. Worman, W.T. Dauer, G.G. Gundersen, and G.W.G. Luxton.

2017. TorsinA controls TAN line assembly and the retrograde flow of dorsal perinuclear actin cables during rearward nuclear movement. J. Cell Biol. 216: 657–674.

93. Chen, Y., J. Johnson, P. Macdonald, B. Wu, and J.D. Mueller. 2010. Observing protein interactions and their stoichiometry in living cells by brightness analysis of fluorescence fluctuation experiments. Methods Enzymol. 472: 345–363.

94. Macdonald, P., J. Johnson, E. Smith, Y. Chen, and J.D. Mueller. 2013. Brightness analysis. Methods Enzymol. 518: 71–98.

95. Ellgaard, L., and A. Helenius. 2003. Quality control in the endoplasmic reticulum. Nat. Rev. Mol. Cell Biol. 4: 181–191.

96. Horn, H.F., D.I. Kim, G.D. Wright, E.S.M. Wong, C.L. Stewart, B. Burke, and K.J. Roux. 2013. A mammalian KASH domain protein coupling meiotic chromosomes to the cytoskeleton. J. Cell Biol. 202: 1023–1039.

97. Liu, Q., N. Pante, T. Misteli, M. Elsagga, M. Crisp, D. Hodzic, B. Burke, and K.J. Roux. 2007. Functional association of Sun1 with nuclear pore complexes. J. Cell Biol. 178: 785–798.

98. Lottersberger, F., R.A. Karssemeijer, N. Dimitrova, and T. de Lange. 2015. 53BP1 and the LINC Complex Promote Microtubule-Dependent DSB Mobility and DNA Repair. Cell. 163: 880–893.

99. Kracklauer, M.P., J. Link, and M. Alsheimer. 2013. LINCing the nuclear envelope to gametogenesis. Curr. Top. Dev. Biol. 102: 127–157.

100. Bag, N., and T. Wohland. 2014. Imaging fluorescence fluctuation spectroscopy: new tools for quantitative bioimaging. Annu. Rev. Phys. Chem. 65: 225–248.

101. Chen, Y., M. Tekmen, L. Hillesheim, J. Skinner, B. Wu, and J.D. Müller. 2005. Dual-Color Photon-Counting Histogram. Biophys. J. 88: 2177–2192.

102. Wu, B., Y. Chen, and J.D. Müller. 2006. Dual-Color Time-Integrated Fluorescence Cumulant Analysis. Biophys. J. 91: 2687–2698.

103. Göb, E., J. Schmitt, R. Benavente, and M. Alsheimer. 2010. Mammalian Sperm Head Formation Involves Different Polarization of Two Novel LINC Complexes. PLoS ONE. 5.

104. Sosa, B.A., U. Kutay, and T.U. Schwartz. 2013. Structural insights into LINC complexes. Curr. Opin. Struct. Biol. 23: 285–291.

105. Calvi, A., A.S.W. Wong, G. Wright, E.S.M. Wong, T.H. Loo, C.L. Stewart, and B. Burke. 2015. SUN4 is essential for nuclear remodeling during mammalian spermiogenesis. Dev. Biol. 407: 321–330.

106. Shao, X., H.A. Tarnasky, J.P. Lee, R. Oko, and F.A. van der Hoorn. 1999. Spag4, a novel sperm protein, binds outer dense-fiber protein Odf1 and localizes to microtubules of manchette and axoneme. Dev. Biol. 211: 109–123.

107. Tarnasky, H., D. Gill, S. Murthy, X. Shao, D.J. Demetrick, and F.A. van der Hoorn. 1998. A novel testis-specific gene, SPAG4, whose product interacts specifically with outer dense fiber protein ODF27, maps to human chromosome 20q11.2. Cytogenet. Cell Genet. 81: 65–67.

108. Digman, M.A., R. Dalal, A.F. Horwitz, and E. Gratton. 2008. Mapping the Number of Molecules and Brightness in the Laser Scanning Microscope. Biophys. J. 94: 2320–2332.

109. Wiseman, P.W., J.A. Squier, M.H. Ellisman, and K.R. Wilson. 2000. Two-photon image correlation spectroscopy and image cross-correlation spectroscopy. J. Microsc. 200: 14–25.

110. Sankaran, J., M. Manna, L. Guo, R. Kraut, and T. Wohland. 2009. Diffusion, Transport, and Cell Membrane Organization Investigated by Imaging Fluorescence Cross-Correlation Spectroscopy. Biophys. J. 97: 2630–2639.

111. Kim, D.I., K.C. Birendra, and K.J. Roux. 2015. Making the LINC: SUN and KASH protein interactions. Biol. Chem. 396: 295–310.

112. Wang, Q., X. Du, Z. Cai, and M.I. Greene. 2006. Characterization of the structures involved in localization of the SUN proteins to the nuclear envelope and the centrosome. DNA Cell Biol. 25: 554–562.

113. Jameson, D.M., J.A. Ross, and J.P. Albanesi. 2009. Fluorescence fluctuation spectroscopy: ushering in a new age of enlightenment for cellular dynamics. Biophys. Rev. 1: 105–118.

114. Petersen, N.O., P.L. Höddelius, P.W. Wiseman, O. Seger, and K.E. Magnusson. 1993. Quantitation of membrane receptor distributions by image correlation spectroscopy: concept and application. Biophys. J. 65: 1135–1146.

115. Schwille, P., F.J. Meyer-Almes, and R. Rigler. 1997. Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution. Biophys. J. 72: 1878–1886.

116. Wiedenmann, J., F. Oswald, and G.U. Nienhaus. 2009. Fluorescent proteins for live cell imaging: opportunities, limitations, and challenges. IUBMB Life. 61: 1029–1042.

117. Chen, Y., L.-N. Wei, and J.D. Müller. 2005. Unraveling protein-protein interactions in living cells with fluorescence fluctuation brightness analysis. Biophys. J. 88: 4366–4377.

118. Dauer, W.T., and H.J. Worman. 2009. The nuclear envelope as a signaling node in development and disease. Dev. Cell. 17: 626–638.

119. Lombardi, M.L., D.E. Jaalouk, C.M. Shanahan, B. Burke, K.J. Roux, and J. Lammerding. 2011. The interaction between nesprins and sun proteins at the nuclear envelope is critical for force transmission between the nucleus and cytoskeleton. J. Biol. Chem. 286: 26743–26753.

120. Wang, W., Z. Shi, S. Jiao, C. Chen, H. Wang, G. Liu, Q. Wang, Y. Zhao, M.I. Greene, and Z. Zhou. 2012. Structural insights into SUN-KASH complexes across the nuclear envelope. Cell Res. 22: 1440–1452.

121. Hur, K.-H., Y. Chen, and J.D. Mueller. 2016. Characterization of Ternary Protein Systems In Vivo with Tricolor Heterospecies Partition Analysis. Biophys. J. 110: 1158–1167.

122. Wu, B., Y. Chen, and J.D. Müller. 2009. Fluorescence fluctuation spectroscopy of mCherry in living cells. Biophys. J. 96: 2391–2404.

123. Liu, Z., A. Zolkiewska, and M. Zolkiewski. 2003. Characterization of human torsinA and its dystonia-associated mutant form. Biochem. J. 374: 117–122.

124. Bacia, K., and P. Schwille. 2007. Practical guidelines for dual-color fluorescence cross-correlation spectroscopy. Nat. Protoc. 2: 2842–2856.

125. Patterson, G., R.N. Day, and D. Piston. 2001. Fluorescent protein spectra. J. Cell Sci. 114: 837–838.

126. Bacia, K., I.V. Majoul, and P. Schwille. 2002. Probing the endocytic pathway in live cells using dual-color fluorescence cross-correlation analysis. Biophys. J. 83: 1184–1193.

127. Worman, H.J., C. Östlund, and Y. Wang. 2010. Diseases of the Nuclear Envelope. Cold Spring Harb. Perspect. Biol. 2.

128. Cain, N.E., and D.A. Starr. 2015. SUN proteins and nuclear envelope spacing. Nucl. Austin Tex. 6: 2–7.

129. Hennen, J., K.-H. Hur, and J.D. Mueller. 2019. Quantitative modeling of self-oligomerization of proteins in the nuclear envelope by fluorescence fluctuation analysis. Anal. Biochem. 582: 113359.

130. Hieda, M. 2017. Implications for Diverse Functions of the LINC Complexes Based on the Structure. Cells. 6.