Dissertation

Submitted to the

Combined Faculties for Natural Sciences and for Mathematics

of Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

M.Sc. Molecular Biosciences

Iker Valle Aramburu

San Sebastian, Spain

27.11.2017

FG-Nucleoporins and the nucleocytoplasmic transport; two distinct molecular mechanisms of multivalent interactions

Referees: Dr. Christoph Müller

Prof. Dr. Frauke Melchior

Summary

FG-Nucleoporins (FG-Nups) are intrinsically disordered proteins (IDPs) located at the nuclear pore complex (NPC) where they form the permeability barrier of the NPC. The selective transport of cargoes, with a molecular weight above 40 kDa (>4 nm), across the NPC is mediated by nuclear transport receptors (NTRs). NTR mediated nucleocytoplasmic transport requires the direct interaction between NTRs and the phenylalanine-glycine motifs (FG-motifs) present in the FG-Nups of the NPC barrier.

The selective crossing of NTRs through the ~30 nm permeability barrier of NPC has been shown to occur on the millisecond timescale. However, these fast transport times don't seem to correlate with the high specificities reported for several FG-Nup/NTR complexes, which are often associated with long lasting complexes. Thus, the understanding of the fast and selective nucleocytoplasmic transport resides in deciphering the molecular basis of the interaction between FG-Nups and NTRs.

In my PhD thesis I have focused on understanding the binding mechanism between FG-Nups and NTRs by studying the structure, dynamics and kinetics of multiple FG-Nups upon binding to NTRs. I have shown, with different biophysical techniques like; multiparameter single-molecule fluorescence resonance energy transfer (smFRET), fluorescence correlation spectroscopy (FCS) and stopped-flow spectroscopy that many FG-Nups are able to bind NTRs forming highly dynamic complexes. Moreover, most FG-Nups engage with different NTRs without undergoing a conformational which I have shown that it is linked to diffusion limited binding not driven by long-range electrostatic interactions. The work presented in this thesis contributed to the proposal of a model where FG-Nups interact with multiple minimalistic low affinity binding motifs with different NTR binding pockets with ultrafast kinetics allowing the fast NPC crossing of NTRs. This novel binding mechanism is due to the high dynamics of IDPs and can also explain why IDPs have enriched the development of higher organisms.

Exceptionally, Nup214FG which is not forming part of the central FG-Nup barrier binds to NTRs like CRM1 or Importinβ with a different binding mechanism, where the binding is coupled to a conformational change. Moreover, these different FG-Nup/NTR binding mechanisms don't seem to be affected by the glycosylation of FG-Nups, which is a highly abundant posttranslational modification in the metazoan FG-Nups located at the NPC.

In addition, motivated by the possibility of translating the *in vitro* results to a physiological environment and be able to study the structure of FG-Nups *in situ* at the NPC, I am developing a strategy that would improve the study of the structure and dynamics of the disordered regions of the FG-Nups in cells, paving the way toward the intracellular study of IDPs.

Zusammenfassung

FG-Nukleoporine (FG-Nups) sind intrinsisch ungeordnete Proteine (IDPs), die sich im Kernporenkomplex (NPC) befinden, und dort die Permeabilitätsbarriere des NPCs formen. Der selektive Transport von Kargomolekülen grösser als ~40 kDa durch den NPC, wird durch Kerntransportrezeptoren (NTRs) bewerkstelligt. NTRs abhängiger nukleocytoplasmatischer Transport erfordert eine direkte Interaktion von NTRs mit Phenylalanin-Glycin Motiven (FG-Motiven), die in den FG-Nups der NPC-Barriere vorhanden sind. Es wurde gezeigt, dass der selektive Übergang der NTRs durch die ~ 30 nm NPC-Barriere im Millisekundenbereich abläuft. Diese schnellen Transportzeiten stehen im Konflikt mit den hohen Spezifitäten, die für einige FG-Nup/NTR-Komplexe berichtet wurden. Um nun diesen schnellen und selektiven nukleocytoplasmatischen Transport zu verstehen, ist die Entschlüsselung der molekularen Grundlage der Interaktion zwischen FG-Nups und NTRs notwendig.

Während meiner Doktorarbeit habe ich mich darauf fokusiert, den Bindungsmechanismus zwischen FG-Nups und NTRs durch die Untersuchung der Struktur, Dynamik und Kinetik verschiedener FG-NUPs bei der Bindung an verschiedene NTRs, zu verstehen. Durch die Verwendung verschiedener biophysikalischer Techniken, wie Multiparameter Einzelmolekül-Fluoreszenz-Resonanz-Energietransfer (smFRET), Fluoreszenzkorrelationsspektroskopie (FCS) und Stopped-Flow-Spektroskopie, habe ich gezeigt, dass viele FG-Nups durch die Bildung eines hochgradig dynamischen Komplexes in der Lage sind NTRs zu binden. Ausserdem binden die meisten FG-Nups mit verschiedenen NTRs ohne eine signifikante Konformationsänderung zu vollziehen. Ich konnte zeigen, dass die Bindungsgeschwindigkeit diffussions-limitiert ist und nicht durch elektrostatische Interaktionen verursacht wird. Die Ergebnisse dieser Doktorarbeit trugen zu einem Modellvorschlag bei, bei dem FG-Nups mit mehrfach minimalistischen gering affinen Bindungsmotiven mit verschiedenen Bindungstaschen der NTRs mit ultraschnellen Kinetiken interagieren und damit den schnellen aber spezifischen Transport der NTRs durch den NPC erlauben. Dieser neuartige Bindungsmechanismus kommt durch die hohe Dynamik von IDPs zu Stande, und kann auch erklären warum IDPs sich in der Entwicklung höherartiger Organismen angereichert haben.

Wie ich weiterhin zeigen konnte, bindet Nup214FG, welches nicht Teil der zentralen FG-Barriere ist, NTRs, wie CRM1 oder Importinβ, dagegen durch einen anderen Bindungsmechanismus. Bei diesem ist die Bindung an eine Konformationsänderung des ungefalteten Proteins gekoppelt. Des Weiteren, scheinen diese unterschiedlichen FG-Nup/NTR Bindungsmechanismen nicht durch die Glykosylierung der FG-Nups, einer sehr häufig vorkommenden posttranslationalen Modifikation von Metazoa FG-Nups, beeinflusst zu werden.

Um die *in vitro* Ergebnisse in eine physiologische Umgebung übertragen zu können, bedarf es neuartiger Techniken um die Struktur der FG-Nups *in situ* am NPC zu untersuchen. Ich entwickelte daher eine Strategie, die die Studie der Struktur und der Dynamiken von ungeordneten Regionen von FG-Nups in Zellen möglichen machen wird, um den Weg zu intrazellulären Studien von IDPs zu ebnen.

Publication list

1. Mercadante, D.; Wagner, J.A.; **Aramburu, I.V.**; Lemke, E.A.; Grater, F., Sampling long versus short range interactions defines the ability of force fields to reproduce the dynamics of intrinsically disordered proteins. *Journal of chemical theory and computation* 2017.
2. **Aramburu, I.V.**; Lemke, E.A., Floppy but not sloppy: Interaction mechanism of FG-nucleoporins and nuclear transport receptors. *Seminars in cell & developmental biology* 2017, *68*, 34-41.
3. Kozma, E.; Nikic, I.; Varga, B.R.; **Aramburu, I.V.**; Kang, J.H.; Fackler, O.T.; Lemke, E.A.; Kele, P., Hydrophilic trans-cyclooctenylated noncanonical amino acids for fast intracellular protein labeling. *Chembiochem : a European journal of chemical biology* 2016, *17*, 1518-1524.
4. Milles, S*.; Mercadante, D*.; **Aramburu, I.V.***; Jensen, M.R.; Banterle, N.; Koehler, C.; Tyagi, S.; Clarke, J.; Shammas, S.L.; Blackledge, M., *et al.*, Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors. *Cell* **2015**, *163*, 734-745.
5. Hoffmann, J.E.; Plass, T.; Nikic, I.; **Aramburu, I.V.**; Koehler, C.; Gillandt, H.; Lemke, E.A.; Schultz, C., Highly stable trans-cyclooctene amino acids for live-cell labeling. *Chemistry* **2015**, *21*, 12266-12270.
6. Nikic, I.; Kang, J.H.; Girona, G.E.; **Aramburu, I.V.**; Lemke, E.A., Labeling proteins on live mammalian cells using click chemistry. *Nature protocols* **2015**, *10*, 780-791.
7. Lemke, E.A.; Schultz, C.; Plass, T.;Nikic, I.; Hoffmann, J.E.; **Aramburu, I.V.**;, Multiple cycloaddition reactions for labeling of molecules. US Patent App. 15/111,138, 2015

*Equal contribution

Table of content

List of figures

List of abbreviations

aa	Amino acid
AcF	p-Acetylphenylalanine
ACN	Acetonitrile
BME	β-mercaptoethanol
BOC-Lys	t-butyloxycarbonyl-lysine
BSA	Bovine serum albumin
BV	Bead volume
CBD	Chitin binding domain
CXH	Cycloheximide
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FCS	Fluorescence correlation spectroscopy
FG	Phenylalanine-glycine
FG-Nup	FG-Nucleoporin
FRET	Fluorescence resonance energy transfer
GCE	Genetic code expansion
GdmHCl	Guanidimium hydrochloride
GFP	Green fluorescent protein
GTP	Guanosine-5'-triphosphat
IDP	Intrinsically disordered protein
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	Kilo Dalton
LB medium	Lysogeny broth
M	Molar
MD	Molecular dynamics
MgAc	Magnesium acetate
ncAA	Non-canonical amino acid
NES	Nuclear export signal
NLS	Nuclear localization signal
NPC	Nuclear pore complex
NTR	Nuclear transport receptor
OD	Optical density
PBS	Phosphate buffer saline
PDB	Protein data bank
PEI	Polyethylenimine
PFA	Paraformaldehyde
PMSF	Phenylmethanesulfonyl fluoride
PMSF	Phenyl methane sulfonyl fluoride
PylRS	Pyrrolysine amino-acyl tRNA synthetase
RT	Room temperature
SDS	Sodium dodecyl sulfate
smFRET	Single molecule FRET
SRM	Super-resolution microscopy
TB medium	Terrific broth
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
UDP-GlcNAc	Uridine 5′-diphospho-N-acetylglucosamine
WB	Western blot

Chapter 1

Introduction

1. Introduction

1.1. Intrinsically disordered proteins, key constituents of the cell

1.1.1. Characteristics of intrinsically disordered proteins

The thermodynamic hypothesis known as Anfinsen's postulate proposed that the three dimensional structure of a native protein under physiological conditions is the one with lowest Gibbs free energy and that this conformation is determined by the amino acid sequence and its inter-atomic interactions in a given environment (Anfinsen, 1973). The technological and biochemical advances achieved at that time led to the structural determination of proteins like myoglobin, hemoglobin and lysozyme (Blake et al., 1965; Kendrew et al., 1958; Perutz et al., 1960), which were key for postulating the structure-function paradigm. The structure-function paradigm hypothesizes that the function of a protein is related to its structure which is encoded in the amino acid sequence. Nevertheless, since more than two decades ago the increase in the identification and biochemical characterization of unstructured proteins has revealed that they are functionally relevant cells (Boesch et al., 1978; Kriwacki et al., 1996; Nolte et al., 1998; Torchia et al., 1997). In addition, it has been shown that a large number of genes contain sequences that do not code for folded proteins but for long amino acid stretches that were suggested to be unfolded in solution or to have an unknown conformation (Dunker et al., 1998; Dunker et al., 2000). These stretches have been termed as intrinsically disordered proteins (IDPs) or intrinsically disordered domains (IDDs): for the sake of simplicity throughout this PhD thesis I will be referring to all as IDPs (Dunker et al., 2013; Wright and Dyson, 1999). Genome-wide bioinformatic studies have shown that IDPs are present in all branches of the life kingdom (Dunker et al., 2000; Ward et al., 2004) and that the proportion of proteins with disordered regions increases with increasing complexity of the organisms (Dunker et al., 2002; Ward et al., 2004). For example, around 40% of the human proteome has been reported to contain long disordered regions (Pentony et al., 2010) whereas only ~8% is present in the proteome of *E.coli* indicating that IDPs have been favored throughout evolution (Ward et al., 2004).

The structural and biophysical characteristics of IDPs and folded proteins are encoded in the amino acid sequence. Polypeptides that are folded under physiological conditions usually contain a large fraction of hydrophobic residues. However, a characteristic feature of the polypeptide sequence of IDPs is a low proportion of hydrophobic residues along with a high proportion of charged or polar amino acids, which enables these sequences to populate an ensemble of different conformations under physiological conditions (Figure 1.1) (Uversky et al., 2000). Moreover, sequence determinants such as fraction of polar residues, net charge content and charge distribution can modulate the degree of extension or collapse of

the disordered polypeptides. For example, proteins that have a high net charge are usually populating a more expanded conformational ensemble. However, depending on the charge patterning, the attraction of opposite charges can also lead to the collapse of the polypeptide chain (Das and Pappu, 2013; Mao et al., 2010; Muller-Spath et al., 2010). To better understand the disordered ensemble of conformations that an IDP can populate we could compare IDPs and folded proteins in terms of their free energy landscapes. The free energy landscape of a folded protein is often defined by one or more local minima which are clearly separated from the less energetically favorable states by large free energy barriers, kinetically trapping the protein in the most stable conformation (Figure 1.1). The free energy landscape of IDPs however, will contain multiple local energy minima where the free energy barrier separating the different local minima will be low, leading to the formation of an ensemble of rapidly interconverting conformations (Boehr et al., 2009; Uversky, 2013).

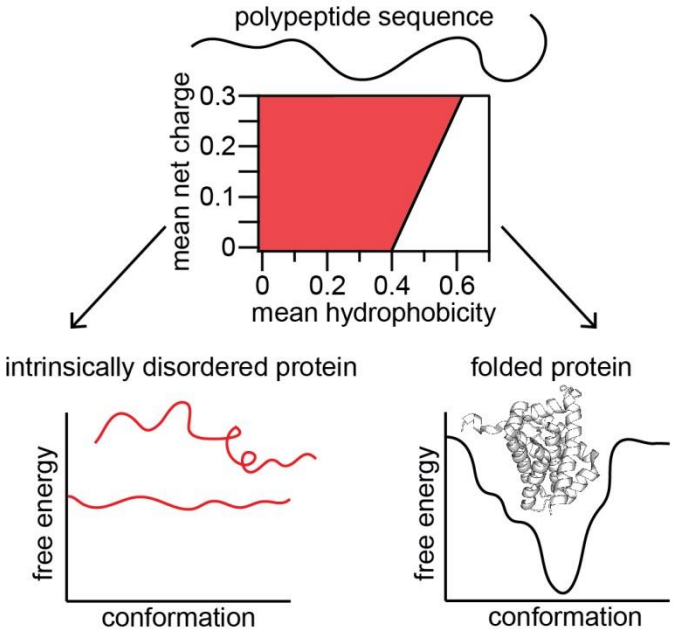

Figure 1.1. Scheme representing the sequence-structure relationship between folded proteins and IDPs: Depending on the sequence composition and distribution of the different amino acids a polypeptide sequence will adopt a folded state under physiological conditions or it will remain unfolded. High net charge and low mean hydrophobicity is associated with IDPs (red area). High mean hydrophobicity values are associated with folded proteins (white area). Representation of free energy landscape for an IDP contains multiple local minima separated by small free energy barriers (red) and a folded protein (black) often have large free energy barriers between different local minima maintaining the protein on a stable conformation.

1.1.2. Binding modes of intrinsically disordered proteins

Generally, protein-protein interactions take place spontaneously when the process is energetically and entropically favored (negative Gibbs free energy values). This means that the difference in the Gibbs free energy values of the bound form compared to the unbound form should be negative. The change in the

free energy is related to two thermodynamic parameters (besides temperature, T) as described in the Gibbs-Helmholz equation:$\Delta G = \Delta H - T\Delta S$. The relative enthalpy (ΔH) represents the difference in the energy of the system (gained or lost) due to the formation or breakage of non-covalent bonds and electrostatic interactions. The relative entropy (ΔS), describes the change in randomness or disorder of the system. In the case of proteins entropy is mainly contributed by the conformational freedom of the polypeptide and the rearrangement of solvated water molecules.

The unbound form of an IDP will populate an ensemble of multiple conformational states (high entropy). Upon binding to a folded partner, the IDP will decrease the set of accessible conformations and this will have an entropic cost (decrease of entropy). In order to achieve a favorable binding the IDP can compensate the loss in entropy by increasing the number of contacts with the binding partner decreasing the ΔH or/and by maintaining some degree of flexibility in the bound form that will reduce the entropic penalty. Depending on the strategy adopted by the IDP the complexes formed with the folded partners can be classified in two main groups: coupled folding-binding complexes and fuzzy complexes (Figure 1.2) (Dyson and Wright, 2002; Fuxreiter and Tompa, 2009; Sharma et al., 2015; Tompa and Fuxreiter; Wright and Dyson, 1999; Wright and Dyson, 2009).

In complexes formed between IDPs and folded proteins via a coupled folding-binding mechanism, the IDP undergoes a disorder to order transition upon binding. Given this, one can ask if the folding of the IDP occurs before or after binding the folded binding partner. These two extremes of coupled folding-binding mechanisms are known as conformational selection and induced fit respectively. In the conformational selection mechanism the binding conformation is pre-existing in the disordered ensemble of conformations of the IDP in the unbound form. The folded binding partner will then selectively bind to this particular binding competent conformation. The unbound conformers will reestablish a new conformational equilibrium generating more binding competent conformers. In the induced fit binding mechanism first the IDP binds the folded partner and then it undergoes a conformational change where the IDP adapts to the binding partner's binding pocket. Conformational selection and induced fit are two binding mechanisms that do not need to be exclusive from each other and may even operate simultaneously. For example, a folded protein could bind a particular conformation of an IDP and then the IDP could still adapt and undergo a conformational change in the bound state.

Alternative to the coupled folding-binding scenario, where the change in enthalpy should compensate for the entropic cost, is the formation of fuzzy complexes. Fuzzy complexes refer to complexes formed by an IDP and a folded binding partner where the IDP is bound but does not undergo a complete disorder to order structural transition, maintaining some degree of conformational heterogeneity and thus enabling a lower entropic penalty upon binding to a folded partner (Sharma et al., 2015; Tompa and Fuxreiter).

Fuzzy regions can for example be found flanking an interacting site, generating additional dynamic contacts with the binding partner but keeping the intrinsic disorder (Adams et al., 2007; Billeter et al., 1993) or could also be acting as flexible linkers between two folded domains or motifs providing the protein with a higher degree of binding plasticity (Clerici et al., 2009; Graham et al., 2001). In summary, the intrinsic flexibility of IDPs allows the fine tuning of ΔH and ΔS modulating protein-protein interactions (Marlow et al., 2010; Tzeng and Kalodimos, 2012).

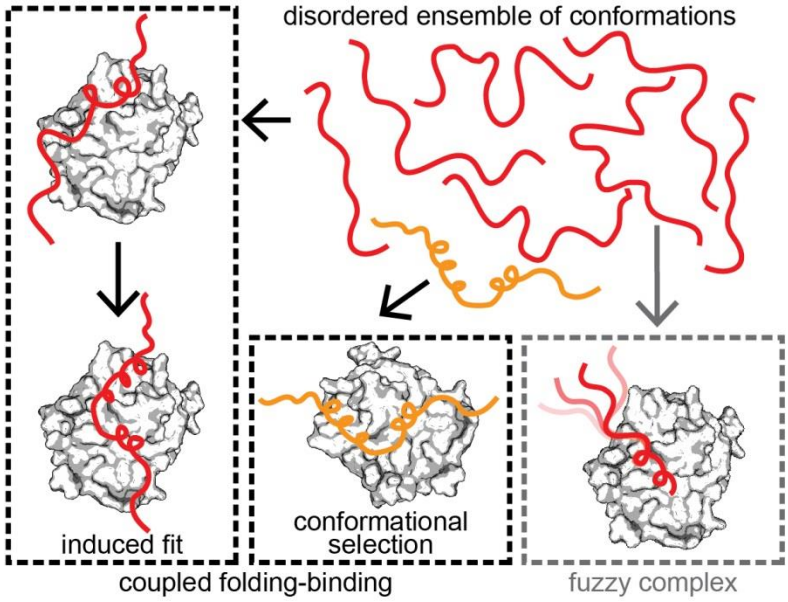

Figure 1.2. IDP binding mechanisms: An IDP populating an ensemble of multiple interconverting conformations is able to bind folded proteins following a coupled folding-binding mechanism (black dashed box) or forming a fuzzy complex (gray dashed box). The folded protein (gray surface protein structure PDB:2ROC) interacts with the IDP (red curved lines). In induced fit bindings the IDP first interacts with the folded protein and then folds adapting to the binding site. In conformational selection, a pre-existing conformation (orange IDP) is able to bind the folded partner. IDPs that form fuzzy complexes are able to retain some degree of flexibility when bound (gray dashed square).

1.1.3. Functional advantages of intrinsically disordered proteins

Disordered polypeptide sequences frequently contain short linear motifs (sLiMs) (Dinkel et al., 2014; Fuxreiter et al., 2007; Kalderon et al., 1984). SLiMs are often sites that are exposed on the surface of the protein in contact with the solvent that are used as functional modules. Due to the exposure of the sequence IDPs and sLiMs are perfect candidates for the addition and removal of post-translational modifications (PTMs). PTMs can affect the structure and dynamics of IDPs. For example, phosphorylation of Ser or Thr residues incorporate a negative charge on the IDP sequence influencing its chemical composition, which could lead to a structural change that can modulate the partners the IDP is binding to (Parker et al., 1999) (Boehr et al., 2009; Kriwacki et al., 1996; Xie et al., 1998). Being major cellular PTM targets makes IDPs particularly ideal components in mediating cell signaling processes

where a detected signal is rapidly transduced downstream in a signaling cascade (Diella et al., 2008; Miller et al., 2008; Wright and Jane Dyson, 2015; Zhou, 2010).

The conformational heterogeneity or protein plasticity of IDPs is related to their functional versatility. For example, IDP regions present between two folded domains will increase the conformational heterogeneity that the protein can adopt, increasing its functional versatility (Clerici et al., 2009; Hurley et al., 2007; Marsh et al., 2010; Sharma et al., 2015). In addition, the protein plasticity of IDPs enables the interaction of the same protein with multiple binding partners (one-to-many) and it is exploited to adapt to different binding sites and tune the affinity towards different binding partners (Vuzman et al., 2012; Vuzman and Levy, 2012). Alternatively to one-to-many binding of IDPs there is also the many-to-one binding scenario in which different IDPs bind the same binding partner (Oldfield et al., 2008).

The ability of IDPs to act as protein interaction and signaling hubs makes possible the rewiring of protein-protein interaction and signaling cascade networks by the alternative splicing of IDP segments from the transcripts. Thus, the incorporation or removal of an IDP segment in a protein can change the type and number of binding partners the protein can bind to through the IDP segment creating or deleting protein-protein interactions and signaling networks in different species or in a tissue specific manner, using the IDP sequences as connecting modules between networks (Buljan et al., 2013; Colak et al., 2013; Ellis et al., 2012).

The capacity of IDPs to interact via multiple binding sites by homo or hetero-oligomerization is a characteristic that has been linked to their ability in some cases to form supramolecular structures.

1.1.4. IDPs and membrane-less organelles

In the recent years, IDPs containing repetitive sequences have been shown to engage via multivalent interactions and result in the formation of supramolecular assemblies such as liquid-like droplets, hydrogels or amyloid fibers (Brangwynne et al., 2009; Labokha et al., 2013; Patel et al., 2015). This phenomenon of liquid-liquid phase transition has also been observed with repetitive folded proteins connected by flexible linkers (Banani et al., 2017; Banjade and Rosen, 2014) like it is the case for Nck/N-WASP/nephrin assemblies. The formation of liquid-like droplets associated to IDPs has been shown to be caused by the interaction of weakly interacting sites that act as a transient molecular glue between different binders. The multiple contacts between molecules will decrease the solubility of the proteins leading to a liquid-liquid phase transition of the molecules. Liquid-like droplets act as intracellular compartments, which often contain other molecules like RNA or proteins, without being surrounded by a lipid membrane. These organelles have been previously observed as cytoplasmic or nucleoplasmic granules including: nuclear bodies, stress granules, nucleoli and Cajal bodies among others (Brangwynne

et al., 2009; Brangwynne et al., 2011; Feric et al., 2016; Saha et al., 2016). The assembly and disassembly of liquid-like droplets is often regulated by PTMs (Nott et al., 2015). However, not much is known about the regulation of these intracellular structures (Burke et al., 2015; Molliex et al., 2015; Nott et al., 2015). Recent studies have shown that aging of the liquid droplets can lead to the formation of irreversible aberrant aggregates containing amyloid fibers (Patel et al., 2015).

In summary, the structural and physicochemical properties of IDPs have enabled them to efficiently engage in different cellular processes like; transcriptional regulation and RNA processing, cell cycle regulation, DNA damage and repair response, signal transduction and nucleocytoplasmic transport regulation among others (Oldfield and Dunker, 2014; van der Lee et al., 2014). During the course of my PhD I have focused on understanding a subset of IDPs involved in the nucleocytoplasmic transport.

1.2. The Nuclear pore complex and the nucleocytoplasmic transport

1.2.1. The nuclear pore complex structure

The nuclear envelope is the physical barrier between the nucleus and the cytoplasm. Cellular homeostasis requires the constant exchange of molecules between these two compartments. Transport of molecules between the nucleus and the cytoplasm is only achieved through a macromolecular complex known as the nuclear pore complex (NPC). The structure of the NPC can be divided in the inner pore ring at the center and the nuclear and the cytoplasmic rings from which the nuclear basket and the cytoplasmic filaments emanate respectively. The NPC has an approximate molecular mass of 120 MDa in humans and it is formed of around 30 different proteins called nucleoporins (Nups) (Beck and Hurt, 2017; Goldberg and Allen, 1996; Rout and Blobel, 1993). The different Nups can be classified as scaffold, transmembrane, peripheral and barrier Nups. Around 1/3 of the total nucleoporins of the NPC form the selective permeability barrier, these Nups are known as FG-Nucleoporins (FG-Nups) (Hurt, 1988; Ori et al., 2013; Patel et al., 2007; Strawn et al., 2004; Wente et al., 1992). The correct functioning of the nucleocytoplasmic transport and the selectivity of the NPC barrier relies on the interaction between the barrier forming disordered FG-Nups and nuclear transport receptors (NTRs), which specifically interact with FG-Nups in order to cross through the NPC barrier.

1.2.2. The FG-Nucleoporins

FG-Nups are formed of at least a folded NPC anchoring domain and an intrinsically disordered domain (Denning et al., 2003). In contrast to other IDPs, the sequence of FG-Nups contains a mean hydrophobicity comparable to the one present in globular proteins and a mean net charge which is lower than the ones usually present in IDPs (Schmidt and Görlich, 2016; Schmidt et al., 2015). FG-Nups are

characterized by containing multiple F residues across their sequence (Radu et al., 1995; Wente and Blobel, 1994; Wente et al., 1992). These residues are often neighbored by a G residue. According to the frequency of occurrence of different amino acids flanking the F residues different FG-motifs (PxFG. FxFG, GLFG, SAFG and FG) have been classified in the FG-Nup sequences (Cushman et al., 2006; Hallberg et al., 1993; Radu et al., 1995; Wente and Blobel, 1994; Wente et al., 1992). FG-motifs are highly abundant across the disordered regions of the FG-Nups (Figure 1.3) and engage via hydrophobic interactions with diverse NTRs, which is required for the NPC crossing of the NTR bound cargo molecules (see section 1.1.4) (Radu et al., 1995; Rexach and Blobel, 1995).

FG-Nups as other IDPs, have a high amino acid substitution rate which leads to low sequence conservation across species (Denning and Rexach, 2007). Nevertheless, different FG-Nups share common features like the intrinsic disorder of their sequence or the high content of hydrophilic amino acids present in the inter-motif sequences (10-20 amino acids between FG-motifs). Despite the high amino acid substitution rate across the FG-Nup sequences, FG-Nups contain discrete conserved sequences which correspond to known FG-Nup/NTR binding sites (Denning and Rexach, 2007).

These physicochemical properties of FG-Nups, equally to other IDPs, can be influenced by the addition or removal of PTMs. It has been shown that multiple Nups and FG-Nups can be phosphorylated, glycosylated, SUMOylated, ubiquitinated and acetylated (Choudhary et al., 2009; Favreau et al., 1996; Golebiowski et al., 2009; Miller et al., 1999; Starr and Hanover, 1990). Some of these modifications may interplay at the same residue position like glycosylation and phosphorylation or SUMOylation and acetylation (Choudhary et al., 2009; Wang et al., 2010). Independently of whether the PMTs occur at the same position, it has been reported that some PTMs are more abundant in Nups at different stages of the cell cycle. (Laurell et al., 2011; Lubas et al., 1995; Miller et al., 1999).The exact role that the different PTMs play or their exact location is still not well known. For example 16 out of 30 Nups are known to be substrates of the enzyme O-GlcNAc glycosyltransferase (OGT) and it is known that at the NPC the lectin wheat germ agglutinin (WGA) binds to the glycosylated FG-Nups that are necessary for the nucleocytoplasmic transport and when bound the transport through the NPC is inhibited. The reduction of the O-GlcNAcylation levels of FG-Nups influences the stability of the FG-Nups at the NPC by an increase of their ubiquitination and proteosomal degradation (Zhu et al., 2016). Nevertheless, whether O-GlcNAc can influence the NPC function affecting the interaction between FG-Nups and NTRs it is so far unknown.

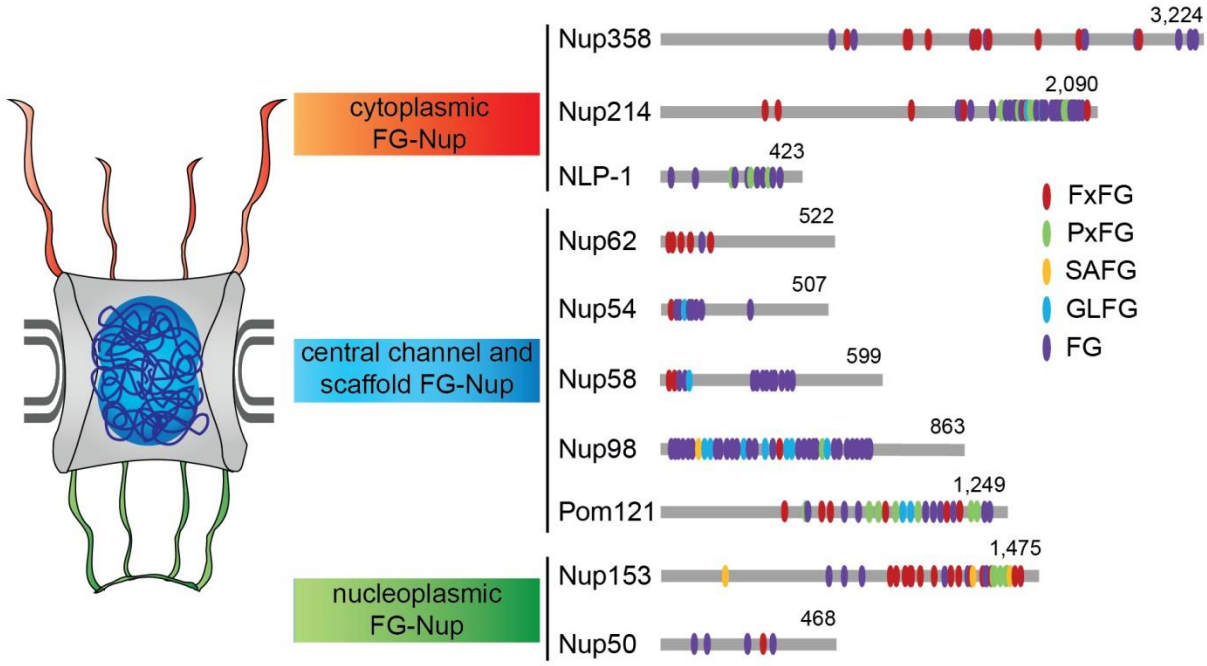

Figure 1.3. Cartoon of the NPC with different FG-Nups: The different FG-Nups are grouped depending on their localization at the NPC (cytoplasmic side, central channel, nucleoplasmic side). The FG-motifs are color coded and distributed across the FG-Nup (the numbers correspond to the length of the FG-Nups).

The physicochemical characteristics of FG-Nups are linked to the capacity of FG-Nups to interact with each other, which is directly related to the structure of the permeability barrier. Hydrophilic, electrostatic and hydrophobic interactions may take place between different FG-Nups. However, how could these interactions affect the macromolecular structure of the selective barrier is a highly debated topic in the field and different models have been proposed based on the different inter-molecular FG-Nup interaction strengths(Frey and Gorlich, 2009; Lim et al., 2015; Lim et al., 2006; Peters, 2009; Rout et al., 2003; Walde and Kehlenbach, 2010; Yamada et al., 2010). A systematic study, where immobilized FG-Nups were probed against different soluble fluorescent FG-Nup fusions showed weak hydrophobic inter-chain FG-Nup interactions, which were disrupted by the addition of the aliphatic alcohol 1,6-Hexanediol (Patel et al., 2007). These results are in line with the observation that incubation of cells with 1,6-Hexanediol reversibly disrupts the permeability barrier of the NPC (Ribbeck and Gorlich, 2002).

Inter-chain interactions between FG-Nups have also been suggested to enable FG-Nups to undergo phase separation. Different studies have shown that FG-Nups are able to form phase separated particles, hydrogels and amyloid fibers under certain in vitro experimental conditions (Labokha et al., 2013; Milles et al., 2013; Schmidt et al., 2015). Phase separated FG-Nups in solution formed solid particles that were able to act as a selective barrier by excluding inert molecules of a molecular weight of 65 kDa while at the same time enabling the diffusion of NTRs (Schmidt et al., 2015). FG-Nup hydrogels also showed a

similar selective permeability behavior (Labokha et al., 2013). Nevertheless, NMR experiments of the FG-Nup Nsp1 have shown that the aggregation caused by weak inter-chain interactions between FG-Nups was avoided by performing the measurements inside *E.coli* cells (Hough et al., 2015).

The different studies of the interaction between FG-Nups have led to the proposal of a various models that focus on the structure of the selective barrier. However, all the models seem to agree on the required interaction between FG-Nups and NTRs to cross the selective NPC barrier.

1.2.3. Nuclear transport receptors

The nucleocytoplasmic transport of cargoes is mediated my different transport receptors that are able to recognize and bind to the nuclear localization sequence (NLS) or the nuclear export signal of a cargo and transport it to the nucleus or the cytoplasm respectively. The biggest family of transport receptors is the one formed by the β-karyopherin/Importinβ superfamily (Moroianu, 1998). Out of the 20 different types of NTRs in the β-karyopherin superfamily proteins like Importinβ are involved in mediating the import of molecules. Others like CRM1 or CAS, mediate the export of molecules and some NTRs like Transportin are bi-directional (Shamsher et al., 2002). Members of this superfamily have a high structural similarity but a low sequence identity (between 15-20%) (O'Reilly et al., 2011). β-karyopherins like Importinβ, CRM1 or Transportin are composed of stacked tandem series of HEAT repeats (Cansizoglu et al., 2007; Chook and Blobel, 1999; Cingolani et al., 1999; Monecke et al., 2013). HEAT is the acronym coming from the different proteins where this motif was initially found: Huntingtin, elongation factor 3, the A subunit of the protein phosphatase 2A and the kinase TOR1. HEAT repeats are formed by two amphiphilic α-helices (A-helix and B-helix) of approximately 30-40 amino acids in length (Figure 1.4 A,E). The helices are linked by a turn and are oriented in an antiparallel manner with the hydrophobic surface facing each other helix (Chook and Blobel, 1999; Cingolani et al., 1999). NTRs are composed of multiple stacked HEAT repeats that form a solenoid where the A-helix and the B-helix are forming the concave and convex surface respectively. This structural composition allows NTR to undergo conformational changes where the pitch of the helicoid changes due to small changes in the angle between HEAT repeats, which provide the NTRs with a high structural elasticity (Conti et al., 2006; Fukuhara et al., 2004).

NTRs undergo different conformational changes depending on which protein they are bound to (Figure 1.4E). For example, structural studies on Importinβ bound to RanGTP showed that the NTR has a conformation that inhibits the binding to Importinα via an allosteric mechanism (Cingolani et al., 1999; Lee et al., 2005). Alternatively, the binding of Importinβ to the auto-inhibitory binding domain of Importinα (IBB) causes an increase in the level of compaction of Importinβ (Cingolani et al., 1999). In the case of the exportin CRM1, the conformational change occurring upon RanGTP binding is required

for the formation of the export complex (Monecke et al., 2013). In addition to the conformational changes generated upon binding of binding partners it has been reported that NTRs composed of HEAT repeats also undergo structural fluctuations which allows them to adapt the degree of exposed hydrophobic surface depending on the environment where the NTR is in. For example, in the transport conduit of the NPC the degree of surface hydrophobicity would be higher than in the cytoplasm, where it would have a more hydrophilic surface. This is achieved by small shifts in the angle between HEAT repeats leading to a higher or lower exposure degree of the hydrophobic side of each repeat (Yoshimura and Hirano, 2016; Yoshimura et al., 2014).

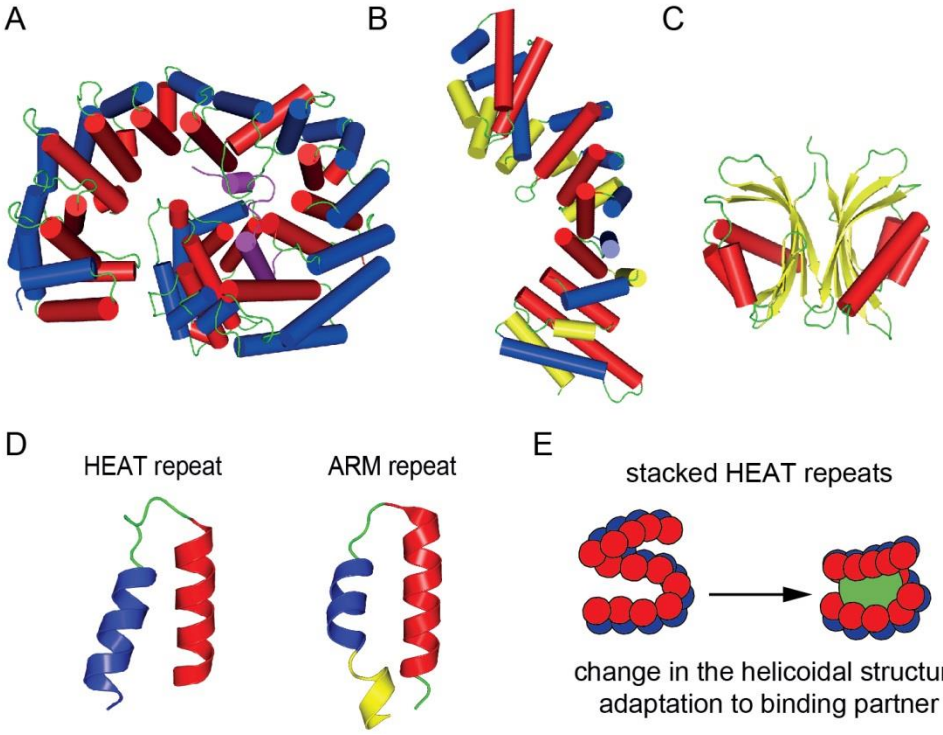

Figure 1.4. Structural features of NTRs: A) Crystal structure of Importinβ bound to IBB domain (pink) of Importinα, A-helices are colored in blue and B-helices in red (PDB:1QGK). B) Crystal structure of Importinα, H1 (yellow), H2 and H3 are colored in blue and red respectively as in Importinβ (PDB:1EE5). C) Crystal structure of NTF2 homo-dimer with α-helixes in red and β-sheets in yellow (PDB:1OUN). D) Structural comparison of the HEAT repeat 5 of Importinβ and the ARM repeat 5 of Importinα. E) Cartoon representation of an NTR containing stacked HEAT repeats forming a helicoidal structure. When the NTR is bound to a protein changes in the angle of the repeats varies the pitch of the helicoid.

The members of the Importinα family are often used as adaptor proteins that recognize the NLS from a cargo and bind to the transporter Importinβ. The most commonly studied nuclear transport receptor/ adaptor protein studied is the one formed by Importinβ/Importinα. Importinα is formed by consecutive armadillo repeats (ARM repeats) (Figure 1.4B,D). ARM repeats are constituted of three α-helices (H1,

H2 and H3) where helices H2 and H3 are placed in an antiparallel manner, as it is the case for the A and B-helix of the HEAT repeats, and the helix H1 is placed slightly perpendicular to H2 and H3 (Conti et al., 1998; Huber et al., 1997). As for NTRs formed of HEAT repeats, the stacking of ARM repeats also confers NTRs with a solenoid structure that can undergo conformational changes (Andrade et al., 2001).

Other NTRs like the nuclear transport factor 2 (NTF2), which is required for the transport of RanGDP inside of the nucleus (Bayliss et al., 1999; Stewart, 2000), are not formed by stacked α-helical repeats. NTF2 is formed by two homo-dimers containing a bent β-sheet and three α-helices, one longer and two shorter ones (Bullock et al., 1996) (Figure 1.4C).

Despite the structural similarities or differences within NTRs, all transport receptors have the common ability to cross the NPC with a cargo by interacting with the FG-motifs of the FG-Nups located at the NPC. In the following sections, I will focus on describing the nucleocytoplasmic transport process and the interaction mechanism between FG-Nups and NTRs that is required for such process.

1.2.4. The nucleocytoplasmic transport

The nucleocytoplasmic transport of molecules across the NPC can be divided in three major steps: cargo recognition and binding to the NTR, NPC translocation and cargo release either in the nucleus or the cytoplasm for import and export processes respectively. The recognition of a cargo by an NTR is mediated by the binding of importins and exportins to the nuclear localization signal (NLS) or the nuclear export signal (NES) respectively. The directionality of the transported molecules is dependent on a RanGTP/GDP gradient formed between the nucleus and the cytoplasm.

The RanGTP/GDP gradient is maintained by the spatial separation of different components involved in the RanGTP/GDP cycle (Kalab et al., 2002). Cytoplasmic RanGDP binds to NTF2 and is transported to the nucleus (Bayliss et al., 1999; Stewart, 2000). The exchange of RanGDP to GTP is catalyzed by the Ran guanine nucleotide exchange factor RCC1 which is localized at the nucleus (Klebe et al., 1995a; Klebe et al., 1995b). RanGTP will then bind to exportins increasing the affinity towards export cargoes or to import complexes causing the cargo release at the nucleus. RanGTP bound to an exportin or an export complex will then be exported to the cytoplasm where the Ran GTPase activating protein (RanGAP), which is enriched in the cytoplasmic RanBP2 NPC filaments, will hydrolyze RanGTP to RanGDP leading to the cytoplasmic export cargo release (Kalab et al., 2002; Ritterhoff et al., 2016; Yudin and Fainzilber, 2009).

In the classical import cycle, a NLS-cargo is recognized by Importinβ or by the adaptor protein Importinα that will then bind to Importinβ via the IBB domain forming the import complex (Cingolani et al., 1999). Then the tertiary complex is imported through the NPC and binds Nup50 at the nuclear basket (Matsuura

and Stewart, 2005). Upon binding of RanGTP to Importinβ the NLS-cargo will be released to the nucleoplasm and Importinβ/RanGTP will then shuttle back to the cytoplasm where RanGTP will be hydrolyzed and Importinβ will be available to form additional import complexes. CAS/RanGTP will then bind to Importinα/Nup50 forming the tertiary complex Importinα/CAS/RanGTP which will be exported to the cytoplasm and upon RanGTP hydrolysis Importinα will be released in the cytoplasm (Sun et al., 2013). In the classical export cycle, CRM1 will bind to RanGTP undergoing a conformational change that will increase the affinity for NES-cargoes (Monecke et al., 2009; Monecke et al., 2013). The export complex formation has been suggested to be facilitated by the cofactor RanBP3 (Englmeier et al., 2001). The export complex NES-Cargo/CRM1/RanGTP is then translocated though the nucleus, where it will bind to Nup214 and Nup358. that the RanGTP hydrolysis on the cytoplasmic side will case the release of the export cargo in the cytoplasm (Hutten and Kehlenbach, 2006, 2007; Koyama and Matsuura, 2010; Koyama et al., 2014; Ritterhoff et al., 2016).

1.2.5. The interaction between FG-Nups and NTRs

The interaction between FG-Nups and NTRs has been known to be a requirement for the correct functioning of the nucleocytoplasmic transport since more than two decades. Early biochemical studies showed that the depletion of FG-Nups from Xenopus nuclei disrupted the ability of importing NLS containing cargoes. Additionally, the retrieval of the subtracted protein content restored the nucleocytoplasmic transport process (Finlay and Forbes, 1990; Lubas et al., 1995; Radu et al., 1995; Starr and Hanover, 1990). Moreover, *in vitro* interaction assays showed than many NTRs from the Importinβ superfamily interact with different FG-Nups like Nup153, Nup358, Nup62 or Nup214 (Gorlich et al., 1996; Hu et al., 1996; Kehlenbach et al., 1999; Moroianu et al., 1995; Seedorf et al., 1999; Yaseen and Blobel, 1999). Crystallographic studies revealed that the FG-motifs engage with different hydrophobic binding pockets of the NTRs (Figure 1.5). The crystal structure of Importinβ with five tandem FxFG repeat from Nsp1p revealed that the FxFG peptide binds to the binding pocket of Importinβ between the A-helix of HEAT repeat 5-6 and 6-7 (Bayliss et al., 2000a; Bayliss et al., 2000b). Moreover, a GLFG peptide was also crystallized bound to the same binding pockets of Importinβ (Bayliss et al., 2002b). Molecular dynamics simulations (MD) of Importinβ binding to different peptides of Nsp1p and Nup116p showed that in addition to the two previously characterized binding sites by X-ray crystallography the NTR has potentially 6 additional sites where FG-motifs could engage with the NTRs (Isgro and Schulten, 2005).

The conformational changes of Importinβ upon for example RanGTP binding can also regulate the number of FG binding sites in the NTR. The crystal structure of Importinβ/RanGTP bound to an FG peptide showed that the A-helix from the HEAT repeat 5 of Importinβ occupies a different position than

in the unbound form and may be used to facilitate the release of the FxFG peptide from the binding groove of Importinβ (Bayliss et al., 2000a). Moreover, single molecule force measurements of the binding between Nup153 and Importinβ showed that different binding pockets have different binding affinities and that the conformational change of the NTR upon RanGTP binding can tune the interaction strength with the FG-Nup (Otsuka et al., 2008).

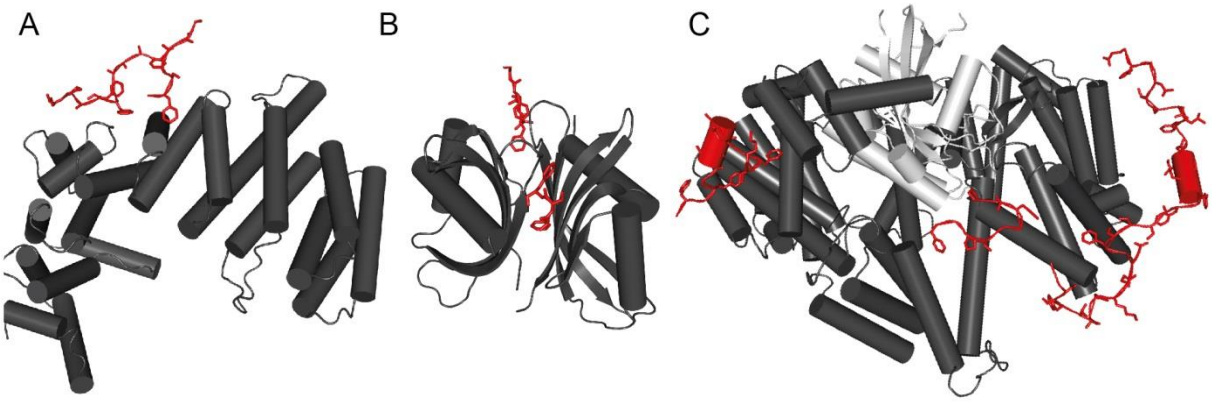

Figure 1.5. Crystal structures of NTRs bound to FG-Nups: A) Crystal structure of Importinβ (1-442 aa) (gray) bound to a FxFG motifs (red) of Nsp1p peptide composed of 5 FxFG repeats (PDB:1F59)..Binding of FxFG motifs takes place between HEAT 5-6 and HEAT 6-7. B) Crystal structure of a homo-dimer of NTF2 (gray) bound to an FxFG peptide (PDB:1GYB). The FxFG motif binds to the hydrophobic groove generated in the binding interface between the two NTF monomers. C) Crystal structure of CRM1/RanGTP (gray/light gray) bound to a 117 aa fragment of the C-terminal site of Nup214 (red) (PDB: 5DIS). Nup214 binds CRM1 at the C- and N-terminal sites binding in different hydrophobic pockets of the outer surface of CRM1.

Studies on NTF2 revealed that the successful import of RanGDP into the nucleus was dependent on the capability of NTF2 to interact with FxFG rich nucleoporins (Bayliss et al., 1999). The crystal structure of NTF2 bound to a short FxFG peptide showed that in this case the FG motifs engage in the hydrophobic pocket formed between the two NTF2 homodimers (Bayliss et al., 2002a) (Figure 1.5B). Moreover, NMR studies revealed that in addition to the binding sites between the F residues and the NTF2 binding groove showed by the crystal structure, a secondary FxFG binding site exists (Morrison et al., 2003).

A recent crystal structure of CRM1/RanGTP bound to a 117 amino acid long disordered Nup214 shows that Nup214 engages with different FG-motifs into the hydrophobic pockets of the outer convex surface of CRM1 (Port et al., 2015) (Figure 1.5C). Remarkably, Nup214 interacts with the N- and the C-terminal sites of CRM1 and the authors suggested that it may act as a molecular clamp stabilizing the export complexes (Port et al., 2015). Interestingly, the crystal structure showed some secondary structure in the FG-Nup and it remains up to investigation if the Nup214FG undergoes a conformational change coupled to the binding of CRM1/RanGTP. In summary, these studies highlight the capability of FG-Nups to engage with multiple binding motifs on the different binding sites of the NTR indicating the multivalent nature of the interaction.

1.2.6. The transport paradox

The transport of molecules across the NPC is considered to be an extremely fast and selective process. Single molecule fluorescence microscopy experiments have monitored the trajectories of molecules crossing the NPC and have measured that the translocation times of molecules shuttling between the nucleus and the cytoplasm are on millisecond timescales (Kubitscheck et al., 2005; Musser and Grunwald, 2016; Sun et al., 2013; Yang et al., 2004). During the NPC crossing, free or cargo-bound NTRs will interact with the different FG-motifs present in the FG-Nups. Considering that various *in vitro* studies on purified FG-Nup and NTR components have reported affinities ranging from nM to µM (Bednenko et al., 2003; Ben-Efraim and Gerace, 2001; Milles and Lemke, 2014; Tetenbaum-Novatt et al., 2012) one would expect that due to the high binding affinities of NTRs and FG-Nups, at the high local concentrations of motifs at the inner channel of the NPC, NTRs would be forming a stable complex with the FG barrier which seems paradoxical since NTRs are migrate through the NPC in fast timescales.

The better understanding of the nucleocytoplasmic transport requires the comprehension of how the fast crossing of NTRs through the NPC is achieved.

1.3. Towards the study of FG-Nups at the NPC

1.3.1. State of the art technology for the study of protein structure and dynamics *in situ*

The controlled environment at which the biochemical study of the structure and dynamics are performed has usually little resemblance to the crowded environment inside the cell. Thus, major efforts are being invested in the development of new technologies and tools to study the structure and dynamics of proteins in their intracellular environment. Techniques like cryo-electron tomography are used study the molecular architecture of protein complexes *in situ*. In addition, sub-tomogram averaging provides high resolution maps that can be used to fit high resolution structures of the components forming the protein complex under study. However, this technique is not suitable when the protein complex under study is present in multiple interconverting conformations or it contains disordered domains. The study of the structure and dynamics of dynamic proteins *in situ* has recently benefited from super-resolution microscopy (SRM), single molecule spectroscopy and NMR spectroscopy techniques (Plitzko et al., 2017).

The biochemical and computational findings of FG-Nup/NTR interactions is enabling a better understanding of the nucleocytoplasmic transport. However, in order to study the structural characteristics, localization and dynamics to further learn about the behavior and localization of the disordered FG-Nups, one has to study them *in situ* in a minimally perturbed system at the NPC. Latest cryo-electron tomography and SRM experiments have provided detailed information regarding the structure and localization of the folded domains of different nucleoporins (Kosinski et al., 2016;

Szymborska et al., 2013; von Appen et al., 2015). However, the disordered regions of FG-Nups at the NPC are currently not easily accessible by state of the art technologies. One approach that can be used to investigate the localization of the FG-Nup region at the NPC is site-specific labeling combined with SRM. One of the most versatile and powerful methods at the moment to label proteins site-specifically is by incorporating non-canonical amino acids (ncAA) in a determined position in the protein of interest (POI) using genetic code expansion technology (GCE) (Noren et al., 1989; Wang et al., 2006) combined with click-chemistry to achieve the incorporation of the desired dye.

1.3.2. Amber suppression technology

The incorporation of a ncAA into a specific position of a protein can be achieved using an orthogonal aminoacyl tRNA synthetase (aaRS) and tRNA pair (aaRS/tRNA) (Noren et al., 1989). The aaRS/tRNA pair system used by the archae *Methanosarcina* that incorporates the amino acid pyrrolysine, a lysine derivative with a pyrroline ring in response to the Amber stop codon (UAG), was shown to be orthogonal in eukaryotes and prokaryotes and permitted the incorporation of a ncAA at the position where the TAG mutation was introduced in the coding region (Srinivasan et al., 2002) (Figure 1.6). Further mutagenesis of the pyrrolysine aaRS (pylRS) enabled the incorporation of bulkier ncAA, which expanded the repertoire of ncAA that could be incorporated (Yanagisawa et al., 2008). The incorporation bulkier ncAA containing cyclooctene and cyclooctyne functional groups enabled site-specific live-cell labeling with synthetic dyes which are suitable for SRM (Nikic et al., 2014; Plass et al., 2011; Plass et al., 2012; Yanagisawa et al., 2008).

In our lab a previous PhD student Dr. Kang developed a methodology to use Amber suppression technology combined with click-chemistry to label site specifically different positions of the IDP region of Nup98. The different positions were imaged using SRM to achieve the mapping of the disordered region of Nup98. However, one of the caveats of the Amber suppression system is that ~80 % of the protein expressed is produced as a truncated product, which will not contain the ncAA. In the case of folded proteins this is often not a problem since many truncated products get eliminated by the cell. With truncated FG-Nups however, when the probed FG-region is located at the C-terminal site, the truncated products will still contain the nuclear targeting domain at the N-terminal site, which will anchor at the NPC despite lacking the FG disordered regions, competing in this way with the other 20% of the POI that did incorporate the ncAA and will get labelled. This means that the truncated POI will be competing with the full length protein containing the ncAA, leading to a low labeling density at the NPC and thus dramatically decreasing the quality of the sample for SRM studies of the disordered NPC barrier *in situ*.

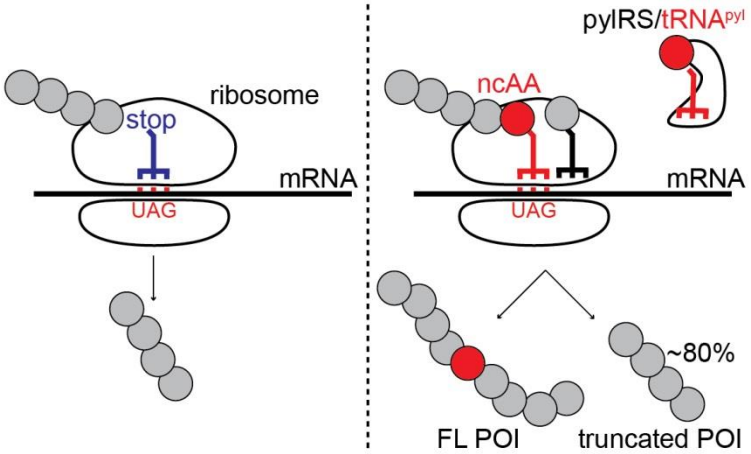

Figure 1.6. Simplified cartoon representation of the incorporation of ncAA: A) During protein translation when the ribosome encounters the UAG stop codon, it is recognized by the release factor (RF1, blue rod) which promotes the hydrolysis of the ester bond that is linking the tRNA and the nascent polypeptide. Subsequently, translation is stopped and release of the synthetized polypeptide will take place. B) Amber suppression technology using the orthogonal pylRS/$tRNA^{pyl}$ pair will recognize the Amber stop codon engineered on a POI and instead of triggering the termination of translation the ncAA coupled to the $tRNA^{pyl}$ will be incorporated. When the ncAA is not incorporated, truncated POI.

.

Chapter 2

Objective

2. Objective

The main aim of my PhD thesis is to study the interaction mechanism between the disordered domain of FG-Nucleoporins (FG-Nups) and nuclear transport receptors (NTRs). With this, I intend to contribute to the understanding of how the fast yet selective transport of NTRs across the nuclear pore complex (NPC) is achieved given that during this process NTRs interact multiple times with FG-motifs of the different FG-Nups that form the selective barrier of the NPC. Additionally, I aim to develop a strategy that will improve the study of the structure of the disordered regions of FG-Nucleoporins *in situ*.

The specific aims are outlined as following:

2.1. *In vitro* study of the structure, kinetics and dynamics of the interaction between FG-Nups and NTRs.

My main objective is to understand how the transport of NTRs across the NPC can occur in a fast and specific manner. I aim to characterize the binding mechanism of diverse disordered FG-Nups and NTRs from a kinetic point of view using stopped-flow spectroscopy. In addition, I aim to study the structure and dynamics of the disordered region of FG-Nups upon binding to different NTRs using single molecule spectroscopy techniques. Furthermore, I aim to study the possible role that the post-translational modification with O-linked N-Acetylglycosamine of FG-Nups may have in the FG-Nup structure and the binding to NTRs.

2.2. Development of a synthetic biology approach to optimize the *in situ* structural studies of FG-Nups.

A long-term goal is to study the structure and dynamics of the disorder region of FG-Nups *in situ* at the NPC in order to reveal the properties of FG-Nups in a physiological system. However, the current technology that enables site-specific labeling of Nups by incorporation of ncAAs via Amber suppression technology lead high amounts of truncated products compromising the power of super-resolution and single molecule studies in the cell. Thus, I aim to develop a novel synthetic biology approach to achieve the selective enrichment of proteins containing the ncAA in cells.

Chapter 3

Materials and Methods

3. Materials and methods

3.1.Methods

3.1.1. Purification of recombinant proteins

The plasmids used for the recombinant expression of the proteins are listed in the materials section 3.2.1

Purification of Nup153FG labeling mutants

Lysis buffer: 2 M Urea, 4xPBS pH 8.5, 5 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP

Wash buffer 1: 2 M Urea, 4xPBS pH 8.5, 10 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP

Wash buffer 2: 2 M Urea, 4xPBS pH 8.5, 15 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP

Elution buffer: 2 M Urea, 4xPBS pH 8.5, 400 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP

TEV cleavage buffer: 50 mM Tris pH 7.5, 150 mM NaCl, 1 mM PMSF, 0.2 mM TCEP

SEC buffer: 2 M Urea, 4xPBS pH 8.5, 1 mM PMSF, 0.2 mM TCEP

Storage buffer: 4M GdmHCl, 1xPBS pH 7.4

A colony of *E.coli* BL21 DE3 (AI) cells containing the pTXB3-6His-TEV-N153-intein-CBD plasmid was inoculated in LB medium containing 50 µg/ml of ampicillin. For single labeled Nup153FG constructs containing a single cysteine residue at positions (1391 or 883) were used. Expression of Nup153 constructs for smFRET measurements the cells were cotransformed with the plasmid pEvol-AcF (resistance to chloramphenicol), which contains the tRNA/RS pair for charging the $tRNA^{TAG}$ with the ncAA AcF. Nup153FG constructs containing a single cysteine and an Amber stop codon, which will incorporate the ncAA AcF, were used (Nup153FG 1313AcF+1391C, 994AcF+883C and 1049AcF+990C). In the case of SmFRET constructs the cells were inoculated in LB medium containing 50 µg/ml of ampicillin and 33 µg/ml of chloramphenicol. A detailed list of the plasmids for each construct can be found in the material section (3.2.1.). The cultures were grown overnight at 37°C under shacking conditions. Then TB expression culture was inoculated with a 1:100 dilution of the overnight inoculate culture. Cells were incubated at 37°C under shacking conditions. Cultures containing the constructs with the Amber mutant were supplemented with 1 mM AcF at OD_{600}=0.2-04. The induction of the protein expression was triggered by the addition of 1 mM IPTG and 0.02 % arabinose at OD_{600}=1. The protein was expressed for 6 h at 37°C under shacking conditions. Cells were harvested by centrifugation in a Beckmann centrifuge, rotor JLA 8.100, 5000 rpm for 20 min at 4°C.

Cells were resuspended with lysis buffer (10ml/L expression) and lysed through 3-5 rounds in the microfluidizer. The lysate was then clarified by centrifugation in a Beckmann centrifuge, rotor JA 25.50

at 18000 rpm for 1 h at 4°C. The clear lysate was then incubated with Ni^{2+}-beads (1ml/ L expression) for 2 h at 4°C under mild rotation. The beads were places on a polypropylene column and were washed with wash buffer 1 and wash buffer 2 (10 ml/ BV per wash buffer). The protein was eluted with elution buffer and incubated with chitin beads (2-5 mL/L expression) overnight at 4°C. The flow through was discarded and the beads were washed (10 mM/ BV) with lysis buffer and then were equilibrated on TEV cleavage buffer. 0.5 mg of TEV protease was added per litter of expression and the cleavage buffer was supplemented with 100 mM BME. Cleavage was performed 6-8 hours at room temperature under mild rotation. Then the flow through containing the cleaved protein was collected and subjected to dialysis against lysis buffer to remove the BME and exchange the buffer to mild denaturing conditions. Then the sample was incubated with Ni-beads to remove the cleaved protein (0.1 ml/L expression). The flow through was then supplemented with 4M GdmHCl and concentrated. The concentrated protein was then subjected to size exclusion chromatography (Superderx 200 10/300). The fractions of interest were then concentrated and buffer exchanged in storage buffer.

Note: Labeling of the Nup153FG samples was done prior to the SEC step. SEC was then used to eliminate degradation products originated from the labeling procedure and to remove the excess of dye.

Purification of Nup153FG WT

HPLC buffer A: ddH_2O, 0.1 % Trifluoroacetic acid (TFA)

HPLC buffer B: Acetonitrile (ACN), 0.1 % TFA

Unlabeled Nup153FG WT was used for the stopped-flow dissociation experiments. The purification of Nup153FG WT is the same as for Nu153FG mutants. Here, I will describe the steps used only in the Nup153FG WT purification protocol. The expression of Nup153FG WT was performed by using E.coli Top10 cells transformed with the plasmid pBAD-N153FG-intein-CBD-12His. Induction of protein expression was done with 0.02 % arabinose. After following the same standard Ni and chitin affinity purification protocol describe above the eluted protein was subjected to high performance liquid chromatography (HPLC).

The eluted protein was mixed with 5 % ACN and 1 % HCl then the sample was filtered. The C18 HPLC column was equilibrated with 95 % buffer A and 5 % buffer B and the sample was then loaded. Elution of Nup153FG WT was performed with a gradient from 5-100 % of buffer B. Elution of Nup153FG takes place between 60-64 % of buffer B. The eluted sample was then lyophilized and stored at -20°C sealed with parafilm. For usage, Nup153FG WT powder was resuspended in storage buffer.

Purification of Nup214FG

Lysis buffer: 2 M Urea, 4xPBS pH 8.5, 5 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP

Wash buffer: 2 M Urea, 4xPBS pH 8.5, 30 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP

Elution buffer: 2 M Urea, 4xPBS pH 8.5, 400 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP

Cleavage buffer: 2 M Urea, 4xPBS pH 8.5, 5 mM Imidazole, 100 mM BME, 1x Compete EDTA free protease inhibitor,1 mM PMSF, 0.2 mM TCEP

SEC buffer: 2 M Urea, 4xPBS pH 8.5, 1 mM PMSF, 0.2 mM TCEP

Storage buffer: 4M GdmHCl, 1xPBS pH 7.4

Expression of Nup214FG (1392-2090 aa numbering with respect to the full length protein Uniprot: P35658). Nup214FG constructs contained a single cysteine and an Amber stop codon, which will incorporate the ncAA AcF (Nup214FG 1095C+2043AcF).

The expression and purification of Nup214FG was performed similarly to Nup153FG. After the incubation of the lysate with Ni beads the beads were washed with wash buffer containing 30 mM imidazole. Elution from the Ni beads was then incubated overnight in cleavage buffer at room temperature with mild rotation. The sample was then dialyzed to remove excess of BME and imidazole and was subjected to an additional step of Ni affinity purification to remove the uncleaved protein. GdmHCl was then added to a final concentration of 4 M and the sample was concentrated and subjected to SEC (Superdex 200, 10/300). The fractions of interest were then buffer exchanged into storage buffer and kept at -80°C.

Purification of Importinβ

Lysis buffer: 50 mM Tris pH 7, 650 mM NaCl, 5 mM $MgCl_2$, 5 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP.

Wash buffer 1: 50 mM Tris pH 7, 650 mM NaCl, 5 mM $MgCl_2$, 10 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP.

Wash buffer 2: 50 mM Tris pH 7, 650 mM NaCl, 5 mM $MgCl_2$, 15 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP.

Elution buffer: 50 mM Tris pH 7, 650 mM NaCl, 5 mM $MgCl_2$, 400 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP.

SEC buffer: 50 mM Tris pH 7, 650 mM NaCl, 5 mM $MgCl_2$, 1 mM PMSF, 0.2 mM TCEP.

LB medium containing 50 µg/ml ampicillin was inoculated with a colony (or from a glycerol stock) of *E.coli* BL21 DE3 (AI) cells containing the plasmid pTXB3-Importin beta WT-intein-12His and was incubated under shaking conditions overnight at 37˚C. Then the overnight culture was added in a 1:100 dilution to TB medium containing 50 µg/ml ampicillin. The cultures where grown at 37˚C under agitation

conditions. Induction of protein expression was performed by the addition of 1mM IPTG and 0.02 % arabinose at OD_{600}=0.4-0.6. The cultures where grown at 30˚C for 6 h. Cells were harvested by centrifugation in a Beckmann centrifuge, rotor JLA 8.1000 at 4.500 rpm for 20 min at 4˚C. Cell pellets were resuspended in lysis buffer (10ml/L expression) and lysed by 3 rounds of microfluidizer. The lysate was clarified by centrifugation of 1 h, at 15.000 rpm and 4˚C in a Beckmann centrifuge with the rotor JA 25.50. Pre-equilibrated Ni-NTA beads with lysis buffer were then incubated with the clarified cell lysate (1 ml beads/L expression) for 2 h at 4˚C on slow rotation. The Ni-beads were then placed on a polypropylene column. Subsequently the beads were washed with lysis buffer (10 bead volume (BV)) followed by 10 BV of wash buffer 1 and 2. Elution was performed with 10 BV of elution buffer. The 12His-tag was cleaved by intein cleavage by incubating the lysate with 100 mM β-mercaptoethanol (BME) and 1x Complete EDTA free protease inhibitor cocktail at room temperature for 8 h or overnight. The sample was then dialyzed against lysis buffer at 4˚C under mild stirring conditions. Then the sample was reincubated with Ni-beads (0.1 ml beads/ L expression) to remove the uncleaved protein for 2 h at 4˚C. The collected flow through was then concentrated and subjected to size exclusion chromatography (SEC) at 4˚C (Superdex S200, 16/60). The selected fractions were then concentrated and stored at -80˚C.

The same purification protocol was used for the purification of Importinβ I178D Y255A double mutant (pTXB3-ImportinBeta-I178D-Y255A-intein-12His).

Purification of CRM1

Lysis buffer: 50 mM Tris pH 8, 650 mM NaCl, 5 mM imidazole, 1 mM PMSF, 0.2 mM TCEP
Wash buffer 1: 50 mM Tris pH 8, 650 mM NaCl, 10 mM imidazole, 1 mM PMSF, 0.2 mM TCEP
Wash buffer 2: 50 mM Tris pH 8, 1.5 M NaCl, 15 mM imidazole, 1 mM PMSF, 0.2 mM TCEP
Wash buffer 3: 50 mM Tris pH 8, 650 mM NaCl, 25 mM imidazole, 1 mM PMSF, 0.2 mM TCEP
Elution buffer: 50 mM Tris pH 8, 650 mM NaCl, 400 mM imidazole, 1 mM PMSF, 0.2 mM TCEP

The overnight inoculate was performed with a single colony of BL21 XL10 Gold cells containing the plasmid pQE60-CRM1-intein-12His. The inoculate was then diluted 1:100 in TB medium containing 50 µg/ml ampicillin. Induction of expression was performed at OD_{600}=0.8 with 1 mM IPTG. Cultures were incubated under shaking conditions for 6 h at 30˚C or overnight at 18˚C. CRM1 purification follows as described for Importinβ.

Purification of TRN1

The overnight inoculate was performed with a single colony of *E.coli* BL21 XL10 Gold cells containing the plasmid pQE60-6His-TRN1. The expression and purification used for TRN1 was the same as for Importinβ.

Purification of NTF2

Lysis buffer: 50 mM Tris pH 8, 0.5 M NaCl, 5 mM $MgCl_2$, 5 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP.

Wash buffer 1: 50 mM Tris pH 8, 0.5 M NaCl, 5 mM $MgCl_2$, 10 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP.

Elution buffer: 50 mM Tris pH 8, 0.5 M NaCl, 5 mM $MgCl_2$, 400 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP.

SEC buffer: 50 mM Tris pH 8, 0.5 M NaCl, 5 mM $MgCl_2$, 1 mM PMSF, 0.2 mM TCEP.

The overnight inoculate was performed with a single colony of *E.coli* BL21 DE3 AI cells containing the plasmid pTXB3-NTF2-intein-6His encoding the hNTF2 protein. The expression of NTF2 was performed on 2xYT medium containing 50 µg/ml ampicillin. NTF2 was purified by a Ni affinity step as described for other NTRs. Then the eluate was subjected to intein cleavage by adding 100 mM BME, the sample was incubated overnight at room temperature under mild rotation. After dialyzing the sample against lysis buffer a second Ni affinity incubation was performed to remove uncleaved proteins (as previously described for the purification of other NTRs. After concentration of the flow through, SEC was performed. The selected fractions were then concentrated and stored at -80˚C.

Purification of RanQ69L (1-180)

Lysis buffer: 30 mM Tris pH 7.4, 200 mM NaCl, 5 mM $MgCl_2$, 5 mM Imidazole, 10 % glycerol, 1 mM PMSF, 0.2 mM TCEP.

Wash buffer: 30 mM Tris pH 7.4, 200 mM NaCl, 5 mM $MgCl_2$, 10 mM Imidazole, 10 % glycerol, 1 mM PMSF, 0.2 mM TCEP.

Elution buffer: 30 mM Tris pH 7.4, 200 mM NaCl, 5 mM $MgCl_2$, 400 mM Imidazole, 10 % glycerol, 1 mM PMSF, 0.2 mM TCEP.

SEC buffer: 30 mM Tris pH 7.4, 200 mM NaCl, 5 mM $MgCl_2$, 10 % glycerol, 1 mM PMSF, 0.2 mM TCEP.

The overnight inoculate was performed with a single colony of *E.coli* BL21 DE3 RIPL cells containing the plasmid pPROEX-6His-RanQ69L (1-180) on LB medium containing 50 µg/ml ampicillin and 33 µg/ml chloramphenicol. TB medium cultures containing ampicillin and chloramphenicol were then inoculated with the overnight culture (1:100). Induction of expression was performed at OD_{600}=0.6-0.8 with 1 mM IPTG and the cells were incubated overnight at 18 ˚C. Harvest, resuspension, lysis and first Ni-bead incubations was performed as previously described for other NTRs. Elution from the Ni-beads was then supplemented with 0.5 mg/L expression of TEV protease and dialyzed to reduce concentration of imidazole. A second Ni affinity step was performed to remove undesired TEV protease and uncleaved

RanQ69L. The flow through from the second Ni incubation step was then concentrated and a SEC was performed (Superdex 75, 10/300). The selected fractions were then concentrated and stored at -80˚C.

Loading of RanQ69L with γ-S-GTP

Buffer 1: 30 mM Tris pH 7.4, 200 mM NaCl, 5 mM $MgCl_2$, 10 % glycerol, 2 mM EDTA, 1 mM PMSF, 0.2 mM TCEP.

The protein RanQ69L was buffer exchanged with buffer 1, this step can be done in the last step of the purification protocol. Then the protein was mixed with the non-hydrolysable γ-S-GTP in a ratio 1:20 (protein: nucleotide) in the presence of 10 mM EDTA and 1 µl of alkaline phosphatase (for 10 nmoles of RanQ69L). The sample was incubated for 2 h at 20˚C. Removal of precipitated protein was performed by centrifugation for 15 min at 13,000 rpm. 15 mM $MgCl_2$ was added to the supernatant to stabilize the nucleotide bound to RanQ69L. Excess of nucleotide can be removed by SEC.

Purification of OGT

Lysis buffer: 50 mM Tris pH 7.4, 150 mM NaCl, 5 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP.
Wash buffer 1: 50 mM Tris pH 7.4, 150 mM NaCl, 15 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP.
Elution buffer: 50 mM Tris pH 7.4, 150 mM NaCl, 400 mM Imidazole, 1 mM PMSF, 0.2 mM TCEP.
SEC buffer: 50 mM Tris pH 7.4, 150 mM NaCl, 1 mM PMSF, 0.2 mM TCEP.

LB medium containing 50 µg/ml ampicillin was inoculated with a colony (or from a glycerol stock) of *E.coli* BL21 DE3 (AI) cells containing the plasmid pET-21d-OGT1-6His (OGT1 gene was recloned from the addgene plasmid # 29760). The culture was incubated under shaking conditions overnight at 37˚C. TB medium containing 50 µg/ml ampicillin were inoculated with the overnight culture (1:100) and incubated under shaking conditions at 37˚C. Induction of protein expression was performed at OD_{600}=0.6 by the addition of 1 mM IPTG and 0.02 % arabinose. The cultures were incubated 6-8 h at 30˚C. Cells were harvested and resuspended in lysis buffer (10 ml/ L expression). Cell lysis was performed by pressure doing 3-5 rounds in the microfluidizer. After centrifugation the clear lysate was incubated for 2 h at 4˚C with Ni-beads. The beads were then washed with wash buffer (10-20 BV) and the elution was performed by adding elution buffer (5-10 BV). The eluate was concentrated and subjected to SEC at 4˚C (Superdex S200, 16/60). OGT elution volume is between 50-60 ml (Figure 3.1).

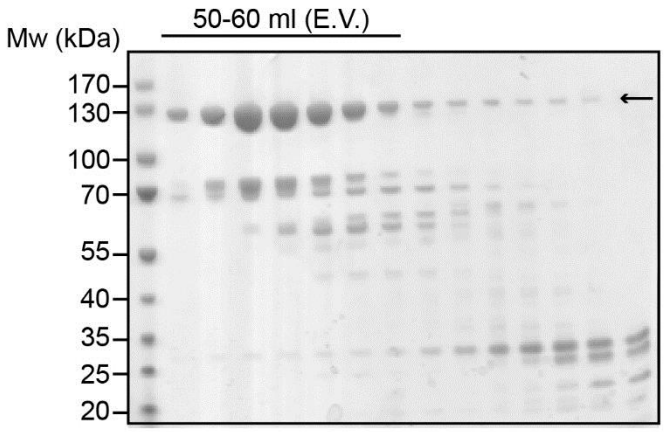

Figure 3.1. SDS-PAGE of the fractions obtained from SEC of OGT1: OGT1 (arrow) at 130 kDa is enriched at a volume between 50 to 60 ml.

3.1.2. In vitro labeling of recombinantly expressed FG-Nups

Maleimide labeling

DTT wash buffer: 4 M GdmHCl, 1xPBS pH 7, 0.2 mM EDTA, 10 mM DTT

TCEP wash buffer: 4 M GdmHCl, 1xPBS pH 7, 0.2 mM EDTA, 0.2 mM TCEP

SEC buffer: 2 M Urea, 4xPBS pH 8.5, 1 mM PMSF, 0.2 mM TCEP

Storage buffer: 4M GdmHCl, 1xPBS pH 7.4

10 nmoles of FG-Nup were buffer exchanged 3x with DTT wash buffer to reduce the Cys residues. Then the sample was buffer exchanged 5x with TCEP wash buffer to remove the excess of DTT. The sample was concentrated to 50 µl and 25-50 nmoles of maleimide dye previously resuspended in DMSO were added to the sample (not more than 10 % DMSO). Then labeling reaction was incubated for 2 h at room temperature. The excess of dye was quenched by washing with DTT wash buffer and the labeled protein was then subjected to SEC. The fractions of interest were then concentrated and buffer exchanged into storage buffer. The labeled protein was stored at -80˚C. All the buffer exchange washes were performed in amicon ultra centrifugal filters 0.4 ml with a filter cut-off of 3 kDa.

Oxime ligation

Oxime ligation buffer: 4 M GdmHCl, 1xPBS pH 4, 50 mM sodium acetate-HCl, 0.2 mM TCEP

SEC buffer: 2 M Urea, 4xPBS pH 8.5, 1 mM PMSF, 0.2 mM TCEP

Storage buffer: 4M GdmHCl, 1xPBS pH 7.4

10 nmoles of FG-Nup were buffer exchanged 3 times with oxime ligation buffer. Then 50 nmoles of hydroxylamine dye (dissolved in oxime ligation buffer) were added to the protein. Then the labeling reaction was incubated overnight at 65˚C. In the case of Nup214FG double labeled samples the oxime ligation was incubated for 48 h at room temperature to reduce the protein degradation. The excess of dye

was removed by SEC and the fractions of interest were concentrated and buffer exchanged into storage buffer. The labeled protein was stored at -80˚C.

3.1.3. *In vitro* glycosylation of FG-Nups

Glycosylation buffer: 50 mM Tris pH 7.5, 200 mM NaCl, 20 mM $MgCl_2$, 0.2 mM TCEP, 1 % Tween20

Small scale of FG-Nup glycosylation was tested on 10 nmoles of FG-Nup in glycosylation buffer containing 5 μM OGT1 and 1 mM UDP-GlcNAc (Sigma U4375). The glycosylated FG-Nup migrates higher on an SDS-PAGE gel compared to the unglycosylated form and is recognized by the antibody RL2 (antibody against O-linked-N-Acetylglycosamine) (Figure 3.2A).

Large scale in vitro glycosylation of FG-Nups was performed on the Ni-beads after the washing step but before the elution of the protein, described on the protocol of the purification of FG-Nups, adapting a previously published protocol for in vitro glycosylation of Nup98 (Labokha et al., 2013). The Ni-beads containing the bound FG-Nup were equilibrated on glycosylation buffer. Then 20x bead volume of glycosylation buffer was added to the beads with 5 μM OGT1 and 1 mM UDP-GlcNAc (Sigma U4375). The sample was incubated overnight at room temperature with mild steering in order to keep the beads suspended. Then the protocol follows with the removal of the flow through, wash of the Ni-beads and elution of the FG-Nup. The following steps are as described in the protocol for the purification of FG-Nups (Figure 3.2B).

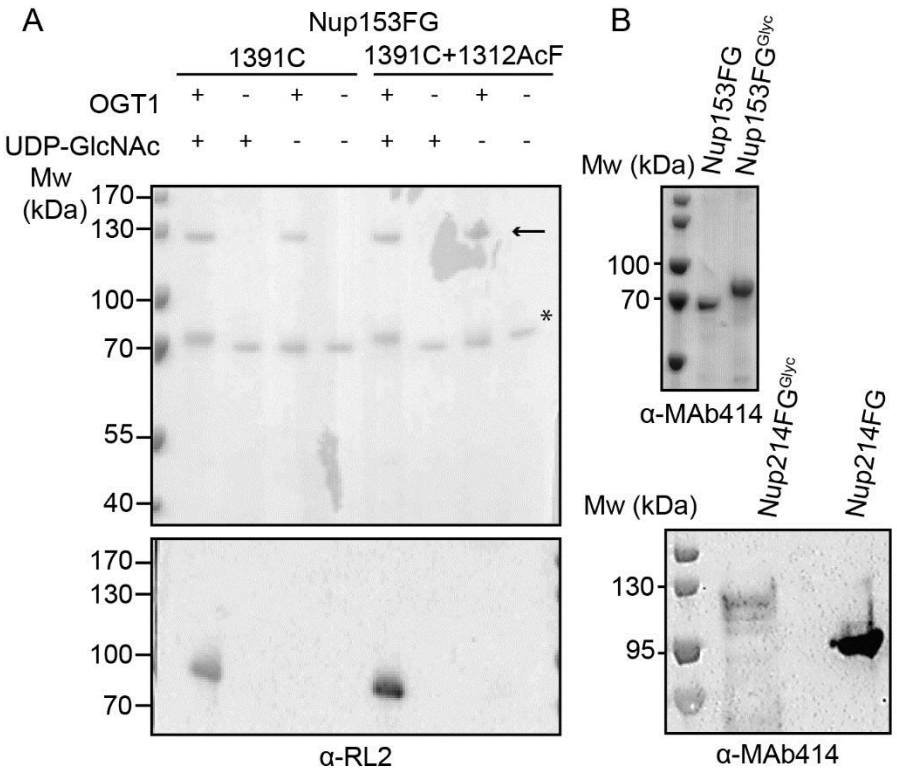

Figure 3.2. In vitro glycosylation of FG-Nups: A) SDS-PAGE (top) and western blot (WB) (bottom) of small scale in vitro glycosylation reactions of Nup153FG 1391C and 1391C+1312AcF mutants, in the presence (+) or absence (-) of OGT and UDP-GlcNAc. RL2 signal is only present when OGT1 and UDP-GlcNAc were present. The arrow indicates the OGT and the (*) the Nup153FG. B) WB against FG regions of samples obtained from large scale on-bead in vitro glycosylation of Nup153FG and Nup214FG compared to unglycosylated FG-Nups.

3.1.4. Single molecule spectroscopy

Single molecule Förster resonance energy transfer (smFRET) enables to investigate the conformational dynamics and the intramolecular distance distributions of single proteins. Moreover, it is a technique that has the capacity to resolve different population which is an advantage when studying samples with conformational heterogeneity as it is the case of IDPs. For smFRET measurements, the POI is site specifically labeled with a donor and an acceptor dye (in this thesis Alexa488 and Alexa594 respectively) with a distance between the dyes <10 nm. If the emission spectrum of the donor dye overlaps significantly with the absorption spectrum of the acceptor dye, when the donor dye is excited, the emitted fluorescence will be absorbed by the acceptor dye which will also return to the ground state by emitting fluorescence. The efficiency of energy transferred from donor to acceptor will depend on the inverse sixth power of the distance between dyes (r) (Equation 3.1) (Lakowicz, 2006); therefore FRET can be used as a molecular ruler.

Equation 3.1

$$E = \frac{R_0^6}{R_0^6 + r^6}$$

The Förster radius (R_0) is the radius at which half of the excitation energy of the donor is transferred to the acceptor fluorophore (Equation 3.2). Thus R_0 is the distance at which the efficiency of energy transferred is 50 %. R_0 depends on the quantum yield of the donor in the absence of acceptor (Q_D), the refractive index of the medium (n) , the dipole angular orientation of each dye (K^2) and the spectral overlap integral of the donor and acceptor pair (J) (Schuler et al., 2012). N_A is the Avogadro's number.

Equation 3.2

$$R_0^6 = \frac{9000\ ln10\kappa^2 Q_D J}{128\pi^5 n^4 N_A}$$

The single molecule fluorescence experiments were performed on a custom-built multiparameter spectrometer confocal setup. Excitation of freely diffusion fluorescently labeled proteins was performed using linearly polarized light from a picosecond laser diode (LDH 485 Picoquant) and from a white light laser (SuperKExtreme, NKT Photonics). Pulsed excitation light in which the orange excitation had a laser delay of 25 ns compared to the green excitation was achieved by a computer controlled multichannel picosecond diode laser driver (Hydraharp 400, Picoquant). In this way the stoichiometry of the dyes can be obtained from the measurement (Kudryavtsev et al., 2012). The excitation beams were filtered through excitations filters (482/18) and 527/15) respectively. Then the sample was excited through a high numerical aperture water immersion objective (60x, 1.27 NA). Emitted fluorescence light was then filtered by a 0.1 mm pinhole and was passed through a multi-band fluorescence bandpass filter (488/568/660). Then the fluorescent signal is split by a polarization beam splitter in parallel and perpendicular polarized light. Following the fluorescence light is spectrally separated by emission filters (525/50) for the green (donor) and (620/60) for the orange (acceptor) fluorescence. Single photon counting detectors (τ-SPAD and PMA Hybrid detectors) were used to detect the light from the acceptor and the donor dyes respectively. The photon signals were acquired with a multichannel time correlated single photon counting module (Hydraharp400, Picoquant).

The identification of single bursts of fluorescence was performed by a burst search algorithm. Burst were first identified through a time filter to discriminate a burst from shot noise (<0.16 ms is considered a burst) .The obtained signal was then smoothened by a Lee filter (Enderlein et al., 1997; Schaffer et al., 1999). The identified burst were then selected with an intensity threshold between 70-100. In addition, burst corresponding to signal from single molecules and not clusters or aggregates was selected by discarding burst longer than 8 ms and brighter than 600 photons.

The fluorescence intensity (I), lifetime (τ) and anisotropy (r) was then obtained from individual burst and subjected to multiparameter fluorescence analysis. The data was analyzed with a custom written program using Igor Pro (Wavemetrics).

The stoichiometry and FRET efficiency (E_{FRET}) of each burst was calculated as indicated in the equation 3.3 and the stoichiometry of the dyes by using the equation 3.4 (Schuler et al., 2012; Sisamakis et al., 2010a). I_{xy} x correspond to the excitation source and y the detection channel (D corresponds to donor laser excitation or donor detection and A acceptor excitation or detection). The γ factor corrects for the detection efficiency differences of the donor and acceptor channel .

Equation 3.3

$$E_{FRET} = \frac{I_{DA}}{\gamma I_{DD} + I_{DA}}$$

Equation 3.4

$$S = \frac{\gamma I_{DD} + I_{DA}}{\gamma I_{DD} + I_{DA} + I_{AA}}$$

One of the powerful features of multiparameter smFRET is the ability to detect donor lifetimes to obtained information about the dynamics of the sample measured. For instance, in a static molecule all the measured molecules will contain the same conformation, this would follow the relationship of E_{FRET} vs and lifetime of the donor (τ_D) will be described by the equation 3.5. When molecules are dynamic, they will be undergoing multiple conformational changes and this will be reflected in the relationship between the donor lifetime and the E_{FRET} (Kalinin et al., 2010).

Equation 3.5

$$E_{static} = 1 - \frac{\tau_{D(A)}}{\tau_{D(0)}}$$

where $\tau_{D(A)}$ and $\tau_{D(0)}$ is the lifetime of the donor in the present and absence of FRET, respectively.

The double labeled samples were measured at 50 pM in solution in 1x PBS pH 7.4 containing 2 mM DTT and 2mM MgAc for 30 minutes. In the cases where very few bursts were detected per measurement, signal from different measurements was accumulated. FG-Nup/NTR smFRET measurements were performed unless stated differently at 1 μM of NTR.

3.1.5. Fluorescence correlation spectroscopy (FCS)

Fluorescence correlation spectroscopy (FCS) is a method that can be used to determine the diffusion the diffusion coefficient of fluorescently labeled molecules. The fluorescence intensity fluctuations generated over time by the diffusion of labeled molecules through a limited detection volume can be autocorrelated (Equation 3.6) to obtain information about the diffusion correlation time (or average dwell time of the

labeled molecule in the confocal volume) and of the number of molecules that were detected (N) during the measurement (Lakowicz, 2006; Serdyuk et al., 2007).

Equation 3.6

$$G(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I\rangle^2}$$

The autocorrelation function compares the fluorescence intensity at a time (I(t)) with the fluorescence intensity at a later time point (I(t+τ)), as a result the autocorrelation function G(τ) is obtained (Lakowicz, 2006; Serdyuk et al., 2007). Since the excitation was performed with a laser beam with a Gaussian profile and the emitted fluorescence was collected through a confocal pinhole, the illumination confocal volume can be described by a three dimensional Gaussian function (Equation 3.7). Where τ_D is the diffusion time required by the fluorescent molecule to diffuse through the confocal volume, N the number of molecules observed during the measurement, r and l are the radial and axial dimensions of the confocal volume also known as structural parameter.

Equation 3.7

$$G(\tau) = \frac{1}{N}\left(\frac{1}{1+\frac{\tau}{\tau_D}}\right)\left(\frac{1}{1+\left(\frac{r}{l}\right)^2\frac{\tau}{\tau_D}}\right)^{1/2}$$

The diffusion time of a molecule will change depending on the size of the confocal volume, which will depend on the wavelenght of the laser used and the optics of the instrument used. However, one can calculate the diffusion coefficient (D), which is constant for a molecule, of a molecule knowing the diffusion time and the size of the confocal volume (Equation 3.8) (Serdyuk et al., 2007)..

Equation 3.8

$$\tau_D = \frac{r^2}{4D}$$

FCS experiments were performed with the diffusion confocal setup previously described (section 3.1.4). Emitted fluorescence signal was detected in parallel and perpendicular polarization directions and the correlation curve obtained was a result of the cross-correlation between both polarized detected channels. FCS measurements were performed in 1x PBS pH 7,4, 2mM DTT and 2mM MgAc. Unless stated the concentration of labeled protein on the FCS experiments was of 10 nM. FCS was used to perform titration binding curves of different FG-Nup/NTR complexes and to obtain the D for FG-Nup and Importinβ. When smFRET FCS is shown (Figure 4.1) the FCS curves were obtained from the intensity fluctuations detected in the acceptor channel upon donor excitation.

3.1.6. Stopped-flow spectroscopy

Stopped-flow spectrophotometer

Stopped-flow kinetic experiments were performed using a 3 syringe system stopped-flow spectrophotometer (SFM-300, Bio-logic) (Figure 3.3A). The advantage of a 3 syringe system is that it allows the automatic titration of one of the reagents over the other one. On a usual binding kinetic measurement, the labeled protein (A) is loaded on S3 and the binding partner on S1 (B), S2 is loaded with buffer. The reaction of binding of A to different concentration of B can be achieved by mixing the content of S1 and S2 in different ratios in the first mixer M1. The concentration of A can then be kept constant by mixing the solution coming from M1 with the solution from S3. Both solutions are mixed in M2. The mixed solution from M2 will fill the observation cuvette and the binding reaction will be monitored.

Due to the requirements of the sample, I optimized the hardware of the spectrophotometer to achieve lower dead times and higher signal with lower sample consumption. I used the uFG-08 micro-cuvette accessory which has a volume of 20 µl and is placed closed to the second mixer (M2) reducing the dead time required to reach for the mixed solution to reach and fill the observation cuvette before the start of the recording of the reaction. In addition, aiming to reduce the amount of labeled FG-Nup required to obtain a good signal to noise ratio, I incorporated a custom polarized laser excitation source (Figure 3.3B). I used a collimated laser diode module at 532 nm and a power of 4.5 mW (CPS532, Thorlabs) followed by a glan-laser polarizer (10 mm, coating 350-700 nm) and a laser clean-up filter (ZET 532/10) and a plano-convex cylindrical lens (H=20, L=22, f=50) to achieve polarized excitation at 532 nm. The cylindrical lens was used to focus the beam on the excitation window of the observation cuvette. When other excitation wavelength was required the Xe-Ar commercial lamp (Bio-logic) was used together with the commercial excitation path. Emitted fluorescence collimated and passed through emission filters before detection by PMT tubes placed at 90˚ from the excitation source. Depending on the detected wavelength different emission filters were used (Cy3B fluorescence: 590/104,Semrock F37-596/ 488: 525/50, Semrock FF01-525250-25/ 594: 620/60 (ET 620/60M)).

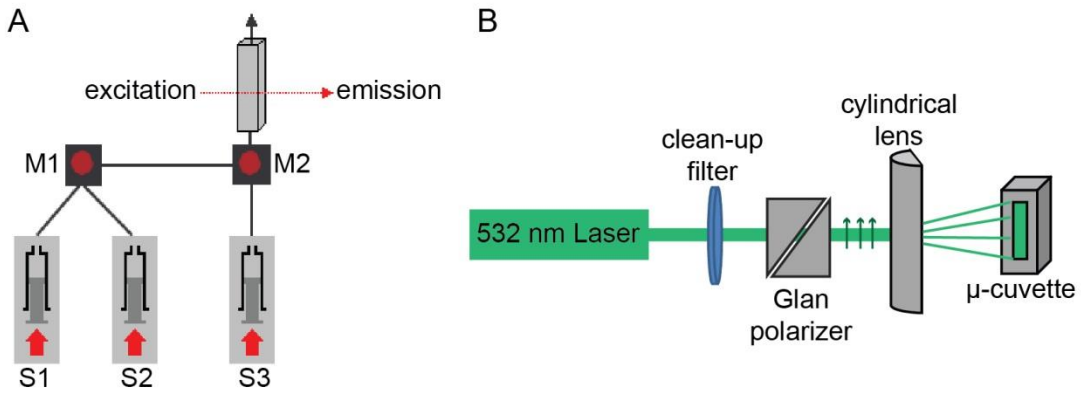

Figure 3.3. Cartoon representation of the stopped-flow spectrophotometer: A) Scheme (adapted from Bio-logic) of the three syringe system (S1, S2 and S3). M1 and M2 correspond to the mixers. B) Scheme of the custom built excitation path.

Fluorescence intentisites were measured to monitor the association between the molecules. For the cases where anisotropy was used to monitor the binding , polarizing filters in parallel (∥) and perpendicular (⊥) position to the excitation source were used. The anisotropy (r) was calculated (Equation 3.9) (Lakowicz, 2006).

Equation 3.9

$$r = \frac{I_{\parallel} - G \cdot I_{\perp}}{I_{\parallel} + G \cdot I_{\perp}}$$

Rate constants for association and disassociation

The binding reaction between two molecules can be described as:

Equation 3.10

$$A + B \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} AB$$

The rate of a binding reaction describes the change in the concentration of the binding molecules (for example [A] and [B]) or of the bound complex [AB] over time. The formation of AB over time will depend on to the association and dissociation rate constant (k_{on} and k_{off} respectively) (Equation 3.11) (Goodrich and Kugel, 2007). In a bimolecular interaction the rate of association is equal to k_{on}[A][B] and the rate of dissociation is equal to k_{off}[AB].

Equation 3.11

$$\frac{d[\mathrm{AB}]}{dt} = k_{\mathrm{on}}[\mathrm{A}][\mathrm{B}] - k_{\mathrm{off}}[\mathrm{AB}]$$

The determination of the rate constant of a reaction can be achieved by monitoring the formation of AB or the disappearance of A or B over time. However, the k_{on} is not directly measured instead the observed rate

constant (k_{obs}) is measured for a determined concentration of A and B. In order to simplify this second order reaction, typically the measurements are performed under pseudo-first order conditions, which involved having one of the reactants at much higher concentrations than the other. In this case for example B will be under pseudo-first order conditions (>10x higher concentration than A) while A will remain constant. By having B at much higher concentrations, during the formation of AB the decrease of B will be negligible being the concentration of free B after AB is formed very close to the initial total concentration ($[B_{free}] \approx [B_{total}]$). Under these conditions the limiting reactant for the formation of AB will be [A] and the relationship between k_{obs}, k_{on}, k_{off} and $[B_{total}]$ is as described with a linear equation (Equation 3.12) (Goodrich and Kugel, 2007).

Equation 3.12

$$k_{\mathrm{obs}} = k_{\mathrm{on}}[\mathrm{B}] + k_{\mathrm{off}}$$

The k_{obs} can be measured by monitoring the formation of AB over time, as shown in the equation 3.4, the k_{obs} values will change with the concentration of B. Thus, when a binding kinetic measurement is performed at increasing concentration of excess of [B], the k_{obs} retrieved from the measured traces will also be faster with higher [B] (Figure 3.4A). The function that describes reaction is usually a single exponential (rise or decay) (Equation 3.13) (Goodrich and Kugel, 2007).

Equation 3.13

$$\mathrm{AB(t)} = \mathrm{Amp}\left(1 - e^{-kobs \cdot t}\right) + \mathrm{AB(t0)}$$

The obtained k_{obs} at different concentration of B, when plotted against [B] from the linear fit the slope will correspond with the k_{on} and the y intercept will be equal to the k_{off} (Figure 3.4B). The y-intercept corresponds to the k_{off} because at very low concentrations of B the reaction will be dominated by the k_{off}. Because the k_{off} obtained from the linear regression will be dependent on the quality of the fit and the y-intercept can vary significantly, it is more reliable to determine the k_{off} experimentally by performing dissociation kinetic experiments (see below).

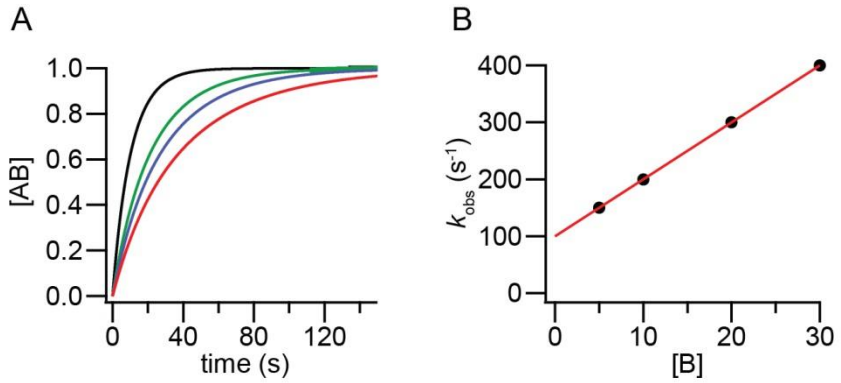

Figure 3.4. Example of a typical kinetic binding experiment: A) Different traces at a constant concentration A and increasing concentration of [B] (red: 30, blue: 40, green: 50 and black: 100 µM). B) Example of k_{obs} vs concentration plot with a k_{on} of 10 $M^{-1}s^{-1}$ and a k_{off} equal to 100 s^{-1}.

Determination of association rate constant

Stopped-flow association experiments were performed using the automatic mixing function (concentration dependent study) with the software Bio-kine32 V.4.72 (Bio-logic). Measurements were performed with 1xPBS pH 7.4, 2 mM DTT and 2 mM $Mg(CH_3COO)_2$ (MgAc). Labeled FG-Nup was loaded on S3 and measured at 20 nM. NTR was mixed at different concentration ranges, in most cases 200-400 nM at 50 nM steps. Single labeled Cy3B FG-Nup was used when the anisotropy was used to monitor the binding. Donor and acceptor intensities, of Nup214FG labeled with Alexa488 and Alexa594, were used to measure the rate of conformational change associated to the binding to CRM1. The measurements were performed at 20˚C. The fitted traces were the results of background substraction and averaging of 20-30 individual traces, depending on the sample measured, replicates. In the case of Nup153FG measurement a double exponential function was used to fit the observed traces (Equation 3.14).

Equation 3.14

$$f(t) = A1 \cdot (1 - e^{-k1obs(ultrafast)\cdot t}) + A2 \cdot (1 - e^{-k2obs\,(fast)\cdot t}) + c$$

The measured traces obtained from association experiments with Nup214FG or $Nup153FG^{Glyc}$ were fitted with a single exponential function. The obtained k_{obs} were then plotted against the NTR concentration and the k_{on} was retrieved from the linear fit as previously described. An experimental error between replicates of ~ 20 % was estimated by performing the same measurements repeatedly over several months of different expression batches. The factors affecting the deviation between replicates are to use the accurate concentration, purity of the used proteins or labeling efficiency of the Nups which will be linked to the amplitude change of the signal upon binding which will affect the quality of the fit.

Determination of dissociation rate constant

The kinetic dissociation experiments were performed by measuring the unbinding of the preformed complex between Nup153FG and Importinβ upon addition of unlabeled Nup153FG. The final concentration of labeled Nup153FG and NTR is the same as the one used on the association kinetic experiments (20 nM Nup and 200 nM NTR) and 50-100x excess of unlabeled protein was used (1-4 μM unlabeled FG-Nup) to measure the complex dissociation. The obtained traces were fitted with an exponential decay function (Equation 3.15). Dissociation reactions were performed at multiple concentrations of excess of unlabeled protein because in these highly dynamic reversible interactions different competitor conditions must be tested to reduce the association reaction.

Equation 3.15

$$f(t) = A1 \cdot e^{-koff \cdot t} + c$$

The k_{off} can be used to calculate the half-life of a complex. The half-life of a complex is the time that it takes for half of the existing bound complex to be dissociated. For a first-order dissociation reaction this time will depend only on the k_{off} thus it will be the same independently of the starting concentration of the bound complex ($[AB]_i$). If we consider that 50 % of the initial bound complex is unbound ($[AB]_t$=0.5), the complex half-life will be equal to 0.693 divided by the k_{off} (Equation 3.16) (Goodrich and Kugel, 2007).

Equation 3.16

$$\frac{[AB]_t}{[AB]_i} = e^{-koff \cdot t_{1/2}} \rightarrow 0.5 = e^{-koff \cdot t_{1/2}} \rightarrow -\ln(0.5) = koff \cdot t_{\frac{1}{2}} \rightarrow t_{\frac{1}{2}} = \frac{0.693}{k_{off}}$$

Determination of diffusion limit

The fastest association rate at which two molecules bind is known as diffusion limited k_{on}. When two molecules bind with diffusion limited kinetics it means that the limiting step for the molecules to interact depends on the diffusion of the molecules. However, not all the contacts between two proteins result in the successful binding, the right orientation between the molecules is also required. The diffusion rate depends on the size and shape of the molecules, the viscosity of the medium and the temperature (Equation 3.17).

Equation 3.17

$$k_{on} = 4\pi DR = N_A \frac{2k_B T}{3\eta} \left(\frac{D_A}{D_B} + \frac{D_B}{D_A} + 2\right)$$

The diffusion limited rate constant was calculated by using an Einstein-Smoluchowski approximation (Equation 3.17.). Where k_B is the Boltzmann constant, T is the temperature, D_A and D_B the diffusion coefficients for Nup153FG and Importinβ and η is the viscosity of the medium.

In a second-order interaction, the association will depend on the concentration of the reactants and the k_{on}. At the same time the k_{on} will depend on the size, number of available interaction sites and the sum of the diffusion of both reactants. Large molecules like proteins will have a slower diffusion than small molecules like ATP. However, even if small molecules diffuse faster it often compensates their small size and causing the k_{on} of a protein-protein and of a protein-small molecule very similar. The diffusion coefficients (D) for Nup153FG and Importinβ labeled with Atto655 where calculated using FCS. The measured D for each component was D_{Nup153} $3.83x10^{-5}$ and $D_{Importinβ}$ $2.34x10^{-5}$ cm^2s^{-1}. The obtained diffusion limited k_{on} for these proteins was of $6.8x10^9$ $M^{-1}s^{-1}$.

Determination of $k_{on,basal}$

The determination of the k_{on} at infinite ionic strength ($k_{on,basal}$) to test whether the diffusion limited k_{on} was influenced by long-range electrostatic interactions, kinetic binding experiments were performed at different ionic strengths. The experiments were performed at 20 mM MOPS pH 7,4, 2 mM DTT and the ionic strength was modified addition of NaCl. The ionic strength at different concentrations of NaCl was calculated using the equation 3.18, being c_i the molar concentration of each ion present in the solution and z_i^2 their corresponding charge.

Equation 3.18

$$I = 1/2 \sum_{i=1}^{n} c_i \cdot z_i^2$$

A k_{on} values was obtained from each kinetics association experiment performed at different ionic strength conditions. The logarithm of the obtained k_{on} values was plotted against the square root of the ionic strength. The data was fit with a Debye-Hückel like approximation previously used for the determination of the $k_{on,basal}$ (Shammas et al., 2014) (Equation 3.19).

Equation 3.19

$$\ln(k_{on}) = \ln\left(k_{on,basal}\right) + \frac{AB}{BR} \cdot \frac{I^{-0.5}}{BR + I^{-0.5}}$$

Where I corresponds to the ionic strength of the buffer and AB and BR are used as free fitting parameters.

3.1.7. Mammalian cell culture techniques

Cell culture

Human embryonic kidney (HEK297T) cells (ATCC CRL-3216) were cultured at 37˚C in 5 % CO_2 atmosphere in Dulbecco's modified Eagle's medium (Life Technologies 41965-039) supplemented with 1 % penicillin-streptomycin (Sigma P0781), 1 % L-Glutamine (Sigma G7513), 1 % sodium pyruvate (Life Technologies 11360) and 10 % FBS (Sigma F7524). The cells were passaged every 2-3 days up to 15-18 passages.

Transfected mammalian cells and expression of POI containing ncAA

Transfections were performed with 70 % confluence. All transfections were performed using PEI, which condenses DNA into positively charged particles that bind to anionic surface residues on the cells and are incorporated via endocytosis. Experiments requiring the Amber suppression system were transfected at a ratio 1:1 with a vector containing the POI^{TAG} and the $tRNA^{Pyl}$/$NESPylRS^{AF}$ (in some cases additional vectors like TEV protease containing vectors were also transfected). 4 h after transfection the medium was exchanged and the ncAA BOC-Lys (t-butyloxycarbonyl-lysine) was added to a final concentration of 250 µM. When required, the translation was stopped by adding 100 µM cycloheximide (CHX).

Immunolabeling and imaging

The medium from cells seeded on a 4-well or 24-well plate with glass bottom was removed and rinsed with 1xPBS. Cell fixation was performed with 2 % PFA incubated for 10 minutes at room temperature. Then the cells were rinsed and permeabilized with 0.5 % Triton-x 100 in PBS for 15 minutes at room temperature. Cells were then rinsed twice and blocked with 3 % BSA in 1x PBS for 1.5 h at RT. Blocking solution was removed and the cells were then incubated with the primary antibody, usually overnight at 4˚C. The cells were rinsed and incubated with the secondary antibody for 1 h at RT. Then the cells were rinsed with 1x PBS and incubated for 10 min with Hoechst for nuclear staining. After rinsing with 1x PBS the cells were imaged using confocal microscopy.

The antibody dilutions are specified on the methods section 3.2.2.

Confocal images were acquired on a Leica SP8 STED 3x microscope with a 63x/1.4 oil immersion objective using 488 nm for the GFP excitation. The FRAP experiments were performed on a Zeiss LSM 780 using a 63x/1.4 oil immersion objective.

Flow-cytometry

Flow-cytometry was used to analyze the fluorescence signal of transiently transfected single cells. Cells were analyzed by flow-cytometry 24-36 h after transfection. The medium was removed and the cells were resuspended in cold 1x PBS pH 7.4. The data was acquired in a LSRFortessa SORP Cell Analyzer (BD) and the data analysis was performed using the FlowJo software (FlowJo). Cells were first gated by cell type using FSC-A and SSC-A scattering parameters and then only the single cells were analyzed using FSC-A and SSC-W scattering information. The fluorescence signal was acquired for GFP and iRFP signal using the 488-530/30 and 640-730/45 channels respectively.

Western-blot

Cells resuspended on RIPA buffer (150 mM NaCl, 1% Triton-x-100, 0.5 % sodium deoxycholate, 0.1 % SDS and 50 mM Tris pH 8) and lysed by sonication. The total amount of protein was quantified using the Pierce 660 nm protein assay with the Ionic detergent compatibility reagent (ThermoScientific 22663). The lysate was mixed with SDS loading buffer and 15-20 µg of total protein were loaded on an SDS-PAGE gel (NuPage 4-12% NP0329). The protein was transferred to a nitrocellulose membrane (Trans-Blot Turbo (64102376) using the BioRAD Trans-Blot Turbo transfer system. The membrane was then blocked for 1 h with 5 % milk or 3 % BSA in 1x PBS pH 7.4. The primary antibody was incubated overnight on rotation at 4˚C. Then the membrane was washed 3 times with 1x PBS 0.2 Tween-20 and incubated with the secondary antibody for 1 h. Following, the membrane was washed and incubated for 1 minute with ECL Western blotting detection (GE RPN2106). The membrane was imaged with a Chemidoc MP (BioRad western blot imager)

3.2.Material

3.2.1.List of frequently used plasmids

Plasmid	Internal number	Labeling site
Bacterial expression		
pTXB3-12His-Importin beta WT	694	
pTXB3-12His-ImportinBeta-I178D-Y255A	691	
pBAD-N153FG-intein-CBD-12His	85	
pTXB3-6His-N153FG-Int-Chitin_ 1312Cys_1391TAG	659	1391 AcF+ 1312C
pTXB3-6His-N153FG-Int-Chitin_ 1391Cys	672	1391C
pTXB3-6His-N153FG-Int-Chitin_Ser990Cys_Ser1094TAG	680	1094 AcF+ 990C
pTXB3-6His-N153FG-Int-Chitin_Ser883Cys_Ser994TAG	1199	990 AcF+ 883C
pTXB3-6His-N153FG-Int-Chitin_Ser883Cys	675	883C
pBAD-Nup214-intein-CBD-12His	1401	2043 AcF+1905C
pTXB3-12His-Importin beta WT	694	
pTXB3-12His-ImportinBeta-I178D-Y255A	691	
pQE60-CRM1-intein-12His	476	
pQE60-6His-TRN1	410	
pTXB3-NTF2-intein-6His	397	
pROEX-RANQ69L (1-180) Human	442	

Plasmid	Internal number
Mammalian cell expression	
pcDNA3.1_CyclinB-Myc-linker-Tev site-iRFP-GFP39TAG-linker-FUS-FLAG	1631
pcDNA3.1_ER50-Myc-linker-Tev site-iRFP-GFP39TAG-linker-FUS-FLAG	1632
pcDNA3.1_UnaG-Myc-linker-Tev site-iRFP-GFP39TAG-linker-FUS-FLAG	1633
pcDNA3.1_DHFR-Myc-linker-Tev site-iRFP-GFP39TAG-linker-FUS-FLAG	1634
pcDNA3.1_FKBP-Myc-linker-Tev site-iRFP-GFP39TAG-linker-FUS-FLAG	1627
pcDNA3.1_FUS-TEV	1628
Addgene 58878 FKBP-NTEV	1629
Addgene 58879 FKBP-CTEV	1630
pcDNA3.1 Lck-FUS-TEV	389
pcDNA3.1 Lck-TEV	390
pcDNA3.1 iRFP-GFP-FUS-CAAX	391
pcDNA3.1 iRFP-GFP-CAAX	392
pcDNA3.1 iRFP-GFP-FUS-CAAX KRas	393
pcDNA3.1 iRFP-GFP-CAAX KRas	394
pcDNA3.1 Lck-FUS-NTEV	-
pcDNA3.1 Lck-NTEV	-

3.2.2. List of antibodies

Target	Type	Source	Label	Dilution	Company/cat.number
mAB414	primary	-		WB: 1:1000	Biolegend #902901
Anti-HA	primary m.	mouse	-	IF: 1:1000 WB:1:10.000/20.000	Sigma H9658
RL2 (GlcNAc)	primary m.	mouse	HRP	WB: 1:1000	ThermoFisher MA1-072
CyclophilinB	primary p.	Rabbit	-	WB: 1:10.000	ThermoFisher PA1-027A
GFP	primary p.	Rabbit	-	WB: 1:1000	Santa Cruz biotech sc-8334
Mouse	secondary	Goat	A647	IF: 1:1000	ThermoFish A32728
Rabbit	secondary	Donkey	HRP	WB: 1:10.000	Jacksonimmuno 711-035-152
Mouse	secondary	-	HRP	WB: 1:10000	Pepcore
Mouse	secondary	Goat	A532	IF: 1:1000	ThermoFisher A-11002

3.2.3. Commercial Kits

Commercial Kit	Company
Pre-diluted protein assay standards (BSA) set	ThermoScientific (23208)
Quick PCR purification kit	Invitrogen (K310002)
Quick Plasmid Miniprep Kit	Invitrogen (K210011)
HiPure Plasmid Maxiprep Kit	Invitrogen (K210007)
Quick Gel extraction kit	Invitrogen (K210012)
BCA Protein Assay Kit	ThermoScientific (23227)
Ionic Detergent Compatibility Reagent for Pierce™ 660nm Protein Assay Reagent	ThermoScientific (22663)
CellTrace™ Calcein Violet, AM, for 405 nm excitation	Invitrogen (MP 34858)
UDP-GloTM Glycosyltransferase assay	Promega (V6961)
ECL Western blotting detection	GE RPN2106

3.2.4. Chemicals/ dyes & ncAA

Frequently used chemicals	
Imidazole	Sigma (I5513)
Hoechst 33342	Sigma (B2261)
FBS (Fetal Bovine Serum)	Sigma (F7524)
1,4- Dithiothreitol (DDT)	biomol (04010)
2-Mercaptoethanol	Aldrich (M6250)
Complete EDTA free (Protease inhibitor cocktail tablets)	Roche (11873580001)
Magnesium acetate tetrahydrate	Sigma-Aldrich (M2545)
Paraformaldehyde	Sigma 158127
Triton X-100	AppliChem (A1388.0500)
Urea	Merck (1.08487.1000)

Chloramphenicol	AppliChem (A6435)
Dulbecco's Eagle medium	Life Technologies 41965-039
Sodium pyruvate	Life Technologies 11360
L-Glutamine	Sigma (G7513)
Ampicillin-sodium salt	AppliChem (A08390025) 25 g
Guanosine 5′-triphosphate sodium salt	Sigma (G8877)
Penicillin-streptomycin	Sigma (P0781)
Chitin Resin	NEB (S6651)
His-Pur Ni-NTA Resin	ThermoScientific (88223)
Uridine 5′-diphospho-N-acetylglucosamine sodium salt (UDP-GlcNAc)	Sigma (U4375)

single molecule buffer reagents	
EDTA (0.5 M EDTA, pH 8)	Ambion (AM9261)
Guanidine hydrochloride	USB (75823 500 GM)
Magnesium acetate tetrahydrate	Merck (1.05819.0250)
PBS 10x pH 7.4	Gibco (70011-036)
Urea	Affymetrix (75826)

Synthetic dyes		
Alexa Fluor 488	Hydroxylamine	Invitrogen
Alexa Fluor 488	Maleimide	Invitrogen
Alexa Fluor 594	Maleimide	Invitrogen
Cy3B	Maleimide	GE Healthcare
Atto 655	Maleimide	Atto-tec

ncAA	
AcF	Synchem
H-L-Lys(BOC)-OH	IRIS Biotech 2418-95-3

3.2.5.Cell lines

Cell type	characteristic	company
Bacteria strains		
E.coli BL21 (DE3) AI	Arabinose inducible promoter araBAD upstream of T7 RNA pol. Gene used in protein expression	Invitrogen
E.coli XL10 Gold cells	High transformation with large plasmids, expression host strain for large plasmids, for protein expression	Stratagene
E.coli BL21 (DE3) RIPL	Extra copies of argU,ileY,proL,leuW tRNA for protein expression	-

E.coli Top 10	Used for plasmid amplification and cloning	Invitrogen
	Eukaryotic strains	
HEK293T	highly transfectable cell line	ATCC

3.2.6. List of software

List of special software	
Igor Pro	Wavemetrics
BioKine32	BioLogic
ImageJ	NIH
SymPhoTime	Picoquant
LAS AF Lite	Leica
PyMOL	Schrödinger
ZEN 2.3 Lite	Zeiss
FlowJo	FlowJo LLC
EndNote	Clarivate Analytics
SnapGene	GSL Biotech LLC
Image Lab 5.2.1	Bio-RAD

Chapter 4

Results

4. Results

The results chapter is divided in two main sections; the first section focuses on the *in vitro* structural and kinetic study of the interaction mechanism between FG-Nups and NTRs (section 4.1). The second section introduces the synthetic biology approach that I am developing to achieve the study of multiple disordered regions of FG-Nups at the NPC *in cellulo* (section 4.2).

4.1. Biochemical study of the interaction between FG-Nups and NTRs

In this section I will show the results obtained from the biochemical study of the interaction between FG-Nups and NTRs. I will first show the results obtained from the study of the structure and dynamics of FG-Nups upon binding to NTRs using multiparameter single-molecule spectroscopy. Following, I will show the results from kinetic studies using stopped-flow spectroscopy, technique that I have established in our laboratory to study the kinetics of binding between FG-Nups and NTRs. Additionally, together with Dr. Piau Siong Tan I have worked on studying the interaction between FG-Nups and NTRs of a particular disordered region of Nup214 (Nup214FG), which seems to have a distinct binding mechanism to the other FG-Nups and NTRs. In addition, motivated by the importance of PTMs in IDPs and knowing that FG-Nups are heavily glycosylated in vertebrates I studied the effect of glycosylation in the NTR/FG-Nup binding.

4.1.1. Single molecule studies of the binding of NTRs to FG-Nups

Multiparameter single-molecule FRET spectroscopy (smFRET) enables the study of the structure and dynamics of a double-labeled sample containing a FRET dye pair. The FRET dye of choice requires a spectral overlap between the emission spectrum of the donor and the absorption spectrum of the acceptor. The energy transferred between the donor and acceptor dye can be monitored in a range from 2-10nm allowing this technique to be used as a molecular ruler. Based on previous smFRET work from out lab I started working with the disordered region of Nup153, here termed as Nup153FG (875-1475 aa of the full length Nup153 UniProt:P49790). Recombinantly expressed Nup153FG containing a single cysteine and the ncAA AcF was site-specifically labeled conjugating Alexa594-maleimide to the thiol group of the cysteine (acceptor dye) and reacting the ketone group of the AcF with Alexa488-hydroxylamine (donor dye) (Figure 4.1). smFRET was used to monitor any possible conformational change that would take place in the FG-Nups upon binding to the folded NTRs that would cause a change in the efficiency of energy transfer from the donor to the acceptor dye. The multiparameter single-molecule detection combined with burst-wise analysis enables the calculation of FRET efficiencies (E_{FRET}) and fluorescence lifetimes (τ) of the freely diffusing molecules. The analysis of the 2D-histogram of fluorescence τ vs

E_{FRET} can be used to reveal information about the dynamics of the molecules crossing the confocal volume (details of smFRET and the analysis will be extended in the materials and methods chapter section 3.1.4). The dynamics of the measured system can easily be detected by plotting the static FRET line, which shows the relationship between τ vs E_{FRET} for static molecules. Molecules having a distribution of the FRET population not following this line, due to a variable τ and E_{FRET} caused by the different distances monitored within the multiple adopted conformations is indicative of protein dynamics (Dimura et al., 2016; Kalinin et al., 2010; Sisamakis et al., 2010b).

The smFRET measurement of Nup153FG (labeled at positions 1312AcF and 1391C with A488 and A594 respectively) revealed an $E_{FRET=}$ 0.6 (Figure 4.1A). In addition, the FRET population stays above the static FRET line indicating that unbound Nup153FG is a dynamic protein. Upon addition of unlabeled Importinβ, which was previously reported to bind preferentially in the PxFG-rich region of Nup153FG located between the FRET dye pairs (Milles and Lemke, 2014), no substantial change in the E_{FRET} values or in the width of the histogram was detected (Figure 4.1A). This indicates that there is no major change in the distance distribution of Nup153FG upon binding to Importinβ. Similarly to the unbound state, the bound Nup153FG/Importinβ FRET population was also located above the static FRET line indicating that bound Nup153FG is also dynamic and is populating an ensemble of interconverting conformations also in the bound form (Figure 4.1A). Fluorescence correlation spectroscopy (FCS) can be used to measure the translational diffusion time of molecules by monitoring and correlating the fluorescence intensity fluctuations though a confocal volume. FCS traces were obtained by analyzing the signal obtained from the acceptor detector channel upon donor laser excitation showed that in the presence of Importinβ Nup153FG diffuses slower (green FCS trace) than in the unbound form (black FCS trace) confirming that the complex is formed (Figure 4.1B).

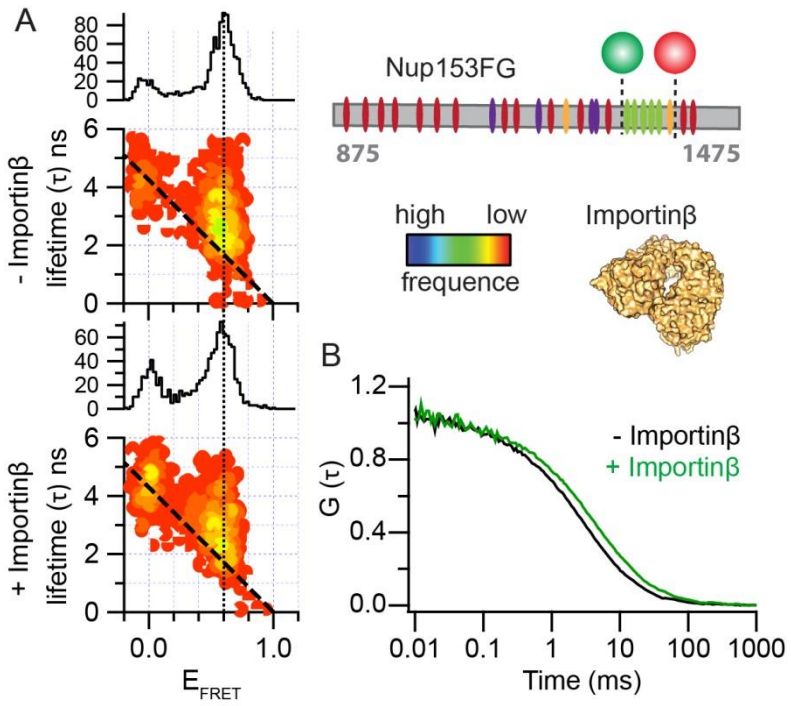

Figure 4.1. Nup153FG binds Importinβ dynamically without affecting the ensemble of conformations: A) FRET efficiency (E_{FRET}) versus fluorescence lifetime (τ) histograms of Nup153FG labeled with Alexa488 and Alexa594 at positions (1312AcF and 1391C respectively) in the absence and presence of Importinβ. The histograms are color coded according to the frequency of occurrence. The dotted line marks the center of the FRET population and highlights that the unbound and the bound population have the same E_{FRET} values. The dashed diagonal line indicates the relationship between τ and E_{FRET} for static systems. Since the FRET population (E_{FRET}=0.6) is not within this line, Nup153FG is dynamic in the unbound and bound form. The population with E_{FRET}=0 corresponds with the detected molecules that contained only donor dye. B) FCS obtained from the single-molecule measurement using the acceptor signal obtained from donor excitation. The translational diffusion of Nup153FG is slower in the presence of Importinβ, indicating that Nup153FG is bound to Importinβ.

In order to determine if this effect is a general mechanism present in the interaction between FG-Nups and NTRs or if it is a characteristic of the binding between the PxFG-rich region of Nup153FG and Importinβ, the same experiments were repeated probing the conformation and dynamics for two different FxFG-rich region of Nup153FG (Nup153FG$^{FxFG(I)}$ and Nup153FG$^{FxFG(II)}$) and for the GLFG-rich region of the yeast Nup49 upon binding to different NTRs (Figure 4.2). smFRET binding experiments of the different FG-Nup FRET mutants was tested for Importinβ, the major exportin CRM1, the GDP import receptor NTF2 and the NTR transportin-1 (TRN1) that is involved in the import of cargoes containing the M9 NLS and the export of RNPs (Figure 4.2). In all the possible tested cases and despite the different NTR structures the smFRET indicated a binding mechanism where FG-Nups and NTRs form dynamic complexes and the FG-Nup is able to bind the NTR in an ensemble of multiple conformations without requiring any change of the distance distribution from the unbound form. FCS measurements of each binding combination was performed at ensemble conditions (10 nM labeled FG-Nup) in order to test the binding to the different NTRs (Figure 4.2).

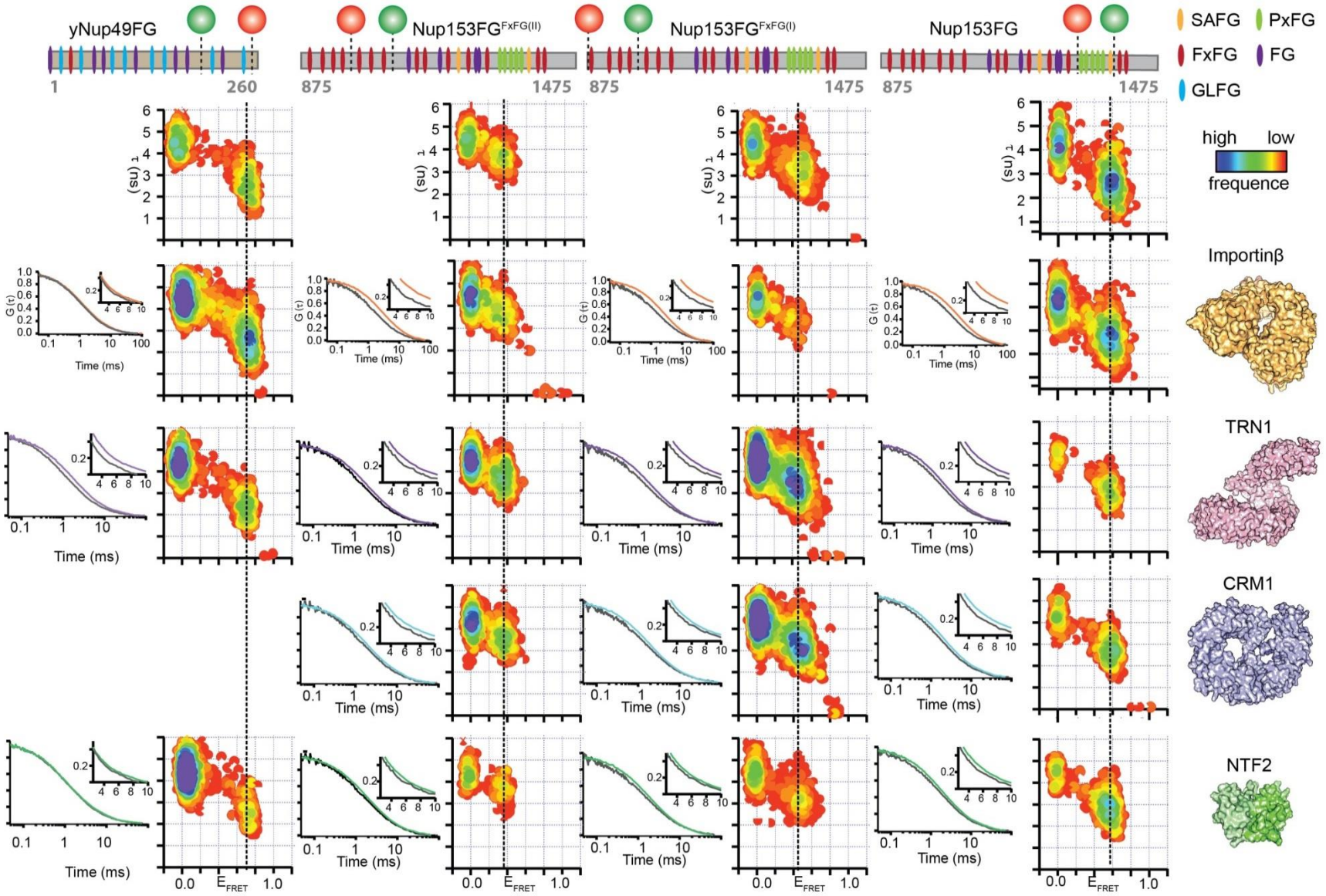
yNup49FG
1
260
Nup153FG^FxFG(II)
875
1475
Nup153FG^FxFG(I)
875
1475
Nup153FG
875
1475
SAFG
PxFG
FxFG
FG
GLFG
high
low
frequence
Importinβ
TRN1
CRM1
NTF2
τ (ns)
G (τ)
Time (ms)
E_FRET

Figure 4.2. smFRET structural study of the binding between FG-Nups and NTRs: τ vs E_{FRET} histograms of double labeled (Alexa488-hydroxilamine and Alexa594-maleimide at the AcF and single cysteine residues respectively) Nup153FG (1312AcF, 1391C), Nup153FG$^{FxFG(I)}$ (994AcF, 883C), Nup153FG$^{FxFG(II)}$ (1049AcF, 990C) and yNup49FG (191AcF 250C) were measured in the absence and in the presence of Importinβ (orange), TRN1 (purple) CRM1 (blue) and NTF2 (green). The histograms are color coded according to the frequency of occurrence. The FCS traces retrieved measured at ensemble concentrations were used to confirm that binding between the FG-Nups and NTRs had taken place, the trace obtained from the unbound protein is shown in gray and the colored trace corresponds to the one obtained in the presence of NTR.

The maintenance of the unbound conformational heterogeneity in the bound form is an unusual and novel phenomenon for the binding between disordered and folded proteins. Its understanding benefitted from all-atom molecular dynamics (MD) simulation experiments (Dr. Mercadante, Gräter laboratory at the HITS institute), which provided an atomic resolution of the binding process.MD simulation experiments of the 82 aa long PxFG-rich region of Nup153FG (Nup153FG^{PxFG}) and a fragment of Importinβ (from the residues 1-462 termed Importinβ^{N} (Bayliss et al., 2000b)) revealed that Nup153FG^{PxFG} was able to bind Importinβ in multiple conformations as indicated by smFRET showing similar end-to-end distance distributions as in the unbound form. This may be possible because the phenylalanine residues of the FG-motifs, responsible of engaging in the FG-Nup/NTR interaction, were already exposed and able to bind the NTR. Moreover, MD simulations showed that the association between Nup153FG^{PxFG} and Importinβ^{N} was occurring in the nanosecond timescale, suggesting a very fast protein-protein association

4.1.2. Kinetic study of the FG-Nup/NTR interaction

Motivated by the MD results that showed very fast binding between the Nup153FG^{PxFG} fragment and Importinβ^{N}, I was prompted to perform a kinetic study of the FG-Nup/NTR interaction. After being trained in the laboratory of Jane Clarke (University of Cambridge), who is the leading expert on the kinetic characterization of IDP folding and binding, I established the pipeline and adjusted the stopped-flow spectrophotometer for the kinetic measurement of fast binding interactions in our lab (detailed description in the materials and methods section 3.1.6) and used stopped-flow spectroscopy to monitor the binding between FG-Nups and NTRs.

In the stopped-flow association experiments, single site-specifically labeled Nup153FG with Cy3B was loaded on one syringe and rapidly mixed with an NTR loaded on a different syringe. The Cy3B dye has a short linker that reports detectable changes in the rotation of the molecule, thus upon binding to the NTR, the increase in anisotropy was monitored (Milles and Lemke, 2014) (Figure 4.3). Measurements were carried out under pseudo-first order conditions (typically considered at a concentration of unlabeled binding partner 10 fold in excess compared to the labeled) at different concentrations of NTR (see methods section 3.1.6). Each trace obtained from monitoring the increase of anisotropy over time was fitted to obtain the observed association rate (k_{obs}) (Figure 4.3).

A single-exponential function did not describe the observed anisotropy changes therefore I used a double exponential function (Figure 4.3C). This phenomenon may be caused by the ability of FG-Nups to engage with multiple FG-motifs to the NTRs and by the ability of NTRs to engage with multiple FG-motifs resulting in a multivalent interaction. To calculate the association rate constant (k_{on}) I used the k_{obs} obtained from the first component of the fitting, with amplitude of 70% of the recorded signal (Figure 4.3G). In addition, I used the k_{obs} from the first component because we are interested in monitoring the k_{on} for the first binding events that occur in the FG-Nup/NTR complex formation. The k_{on} values were retrieved from the slope of the linear fit from the k_{obs} vs NTR concentration plots. The k_{obs} from the first component ($k_{obs,ultrafast}$) and from the second component ($k_{obs,fast}$) were used to calculate the corresponding $k_{on,ultrafast}$ and $k_{on,fast}$ respectively (Figure 4.3 E-F).

The interaction between Nup153FG labeled at position 1391 and Importinβ had a $k_{on,ultrafast}$ of $1.5x10^{9}$ $M^{-1}s^{-1}$, which is in the diffusion limit range. Thus, I calculated the theoretical diffusion limit for Nup153FG and Importinβ. FCS measurements of Atto655 labeled Nup153FG and Importinβ allowed the calculation of the diffusion coefficient (D) of each protein using the known D of the dye as a reference. The D for Nup153FG and Importinβ was of $3.83x10^{-5}$ and $2.34x10^{-5}$ cm^2s^1. The measured D was then used to estimate the theoretical diffusion limit using the Einstein-Smoluchowski approximation (Materials and methods section 3.1.6). The estimated diffusion-limited association for Nup153FG and Importinβ was of $6.8x10^{9}$ $M^{-1}s^{-1}$.

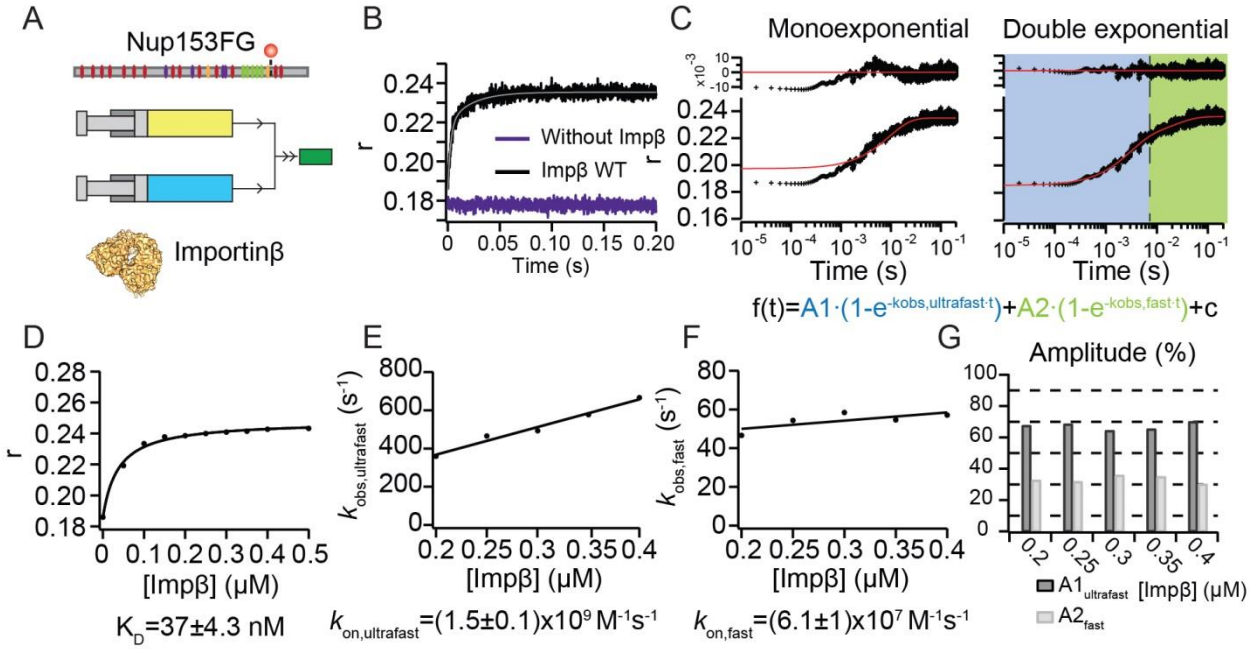

Figure 4.3. Stopped-flow association experiment of Nup153FG binding to Importinβ: A) Cartoon representation of a stopped-flow with two syringes, one contained single-labeled Nup153FG and the other one containing a NTR, in this case Importinβ. The content of the syringes is injected simultaneously and rapidly mixed.

Once the observation head is filled the binding reaction is monitored over time. B) Example of an obtained trace in the absence (purple) and presence (black) of Importinβ. C)The obtained traces were fitted to a double exponential function (first exponential component in blue and second in green) to obtain the $k_{obs,ultrafast}$ and $k_{obs,fast}$ corresponding to the one of the first and second exponential component. D) Saturation binding curve performed at the stopped-flow my measuring the final anisotropy values of a titration experiment. E-F) k_{obs} vs NTR concentration plot of the ultrafast and the fast component respectively. The slope from the linear fit corresponds with the k_{on} values. G) The amplitudes corresponding with the first (~70%) and second coefficient (~30%) of the double exponential function are plotted in black and gray respectively.

Fast k_{on} values are frequently associated to long-range electrostatic attraction between molecules. A classic example is provided by the interaction between the bacterial ribonuclease/inhibitor system of barnase and barstar (Schreiber and Fersht, 1996). Schreiber and Fersht reported a k_{on} for barnase and barstar of $5x10^9$ $M^{-1}s^{-1}$ and they showed that at infinite ionic strength i.e., condition at which the electrostatic forces are shielded, the k_{on} was reduced four orders of magnitude (Schreiber and Fersht, 1996). For IDPs most of the reported k_{on} values at high ionic strength conditions are in the range between 10^4-10^6 $M^{-1}s^{-1}$ (Shammas et al., 2013).

I then performed a series of stopped-flow association experiments and determined the k_{on} of the binding between Nup153FG and Importinβ under different ionic strength conditions by performing a NaCl titration ranging from 0.05 to 1M ionic strength (detailed description in the methods section 3.1.6). The k_{on} values obtained from the association experiments at the different ionic strength conditions were then plotted against the ionic strength. The $k_{on,basal}$ which is the k_{on} in the absence of electrostatic forces, can then be estimated by fitting the obtained data to a Debye-Hückel-like approximation (Shammas et al., 2014). The estimated $k_{on,basal}$ for Nup153FG and Importinβ is of $2.9x10^8$ $M^{-1}s^{-1}$ (Figure 4.4).

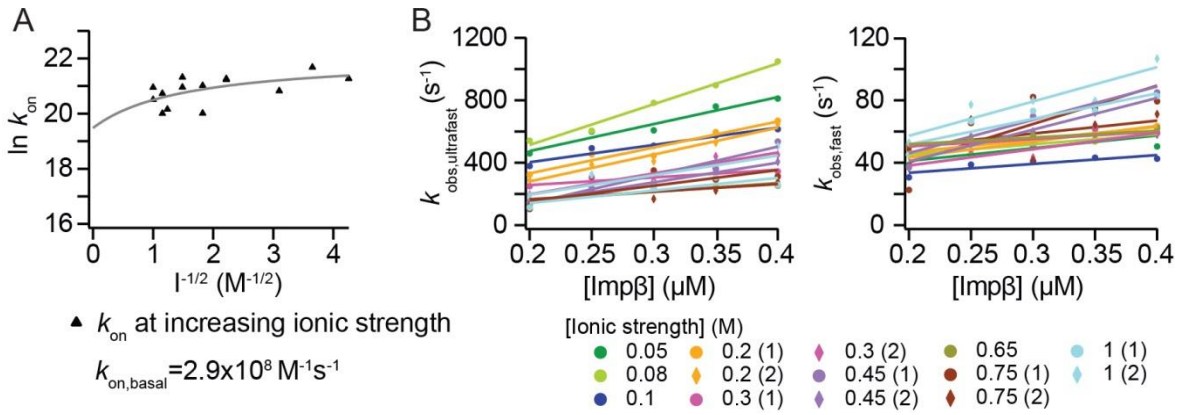

Figure 4.4. Determination of $k_{on,basal}$ for Nup153FG and Importinβ at infinite ionic strength: A) ln k_{on} vs square root of the ionic strength, each triangle corresponds to the k_{on} obtained from a stopped-flow association experiment at different ionic strength conditions. The data was fitted using a Debye-Hückel-like approximation to calculate the $k_{on,basal}$ which corresponds with the k_{on} at infinite ionic strength where the electrostatic interactions would not have an effect on the k_{on} values. B) k_{obs} vs Importinβ concentration plots for the first fitting coefficient ($k_{obs.ultrafast}$) (left plot) and for the second ($k_{obs,fast}$) (right plot). Each color corresponds to measurement performed at a specific ionic strength.

In order to test if the fast binding kinetics between FG-Nups and NTRs is a specific feature of the binding between the PxFG-rich region of Nup153FG and Importinβ or if it is a general characteristic of FG-Nup/NTR binding, I measured the association kinetics of different FG-Nups/NTR complexes. In addition to the PxFG-rich region, I probed a FxFG-motif rich region of Nup153FG and different NTRs like, NTF2 or TRN1. The anisotropy amplitude change of the unbound/bound labeled FG-Nups varies depending on the labeling site and the NTR used, affecting the accuracy of the obtained k_{obs} and increasing the error of the fitted k_{on}. Nevertheless, for all the measurable cases the k_{on} values were between $1.9x10^{9}$ and $7x10^{8}$ $M^{-1}s^{-1}$ (Figure 4.6).

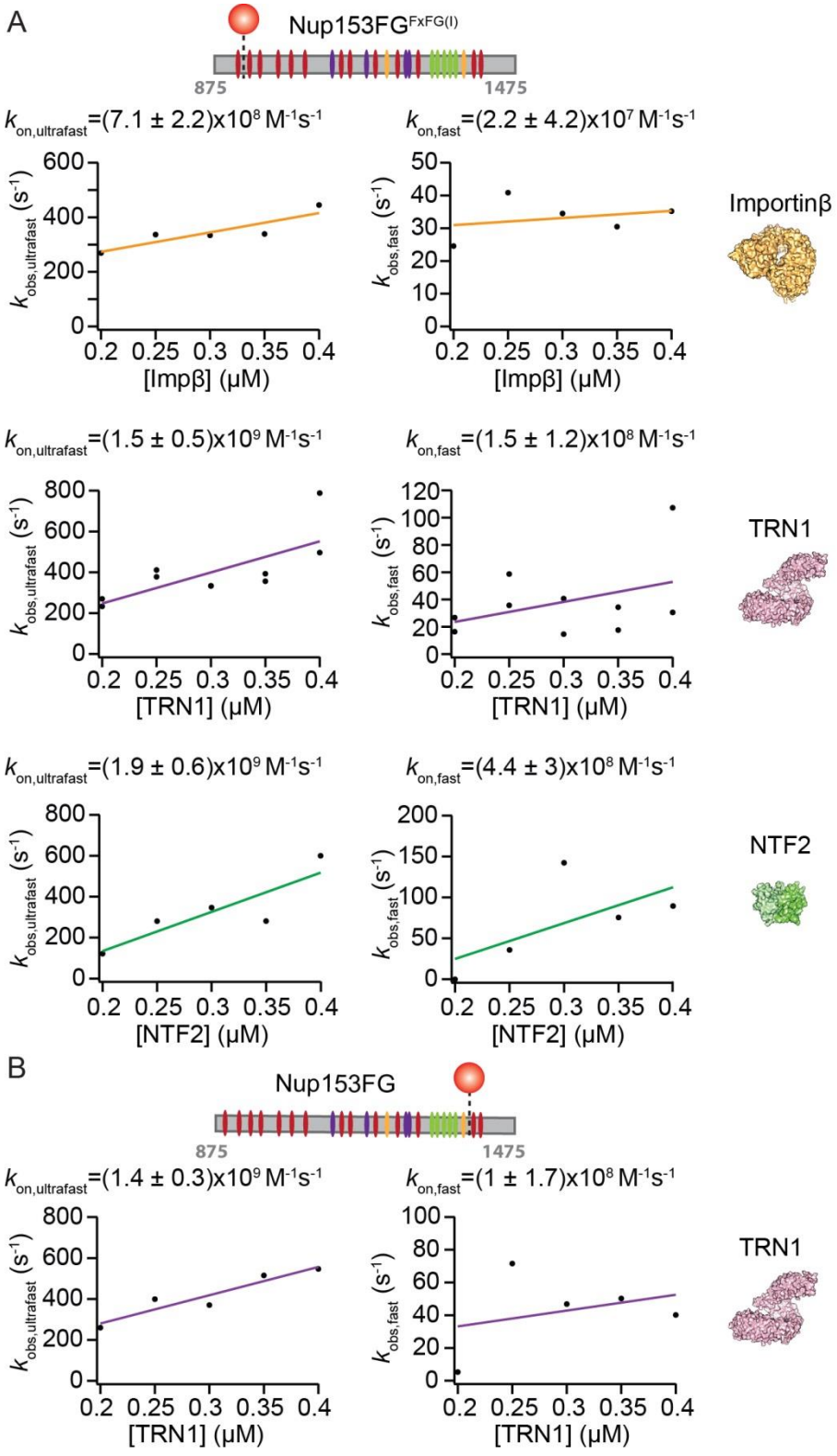

Figure 4.5. Stopped-flow association experiments of different NTRs binding to Nup153FG: The observed rates of the ultrafast and fast component of the double-exponential fit are plotted separately in the k_{obs} vs NTR concentration plots (left and right respectively). Stopped-flow association experiment of Importinβ (orange), TRN1

(purple) and NTF2 (green) to Nup153FG containing the dye at the FxFG region in the position 883. B) Stopped-flow association experiment of TRN1 (purple) to Nup153FG containing the dye at the PxFG region in the position 1391. The obtained k_{on} from the slope of the linear fit are plotted above each graph.

4.1.3. Kinetic dependence of the number of binding sites

Next, I wanted to test how would the kinetics of Nup153FG binding to Importinβ be affected if one of the binding partners contains a less reactive surface that could potentially have an impact on the k_{on}. For this purpose I used Importinβ I178D/Y255A double mutant (Importinβ^{DA}) (Bednenko et al., 2003) (Figure 3.6). Importinβ^{DA} contains the mutations in the HEAT repeat 5 and 7, which correspond to the main FG-Nup binding pockets. Bednenko et al. reported using a solid phase binding assay a decrease in the affinity upon Nup153FG binding of Importinβ^{DA} of more than 60 fold. Therefore, I first tested if the decrease in the affinity was due to a decrease in the k_{on} as a consequence of a less reactive surface in the Importin. The calculated $k_{on,ultrafast}$ was obtained from the $k_{obs,ultrafast}$ coefficient obtained from the first component, which corresponds to the 90% of the measured trace (Figure 4.6C). This is in agreement with the hypothesis of the double exponential fitting being required due to the multivalent nature of the interaction. The $k_{on,ultrafast}$ for Nup153FG and Importinβ^{DA} was of $1x10^{9}$ $M^{-1}s^{-1}$ which is still extremely fast (Figure 4.6A).

Nevertheless, in a multivalent system the number of binding sites, in the contrary to the k_{on}, has been shown to have a big effect in the dissociation rate constant (k_{off}) (Kramer and Karpen, 1998). Thus, I performed stopped-flow competition experiments using excess of unlabeled Nup153FG to determine the k_{off} value of the pre-formed complexes of Nup153FG/Importinβ and Nup153FG/Importinβ^{DA}. In conventional bimolecular protein-protein interactions, the k_{off} is a kinetic parameter that should not be concentration dependent since the dissociation rate will always be the same and the increase of competitor will only generate an increase in the amplitude of the measured detected dissociation (Pollard and De La Cruz, 2013). However, dynamic complexes with fast kinetics can bind and unbind during the dissociation experiment, therefore they often require higher concentrations of competitor and in addition it is convenient to titrate different concentrations in excess to see the dependence of the k_{off} to the concentration of competitor (Shammas et al., 2014) (Methods section 3.1.6). The measured k_{off} for Nup153FG bound to Importinβ^{DA} was over 30 times faster than the one measured for Importinβ WT (Figure 4.6D). Showing that as it was previously reported for other systems, the number of binding sites has a strong effect on the koff of a complex (Kramer and Karpen, 1998).

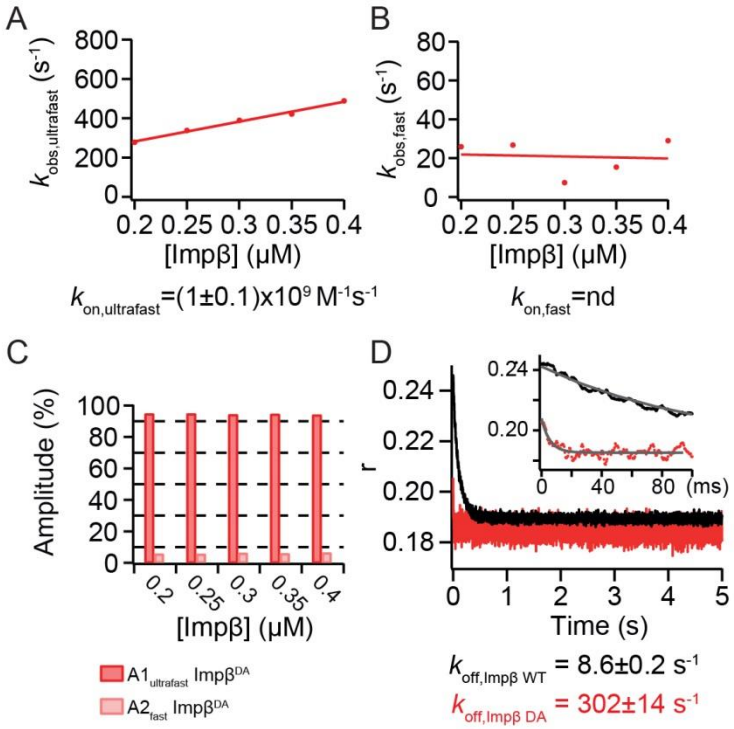

Figure 4.6. Stopped-flow kinetic experiment of Nup153FG/Importinβ^{DA}: Analog association experiments to the ones shown in figures 3.3 and 3.5 measuring the k_{on} between Nup153FG 1391C and Importinβ^{DA} double mutant (I178D/Y255A) affecting the binding sites between FG-motifs and HEAT repeats 5-6 and 6-7. A) $k_{obs,ultrafast}$ vs Importinβ^{DA} concentration plot. B) $k_{obs,fast}$ vs Importinβ^{DA} concentration plot. C) Amplitudes corresponding to the first fitting component ($k_{obs,ultrafast}$) and the second fitting component ($k_{obs,fast}$).D) Stopped-flow dissociation experiment of Nup153FG/Importinβ (black) and Nup153FG/Importinβ^{DA} (red). The pre-formed complex was rapidly mixed with excess of unlabeled Nup153FG protein and the dissociation of the labeled protein was monitored by measuring the anisotropy change. Importinβ^{DA} containing two binding sites less than Importinβ WT shows a faster k_{off}.

During the course of my PhD a crystal structure of CRM1/RanGTP bound to a 117 amino acid long disordered fragment of Nup214FG was published (Port et al., 2015). This crystal structure challenged the ultrafast fuzzy binding mechanism ,described in this thesis, where dynamic complexes are formed with very fast association kinetics and any conformation adopted by the FG-Nup in the disordered ensemble seems to be able to bind the NTR without any change in the ensemble of conformations.

Port et al. showed that the Nup214FG fragment bound in a particular conformation to the outer surface of CRM1 where it seemed to act as a molecular clamp, keeping CRM1 in a closed conformation by linking the N- and the C-terminal sites (Port et al., 2015). Nevertheless, it remains up to investigation if the conformation of Nup214FG bound to CRM1/RanGTP captured by X-ray crystallography is the predominant conformation when the complex is studied in solution or if it is one out of multiple conformations that can be adopted by Nup214FG when bound to CRM1/RanGTP in a dynamic manner.

Thus, I engaged in a project together with Dr. Piau Siong Tan, a postdoc from the lab, in which we studied the structure and kinetics of the interaction between Nup214FG and NTRs.

4.1.4. Coupled folding-binding between FG-Nups and NTRs

Multiparameter smFRET of double labeled Nup214FG revealed that upon binding to CRM1 the disordered ensemble of conformations shifted towards lower E_{FRET} values, indicating an increase of the distance between the donor and the acceptor dyes (Figure 4.7). The binding of CRM1 to Nup214FG showed two populations with different E_{FRET} values at stoichiometry 0.5, this populations correspond to the unbound (E_{FRET}=0.6) and bound (E_{FRET}=0.2) and it is caused by the lower affinity of CRM1 towards Nup214FG in the absence of RanGTP. Binding to CRM1/RanGTP showed a bigger decrease in the E_{FRET} value which is in agreement with the extended conformation that Nup214FG adopts in the crystal structure bound to CRM1/RanGTP (Port et al., 2015).

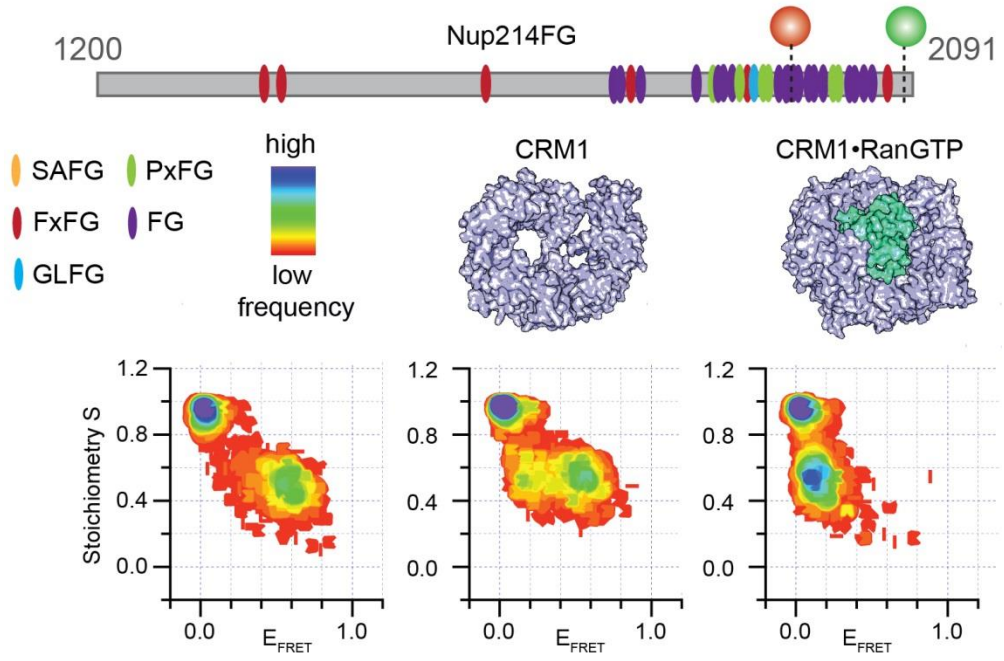

Figure 4.7. Nup214FG undergoes a conformational change when bound to CRM1 and CRM1/RanGTP: Stoichiometry (S) vs FRET efficiency (E_{FRET}) histograms of double labeled Nup214FG at positions 2043AcF and 1905C (with 488-hydroxylamine and 594-maleimide respectively) in the absence and in the presence of CRM1 and CRM1/RanGTP. The histograms are color coded according to the frequency of occurrence. The binding to CRM1 and CRM1/RanGTP causes a change in the ensemble of conformations of Nup214FG to a more extended form.

In order to test whether this is a specific binding mechanism of Nup214FG and CRM1/RanGTP we tested the binding to Importinβ, which has been previously shown to bind Nup214FG (Moroianu et al., 1995). Interestingly, binding to Importinβ also caused a shift in the E_{FRET} value similar to the one detected for Nup214FG/CRM1/RanGTP. The change in the E_{FRET} values was milder when bound to

Importinβ/RanGTP (Figure 4.8). These experiments show that in addition to the ultrafast fuzzy binding mechanism shown for the bindings of NTRs to Nup153FG or yNup49FG, Nup214FG undergoes a conformational change upon binding to CRM1/RanGTP and Importinβ.

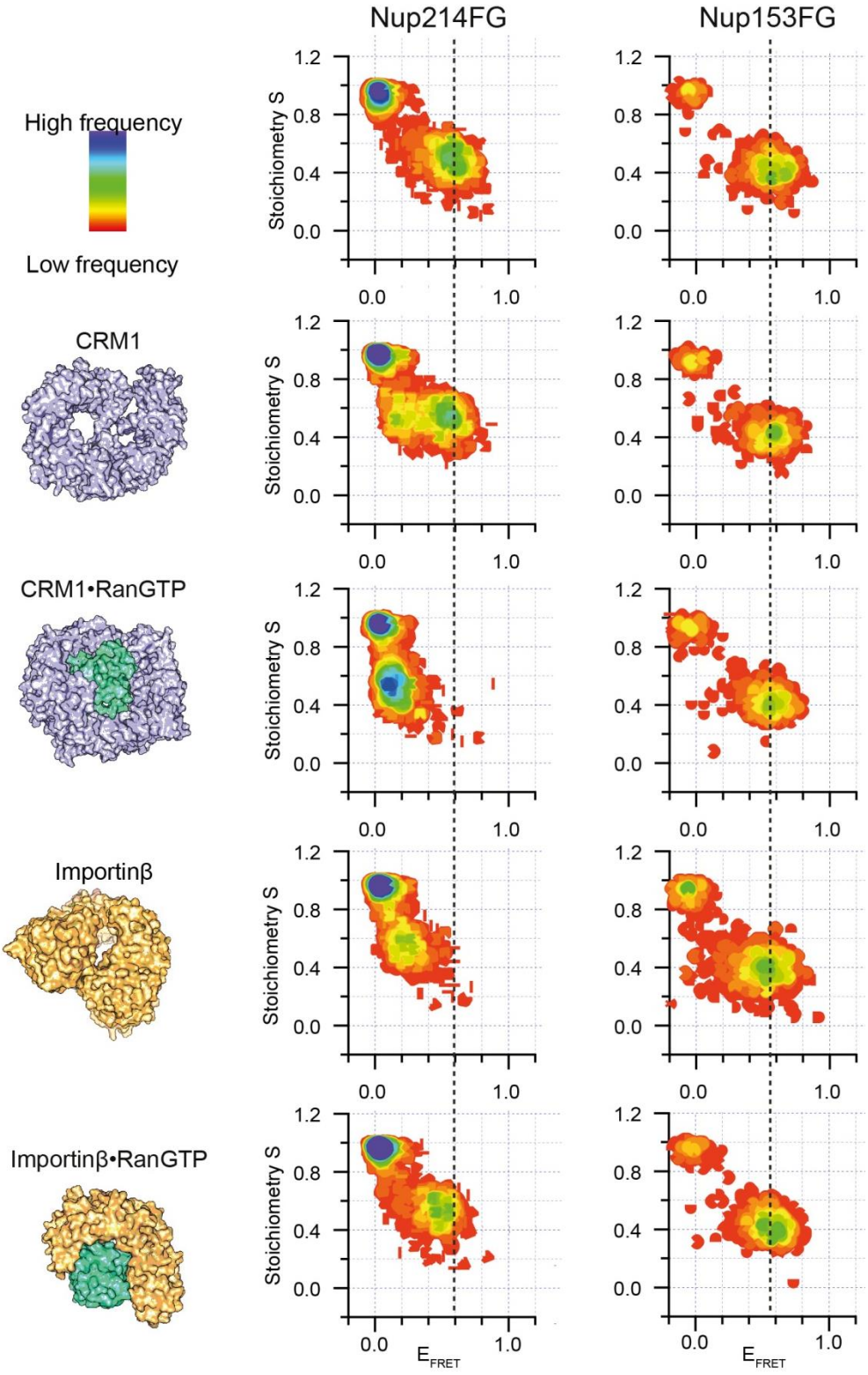

Figure 4.8. smFRET comparison of Nup214FG and Nup153FG binding to NTRs: S vs FRET efficiency (E_{FRET}) histograms of Nup214FG (labeled with 488 and 594 dyes at positions 2043AcF and 1905C respectively) and Nup153FG (488 and 594 labels at residues 1312 and 1391 respectively) in the absence of NTR and in the presence of CRM1, CRM1/RanGTP, Importinβ and Importinβ/RanGTP.

Another key feature of the ultrafast fuzzy binding mechanism between FG-Nups and NTRs is the ultrafast association kinetics. Thus, I tested whether the association kinetics for Nup214FG and CRM1 differ from the ultrafast fuzzy binding mechanism, which is associated with k_{on} values ~$10^9 M^{-1}s^{-1}$. Stopped-flow association experiments of Nup214FG and CRM1 or CRM1/RanGTP were performed under pseudo-first order conditions (Figure 4.9).

In addition to monitoring the complex formation by measuring anisotropy, taking advantage of the change in the E_{FRET} of Nup214FG upon binding to CRM1/RanGTP, I also measured the change in fluorescence of donor and acceptor dye (donor dye increases and acceptor dye decreases due to lower E_{FRET}) (Figure 4.9C). In the case of Nup214FG the measured data was described by a single exponential function which was used to obtain the k_{obs} at different NTR concentrations. The measured k_{on} values were in the order of 10^7, which is two orders of magnitude slower than the one detected for other FG-Nup/NTR interactions but still fast in comparison the association kinetics of other IDPs. It still remains an open question how the increasing concentrations of ionic strength may affect the kinetic values of the Nup214FG/CRM1 or CRM1/RanGTP interaction.

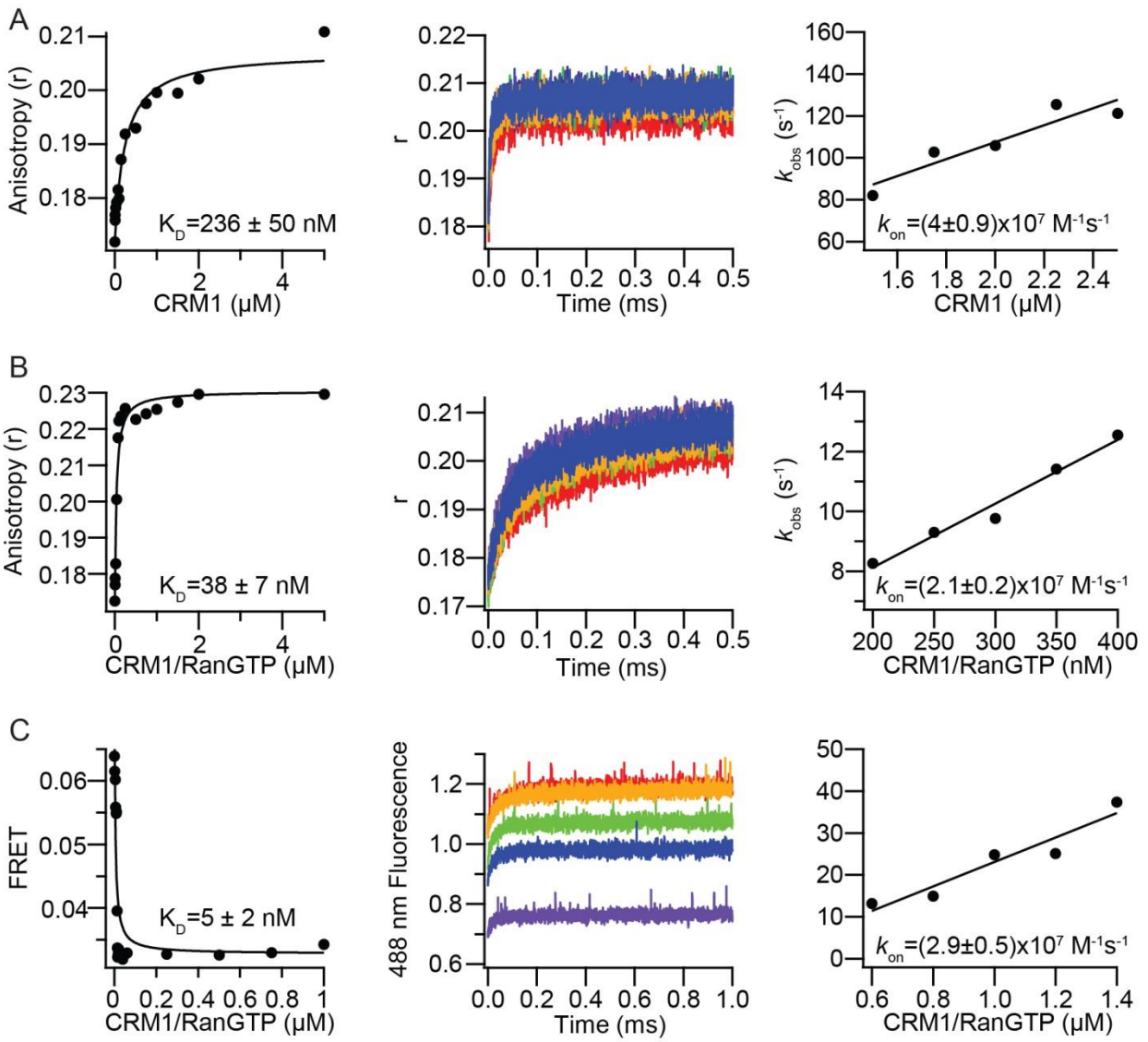

Figure 4.9. Binding curves and kinetic stopped-flow experiments of Nup214FG and CRM1, CRM1/RanGTP: The binding curves and stopped-flow experiments were performed by monitoring the anisotropy change of single labeled Nup214FG at position 1905C with Cy3B-maleimide dye upon binding to CRM1 A) and CRM1/RanGTP B).C) The change in donor fluorescence caused by the decrease of E_{FRET} upon binding to CRM1/RanGTP was used to measure the K_D and the $k_{on.}$

4.1.5. Study of the effect of glycosylation in the FG-Nup/NTR binding mechanism

Given that we have diverse FG-Nup systems in our lab that show two completely different binding mechanisms I wanted to extend these studies by investigating the possible effect that the PTM N-Acetylglycosamine (GlcNAc), abundant PTM present in many FG-Nups of vertebrates at the NPC in the binding to NTRs.

I *in vitro* glycosylated Nup153FG and Nup214FG using recombinantly expressed OGT1, enzyme that covalently attached the GlcNAc to Ser/Thr residues, as it has been previously established for Nup98FG

(Labokha et al., 2013). Glycosylation of the FG-Nups was confirmed by mass spectrometry and western blotting (see methods section 3.1.3). Then glycosylated Nup214FG and Nup153FG (Nup214FG^{Glyc} and Nup153FG^{Glyc}) were used with the same smFRET and stopped-flow experimental pipeline previously shown for unglycosylated FG-Nups. smFRET comparison of unbound unglycosylated and glycosylated FG-Nup revealed that the glycosylated FG-Nups have a more extended conformation than the unglycosylated ones (Figure 4.10). This effect seems to be more prominent in Nup214FG.

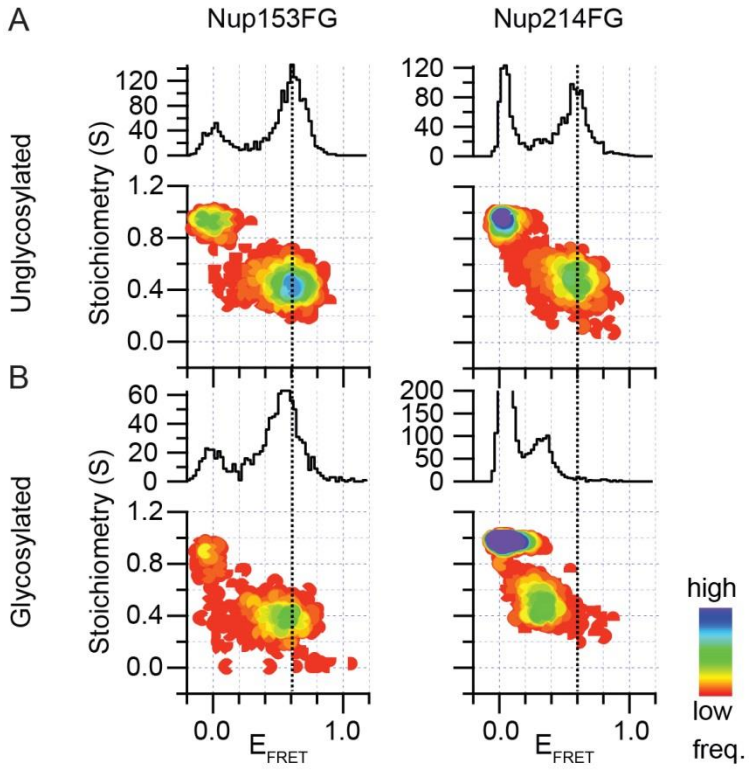

Figure 4.10. smFRET comparison of unglycosylated and glycosylated Nup153FG and Nup214FG: S vs FRET efficiency (E_{FRET}) histograms of unglycosylated A) and glycosylated B) Nup153FG (1312AcF and 1391C) and Nup214FG (1905C and 2043AcF). The histograms are color coded according to the frequency of occurrence.

The smFRET of glycosylated FG-Nups was also tested upon binding to different NTRs. Glycosylated FG-Nups did not show any major difference, compared to unglycosylated FG-Nups, upon binding to NTRs (Figure 4.11).

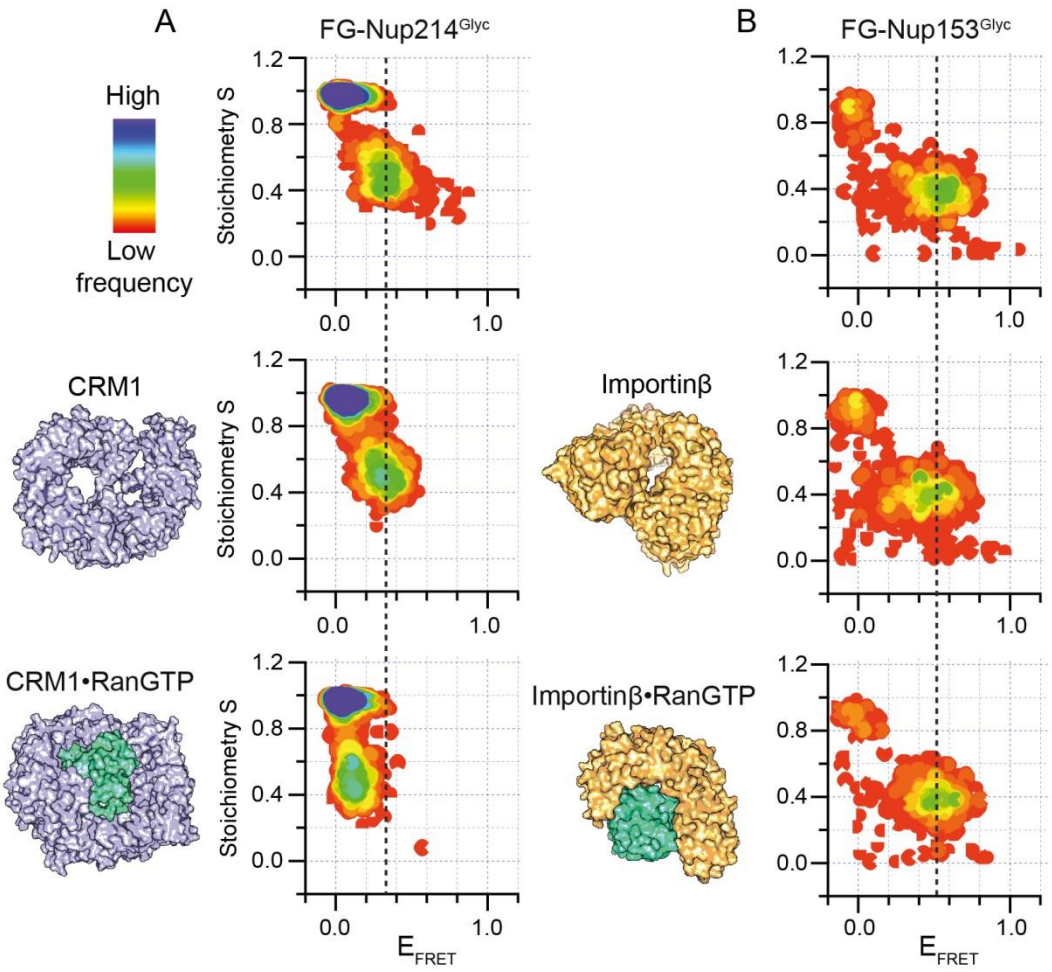

Figure 4.11. smFRET of glycosylated Nup214FG^{Glyc} and Nup153FG^{Glyc}: : S vs FRET efficiency (E_{FRET}) histograms of Nup214FG^{Glyc} A) in the absence and presence of CRM1 and CRM1/RanGTP and of Nup153FG^{Glyc} B) in the presence and absence of Importinβ and Importinβ/RanGTP. The E_{FRET} of the unbound Nup214FG^{Glyc} decreases upon binding to CRM1/RanGTP and for Nup153FG^{Glyc} no change in the E_{FRET} is detected upon NTR binding.

Next, I tested the binding kinetics of the glycosylated FG-Nups. In this case the obtained traces were fitted to a single-exponential function. The obtained k_{on} values resemble the ones obtained for the unglycosylated FG-Nups (Figure 4.12) The K_D values obtained from the unglycosylated and the glycosylated titrations are also comparable (Figure 4.9 B-C and 4..

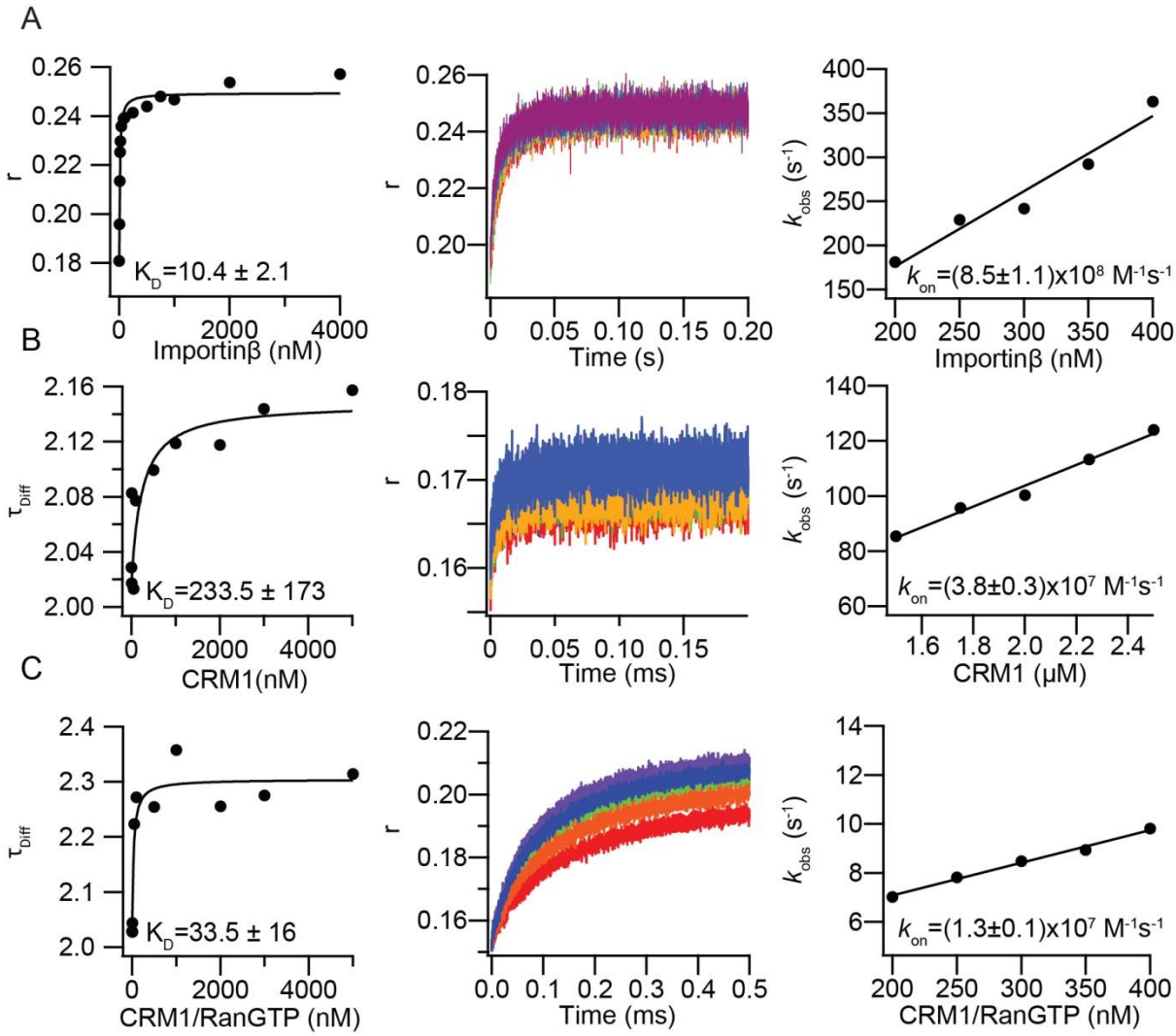

Figure 4.12. Binding curves and kinetic stopped-flow experiments of Nup153FG^{Glyc} and Nup214FG^{Glyc} binding to NTRs: *A*) Nup153FG^{Glyc} binding to Importinβ. B) Nup214FG^{Glyc} binding to CRM1. C) Nup214FG^{Glyc} binding to CRM1/RanGTP. Binding curves indicating the apparent K_D between glycosylated FG-Nups and NTRs (left). In B and C the binding curves were obtained by FCS whereas in A it was obtained by anisotropy. Traces obtained from the stopped-flow experiments (middle). Fitting of the traces gives a k_{obs} at each NTR concentration measured. The slope of the fit k_{obs} vs NTR concentration plot corresponds with the k_{on} (right).

4.2. Development of a synthetic biology approach to optimize the study of FG-Nups *in situ*

The information obtained from the different *in vitro* techniques is extremely useful to understand the fundamental interaction mechanism between FG-Nups and NTRs. However, very little is known about how FG-Nups operate during the nucleocytoplasmic transport or how they are located at the NPC.

Dr. Kan, a previous PhD from the lab, combined Amber suppression technology with click-chemistry in order to achieve *in situ* site-specific labeling of the disordered regions of the FG-Nups. Amber suppression technology has the limitation of the formation of truncated products when the ncAA is not being incorporated and the TAG codon is recognized as a stop signal. When studying the disordered regions of the FG-Nups, like Nup153, if the nuclear targeting and binding domain of the FG-Nup is at the N- and the FG-region of interest at the C-terminal site, the truncated FG-Nup will be able to incorporate at the NPC but it will not contain a ncAA this it will compete with the FL FG-Nups compromising the power of super-resolution and single molecule studies.

Motivated by the study of FG-Nups *in situ*, in order to overcome the accumulation of truncated proteins as a byproduct of an inefficient Amber suppression system, I have developed an intracellular system using a synthetic biology approach to selectively degrade the truncated proteins, while enriching the full length ones.

This selective degradation system must fulfill two requirements; on the one hand it must degrade the truncated products that did not incorporate the ncAA and at the same time degradation of the full length protein of interest (POI) must be avoided.

In order to achieve this, I have used a reporter iRFP-GFP$^{39TAG \rightarrow ncAA}$ that when the ncAA is not incorporated generates iRFP and when it is, it leads to the formation of iRFP-GFP (Figure 4.13). The fusion of a degradation sequence at the N-terminal site of the POI will cause the degradation of the truncated and full length POI (FL POI). However, if the POI contains a polypeptide sequence at the C-terminal site that only gets expressed when the ncAA has been incorporated and this sequence is able to avoid the degradation of the FL POI, FL POI will not be degraded.

In addition, it is also required that the N-terminal degradation domain and the C-terminal domain fused to the POI are removed, thus they are connected to the POI through a linker containing a protease cleavage site (dashed line, or -X-). However, it is necessary that the protease only cleaves off the sequences from the LF POI in order to maintain the degradation of the truncated one.

Aiming to achieve this, I have applied the recent knowledge generated from studies on phase-separated membrane-less organelles (Hyman et al., 2014; Patel et al., 2015) and combined them with targeted

protein degradation in the following way (Figure 4.13): I have fused to the C-terminal site of the POI a disordered domain (the disordered protein FUS), which forms liquid-like droplets in the cell (purple rectangle) (Figure 4.13). When the ncAA is incorporated, the POI will contain the degradation domain at the N-terminal site and a droplet forming domain at the C-terminal site, in both cases connected to the POI by a protease cleavage site.

In addition, to achieve that the protease only cleaves POI containing the ncAA, I have fused the protease also with a FUS droplet forming domain. Thus, the cell will form mixed droplets containing the protease and the FL POI. Ideally, the protease will cleave the cleavage sequences that link the POI to the N-terminal degradation signal and the C-terminal liquid forming domain releasing the POI containing the ncAA into the cytoplasm.

On the other side, when the ncAA is not incorporated, the POI will not contain the droplet forming domain therefore it will not be able to mix with the protease containing droplets and the degradation tag will not be cleaved by the protease, leading to the degradation of the truncated product.

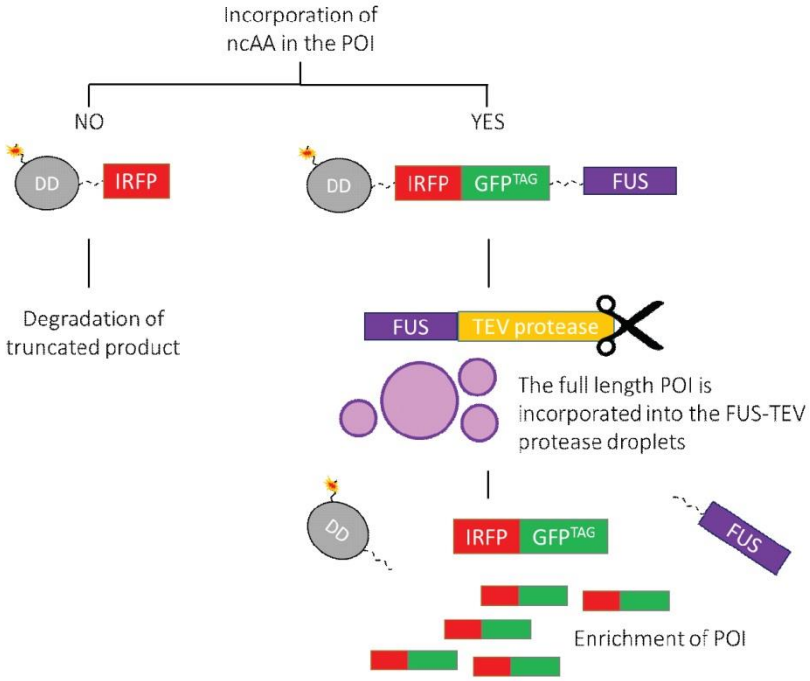

Figure 4.13. Scheme representation of the approach used to selectively enrich the POI containing ncAA and degrade the truncated products: Truncated proteins (iRFP, red box) fused at the N-terminal site by a degradation domain (DD gray sphere) linked by a TEV cleavage site (dashed line) forming DD-iRFP (left). Full length protein also contain the DD fused by the TEV cleavage site is expressed fused to GFP^{39TAG} (greed box) and a FUS droplet forming domain (purple box) also connected to the GFP by a TEV cleavage site (right). The TEV protease (yellow box) is also fused to a FUS droplet forming domain. Selective cleave by TEV protease of FL POI should majorly occur at the mixed FUS-TEV/DD-iRFP-GFP^{39TAG}-FUS droplets, leading to the enrichment of iRFP-GFP^{39TAG} and the degradation of iRFP.

4.2.1. Selection of a degradation system

In cell biology there are different degradation domains or sequences (DD) that are used for targeting the POI to degradation, reducing in this way the protein half-life in the cell. As an initial approach I tested five different DD that have been reported to constitutively increase the degradation rate of their fused proteins. The DD used include: the D-box sequence from cyclin B1 (Yang et al., 2013), a domain derived from the human estrogen receptor (Miyazaki et al., 2012), a mutant of the rapamycin-binding protein (FKBP) (Banaszynski et al., 2006), a mutant of dihydrofolate reductase (DHFR) from *E.coli* (Iwamoto et al., 2010) and the bilirubin-inducible fluorescent protein UnaG (Navarro et al., 2016). The DDs were fused to the N-terminus of the POI (in this case an iRFP-GFP^{39TAG} reporter). In the presence of ncAA, the DD-X-iRFP-GFP^{39TAG}-X-FUS (X represents TEV cleavage site) will be expressed. Truncated proteins will contain the DD-X-iRFP.

I tested the reporter levels in the presence and in the absence of TEV protease for the different DD constructs. The relative comparison of each reporter with and without the TEV construct can indicate the different degrees of degradation efficiencies of each DD (Figure 4.14). Flow cytometry measurements of resuspended HEK cells was used to monitoring the iRFP and GFP signal of single cells.(represented as Y and X axis respectively in the flow cytometry plots). The iRFP signal obtained from the different truncated reporters (vertical y axis corresponds with iRFP only positive cells) in the presence (Figure 4.14A) and absence of TEV protease (Figure 4.14B) was used to calculate the percentage of iRFP signal for each DD (Figure 4.14D). Considering that 100% corresponds to the signal when the protease is present and cleaves off the degradation tag , as observed by western blot (data not shown). Truncated proteins that did not incorporate the ncAA will give as a read-out iRFP signal (Figure 4.14C, bottom). The quantification analysis revealed that DHFR seems to have a stronger degradation effect decreasing the amount of truncated reporter, as it is indicated by the decrease of the iRFP only population, while full length reporter iRFP-GFP signal was still detected (Figure 4.14).

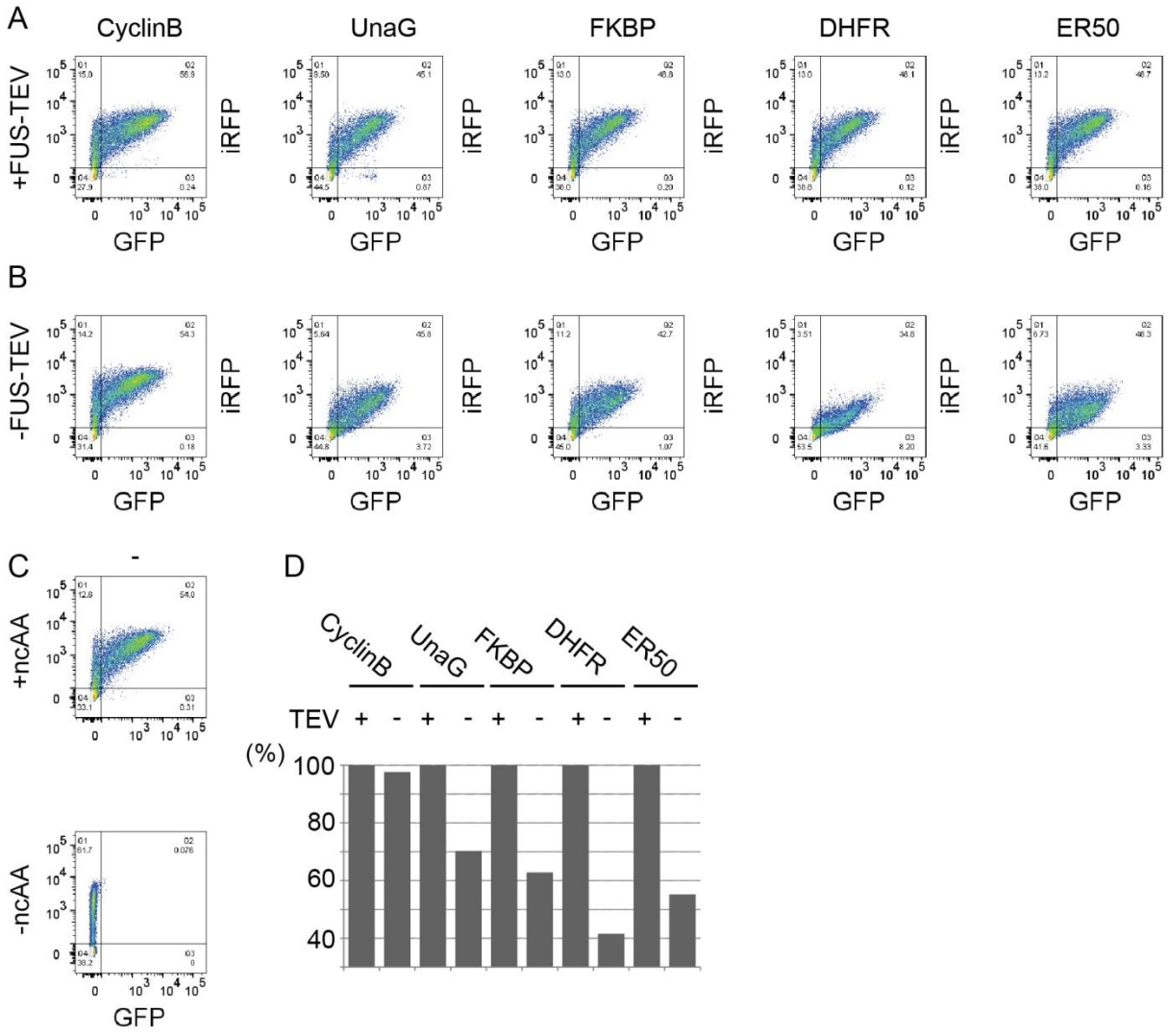

Figure 4.14. Comparison of the different DD-reporter constructs: The iRFP-GFP^{39TAG}-X-FUS reporter was fused at the N-terminal site to different degradation domains (DD) of CyclinB, UnaG, FKBP, DHFR and ER50 (left to right). The constructs were co-transfected to HEK293T cells with constructs containing the tRNA^{pyl}/RS^{pyl,AF} with FUS-TEV (A) or without (B). In both cases the ncAA BOC-Lys was used and cellular fluorescence was analyzed by flow cytometry. C) Control measurement of iRFP-GFP^{39TAG} without a DD in the presence (top) and absence (bottom) of ncAA. The absence of ncAA generates iRFP signal only. D) % iRFP intensity signal of the iRFP signal only population obtained in the absence of FUS-TEV compared to the one obtained in the presence of FUS-TEV.

The full length reporter contains the sequence of the disordered protein, FUS which is able to phase separate. I used fluorescence confocal microscopy to examine the localization of the POI and the morphology of the intracellular structures formed by different DD-reporters (Figure 4.15). All the tested constructs showed cytoplasmic droplet formation of variable number and size.

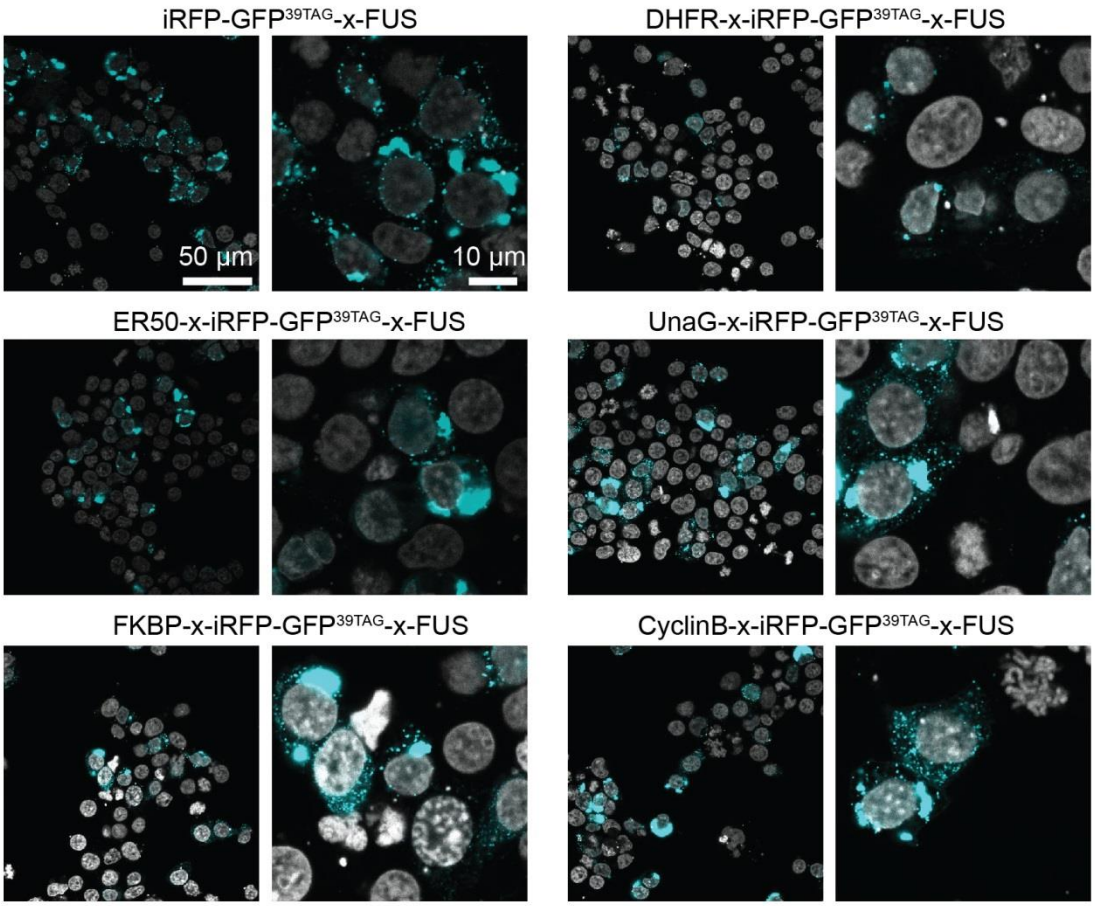

Figure 4.15. Confocal microscopy images of the DD-iRFP-GFP^{39TAG} constructs: Confocal images of HEK293T cells transiently transfected with tRNA^{pyl}/RS^{pyl,AF} and different DD-X-iRFP-GFP^{39TAG}-X-FUS in the presence of the ncAA BOC-Lys. The cells were imaged 24h after transfection (cyan: GFP signal). Scale bar 50 µm and 10 µm.

Based on the flow cytometry results I decided to use DHFR as a DD of my reporter construct. Then I tested how long did it take to the reporter to be degraded from the cytoplasm after stopping translation by performing a cycloheximide treatment (Figure 4.16). Reporter (iRFP-GFP double positive) fluorescence signal from the double positive channel (iRFP-GFP) decrease over time. Moreover, the levels of iRFP only channel showed a very dim signal.

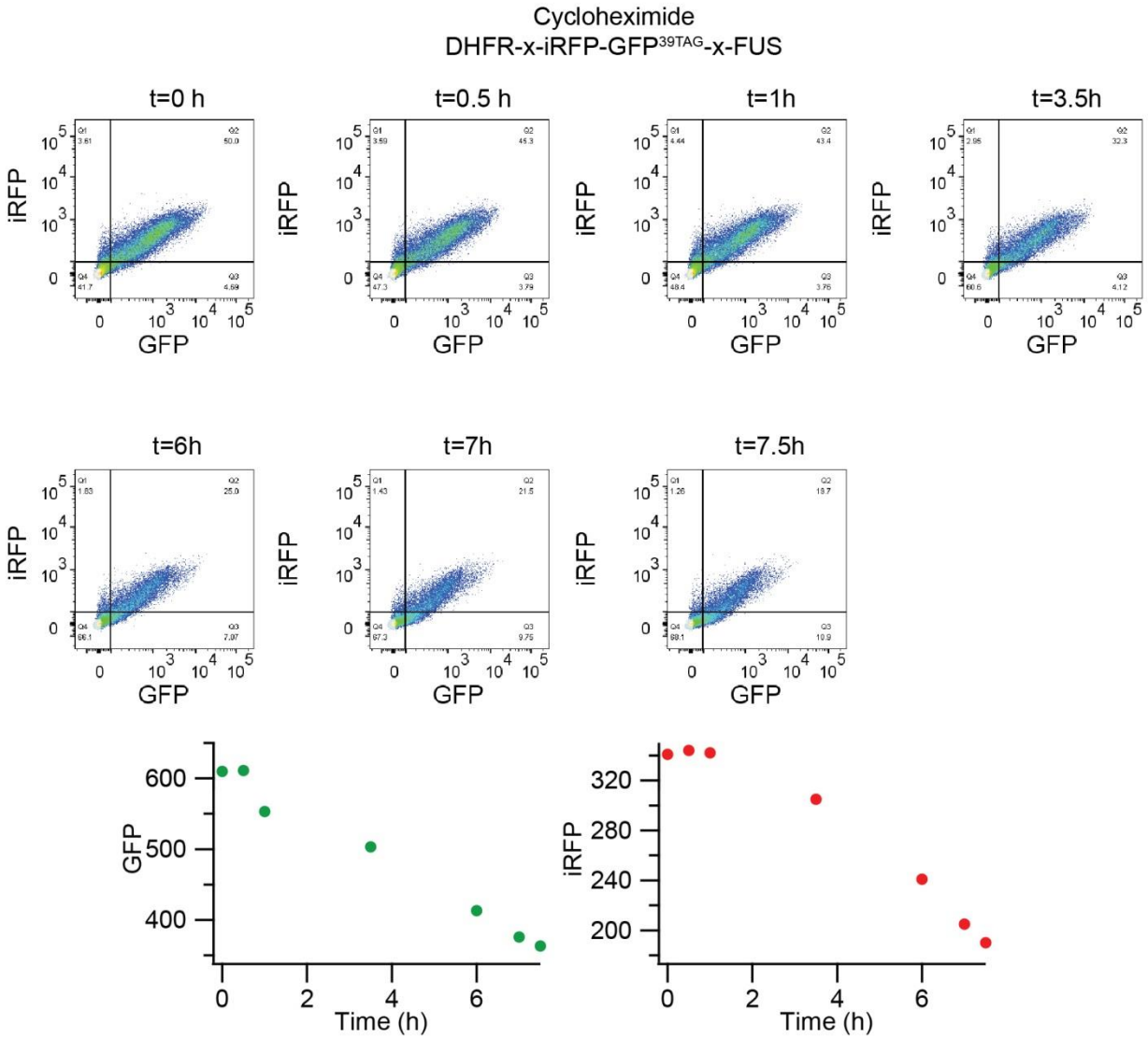

Figure 4.16. Degradation of DHFR-X-iRFP-GFP^{39TAG}-X-FUS: A) Flow cytometry results monitoring the iRFP and GFP signal of HEK293T cells transfected with a plasmid containing the reporter construct DHFR-X-iRFP-GFP^{39TAG}-X-FUS and a plasmid containing tRNA^{pyl}/RS^{pyl,AF} in the presence of the ncAA BOC-Lys. Cycloheximide was applied at different time points to stop protein translation. The intensity signal obtained from GFP and iRFP are plotted individually B) and C) respectively.

4.2.2. Testing the dynamics of the FUS droplet

Cytoplasmic liquid-like synthetic membrane-less organelles are characterized for being highly dynamic structures that are in constant exchange of molecules between the cytoplasm and the droplet. This dynamism is also required in the droplets formed by FUS-TEV and DD-X-iRFP-GFP^{39TAG}-X-FUS to achieve the constant influx of DD-X-iRFP-GFP^{39TAG}-X-FUS reporter that will undergo protease cleavage. In order to test the dynamics of cytoplasmic FUS droplets, I performed fluorescence recovery after photobleaching experiments (FRAP) using a confocal microscope Zeiss LSM 780. Multiple cytoplasmic

droplets of different sizes and shapes were tested. Nevertheless, bleached regions of iRFP-GFP^{39TAG}-X-FUS localized droplets did not achieve the fluorescence recovery expected for dynamic systems (Figure 4.17).

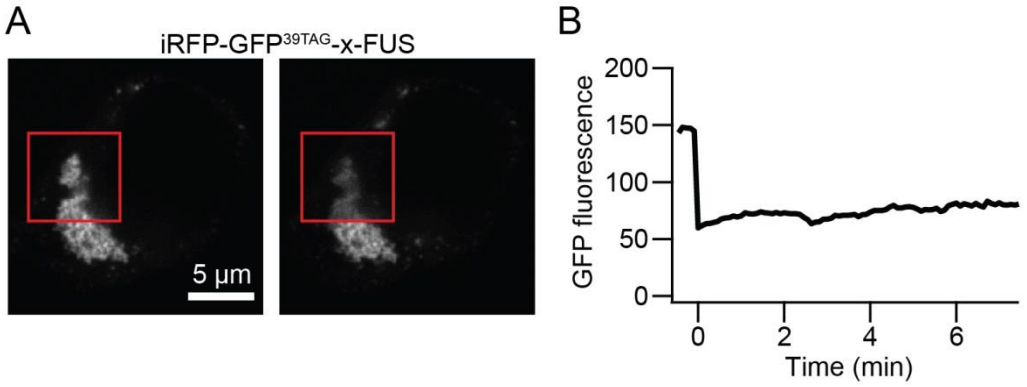

Figure 4.17. FRAP of iRFP-GFP^{39TAG}-X-FUS intracellular structure: A) Representative cytoplasmic structure for the iRFP-GFP^{39TAG}-X-FUS construct transiently transfected in HEK293T cells. Imaging experiments are done 24 h after transfection. The intensity of the bleached region (red square) was measured before and after bleaching over time B). No fluorescence recovery was detected over 8 minutes recorded time. Scale bar; 5 µm.

FRAP experiments indicate that during the first 24h after transfection the FUS forming intracellular structures are not dynamic under these conditions. Motivated by this result I explored different approaches to achieve the selective spatial separation between truncated and full length reporter protein.

4.2.3. Moving the liquid-droplet selection system to the membrane

In order to achieve a better control of the localization of the FUS droplets and maybe avoid the accumulation of big cytoplasmic static structures, I targeted the FUS-TEV construct to the plasma membrane by fusing the 10 aa N-terminal sequence of Lck (lymphocyte-specific protein tyrosine kinase) to the N-terminal site of FUS-TEV. Proteins fused to the Lck sequence will localize at the plasma membrane due to the myristoylation and palmitoylation of the Gly and the two Cys residues in the Lck sequence (Anton et al., 2008; Zlatkine et al., 1997). The expression of Lck-protease constructs with and without FUS revealed that in the presence of Lck-FUS-NTEV (the NTEV fragment was used to avoid cleavage in the reporter) the DD-X-iRFP-GFP^{39TAG}-X-FUS reporter was able to co-localize with the Lck-FUS-NTEV construct, whereas in the absence of FUS no co-localization was detected (Figure 4.18). Further studies revealed that the co-localization between the two FUS containing constructs was only taking place at the membrane when Lck-FUS-protease was present and there was a spatial constrain that led to a small cytoplasmic area or when big cytoplasmic FUS intracellular structures were formed. No sign of homogeneous membrane co-localization was achieved with this approach. In addition, large cell to cell variations were observed due to differences in expression levels that would have a direct impact on the granule-like structures formed in the cytoplasm.

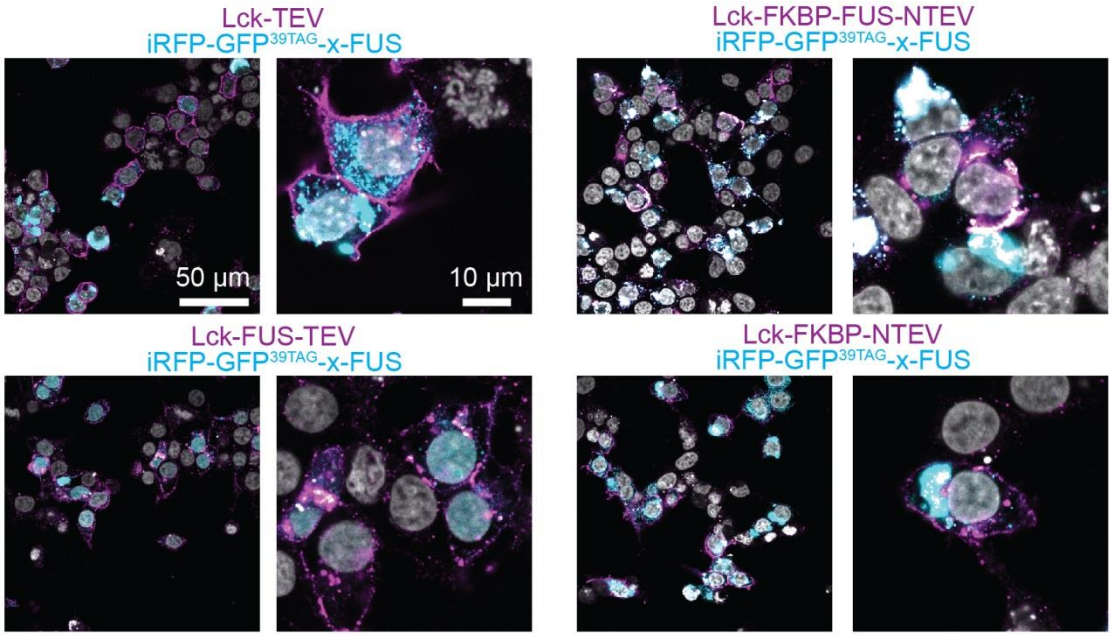

Figure 4.18. Confocal microscopy images of HEK293T cells expressing the reporter construct iRFP-GFP^{39TAG}-X-FUS: Cells co-transfected with tRNA^{pyl}/RS^{pyl,AF} and A) Lck-HA-TEV, B) Lck-HA-FKBP-FUS-NTEV, C)Lck-HA-FUS-TEV, D) Lck-HA-FKBP-NTEV. The constructs containing HA-tag were labeled by immunofluorescence with an anti-HA mouse antibody and an anti-mouse A647 antibody. Cyan: GFP signal, magenta: Anti-mouse A647. Scale bars: 50 and 10 µm.

Analogous to the microscopy images (Figure 4.18), I tested the behavior of the protease targeted to the membrane analyzing by flow cytometry the fluorescent signal for the construct containing the DHFR at the N-terminal site (Figure 4.19). As previously shown (Figure 4.14), when an active protease is produced in the cell independently of if it forms droplets or if it is targeted to the membrane (FUS-TEV/Lck-TEV/Lck-FUS-TEV) there is an increase in the amount of accumulated iRFP only signal meaning an accumulation of truncated reporter that did not incorporate the ncAA. This indicates that the protease is able to cleave off the DD not only to full length reporters but also to truncated ones. Thus, selective cleavage of the full length protein is not optimal.

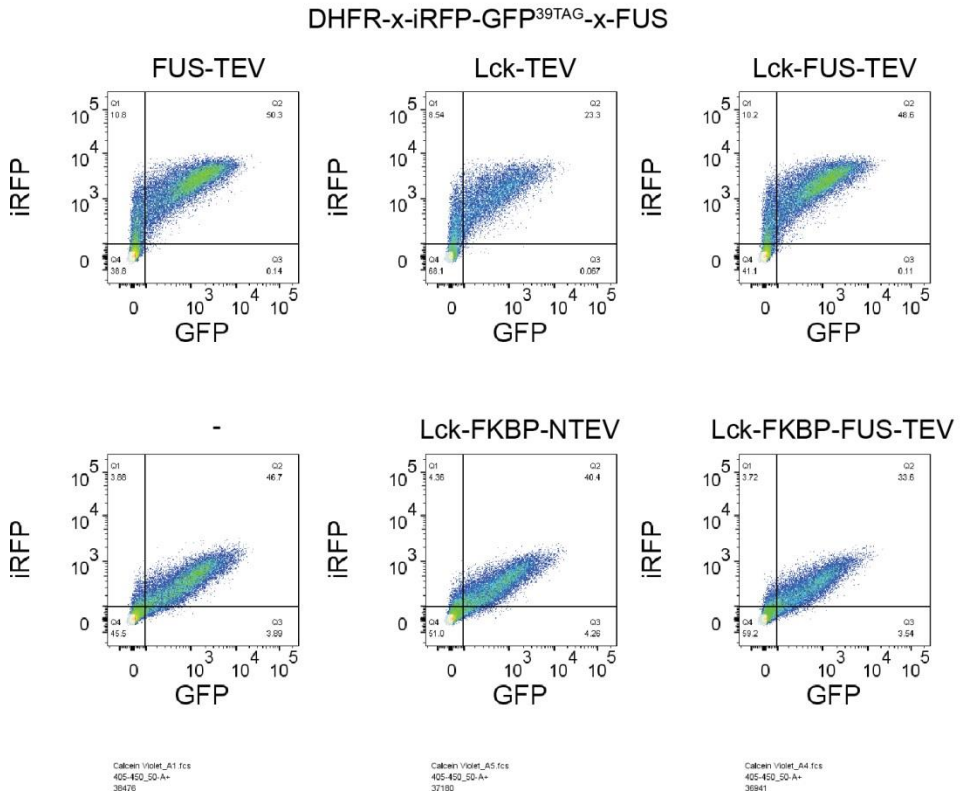

Figure 4.19. Flow cytometry results obtained by measuring the iRFP and GFP signal: HEK293T cells transfected with a plasmid containing the reporter construct DHFR-X-iRFP-GFP^{39TAG}-X-FUS and a plasmid containing the with tRNA^{pyl}/RS^{pyl,AF} in the presence of the ncAA BOC-Lys. The intracellular iRFP-GFP signal was measured in the absence of a protease system (-) and in the absence of a functional protease system (Lck-FKBP-FUS-NTEV, Lck-FKBP-NTEV and -) and in the presence of a functional protease system (Lck-TEV, Lck-FUS-TEV and FUS-TEV).

In order to achieve a complete spatial separation of the selective cleavage mechanism, I decided to target the reporter to the plasma membrane too. I first tested different membrane targeting domains that could be fused at the C-terminal site of my reporter protein (Figure 4.20). These include the C-terminal sequence CIIC derived from the Ras protein and the C-terminal sequence CAAX of the KRas protein (Hancock et al., 1991). Confocal microscopy images show that the constructs containing the C-terminal sequence CAAX are targeted to the membrane more efficiently than the constructs containing the CIIC sequence (Figure 4.20).

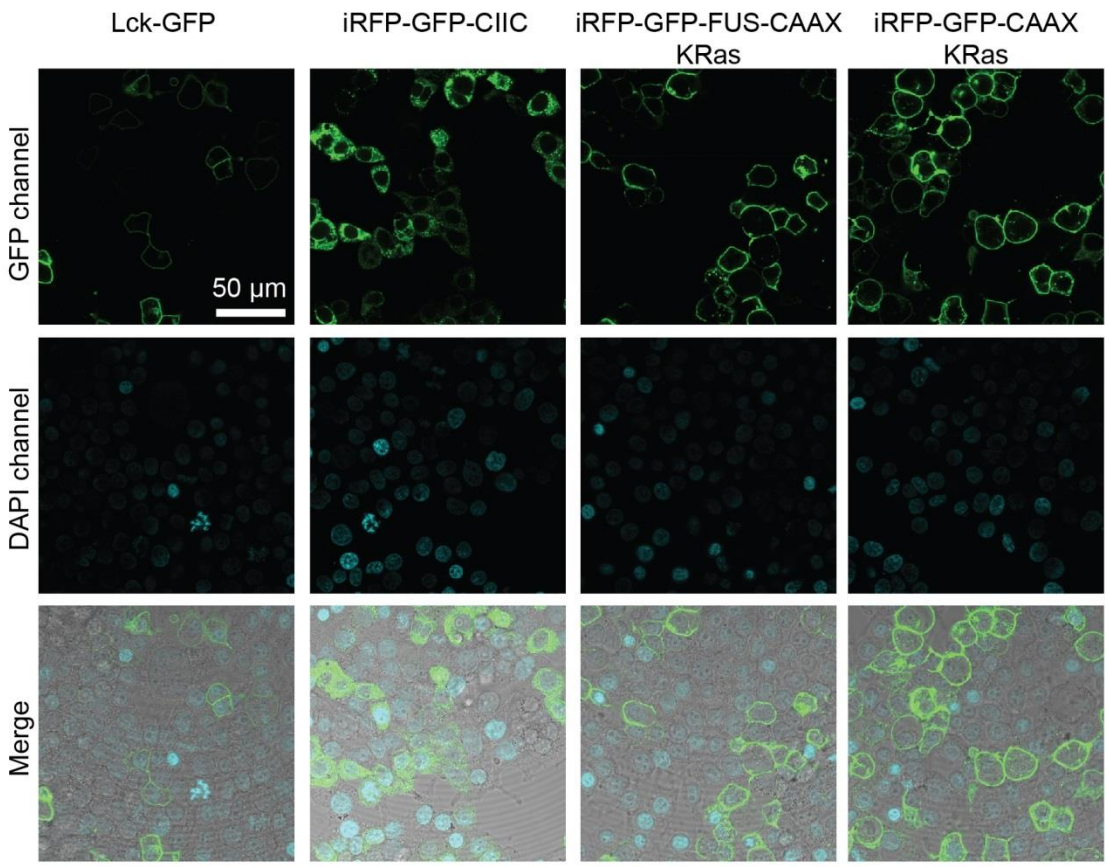

Figure 4.20. Confocal microscopy images of HEK293T cells expressing GFP reporters containing different membrane targeting domains: Lck-GFP was used as a positive control of membrane localization. The reporter containing the CIIC sequence at the C-terminal domain did not show a localization restricted to the plasma membrane. Reporter fused to -FUS-CAAX and -CAAX were localized mainly at the plasma membrane. Scale bar: 50 µm.

Then I tested that the membrane targeted reporter system is also cleavable (Figure 4.21) by cytoplasmic and membrane targeted TEV protease. Membrane targeted reporter with and without FUS was cleaved by membrane targeted TEV proteases. However, membrane targeted TEV protease without FUS generated less cleaved reporter (iRFP-GFP^{39TAG}) and uncleaved reporter (DD-iRFP-GFP^{39TAG}-FUS/-FUS-CAAX) was still detected. On the other hand FUS containing TEV protease only led to the formation of iRFP-GFP^{TAG}. This effect could be due to the FUS domain acting as a flexible linker that is able to better cleave due to a higher mobility of the protease.

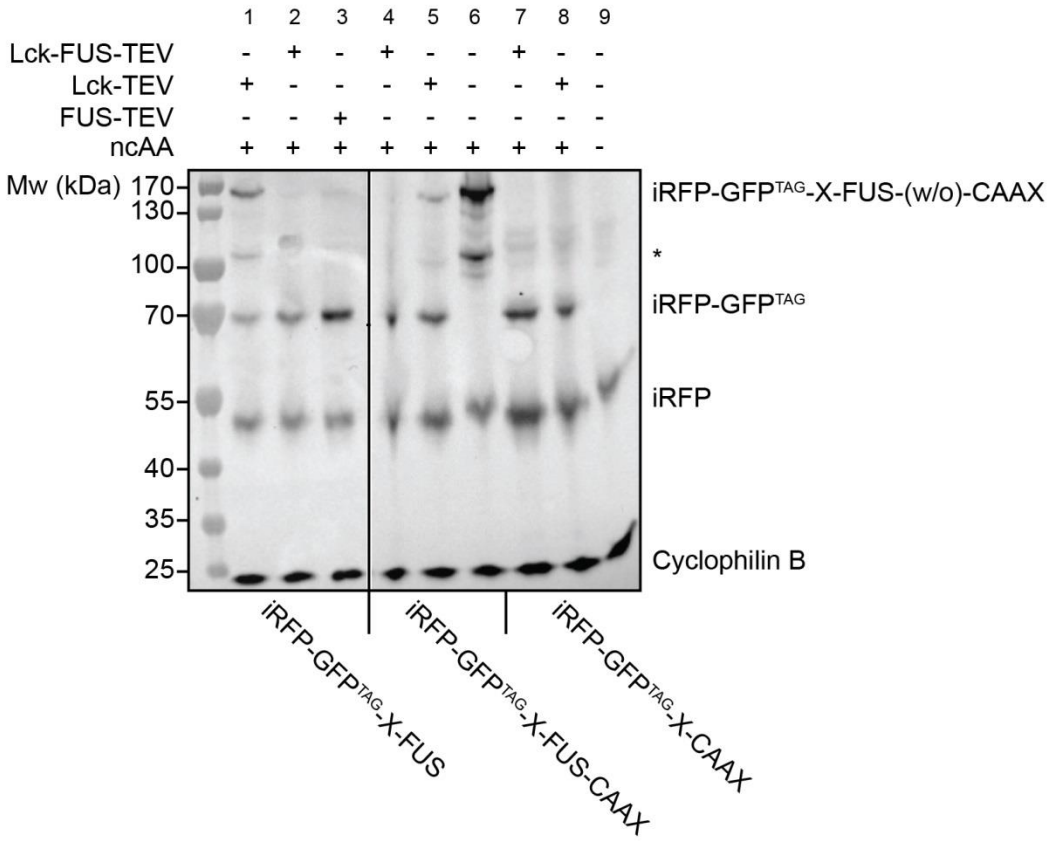

Figure 4.21. Western blot analysis of protease cleavage at the plasma membrane: The reporter in the absence of TEV protease is present in the full length (iRFP-GFP^{39TAG}-X-FUS-(w/o)-CAAX) form and in the truncated from (iRFP) but no cleaved POI is present (iRFP-GFP^{TAG}) (lane #6). (*) corresponds to a reporter that has started in an internal ORF that leads to the formation of GFP-X-FUS-(w/o)-CAAX. The lanes #1) and #5) containing the Lck-TEV protease at the membrane, show show a fraction of uncleaved reporter and of cleaved iRFP-GFP^{TAG}. Lanes #2) and 4) containing the same reporter constructs but in the presence of Lck-FUS-TEV protease show complete clavage of the POI and no uncleaved reporter. Lane #7) and #8) in the absence of FUS in the reporter construct that is targeter to the membrane both Lck-FUS-TEV and Lck-TEV lead the the complete cleavage of the reporter protein.

Chapter 5

Discussion

5. Discussion

I have divided this chapter in three different sections. In the first section (section 5.1) I will discuss the binding mechanisms of FG-Nup/NTR interactions from the point of view of the structure and dynamic of IDPs (corresponding to results section 3.1). In the following section 5.2, I will discuss the transport of NTRs across the NPC based on the binding mechanism discussed in section 5.1. In the final section (corresponding to the result part 3.2), I will discuss the strategy that I have developed to achieve the *in situ* study of the structure and dynamics of FG-Nup.

5.1. Molecular binding mechanism of FG-Nups and NTRs

5.1.1. FG-Nups and NTRs form fuzzy complexes

FG-Nups are key components in the regulation of nucleocytoplasmic transport. The key to elucidate how the fast and specific crossing of the NPC is achieved is to understand the FG-Nup/NTR interaction mechanism. However, the characterization of their structure and dynamics upon binding to NTRs has remained majorly unknown due to the challenges associated with the study of long multivalent disordered proteins.

In this thesis, I have performed smFRET and FCS studies of the interaction between FG-Nups and NTRs which revealed that different FG-Nups are highly dynamic in their bound state and that retain the same conformational heterogeneity as in the unbound state (Figures 4.1 and 4.2). IDPs that when bound to folded proteins can retain a certain degree of conformational heterogeneity are termed as fuzzy complexes (Fuxreiter and Tompa, 2009; Sharma et al., 2015; Tompa and Fuxreiter). I have contributed to show that multiple FG-Nup/NTR complexes occur via the interaction of minimalistic motifs, basically reduced to the F residues of the FG-Nups that engage with different NTR's binding pockets while the dynamics or fuzziness of the FG-Nup is not affected upon binding to the NTR. Moreover, via stopped-flow kinetic measurement, I have shown that FG-Nups engage with the NTRs with diffusion limited kinetics (Figure 4.3-4.6). This suggests that successful binding between FG-Nups and NTRs takes place upon collisional encounter, which is probably enabled by the highly reactive surface in both binding partners.

The reduction of the binding site to the minimum size combined with the multivalent nature of FG-Nups and NTRs enables the formation of multiple contacts which increase the specificity of the molecular recognition between FG-Nups and NTRs paying a minimal entropic cost caused by the structural freedom of the FG-Nup and its fast dynamics. Moreover, no conformational change is required because the F residues are distributed across the sequence, which enables the exposure of binding motifs at the surface

of the FG-Nups making them readily available to engage in the binding to the NTRs (data shown by MD simulations (Milles et al., 2015)). We have termed this novel IDP binding mechanism as ultrafast archetypal multivalent fuzzy complex (Figure 5.1). This phenomenon has also been observed in yeast FG-Nups indicating that it might be a conserved mechanism across different species (Hough et al., 2015; Milles et al., 2015; Raveh et al., 2016).

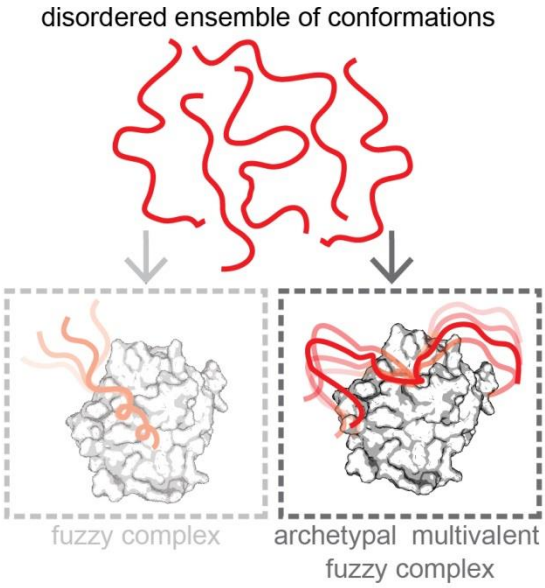

Figure 5.1. Fuzzy complex binding of IDPs to a folded protein: Reported typical fuzzy complexes are formed when an IDP partially undergoes a conformational change and binds a folded binding partner. However, parts of the IDP will remain disordered in the bound state (left, light color). The archetypal multivalent fuzzy complex that we have described is an extreme case of a fuzzy complex where no disorder to order transition takes place upon binding to the folded protein (right, bright color).

5.1.2. Differential binding mechanism between Nup214FG with NTRs

During the course of my PhD a crystal structure of the C-terminal region of Nup214FG in complex with CRM1/RanGTP was reported (Port et al., 2015). The crystal structure showed that Nup214FG binds in three different regions around the outer surface of CRM1 linking the C- and the N-terminal sites of CRM1 when is in complex with RanGTP. Due to the findings revealed by the crystal structure and considering the lack of knowledge on how different functionalities may be encoded in the disordered sequences of FG-Nups, I was prompted to study the structure and kinetics of Nup214FG/NTR binding.

Our results from smFRET showed that the C-terminal region of Nup214FG undergoes a conformational change upon binding to CRM1 or CRM1/RanGTP. Nup214FG adopts a more extended conformation when bound to CRM1 (Figure 4.7). Surprisingly, this was also the case when Nup214FG formed a complex with Importinβ (Figure 4.8). Therefore, it is likely that the amino acid sequence encoded in Nup214FG determines this conformational change-associated binding.

Our stopped-flow kinetic measurements showed that Nup214FG binds to CRM1 and CRM1/RanGTP with very fast kinetics (10^7 $M^{-1}s^{-1}$) that is two orders of magnitude slower than the ones obtained for the ultrafast multivalent fuzzy binding mechanism (Figure 4.9). Moreover, kinetic experiments obtained by monitoring the conformational change (change in donor fluorescence, due to a decrease in the E_{FRET}) revealed similar values as the ones obtained from anisotropy measurements indicating that the conformational change and the binding to the NTR occur at similar rate constants (Figure 4.9). This may suggest that the association kinetics between Nup214FG and CRM1 are limited by the conformational change of Nup214FG.

Interestingly, the crystal structure of the 117 amino acid long Nup214FG with CRM1/RanGTP showed that two α-helices were formed at the two terminal binding sites of Nup214FG. MD simulations (collaboration with Dr. Davide Mercadante, Gräter laboratory) of the crystalized Nup214FG fragment and CRM1/RanGTP revealed that the unbound disordered ensemble of Nup214FG displayed a subset of end-to-end distance distributions that contained helical structures as the ones present in the bound form. Moreover, the FG-motifs neighboring the two α-helices showed higher affinities and were less dynamic than other interacting FG binding motifs in the bound form. Thus, in contrary to the archetypal fuzzy complexes formed by other FG-Nups, Nup214FG undergoes a conformational change upon NTR binding and in addition, it contains certain degree of secondary structure that may influence the binding or stability of the complex. I will now discuss how by exploiting the technology of stopped-flow kinetic measurements additional information could be deciphered regarding the binding mechanism between NTRs and Nup214FG.

Secondary structure formation upon binding of IDPs is known as coupled folding-binding. Coupled folding-binding is a two-step reaction were the IDP undergoes a disorder to order transition upon binding. Depending on if this transition occurs before or after binding it is distinguished between conformational selection and induced fit mechanism respectively. One way of deciphering whether a coupled folding-binding interaction between an IDP and a folded protein takes place by a conformational selection or an induced fit mechanism is by performing kinetic experiments that indicate the limiting step of the two-step binding reaction (Figure 5.2). The studies of one-step kinetic reactions are usually carried out under pseudo-first order conditions (detailed explanation in the material and methods section 3.1.6) (Figure 5.2 left panel). Under these conditions the observed rate (k_{obs}) is equal to the rate constant times the concentration of the reactant that is in excess and the k_{on} is obtained from the slope of the k_{obs} vs concentration of reactant in excess plot. If we now use these same pseudo-first order conditions with a system that binds via a two-step mechanism, like for example an induced fit mechanism, at pseudo-first order conditions (with reactant B in excess and reactant A is the IDP that will undergo a conformational

change after binding) or at inverse pseudo-first order conditions (reactant A in excess and B is the limiting reactant) the k_{obs} vs concentration plot obtained will be the same (Figure 5.2 middle panel). This is because the limiting step in this mechanism is caused by the binding. The main difference from the example of a one-step reaction is that instead of having a linear relationship between the k_{obs} and the concentration it has a hyperbolic relationship as it is the case for two-step reactions.

In the case of conformational selection, one of the reactants is populating an ensemble of different conformations (in this case reactant A). Among the different conformations the ones with higher affinities (A^*) for the reactant B will bind. If now we do the same exercise of measuring the k_{obs} under pseudo-first order conditions with reactant A in excess, the binding will be comparable to a one-step reaction (if the conversion between A and A* is fast). The reaction will be limited by the concentration of B, which will be lower than the concentration of A^* and the relationship between k_{obs} vs concentration will be linear. However, if we now measure the association between A and B under inverse pseudo-first order conditions (B in excess and A is the limiting reactant) only a fraction of A (A^*) will be capable of binding B. Therefore the limiting step will be the disorder to order transition occurring from A to A^*. Thus, among increasing concentration of B the k_{obs} vs concentration plot will have a hyperbolic behavior (Figure 5.2 right panel).

Despite the apparent simplicity of these kinetic experiments, the discrimination between induced fit and conformational selection using kinetic measurements is not trivial and requires the differentiation of the two phases in the measured traces (folding and binding in the case of conformational selection). This can be experimentally challenging when for example the amplitude of one of the phases is too low and it may not be easily distinguished. In addition, if we consider that Nup214FG only contains two short α-helices that only occur in small regions of the disordered sequence, the folding and unfolding kinetics will probably be fast so the ensemble kinetic measurement will detect an average of both conformations and the measured kinetic trace will be dominated by the slower phase, which in this case would be the binding of the two molecules. Thus, a potential conformational selection mechanism could be masked as a simple one-step binding mechanism. In addition to the difficulties that may arise from the quality of the measured traces, labeling CRM1 and achieving a detectable signal change upon binding of Nup214FG is also technically challenging. Moreover, high concentrations of Nup214FG may aggregate, which is an additional factor that would need to be considered. The exact role of these terminal α-helices and their effect on the association or dissociation kinetics will be studied in the future.

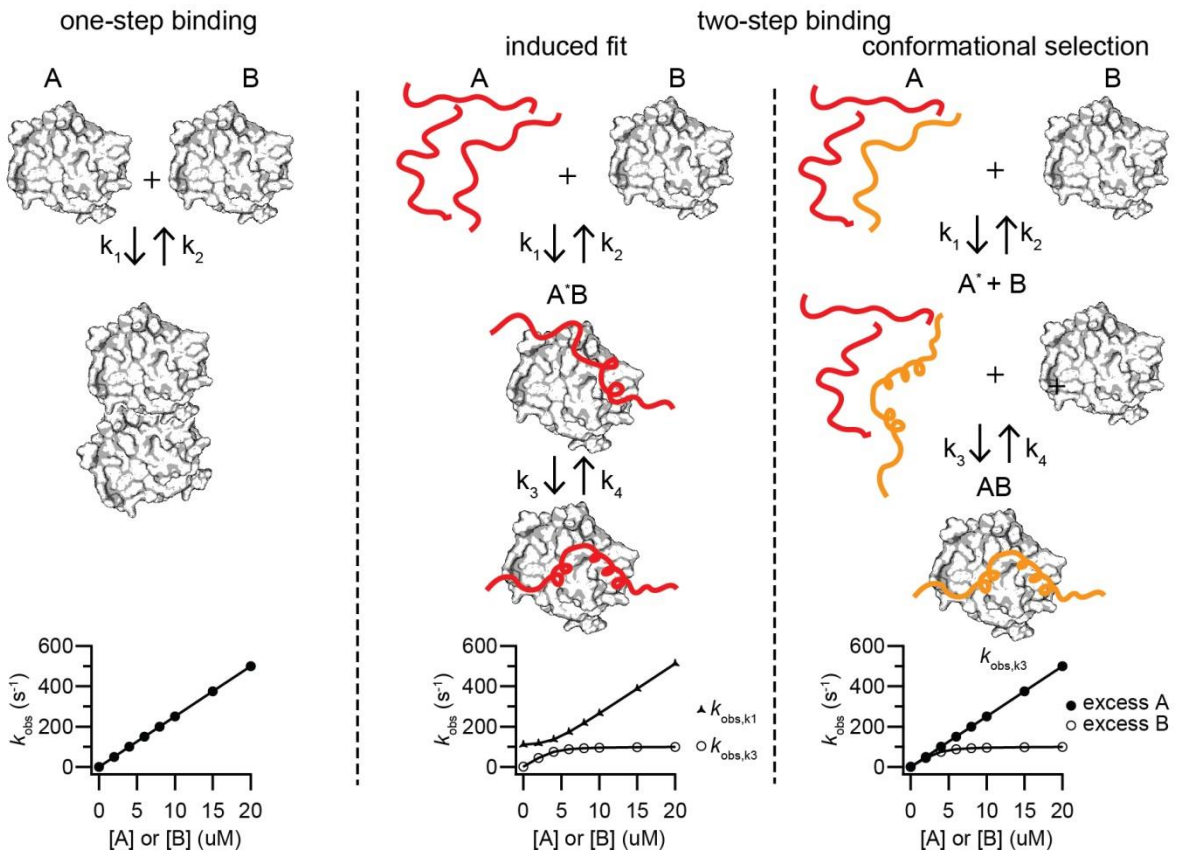

Figure 5.2. Cartoon representation of the binding reactions of a one-step binding and a two-step binding: The reaction rates involved in the one-step binding reaction involve the association rate and the dissociation rate (k_1 and k_2) The k_{obs} vs concentration plot will be linear independently if A or B is in excess. In a two-step binding there are two additional reaction rates (k_3 and k_4). In the case of induced fit k_1 and k_2 also represent the association and dissociation rate. After binding, k_3 represents the folding rate and k_4 the unfolding rate. In induced fit the k_{obs} vs concentration plot will always display a hyperbolic increase with A or B in excess (black triangles k_1 and white circles k_3). In the conformational selection mechanism k_1 and k_2 represent the folding and unfolding rate of the unbound IDP and k_3 and k_4 the binding and unbinding rate of the folded conformer. When the transition from A to A* is fast the dominating rate will be from the second step (k_3). k_{obs} vs concentration plot of the k_{obs} corresponding to the second step (k_3) will behave differently in the excess of A (black circle) or of B (white circle).

5.1.3. Potential sequence characteristics associated to the distinct binding of Nup214FG

Nup214FG interacts with NTRs in a distinctive way compared to the other FG-Nup/NTR interactions studied in this thesis. Nevertheless, it is still not well understood, what exact feature makes Nup214FG behave differently than other FG-Nup for example, Nup153FG. Based on the crystal structure of Nup214FG and CRM1/RanGTP we know that the 117 aa long C-terminal region of Nup214FG engages with 8 FG-motifs distributed in 3 FG-regions. We also know that the FG-region 1 and the FG-region 3 contain small α-helices (located at positions 1923-1932 in region 1 and 2014-2022 in region 2) (Port et al., 2015).The two small helices contain a hydrophobic and a hydrophilic side such that the hydrophobic side is involved in the binding with CRM1 (Figure 5.3).

In order to test the role of these two helices in the Nup214FG/CRM1 binding, the residues A1927P, A2017P and A2019P (Nup214FG AtoP), which are residues present in the hydrophilic side of the helices, were mutated. Saturation binding curves (data not shown) indicated a 5 fold increase on the K_D of the Nup214FG AtoP mutant compared to the wild type, however further experiments are required to confirm the effect of the mutations and the degree of helicity present on the mutant Nup214FG.

Next, I compared the sequence composition of the probed regions by smFRET of Nup153FG^{PxFG}, Nup153FG$^{FxFG(I)}$, Nup153$^{FxFG(II)}$ and of Nup214FG using the online tool EMBOSS Pepinfo (Rice et al., 2000)to identify any particular feature in the sequence that may determine the different binding mechanism. Except for Nup153FG^{PxFG} the rest of the sequences contain charged amino acids across their sequence. Interestingly, in Nup214FG the charged amino acids seem to be predominantly distributed in the FG-region 1 and 3 of Nup214FG (Figure 5.3). Thus, it would be possible that the charges and charge distribution are influencing the binding mechanism of Nup214FG. Hence, by making use of the Nup153FG and Nup214FG models we could advance in the understanding of how FG-Nup sequences that seem similar from their sequence composition are able to encode different functionalities.

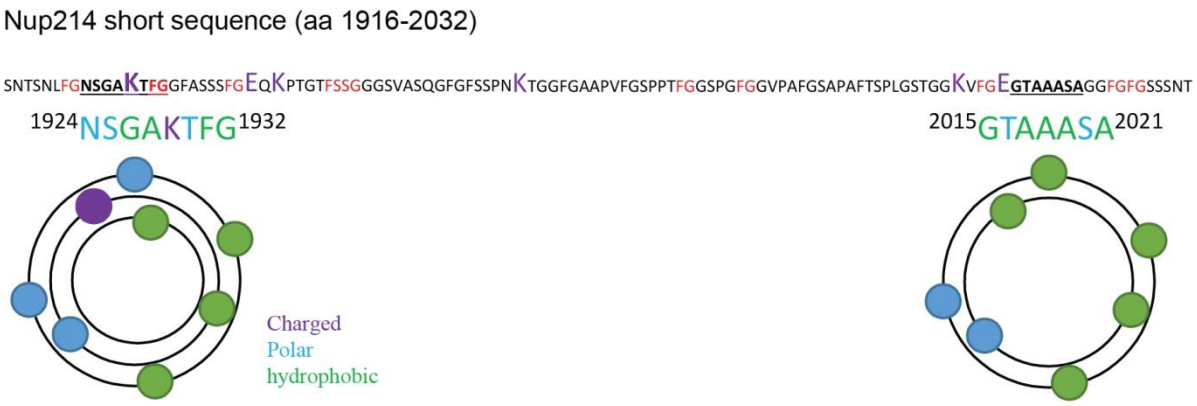

Figure 5.3. Sequence of the CRM1 binding region of Nup214FG^{117aa}: The FG-motifs are colored in red. The charged residues are colored in purple. The two α-helices formed in Nup214FG upon binding CRM1/RanGTP (bold and underlined in the sequence) are represented with concentric circles illustrating a helix turn. The amino acids are color coded (purple: charged, blue: polar and green: hydrophobic) sequence corresponding to 1916-2032.

5.1.4. Effect of glycosylation on FG-Nup/NTR binding

Metazoan FG-Nups are heavily glycosylated with O-linked-GlcNAc (GlcNAc) at the NPC. It has been shown that depletion of glycosylated FG-Nups disrupts the nucleocytoplasmic transport (Favreau et al., 1996) and that glycosylated Nup98FG can form hydrogels that allow a deeper penetration of NTRs, suggesting a possible role in the structural arrangement of FG-Nups when they form supramolecular structures (Labokha et al., 2013). In addition, down-regulation of OGT1, the glycosyltransferase responsible for the GlcNAc glycosylation, caused an increase in ubiquitination and proteosomal

degradation of FG-Nups in mouse embryonic fibroblast cells, affecting the integrity of the NPC (Zhu et al., 2016). However, if and how glycosylation may affect the FG-Nup/NTR binding is poorly understood.

I performed smFRET and kinetic studies on *in vitro* glycosylated FG-Nups. SmFRET measurements of glycosylated and unglycosylated Nup214FG and Nup153FG showed that glycosylated FG-Nups have a more expanded conformation compared to the unglycosylated ones (Figure 4.10). However, this may be a consequence of excessive glycosylation caused by the *in vitro* glycosylation procedure, which lacks the regulatory mechanism that would be present in the cell. Despite the structural change in the ensemble of conformations populated by the IDPs, smFRET showed that the different NTR binding mechanisms of Nup153FG and Nup214FG were conserved in the glycosylated Nups (Figure 4.11). On the other hands, the kinetic values obtained from stopped-flow experiments with glycosylated and unglycosylated FG-Nups upon binding to the NTRs were comparable (Figure 4.12). Our data indicates that the binding mechanism between FG-Nup^{Glyc} and NTRs seems to be consistent with the results obtained using unglycosylated FG-Nups. Thus, for the study of the binding mechanism, unglycosylated FG-Nups appears to be a reasonable mimic.

5.2. FG-Nups/NTR interactions and the nucleocytoplasmic transport

5.2.1. The fast and selective crossing of the NPC requires fuzzy FG-Nup/NTR interactions

The central channel of the NPC is filled by FG-Nups. Based on the latest cryo-electron microscopy data of the NPC structure (Kosinski et al., 2016) I calculated the volume for the central channel of the NPC of 38.200 nm^3 based on the approximated values of 22.5 nm height and 24 nm radius of the of the NPC inner ring. Assuming that the FG-Nups located around the central channel contain the disordered regions within this volume and knowing the FG-Nup stoichiometry at the NPC (Ori et al., 2013), the calculated concentration of FG-motifs is around 160 mM. The concentration increases to 240 mM if the GF-motifs are also taken into consideration and becomes 260mM if all the F residues of the FG-Nups from the central channel are considered. Moreover, the estimated NTR concentration from yeast lysate is ~20 µM (Ghaemmaghami et al., 2003) and the measured Importinβ concentration in neuroblastoma cells was 5 µM (Paradise et al., 2007). If we now consider that reported K_D values for FG-Nup/NTR complexes range from pM to µM (Tetenbaum-Novatt et al., 2012), NTRs would be constantly bound to the FG-barrier. Low K_D values are often associated with long-living complexes which would not be compatible with the fast millisecond crossing times of NTRs measured *in vivo* (Kubitscheck et al., 2005; Sun et al., 2013; Yang et al., 2004). With the results presented in this thesis I aim to contribute to the understanding of this apparent paradox. Next, I will describing how the obtained K_D and kinetic values can be used to explain the mechanism that allows the fast crossing of the NPC.

If we consider the relationship of $K_D=k_{off}/k_{on}$ used in bimolecular interactions and by using the K_D of 40 nM and the k_{on} of $1.5x10^9$ $M^{-1}s^{-1}$ (Figure 4.3), both obtained by monitoring the anisotropy change of the interaction between Nup153FG and Importinβ, the calculated k_{off} value is equal to 60 s^{-1}. Using the obtained k_{off} value we can calculate the half-life of the formed complex (detailed information about the kinetic formulas on materials and methods section 3.1.6). The half-life of a formed complex is equal to $0.693/k_{off}$ thus we would obtain a complex half-life of 11.5ms. Since the reported transport time of molecules across the NPC is between 5-30 ms it would not be possible for a NTR to cross the NPC barrier by undergoing multiple binding and unbinding events with the different FG-motifs. If instead of using the calculated k_{off} from the measured K_D and k_{on} for Nup153FG/Importinβ I use the measured k_{off} value of 8.6 $s,^{-1}$ obtained from the dissociation stopped-flow experiments (Figure 4.6), the FG-Nup/NTR complex has a half-life of 80.5 ms.

The measured k_{off} values that I obtained from the kinetic experiments are all orders of magnitude too slow to achieve the fast NPC transport. However, when we analyze the nature of the interaction between FG-Nups and NTRs we first need to acknowledge that both proteins are multivalent molecules. Different FG-motifs from the same FG-Nup can engage simultaneously with different NTR binding pockets of the same NTR and/or of different NTR molecules. This protein-protein interaction mechanism is complex, the number of binding events occurring between proteins cannot be quantitatively measured and hence makes the interpretation of the results extremely difficult if solely based on conventional protein-protein interaction models.

If for simplification reasons we compare a monovalent and a bivalent molecule in which each individual binding site has the same K_D, k_{off} and k_{on} values and we now focus on the dissociation of the molecule (k_{off}) and of each individual binding site ($k_{off,single}$) (Figure 5.4), the dissociation of the molecule will occur when both binding sites are unbound. The probability of this to occur depends on the affinity of each site and the effective concentration generated when one of the binding sites is bound (Kramer and Karpen, 1998). This causes $k_{off,single}$ to be orders of magnitude faster than the k_{off} of the entire molecule. Thus, the measured k_{off} for Nup153FG/Importinβ will be slower than the $k_{off,single}$ of an individual FG-motif. For this reason dissociation experiments report on the k_{off} value obtained when all the binding motifs have dissociated from the different pockets of their binding partner.

If we consider that when NTRs are bound to some ligands, like Importinβ/RanGTP, some of the FG-binding sites cannot accommodate FG-motifs due to a conformational change of HEAT repeats important for FG-Nup binding (Vetter et al., 1999), the relationship between the k_{off} and the number of binding sites available in the surface of the NTR adds another regulatory level to nucleocytoplasmic transport, which could be used to regulate the dwell times of different transport complexes at the NPC.

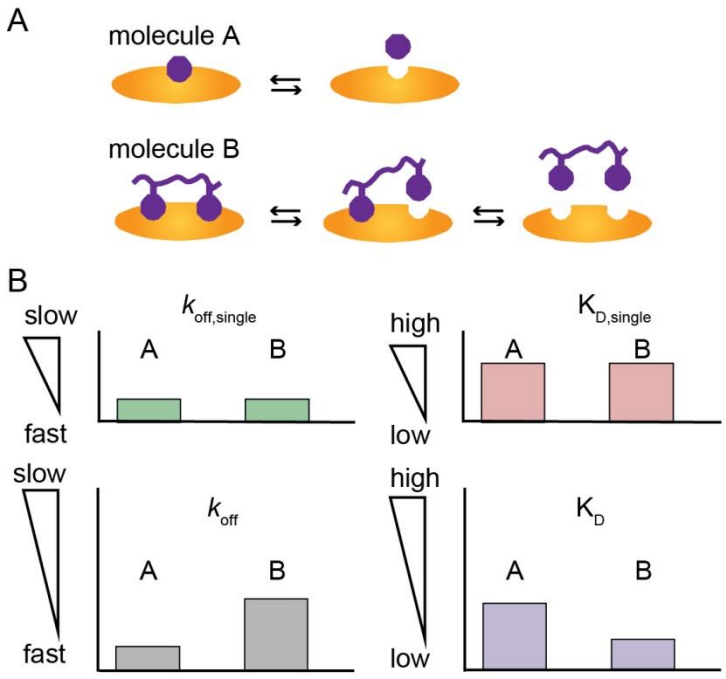

Figure 5.4. Cartoon illustrating the binding relationship between the valency and the k_{off}: A) Single binding bimolecular system (molecule A) and a divalent system (molecule B) the binding receptor (orange) and the binding motif (purple circle). B) Representation of the differences in the $k_{off,single}$ (green bar plot) and $K_{D,single}$ (red bar plot) of a single motif in a monovalent and a divalent molecule (molecule A and B) and the corresponding k_{off} (gray bar plot) and K_D (purple bar plot) plots for the molecules. At the molecule level, molecule A will have a faster dissociation than molecule B due to multivalency.

Another parameter that needs to be differentiated in monovalent and multivalent interactions is the affinity and avidity. Avidity is the result of the combination of individual binding strengths of multiple affinities occurring simultaneously as part of an individual non-covalent binding event. This concept should be differentiated from affinity which refers to an individual binding strength of a single binding event.

So far for all the reported K_D values on this thesis correspond to the avidity of the protein complexes. In the case of the FG-Nup/NTR binding titration curves performed by monitoring anisotropy (Figure 4.9 A-B, Figure 4.12A), I have reported the avidity of the region where the dye is located. This is caused by the segmental motion behavior of FG-Nups, which means that the rotation of one region does not affect the rotation of the other, thus changes in rotation (anisotropy) upon binding of the NTRs will only be sensitive in the proximity of where the dye is located (Milles and Lemke, 2014). FCS binding curves on the other hand, report on the avidity at the molecule level. The component being monitored is the change in diffusion time of the molecules and this will be dependent on how many molecules are bound to your labeled protein independently of where the dye is placed (Figure 4.12B-C).

Biochemical experiments have reported K_D values ranging from pM to µM for different FG-Nup/NTR complexes (Ben-Efraim and Gerace, 2001; Gilchrist et al., 2002; Hough et al., 2015; Lott et al., 2010; Milles and Lemke, 2014; Milles et al., 2015; Otsuka et al., 2008; Pyhtila and Rexach, 2003; Tetenbaum-Novatt et al., 2012). A comparative study of the K_D in the presence and absence of cellular milieu showed that in the presence of cell lysate the K_D values increase several orders of magnitude being close to the mM range (Tetenbaum-Novatt et al., 2012). In addition, other studies have also reported mM avidities (Tu et al., 2013). In the case of Nup153FG, the PxFG-rich region ($Nup153FG^{PxFG}$) located at the C-terminal site was used in NMR experiments carried out by Dr. Sigrid Milles (Blackledge laboratory, IBS, Grenoble). ^{15}N relaxation experiments of the $Nup153FG^{PxFG}$ at different Importinβ concentrations revealed a residue specific K_D of ~2 mM per F residue (Milles et al., 2015). If we calculate the $k_{off,single}$ using the measured k_{on} as an upper limit for a single motif and the 2 mM affinity value obtained by NMR we obtain a complex half-life of 0.2 µs, orders of magnitude smaller than the one calculated using the 40nM measured avidity or the one obtained from the dissociation kinetic experiments. Considering that the transport time of labeled Importinβ across the NPC measured in our lab is on average of 6.9 ms (Milles et al., 2015), with a complex half-life of 0.2 µs, Importinβ can unbind more than 20000 times. In a similar way NMR spectroscopy was used to obtain a K_D value of 36µM for the interaction between Nsp1 and Kap95 (Hough et al., 2015). Even with this lower K_D value, the complex half-life is 12.8 µs.

Based on the kinetic information, we propose a model in which the fast transport of the NPC can be achieved. When an NTR encounters the FG-Nup barrier at the NPC it will engage with the FG-Nups with multiple low affinity interactions that bind and unbind with ultrafast kinetics, allowing the fast crossing of the NPC (Figure 5.5) (Milles et al., 2015).

During the course of this PhD thesis a complementary-Nup/NTR binding mechanism called slide-and-exchange, was presented by Raveh *et al*. based on the results obtained from the combination of all-atom MD simulations with NMR spectroscopy of the interaction between a FxFG-rich peptide and a NTF2 dimer (Raveh et al., 2016). They observed that the FG-motifs were very dynamic when bound to NTF2. The local motion of the Nup in the binding groove of NTF2 enabled the FG-motifs to slide in and out of the hydrophobic groove. In this way FG-Nups were bound from a strong interacting state into a weak interacting state allowing the access of other FG-motifs to bind into the strong interacting site (Raveh et al., 2016). In summary, we could consider the slide-and-exchange mechanism as an ultrafast fuzzy binding mechanism, were the fast exchange of individual binding motifs would allow the NTR to creep through the NPC in fast timescales.

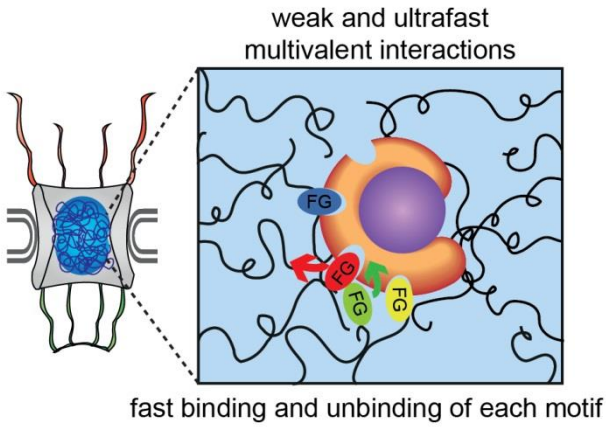

Figure 5.5. Cartoon representation of the crossing of NTRs through the NPC: Illustration of an NTR (orange) crossing the NPC. Multiple FG-motifs (blue, red, green and yellow ovals) are able to interact with the NTR via multivalent binding.

5.2.2. Spatial segregation of distinct binding mechanisms in the nuclear pore complex

Nup153FG and the C-terminal disordered region of Nup214FG are able to bind to CRM1 and to Importinβ. However, Nup153FG engages with the NTRs forming an ultrafast fuzzy complex and Nup214FG undergoes a conformational change upon NTR binding. The binding mechanism showed for FG-Nup214/NTR complexes is not in line with how FG-Nups can facilitate fast transport of NTRs across the NPC. This points to a unique role of the Nup214FG/NTR interaction in the nucleocytoplasmic transport, in line with previous observations(Hutten and Kehlenbach, 2006; Labokha et al., 2013). In the central channel of the NPC, the permeability barrier is formed by high densities of FG-Nups, a tight clamping mechanism will reduce the transport efficiency of cargoes. Nup214FG, is localized at the cytoplasmic side of the NPC and is most likely not a key component of the permeability barrier of the central channel at the NPC but rather may plays a different role in the nucleocytoplasmic transport. We can speculate that the free CRM1 at the cytoplasm that needs to cross the NPC to keep the pool of free CRM1 at the nucleus will be able to bind and unbind with very fast kinetics and subtle conformational changes through Nups in the central channel other than Nup214. However, when CRM1 is forming part of the export complex (bound to RanGTP), it will specifically bind to the C-terminal region of Nup214FG at the cytoplasmic face of the NPC. In this way, Nup214FG may act as a docking station for export complexes which then they may get in close proximity to RanBP2, which has also been shown to bind strongly to CRM1/RanGTP with two FG-regions two hundred amino acids apart(Ritterhoff et al., 2016), in order to undergo GTP hydrolysis and subsequent cargo release(Port et al., 2015).

Moreover, competition experiments have shown that the Ran binding protein, RanBP3, which facilitates the formation of the export complexes at the nucleus replaces Nup214FG from the CRM1•RanGTP complex(Port et al., 2015). Further studies have shown that two FG-regions of Yrb2p, yeast homolog of

RanBP3, located at the disordered domain also bind to the N- and C-terminal site of the yeast homolog of CRM1, Xpo1p (Koyama et al., 2014). This might indicate that RanBP3 and Nup214FG interact with CRM1•RanGTP with a similar binding mechanism as FG-Nup214. Thus, the formation and release of export complexes might be subjected to a tight biochemical control instead that the ultrafast spontaneous dissociation that is required at the central channel.

5.3. Towards the *in situ* stud of FG-Nups

State-of-the-art technological advances like superresolution microscopy, cellular spectroscopy and in-cell NMR are helping to bridge the gap between high resolution *in vitro* and *in cellulo* studies of IDPs (Konig et al., 2015; Sun et al., 2013; Theillet et al., 2016). The study of the structure and dynamics of the disordered regions of FG-Nups has so far mainly been limited to *in vitro* research. The aim of this project is to ultimately study the structure of the disordered region of FG-Nups, like Nup153, at the NPC. In order to achieve this I will make use of a previously established pipeline in the lab that uses Amber suppression technology for site-specific labeling via click-chemistry of a ncAA placed in the disordered region of Nup153FG.

The problem that arises from using this approach is that Amber suppression technology generates big amounts of truncated products which have not incorporated the ncAA. Usually, when the is folded, the truncated product may not be able to fold adequately and it will eventually get degraded. In the case of FG-Nups like Nup153, which has an NPC targeting domain at the N-terminal and the disordered FG-region at the C-terminal site, the truncated proteins will be outcompeting full length ones to get incorporated into the NPC. Thus, I am working on a system that selectively degrades truncated products while enriches the full length protein. Here, I will discuss the current key progress I made as a part of this PhD thesis and describe strategies on how to successfully complete the project in the future.

The selective degradation of truncated POI needs to be working synergistically with a selection mechanism that will discriminate between truncated and FL POI. My initial approach to achieve this was based on the use of a liquid-like synthetic intracellular droplet that would be able to "recognize and incorporate" FL protein and exclude truncated POI (Figure 4.13). Elimination of the truncated protein will then be achieved by fusing a degradation domain to the N-terminal site of the POI. In this way the FL POI will form liquid droplets and the truncated POI will be systematically degraded.

However, with the POI of interest trapped at the intracellular liquid-like droplets there has to be an additional mechanism for the successful release of the POI without the N-terminal degradation tag and the C-terminal droplet forming domain. The fusion of a droplet forming domain to a protease that will cleave specifically the sequence linking the N- and C-terminal domains to the POI will lead to the formation of

mixed droplets of FL POI and protease. If the phase-separated droplet is dynamic, the protease will cleave off the degradation and droplet forming domain and the POI containing the ncAA will be able to diffuse to the cytoplasm. Newly formed FL POI will be able to mix with the protease droplet, leading to a protease active liquid droplet that can selectively cleave the POI. Because the truncated product will not contain the droplet forming domain, it will not be able to be incorporated in the phase separated cytoplasmic droplets. Thus, it will contain the N-terminal degradation domain that would cause the continued degradation of the truncated protein.

Here, I will first discuss the findings regarding the degradation sequences followed by the different strategies to achieve the selective spatial separation of POI containing the ncAA.

5.3.1. Tuning protein degradation

In order to determine the ideal components to establish a selective degradation system I first designed a reporter POI formed by iRFP-GFP$^{39TAG \rightarrow ncAA}$, which can give a clear fluorescence readout of truncated (iRFP signal) and full length POI (FL POI) (iRFP-GFP signal). The selective degradation system needs to degrade truncated proteins while keeping the full length ones stable.

The degradation step is crucial for the elimination of the truncated products originated from the use of Amber suppression technology with a POI. Among the different degradation domains that I tested, the engineered DHFR protein from *E.coli* is the one that caused a bigger degradation effect compared to the other degradation sequences (Figure 4.14) (Iwamoto et al., 2010). DHFR constructs had a minimal signal on the iRFP channel only (single positive for iRFP). This indicates that probably DHFR is able to degrade the truncated proteins very efficiently.

Degradation of the reporter containing DHFR, after stopping the translation with cycloheximide, showed that the signal coming from the truncated POI (iRFP channel only) was reduced together with the progressive decrease of the full length reporter. Nevertheless, after 7.5 h 50 % of the initial iRFP-GFP^{39TAG} double positive signal was remaining (Figure 4.16). This indicates that full length iRFP-GFP^{39TAG}-FUS signal despite the droplet formation can get degraded but less efficiently that the truncated reporter (iRFP) which only shows some residual signal alto at time 0 (Figure 4.16)

Despite the apparent more efficient degradation of truncated reporter products, in an ideal case, the degradation tag will cause degradation of the fused POI in shorter timescales to reduce the probability of having truncated POI for long periods of time in the cytoplasm, which in the case of Nup153 may be enough time to bind at the NPC. Thus, it might be beneficial to incorporate an inducible degradation system which would allow temporal control of protein degradation. It would be beneficial to incorporate a degradation system that reduces drastically the protein half-life. One possible system that would be

orthogonal in mammalian cells to the intracellular degradation system and is inducible is the auxin induced degradation (AID) system (Holland et al., 2012; Nishimura et al., 2009).

The AID system is based on two different components; the degradation sequence formed of the auxin inducible repressor, which is fused to the POI targeted for degradation and the F-box transport inhibitor response protein TIR1. In presence of auxin, the F-box TIR1 will bind to the E3 ubiquitin ligase and the auxin inducible repressor degradation sequence. Then the SFC-TIR1 will recruit the E2 ubiquitin conjugating enzyme, which will polyubiquitinate the auxin inducible repressor sequence leading to the proteosomal degradation of the POI. This degradation system can be used in mammalian cells. Moreover, the AID system has shown a protein half-life for different POI to be in minute timescale (Holland et al., 2012). This system will enable temporal control of the protein degradation which will allow the fine tuning of the time required to achieve the best balance between degradation time and the time required by the FL POI to spatially separate in order to avoid protein degradation.

5.3.2. Making use of the cellular phase separation process to discriminate between truncated and FL POI

I used the droplet forming disordered protein FUS to achieve droplets containing the POI and TEV protease. I measured the dynamics of cytoplasmic FUS droplets containing the POI, by performing fluorescence recovery after photobleaching (FRAP) experiments. However, none of the tested cases had an indication of protein dynamics within the droplet (Figure 4.17). This discourages the current synthetic biology strategy of using cytoplasmic liquid-like droplets as a mechanism to selectively enrich POI containing the ncAA. The FUS droplet forming region has been shown to form gels and aberrant structures over time after the liquid droplets are formed (Patel et al., 2015). Thus, it is crucial to further study the dynamics over time of the liquid droplets to pursue this approach. One way of overcoming this problem could be by testing the droplet dynamics at different time points of other droplet forming domains like EWSR or hnRNP, which like FUS are known components of stress granules involved in protein-RNA interactions. In addition, it would also be interesting to test the potential of liquid droplets formed out of folded multidomain polypeptides, which have been reported to phase separate, like N-WASP and Nck (Banjade and Rosen, 2014), or the ones formed by the multivalent interaction of LAT, Grb2 and Sos1 (Su et al., 2016). This approach may enable to engineer and test the relationship between a different number of multidomains and their effect on the droplet size and dynamics and adjust these parameters to the needs of our selective degradation mechanism.

5.3.3. Alternative discrimination mechanism of truncated POI over FL

A different mechanism to achieve the selective enrichment of FL POI spatially separated from truncated products is by targeting the protein to the plasma membrane. In this case, the protease must also be targeted to the membrane in order to selectively cleave off the membrane targeting domain and the degradation domain of the FL POI. I achieved the membrane localization of the TEV protease by fusing an N-terminal Lck sequence (Figure 4.18). Then I tested if the presence of FUS at the plasma membrane fused to the TEV protease was enough to cause the membrane localization of FL POI via FUS/FUS interaction. However, when FUS is present at the POI it does not necessarily form mixed droplets with Lck-FUS-TEV since the droplets of the POI are formed in the cytoplasm while Lck-FUS-TEV will predominantly be located at the membrane, resulting in an inefficient system. Nevertheless, it seems that when FUS is present in both the protease and the POI construct the protease cleavage was more efficient than when only the FL POI contained the droplet forming domain.

Alternatively, I also targeted the FL POI to the plasma membrane by incorporating the C-terminal membrane targeting domain from KRas. Membrane targeting of the FL POI also seems to achieve a better cleavage of the FL POI when both constructs contain FUS domains (Figure 3.21). However, more experiments need to be done to systematically compare the efficiencies of the different systems.

Nevertheless, the main problem that can be interpreted from the different datasets is that in the presence of TEV protease independenly of the localization (on droplets or at the membrane) there is a very big increase of iRFP only signal (Figure 4.19), indicating that TEV is cleaving the degradation domain out of truncated DD-iRFP proteins too. In the cases where the protease contains the FUS droplet domain, it would be expected that some free FUS-TEV will be at the cytoplasm since the droplets are formed by a concentration gradient of FUS across the cytoplasm. Thus freely diffusing TEV will be able to cleave tre truncated DD out of the DD-iRFP. In addition, when the protease is localized at the membrane by the Lck peptides, since it requires myristoylation and palmitoylation of the Gly and Cys residues to localize at the membrane, potential unmodified cytoplasmic protease will be able to cleave DD-iRFP disrupting the selectivity of the system. Thus, one needs to further improve the efficiency of spatial separation of the protease construct.

An improvement to the system to reduce the potential mislocalized protease activity would be to incorporate a split TEV protease which is only active after the induction of the dimerization of the two split parts. One could target one part of the split TEV to the membrane and let the other part in the cytoplasm, using the FKBP/FRB Rapamycin induced dimerization system fused to each part of the split TEV (Wehr et al., 2006), the protease will only be active upon Rapamycin induction. This enables temporal control of the proteolytic cleavage, which adds an additional tunable tool to the system.

However, inducible proteolysis does not solve the problem of possible cytoplasmic active TEV that may be generated. To really decrease the amount of cytoplasmic TEV, which relies on posttranslational modifications in order to localize at the membrane, a transmembrane domain could be used (Wehr et al., 2006). Due to the presence of the transmembrane domain the protein will be inserted more stably into the plasma membrane decreasing the cytosolic protease levels.

The selective spatial separation of the FL POI from the truncated POI independently of the droplet forming domains combined with the degradation of the truncated POI will already provide for a system that enriches for the FL POI, enabling the use of Amber suppression technology and click-chemistry to site-specifically study the structure and dynamics of the disordered regions of FG-Nup *in situ*.

Acknowledgements

First of all, I would like to thank my supervisor Dr. Edward Lemke for his support and mentorship during the last years. He has been a great source of inspiration and I am specially thankful for the moments of shared excitement towards promising preliminary results even if afterwards it was just some artifact.

I would also like to thank the members of my thesis advisory committee Prof. Dr. Frauke Melchior, Dr. Christoph Müller and Dr. Jonas Ries. Thanks for the helpful advice throughout these years and the constructive TAC meetings it is really much appreciated. Special thanks to Prof Dr. Frauke Melchior who mentored me during the masters.

Many others, who are not part of the Lemke lab, have contributed to this work that without them would not have been possible. I would like to thank Prof. Dr. Frauke Gräter and Dr. Davide Mercadante for the always friendly and fruitful collaborations we have had in the last years and the inspiring discussions generated by their computational experiments. I am very grateful to Prof. Dr. Jane Clarke and Prof. Dr. Sarah Shammas for welcoming me in their lab and for the training and support that I received about stopped-flow kinetics.

A big thanks goes to all past and present members of the Lemke lab. I am grateful for having been able to work in an environment where people help each other and share each other's failures and successes. I would like to thank Sigrid for showing me introducing me in the world ofsingle molecule spectroscopy. Niccolo, for being so helpful in explaining complex things in a very easy way and for being an endless source of 'basic biological questions' during coffee breaks to which nobody had a definite answer. Swati, for being so caring and late evening discussions. Ivanna, for being always supportive and a source of inspiration. Jun for always keeping a smile and for all the moments that we have shared together and that definitely shaped this journey. Giulia, for triggering many discussions in the corner that always make me think and for being my swimming, bouldering and German partner. Sofya, for always being so positive. Gemma, for always providing helpful tips that make protocols more efficient, GOT nights and the Chitchat moments. Christine, for being such an enormous support when it is needed the most. I would also like to thank Giorgia for standing all the questions about microfluidic devices that I came up with; it is always pleasant to work with such a positive person around. Aritra thanks for your creative suggestions and our useful discussions. Nataliia, for bringing a little bit of chemistry to every group meeting and for making it understandable. Paul, thanks for the patience and guidance in the bouldering hall, Gustavo, for your healthy birthday cakes and the helpful biochemistry tips. Joanna, thanks for contributing to the great atmosphere. Christopher, thanks for all the new cloning strategies that all of a sudden we have all started

to use. Daniel, thanks for being such a great help with all the software we use. And thanks for being such a nice person to work with. Last but not least, I would like to thank Piau Siong (TPS) with whom I had the pleasure to work with in a very exciting project. Thanks for being a mentor, for the scientific discussions and for sharing the excitement and the frustration of trying to make things better.

I am also indebted will all the good people whom I feel privileged to have by my side;.

I'm extremely thankful to my friends Bea, David, Arantxa, Marta et.al thanks for being always there and make time and distance not matter. Jesus, Ane, Ander It has been great to share our PhD stories every now and then, you have been a source of inspiration and I wish you the best in your future. Aitor, Zaloa, Desi I always recharge with optimism every time we have the chance to spend some time together. Matti for your always inspiring discussions and your vision of science, which is always contagious and refreshing. Marc for being there since the masters making me a better scientist and for the always encouraging "we will manage"

I am extremely grateful to the Zistler family for welcoming me as a member of their family and for being so caring and warm. Being able to enjoy many family moments with them and the Gengler family has been a privilege these years abroad.

I am wholeheartedly thankful to my family my parents Lourdes Aramburu y Ciriaco Valle. Gracias por apoyarme siempre y motivarme para que hiciera lo que más me gustara. No sería como soy hoy si no fuera por vosotros. También quiero agradecer a mi hermano Jon Valle su apoyo especialmente en momentos duros donde con frases cortas siempre me hace pensar.

Katharina, special thanks for being how you are, for being always there and for the adventures that we still have ahead.

Bibliography

Adams, V.H., McBryant, S.J., Wade, P.A., Woodcock, C.L., and Hansen, J.C. (2007). Intrinsic disorder and autonomous domain function in the multifunctional nuclear protein, MeCP2. J Biol Chem *282*, 15057-15064.

Andrade, M.A., Petosa, C., O'Donoghue, S.I., Muller, C.W., and Bork, P. (2001). Comparison of ARM and HEAT protein repeats. J Mol Biol *309*, 1-18.

Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. Science *181*, 223-230.

Anton, O., Batista, A., Millan, J., Andres-Delgado, L., Puertollano, R., Correas, I., and Alonso, M.A. (2008). An essential role for the MAL protein in targeting Lck to the plasma membrane of human T lymphocytes. The Journal of experimental medicine *205*, 3201-3213.

Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: organizers of cellular biochemistry. Nature Reviews Molecular Cell Biology *18*, 285-298.

Banaszynski, L.A., Chen, L.C., Maynard-Smith, L.A., Ooi, A.G., and Wandless, T.J. (2006). A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. Cell *126*, 995-1004.

Banjade, S., and Rosen, M.K. (2014). Phase transitions of multivalent proteins can promote clustering of membrane receptors. Elife *3*.

Bayliss, R., Kent, H.M., Corbett, A.H., and Stewart, M. (2000a). Crystallization and initial X-ray diffraction characterization of complexes of FxFG nucleoporin repeats with nuclear transport factors. Journal of structural biology *131*, 240-247.

Bayliss, R., Leung, S.W., Baker, R.P., Quimby, B.B., Corbett, A.H., and Stewart, M. (2002a). Structural basis for the interaction between NTF2 and nucleoporin FxFG repeats. The EMBO journal *21*, 2843-2853.

Bayliss, R., Littlewood, T., and Stewart, M. (2000b). Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. Cell *102*, 99-108.

Bayliss, R., Littlewood, T., Strawn, L.A., Wente, S.R., and Stewart, M. (2002b). GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta. J Biol Chem *277*, 50597-50606.

Bayliss, R., Ribbeck, K., Akin, D., Kent, H.M., Feldherr, C.M., Gorlich, D., and Stewart, M. (1999). Interaction between NTF2 and xFxFG-containing nucleoporins is required to mediate nuclear import of RanGDP. J Mol Biol *293*, 579-593.

Beck, M., and Hurt, E. (2017). The nuclear pore complex: understanding its function through structural insight. Nature reviews Molecular cell biology *18*, 73-89.

Bednenko, J., Cingolani, G., and Gerace, L. (2003). Importin beta contains a COOH-terminal nucleoporin binding region important for nuclear transport. The Journal of cell biology *162*, 391-401.

Ben-Efraim, I., and Gerace, L. (2001). Gradient of increasing affinity of importin beta for nucleoporins along the pathway of nuclear import. The Journal of cell biology *152*, 411-417.

Billeter, M., Qian, Y.Q., Otting, G., Muller, M., Gehring, W., and Wuthrich, K. (1993). Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex. J Mol Biol *234*, 1084-1093.

Blake, C.C., Koenig, D.F., Mair, G.A., North, A.C., Phillips, D.C., and Sarma, V.R. (1965). Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. Nature *206*, 757-761.

Boehr, D.D., Nussinov, R., and Wright, P.E. (2009). The role of dynamic conformational ensembles in biomolecular recognition. Nature chemical biology *5*, 789-796.

Boesch, C., Bundi, A., Oppliger, M., and Wuthrich, K. (1978). 1H nuclear-magnetic-resonance studies of the molecular conformation of monomeric glucagon in aqueous solution. European journal of biochemistry *91*, 209-214.

Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C., Gharakhani, J., Julicher, F., and Hyman, A.A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science *324*, 1729-1732.

Brangwynne, C.P., Mitchison, T.J., and Hyman, A.A. (2011). Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. Proc Natl Acad Sci U S A *108*, 4334-4339.

Buljan, M., Chalancon, G., Dunker, A.K., Bateman, A., Balaji, S., Fuxreiter, M., and Babu, M.M. (2013). Alternative splicing of intrinsically disordered regions and rewiring of protein interactions. Curr Opin Struct Biol *23*, 443-450.

Bullock, T.L., Clarkson, W.D., Kent, H.M., and Stewart, M. (1996). The 1.6 angstroms resolution crystal structure of nuclear transport factor 2 (NTF2). J Mol Biol *260*, 422-431.

Burke, K.A., Janke, A.M., Rhine, C.L., and Fawzi, N.L. (2015). Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. Mol Cell *60*, 231-241.

Cansizoglu, A.E., Lee, B.J., Zhang, Z.C., Fontoura, B.M., and Chook, Y.M. (2007). Structure-based design of a pathway-specific nuclear import inhibitor. Nature structural & molecular biology *14*, 452-454.

Chook, Y.M., and Blobel, G. (1999). Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp. Nature *399*, 230-237.

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V., and Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science *325*, 834-840.

Cingolani, G., Petosa, C., Weis, K., and Muller, C.W. (1999). Structure of importin-beta bound to the IBB domain of importin-alpha. Nature *399*, 221-229.

Clerici, M., Mourao, A., Gutsche, I., Gehring, N.H., Hentze, M.W., Kulozik, A., Kadlec, J., Sattler, M., and Cusack, S. (2009). Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. The EMBO journal *28*, 2293-2306.

Colak, R., Kim, T., Michaut, M., Sun, M., Irimia, M., Bellay, J., Myers, C.L., Blencowe, B.J., and Kim, P.M. (2013). Distinct types of disorder in the human proteome: functional implications for alternative splicing. PLoS computational biology *9*, e1003030.

Conti, E., Muller, C.W., and Stewart, M. (2006). Karyopherin flexibility in nucleocytoplasmic transport. Curr Opin Struct Biol *16*, 237-244.

Conti, E., Uy, M., Leighton, L., Blobel, G., and Kuriyan, J. (1998). Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. Cell *94*, 193-204.

Cushman, I., Palzkill, T., and Moore, M.S. (2006). Using peptide arrays to define nuclear carrier binding sites on nucleoporins. Methods *39*, 329-341.

Das, R.K., and Pappu, R.V. (2013). Conformations of intrinsically disordered proteins are in fl uenced by linear sequence distributions of oppositely charged residues. Proceedings of the National Academy of Sciences *110*, 13392-13397.

Denning, D.P., Patel, S.S., Uversky, V., Fink, A.L., and Rexach, M. (2003). Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. Proc Natl Acad Sci U S A *100*, 2450-2455.

Denning, D.P., and Rexach, M.F. (2007). Rapid evolution exposes the boundaries of domain structure and function in natively unfolded FG nucleoporins. Mol Cell Proteomics *6*, 272-282.

Diella, F., Haslam, N., Chica, C., Budd, A., Michael, S., Brown, N.P., Trave, G., and Gibson, T.J. (2008). Understanding eukaryotic linear motifs and their role in cell signaling and regulation. Frontiers in bioscience : a journal and virtual library *13*, 6580-6603.

Dimura, M., Peulen, T.O., Hanke, C.A., Prakash, A., Gohlke, H., and Seidel, C.A. (2016). Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. Curr Opin Struct Biol *40*, 163-185.

Dinkel, H., Van Roey, K., Michael, S., Davey, N.E., Weatheritt, R.J., Born, D., Speck, T., Kruger, D., Grebnev, G., Kuban, M., *et al.* (2014). The eukaryotic linear motif resource ELM: 10 years and counting. Nucleic acids research *42*, D259-266.

Dunker, A.K., Babu, M.M., Barbar, E., Blackledge, M., Bondos, S.E., Dosztanyi, Z., Dyson, H.J., Forman-Kay, J., Fuxreiter, M., Gsponer, J., *et al.* (2013). What's in a name? Why these proteins are intrinsically disordered: Why these proteins are intrinsically disordered. Intrinsically disordered proteins *1*, e24157.

Dunker, a.K., Brown, C.J., and Obradovic, Z. (2002). Identification and functions of usefully disordered proteins. Advances in Protein Chemistry *62*, 25-49.

Dunker, A.K., Garner, E., Guilliot, S., Romero, P., Albrecht, K., Hart, J., Obradovic, Z., Kissinger, C., and Villafranca, J.E. (1998). Protein disorder and the evolution of molecular recognition: theory, predictions and observations. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing, 473-484.

Dunker, a.K., Obradovic, Z., Romero, P., Garner, E.C., and Brown, C.J. (2000). Intrinsic protein disorder in complete genomes. Genome informatics Workshop on Genome Informatics *11*, 161-171.

Dyson, H.J., and Wright, P.E. (2002). Coupling of folding and binding for unstructured proteins. Current Opinion in Structural Biology *12*, 54-60.

Ellis, J.D., Barrios-Rodiles, M., Colak, R., Irimia, M., Kim, T., Calarco, J.A., Wang, X., Pan, Q., O'Hanlon, D., Kim, P.M., *et al.* (2012). Tissue-specific alternative splicing remodels protein-protein interaction networks. Mol Cell *46*, 884-892.

Enderlein, J., Goodwin, P.M., VanOrden, A., Ambrose, W.P., Erdmann, R., and Keller, R.A. (1997). A maximum likelihood estimator to distinguish single molecules by their fluorescence decays. Chem Phys Lett *270*, 464-470.

Englmeier, L., Fornerod, M., Bischoff, F.R., Petosa, C., Mattaj, I.W., and Kutay, U. (2001). RanBP3 influences interactions between CRM1 and its nuclear protein export substrates. EMBO Rep *2*, 926-932.

Favreau, C., Worman, H.J., Wozniak, R.W., Frappier, T., and Courvalin, J.C. (1996). Cell cycle-dependent phosphorylation of nucleoporins and nuclear pore membrane protein Gp210. Biochemistry *35*, 8035-8044.

Feric, M., Vaidya, N., Harmon, T.S., Mitrea, D.M., Zhu, L., Richardson, T.M., Kriwacki, R.W., Pappu, R.V., and Brangwynne, C.P. (2016). Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell *165*, 1686-1697.

Finlay, D.R., and Forbes, D.J. (1990). Reconstitution of biochemically altered nuclear pores: transport can be eliminated and restored. Cell *60*, 17-29.

Frey, S., and Gorlich, D. (2009). FG/FxFG as well as GLFG repeats form a selective permeability barrier with self-healing properties. The EMBO journal *28*, 2554-2567.

Fukuhara, N., Fernandez, E., Ebert, J., Conti, E., and Svergun, D. (2004). Conformational variability of nucleo-cytoplasmic transport factors. The Journal of biological chemistry *279*, 2176-2181.

Fuxreiter, M., and Tompa, P. (2009). Fuzzy interactome: the limitations of models in molecular biology. Trends in Biochemical Sciences *34*, 3-3.

Fuxreiter, M., Tompa, P., and Simon, I. (2007). Local structural disorder imparts plasticity on linear motifs. Bioinformatics *23*, 950-956.

Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. Nature *425*, 737-741.

Gilchrist, D., Mykytka, B., and Rexach, M. (2002). Accelerating the rate of disassembly of karyopherin.cargo complexes. J Biol Chem *277*, 18161-18172.

Goldberg, M.W., and Allen, T.D. (1996). The nuclear pore complex and lamina: three-dimensional structures and interactions determined by field emission in-lens scanning electron microscopy. J Mol Biol *257*, 848-865.

Golebiowski, F., Matic, I., Tatham, M.H., Cole, C., Yin, Y., Nakamura, A., Cox, J., Barton, G.J., Mann, M., and Hay, R.T. (2009). System-wide changes to SUMO modifications in response to heat shock. Science signaling *2*, ra24.

Goodrich, J.A., and Kugel, J.F. (2007). Binding and kinetics for molecular biologists (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press).

Gorlich, D., Pante, N., Kutay, U., Aebi, U., and Bischoff, F.R. (1996). Identification of different roles for RanGDP and RanGTP in nuclear protein import. The EMBO journal *15*, 5584-5594.

Graham, T.A., Ferkey, D.M., Mao, F., Kimelman, D., and Xu, W. (2001). Tcf4 can specifically recognize beta-catenin using alternative conformations. Nature structural biology *8*, 1048-1052.

Hallberg, E., Wozniak, R.W., and Blobel, G. (1993). An integral membrane protein of the pore membrane domain of the nuclear envelope contains a nucleoporin-like region. The Journal of cell biology *122*, 513-521.

Hancock, J.F., Cadwallader, K., Paterson, H., and Marshall, C.J. (1991). A CAAX or a CAAL motif and a second signal are sufficient for plasma membrane targeting of ras proteins. The EMBO journal *10*, 4033-4039.

Holland, A.J., Fachinetti, D., Han, J.S., and Cleveland, D.W. (2012). Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. Proc Natl Acad Sci U S A *109*, E3350-3357.

Hough, L.E., Dutta, K., Sparks, S., Temel, D.B., Kamal, A., Tetenbaum-Novatt, J., Rout, M.P., and Cowburn, D. (2015). The molecular mechanism of nuclear transport revealed by atomic-scale measurements. eLife *4*, 1-23.

Hu, T., Guan, T., and Gerace, L. (1996). Molecular and functional characterization of the p62 complex, an assembly of nuclear pore complex glycoproteins. The Journal of cell biology *134*, 589-601.

Huber, A.H., Nelson, W.J., and Weis, W.I. (1997). Three-dimensional structure of the armadillo repeat region of β- catenin. Cell *90*, 871-882.

Hurley, T.D., Yang, J., Zhang, L., Goodwin, K.D., Zou, Q., Cortese, M., Dunker, A.K., and DePaoli-Roach, A.A. (2007). Structural basis for regulation of protein phosphatase 1 by inhibitor-2. J Biol Chem *282*, 28874-28883.

Hurt, E.C. (1988). A novel nucleoskeletal-like protein located at the nuclear periphery is required for the life cycle of Saccharomyces cerevisiae. The EMBO journal *7*, 4323-4334.

Hutten, S., and Kehlenbach, R.H. (2006). Nup214 is required for CRM1-dependent nuclear protein export in vivo. Molecular and cellular biology *26*, 6772-6785.

Hutten, S., and Kehlenbach, R.H. (2007). CRM1-mediated nuclear export: to the pore and beyond. Trends in cell biology *17*, 193-201.

Isgro, T.A., and Schulten, K. (2005). Binding Dynamics of Isolated Nucleoporin Repeat Regions to Importin-β. Structure *13*, 1869-1879.

Iwamoto, M., Bjorklund, T., Lundberg, C., Kirik, D., and Wandless, T.J. (2010). A general chemical method to regulate protein stability in the mammalian central nervous system. Chemistry & biology *17*, 981-988.

Kalab, P., Weis, K., and Heald, R. (2002). Visualization of a Ran-GTP gradient in interphase and mitotic Xenopus egg extracts. Science (New York, NY) *295*, 2452-2456.

Kalderon, D., Roberts, B.L., Richardson, W.D., and Smith, A.E. (1984). A Short Amino-Acid Sequence Able to Specify Nuclear Location. Cell *39*, 499-509.

Kalinin, S., Valeri, A., Antonik, M., Felekyan, S., and Seidel, C.A. (2010). Detection of structural dynamics by FRET: a photon distribution and fluorescence lifetime analysis of systems with multiple states. The journal of physical chemistry B *114*, 7983-7995.

Kehlenbach, R.H., Dickmanns, A., Kehlenbach, A., Guan, T., and Gerace, L. (1999). A role for RanBP1 in the release of CRM1 from the nuclear pore complex in a terminal step of nuclear export. The Journal of cell biology *145*, 645-657.

Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H., and Phillips, D.C. (1958). A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature *181*, 662-666.

Klebe, C., Bischoff, F.R., Ponstingl, H., and Wittinghofer, A. (1995a). Interaction of the nuclear GTP-binding protein Ran with its regulatory proteins RCC1 and RanGAP1. Biochemistry *34*, 639-647.

Klebe, C., Prinz, H., Wittinghofer, A., and Goody, R.S. (1995b). The kinetic mechanism of Ran--nucleotide exchange catalyzed by RCC1. Biochemistry *34*, 12543-12552.

Konig, I., Zarrine-Afsar, A., Aznauryan, M., Soranno, A., Wunderlich, B., Dingfelder, F., Stuber, J.C., Pluckthun, A., Nettels, D., and Schuler, B. (2015). Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. Nature methods *12*, 773-779.

Kosinski, J., Mosalaganti, S., von Appen, A., Teimer, R., DiGuilio, A.L., Wan, W., Bui, K.H., Hagen, W.J., Briggs, J.A., Glavy, J.S., *et al.* (2016). Molecular architecture of the inner ring scaffold of the human nuclear pore complex. Science *352*, 363-365.

Koyama, M., and Matsuura, Y. (2010). An allosteric mechanism to displace nuclear export cargo from CRM1 and RanGTP by RanBP1. The EMBO journal *29*, 2002-2013.

Koyama, M., Shirai, N., and Matsuura, Y. (2014). Structural insights into how Yrb2p accelerates the assembly of the Xpo1p nuclear export complex. Cell Reports *9*, 983-995.

Kramer, R.H., and Karpen, J.W. (1998). Spanning binding sites on allosteric proteins with polymer-linked ligand dimers. Nature *395*, 710-713.

Kriwacki, R.W., Hengst, L., Tennant, L., Reed, S.I., and Wright, P.E. (1996). Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. Proc Natl Acad Sci U S A *93*, 11504-11509.

Kubitscheck, U., Grünwald, D., Hoekstra, A., Rohleder, D., Kues, T., Siebrasse, J.P., and Peters, R. (2005). Nuclear transport of single molecules. The Journal of cell biology *168*.

Kudryavtsev, V., Sikor, M., Kalinin, S., Mokranjac, D., Seidel, C.A., and Lamb, D.C. (2012). Combining MFD and PIE for accurate single-pair Forster resonance energy transfer measurements. Chemphyschem : a European journal of chemical physics and physical chemistry *13*, 1060-1078.

Labokha, A.A., Gradmann, S., Frey, S., Hü Lsmann, B.B., Urlaub, H., Baldus, M., and Gö Rlich, D. (2013). Systematic analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes. The EMBO journal *32302*, 204-218.

Lakowicz, J.R. (2006). Principles of fluorescence spectroscopy, 3rd edn (New York: Springer).

Laurell, E., Beck, K., Krupina, K., Theerthagiri, G., Bodenmiller, B., Horvath, P., Aebersold, R., Antonin, W., and Kutay, U. (2011). Phosphorylation of Nup98 by multiple kinases is crucial for NPC disassembly during mitotic entry. Cell *144*, 539-550.

Lee, S.J., Matsuura, Y., Liu, S.M., and Stewart, M. (2005). Structural basis for nuclear import complex dissociation by RanGTP. Nature *435*, 693-696.

Lim, R.Y., Huang, B., and Kapinos, L.E. (2015). How to operate a nuclear pore complex by Kap-centric control. Nucleus *6*, 366-372.

Lim, R.Y., Huang, N.P., Koser, J., Deng, J., Lau, K.H., Schwarz-Herion, K., Fahrenkrog, B., and Aebi, U. (2006). Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport. Proc Natl Acad Sci U S A *103*, 9512-9517.

Lott, K., Bhardwaj, A., Mitrousis, G., Pante, N., and Cingolani, G. (2010). The importin beta binding domain modulates the avidity of importin beta for the nuclear pore complex. J Biol Chem *285*, 13769-13780.

Lubas, W.A., Smith, M., Starr, C.M., and Hanover, J.A. (1995). Analysis of nuclear pore protein p62 glycosylation. Biochemistry *34*, 1686-1694.

Mao, A.H., Crick, S.L., Vitalis, A., Chicoine, C.L., and Pappu, R.V. (2010). Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. Proceedings of the National Academy of Sciences of the United States of America *107*, 8183-8188.

Marlow, M.S., Dogan, J., Frederick, K.K., Valentine, K.G., and Wand, A.J. (2010). The role of conformational entropy in molecular recognition by calmodulin. Nature chemical biology *6*, 352-358.

Marsh, J.A., Dancheck, B., Ragusa, M.J., Allaire, M., Forman-Kay, J.D., and Peti, W. (2010). Structural diversity in free and bound states of intrinsically disordered protein phosphatase 1 regulators. Structure *18*, 1094-1103.

Matsuura, Y., and Stewart, M. (2005). Nup50/Npap60 function in nuclear protein import complex disassembly and importin recycling. The EMBO journal *24*, 3681-3689.

Miller, M.L., Jensen, L.J., Diella, F., Jorgensen, C., Tinti, M., Li, L., Hsiung, M., Parker, S.A., Bordeaux, J., Sicheritz-Ponten, T., *et al.* (2008). Linear motif atlas for phosphorylation-dependent signaling. Science signaling *1*, ra2.

Miller, M.W., Caracciolo, M.R., Berlin, W.K., and Hanover, J.A. (1999). Phosphorylation and glycosylation of nucleoporins. Archives of biochemistry and biophysics *367*, 51-60.

Milles, S., Huy Bui, K., Koehler, C., Eltsov, M., Beck, M., and Lemke, E.a. (2013). Facilitated aggregation of FG nucleoporins under molecular crowding conditions. EMBO reports *14*, 178-183.

Milles, S., and Lemke, E.A. (2014). Mapping Multivalency and Differential Affinities within Large Intrinsically Disordered Protein Complexes with Segmental Motion Analysis. Angewandte Chemie International Edition *53*, 7364-7367.

Milles, S., Mercadante, D., Aramburu, I.V., Jensen, M.R.b., Banterle, N., Koehler, C., Tyagi, S., Clarke, J., Shammas, S.L., Blackledge, M., *et al.* (2015). Plasticity of an Ultrafast Interaction between Nucleoporins and Nuclear Transport Receptors. Cell *163*, 734-745.

Miyazaki, Y., Imoto, H., Chen, L.C., and Wandless, T.J. (2012). Destabilizing domains derived from the human estrogen receptor. Journal of the American Chemical Society *134*, 3942-3945.

Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., and Taylor, J.P. (2015). Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell *163*, 123-133.

Monecke, T., Guttler, T., Neumann, P., Dickmanns, A., Gorlich, D., and Ficner, R. (2009). Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. Science *324*, 1087-1091.

Monecke, T., Haselbach, D., Voss, B., Russek, A., Neumann, P., Thomson, E., Hurt, E., Zachariae, U., Stark, H., Grubmuller, H., *et al.* (2013). Structural basis for cooperativity of CRM1 export complex formation. Proc Natl Acad Sci U S A *110*, 960-965.

Moroianu, J. (1998). Distinct nuclear import and export pathways mediated by members of the karyopherin beta family. Journal of cellular biochemistry *70*, 231-239.

Moroianu, J., Hijikata, M., Blobel, G., and Radu, A. (1995). Mammalian karyopherin alpha 1 beta and alpha 2 beta heterodimers: alpha 1 or alpha 2 subunit binds nuclear localization signal and beta subunit interacts with peptide repeat-containing nucleoporins. Proc Natl Acad Sci U S A *92*, 6532-6536.

Morrison, J., Yang, J.C., Stewart, M., and Neuhaus, D. (2003). Solution NMR study of the interaction between NTF2 and nucleoporin FxFG repeats. Journal of Molecular Biology *333*, 587-603.

Muller-Spath, S., Soranno, A., Hirschfeld, V., Hofmann, H., Ruegger, S., Reymond, L., Nettels, D., and Schuler, B. (2010). From the Cover: Charge interactions can dominate the dimensions of intrinsically disordered proteins. Proc Natl Acad Sci U S A *107*, 14609-14614.

Musser, S.M., and Grunwald, D. (2016). Deciphering the Structure and Function of Nuclear Pores Using Single-Molecule Fluorescence Approaches. J Mol Biol *428*, 2091-2119.

Navarro, R., Chen, L.C., Rakhit, R., and Wandless, T.J. (2016). A Novel Destabilizing Domain Based on a Small-Molecule Dependent Fluorophore. ACS chemical biology *11*, 2101-2104.

Nikic, I., Plass, T., Schraidt, O., Szymanski, J., Briggs, J.A., Schultz, C., and Lemke, E.A. (2014). Minimal tags for rapid dual-color live-cell labeling and super-resolution microscopy. Angewandte Chemie *53*, 2245-2249.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nature methods *6*, 917-922.

Nolte, R.T., Wisely, G.B., Westin, S., Cobb, J.E., Lambert, M.H., Kurokawa, R., Rosenfeld, M.G., Willson, T.M., Glass, C.K., and Milburn, M.V. (1998). Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. Nature *395*, 137-143.

Noren, C.J., Anthony-Cahill, S.J., Griffith, M.C., and Schultz, P.G. (1989). A general method for site-specific incorporation of unnatural amino acids into proteins. Science *244*, 182-188.

Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., *et al.* (2015). Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. Mol Cell *57*, 936-947.

O'Reilly, A.J., Dacks, J.B., Field, M.C., Brenner, S.E., and Chervitz, S.A. (2011). Evolution of the Karyopherin-β Family of Nucleocytoplasmic Transport Factors; Ancient Origins and Continued Specialization. PLoS ONE *6*, e19308-e19308.

Oldfield, C.J., and Dunker, A.K. (2014). Intrinsically disordered proteins and intrinsically disordered protein regions. Annual review of biochemistry *83*, 553-584.

Oldfield, C.J., Meng, J., Yang, J.Y., Yang, M.Q., Uversky, V.N., and Dunker, A.K. (2008). Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. BMC genomics *9 Suppl 1*, S1.

Ori, A., Banterle, N., Iskar, M., Andres-Pons, A., Escher, C., Khanh Bui, H., Sparks, L., Solis-Mezarino, V., Rinner, O., Bork, P., *et al.* (2013). Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. Molecular systems biology *9*, 648.

Otsuka, S., Iwasaka, S., Yoneda, Y., Takeyasu, K., and Yoshimura, S.H. (2008). Individual binding pockets of importin-beta for FG-nucleoporins have different binding properties and different sensitivities to RanGTP. Proceedings of the National Academy of Sciences of the United States of America *105*, 16101-16106.

Paradise, A., Levin, M.K., Korza, G., and Carson, J.H. (2007). Significant proportions of nuclear transport proteins with reduced intracellular mobilities resolved by fluorescence correlation spectroscopy. J Mol Biol *365*, 50-65.

Parker, D., Rivera, M., Zor, T., Henrion-Caude, A., Radhakrishnan, I., Kumar, A., Shapiro, L.H., Wright, P.E., Montminy, M., and Brindle, P.K. (1999). Role of secondary structure in discrimination between constitutive and inducible activators. Molecular and cellular biology *19*, 5601-5607.

Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M., *et al.* (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell *162*, 1066-1077.

Patel, S.S., Belmont, B.J., Sante, J.M., and Rexach, M.F. (2007). Natively Unfolded Nucleoporins Gate Protein Diffusion across the Nuclear Pore Complex. Cell *129*, 83-96.

Pentony, M.M., Ward, J., and Jones, D.T. (2010). Computational resources for the prediction and analysis of native disorder in proteins. Methods in molecular biology *604*, 369-393.

Perutz, M.F., Rossmann, M.G., Cullis, A.F., Muirhead, H., Will, G., and North, A.C. (1960). Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-A. resolution, obtained by X-ray analysis. Nature *185*, 416-422.

Peters, R. (2009). Translocation through the nuclear pore: Kaps pave the way. Bioessays *31*, 466-477.

Plass, T., Milles, S., Koehler, C., Schultz, C., and Lemke, E.A. (2011). Genetically encoded copper-free click chemistry. Angewandte Chemie *50*, 3878-3881.

Plass, T., Milles, S., Koehler, C., Szymanski, J., Mueller, R., Wiessler, M., Schultz, C., and Lemke, E.A. (2012). Amino acids for Diels-Alder reactions in living cells. Angewandte Chemie *51*, 4166-4170.

Plitzko, J.M., Schuler, B., and Selenko, P. (2017). Structural Biology outside the box-inside the cell. Curr Opin Struct Biol *46*, 110-121.

Pollard, T.D., and De La Cruz, E.M. (2013). Take advantage of time in your experiments: a guide to simple, informative kinetics assays. Molecular biology of the cell *24*, 1103-1110.

Port, S.A., Monecke, T., Dickmanns, A., Spillner, C., Hofele, R., Urlaub, H., Ficner, R., and Kehlenbach, R.H. (2015). Structural and Functional Characterization of CRM1-Nup214 Interactions Reveals Multiple FG-Binding Sites Involved in Nuclear Export. Cell Reports.

Pyhtila, B., and Rexach, M. (2003). A gradient of affinity for the karyopherin Kap95p along the yeast nuclear pore complex. The Journal of biological chemistry *278*, 42699-42709.

Radu, A., Moore, M.S., and Blobel, G. (1995). The peptide repeat domain of nucleoporin Nup98 functions as a docking site in transport across the nuclear pore complex. Cell *81*, 215-222.

Raveh, B., Karp, J.M., Sparks, S., Dutta, K., Rout, M.P., Sali, A., and Cowburn, D. (2016). Slide-and-exchange mechanism for rapid and selective transport through the nuclear pore complex. Proc Natl Acad Sci U S A *113*, E2489-2497.

Rexach, M., and Blobel, G. (1995). Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. Cell *83*, 683-692.

Ribbeck, K., and Gorlich, D. (2002). The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. The EMBO journal *21*, 2664-2671.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. Trends in genetics : TIG *16*, 276-277.

Ritterhoff, T., Das, H., Hofhaus, G., Schröder, R.R., Flotho, A., and Melchior, F. (2016). The RanBP2/RanGAP1*SUMO1/Ubc9 SUMO E3 ligase is a disassembly machine for Crm1-dependent nuclear export complexes. Nature Communications *7*, 11482-11482.

Rout, M.P., Aitchison, J.D., Magnasco, M.O., and Chait, B.T. (2003). Virtual gating and nuclear transport: the hole picture. Trends in cell biology *13*, 622-628.

Rout, M.P., and Blobel, G. (1993). Isolation of the yeast nuclear pore complex. The Journal of cell biology *123*, 771-783.

Saha, S., Weber, C.A., Nousch, M., Adame-Arana, O., Hoege, C., Hein, M.Y., Osborne-Nishimura, E., Mahamid, J., Jahnel, M., Jawerth, L., *et al.* (2016). Polar Positioning of Phase-Separated Liquid Compartments in Cells Regulated by an mRNA Competition Mechanism. Cell *166*, 1572-1584 e1516.

Schaffer, J., Volkmer, A., Eggeling, C., Subramaniam, V., Striker, G., and Seidel, C.A.M. (1999). Identification of single molecules in aqueous solution by time-resolved fluorescence anisotropy. J Phys Chem A *103*, 331-336.

Schmidt, H.B., and Görlich, D. (2016). Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles. Trends in Biochemical Sciences *41*, 46-61.

Schmidt, H.B., Görlich, D., Kohler, A., Bradatsch, B., Bassler, J., Hurt, E., Reza, R., Acheson, J., Krishnan, V.V., Newsam, S., *et al.* (2015). Nup98 FG domains from diverse species spontaneously phase-separate into particles with nuclear pore-like permselectivity. eLife *4*, 51-62.

Schreiber, G., and Fersht, A.R. (1996). Rapid, electrostatically assisted association of proteins. Nature structural biology *3*, 427-431.

Schuler, B., Muller-Spath, S., Soranno, A., and Nettels, D. (2012). Application of confocal single-molecule FRET to intrinsically disordered proteins. Methods in molecular biology *896*, 21-45.

Seedorf, M., Damelin, M., Kahana, J., Taura, T., and Silver, P.A. (1999). Interactions between a nuclear transporter and a subset of nuclear pore complex proteins depend on Ran GTPase. Molecular and cellular biology *19*, 1547-1557.

Serdyuk, I.N., Zaccai, N.R., and Zaccai, G. (2007). Methods in molecular biophysics : structure, dynamics, function (Cambridge ; New York: Cambridge University Press).

Shammas, S.L., Travis, A.J., and Clarke, J. (2013). Remarkably fast coupled folding and binding of the intrinsically disordered transactivation domain of cMyb to CBP KIX. Journal of Physical Chemistry B *117*, 13346-13356.

Shammas, S.L., Travis, A.J., and Clarke, J. (2014). Allostery within a transcription coactivator is predominantly mediated through dissociation rate constants. Proc Natl Acad Sci U S A *111*, 12055-12060.

Shamsher, M.K., Ploski, J., and Radu, A. (2002). Karyopherin beta 2B participates in mRNA export from the nucleus. Proc Natl Acad Sci U S A *99*, 14195-14199.

Sharma, R., Raduly, Z., Miskei, M., and Fuxreiter, M. (2015). Fuzzy complexes: Specific binding without complete folding. FEBS Letters.

Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P.J., and Seidel, C.A. (2010a). Accurate single-molecule FRET studies using multiparameter fluorescence detection. Methods in enzymology *475*, 455-514.

Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P.J., and Seidel, C.A.M. (2010b). Accurate Single-Molecule Fret Studies Using Multiparameter Fluorescence Detection. Method Enzymol *475*, 455-514.

Srinivasan, G., James, C.M., and Krzycki, J.A. (2002). Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. Science *296*, 1459-1462.

Starr, C.M., and Hanover, J.A. (1990). Glycosylation of nuclear pore protein p62. Reticulocyte lysate catalyzes O-linked N-acetylglucosamine addition in vitro. J Biol Chem *265*, 6868-6873.

Stewart, M. (2000). Insights into the molecular mechanism of nuclear trafficking using nuclear transport factor 2 (NTF2). Cell structure and function *25*, 217-225.

Strawn, L.A., Shen, T., Shulga, N., Goldfarb, D.S., and Wente, S.R. (2004). Minimal nuclear pore complexes define FG repeat domains essential for transport. Nature cell biology *6*, 197-206.

Su, X., Ditlev, J.A., Hui, E., Xing, W., Banjade, S., Okrut, J., King, D.S., Taunton, J., Rosen, M.K., and Vale, R.D. (2016). Phase separation of signaling molecules promotes T cell receptor signal transduction. Science *352*, 595-599.

Sun, C., Fu, G., Ciziene, D., Stewart, M., and Musser, S.M. (2013). Choreography of importin-alpha/CAS complex assembly and disassembly at nuclear pores. Proc Natl Acad Sci U S A *110*, E1584-1593.

Szymborska, A., de Marco, A., Daigle, N., Cordes, V.C., Briggs, J.A., and Ellenberg, J. (2013). Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging. Science *341*, 655-658.

Tetenbaum-Novatt, J., Hough, L.E., Mironska, R., McKenney, A.S., and Rout, M.P. (2012). Nucleocytoplasmic transport: a role for nonspecific competition in karyopherin-nucleoporin interactions. Mol Cell Proteomics *11*, 31-46.

Theillet, F.X., Binolfi, A., Bekei, B., Martorana, A., Rose, H.M., Stuiver, M., Verzini, S., Lorenz, D., van Rossum, M., Goldfarb, D., *et al.* (2016). Structural disorder of monomeric alpha-synuclein persists in mammalian cells. Nature *530*, 45-50.

Tompa, P., and Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions.

Torchia, J., Rose, D.W., Inostroza, J., Kamei, Y., Westin, S., Glass, C.K., and Rosenfeld, M.G. (1997). The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. Nature *387*, 677-684.

Tu, L.-C., Fu, G., Zilman, A., and Musser, S.M. (2013). Large cargo transport by nuclear pores: implications for the spatial organization of FG-nucleoporins. The EMBO journal *32*, 3220-3230.

Tzeng, S.R., and Kalodimos, C.G. (2012). Protein activity regulation by conformational entropy. Nature *488*, 236-240.

Uversky, V.N. (2013). Unusual biophysics of intrinsically disordered proteins. Biochimica et biophysica acta *1834*, 932-951.

Uversky, V.N., Gillespie, J.R., and Fink, A.L. (2000). Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins *41*, 415-427.

van der Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D.T., *et al.* (2014). Classification of intrinsically disordered regions and proteins. Chem Rev *114*, 6589-6631.

Vetter, I.R., Arndt, A., Kutay, U., Gorlich, D., and Wittinghofer, A. (1999). Structural view of the Ran-Importin beta interaction at 2.3 A resolution. Cell *97*, 635-646.

von Appen, A., Kosinski, J., Sparks, L., Ori, A., DiGuilio, A.L., Vollmer, B., Mackmull, M.T., Banterle, N., Parca, L., Kastritis, P., *et al.* (2015). In situ structural analysis of the human nuclear pore complex. Nature *526*, 140-143.

Vuzman, D., Hoffman, Y., and Levy, Y. (2012). Modulating protein-DNA interactions by post-translational modifications at disordered regions. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing, 188-199.

Vuzman, D., and Levy, Y. (2012). Intrinsically disordered regions as affinity tuners in protein-DNA interactions. Molecular bioSystems *8*, 47-57.

Walde, S., and Kehlenbach, R.H. (2010). The Part and the Whole: functions of nucleoporins in nucleocytoplasmic transport. Trends in cell biology *20*, 461-469.

Wang, L., Xie, J., and Schultz, P.G. (2006). Expanding the genetic code. Annual review of biophysics and biomolecular structure *35*, 225-249.

Wang, Z., Udeshi, N.D., Slawson, C., Compton, P.D., Sakabe, K., Cheung, W.D., Shabanowitz, J., Hunt, D.F., and Hart, G.W. (2010). Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates cytokinesis. Science signaling *3*, ra2.

Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F., and Jones, D.T. (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J Mol Biol *337*, 635-645.

Wehr, M.C., Laage, R., Bolz, U., Fischer, T.M., Grunewald, S., Scheek, S., Bach, A., Nave, K.A., and Rossner, M.J. (2006). Monitoring regulated protein-protein interactions using split TEV. Nature methods *3*, 985-993.

Wente, S.R., and Blobel, G. (1994). NUP145 encodes a novel yeast glycine-leucine-phenylalanine-glycine (GLFG) nucleoporin required for nuclear envelope structure. The Journal of cell biology *125*, 955-969.

Wente, S.R., Rout, M.P., and Blobel, G. (1992). A new family of yeast nuclear pore complex proteins. The Journal of cell biology *119*, 705-723.

Wright, P.E., and Dyson, H.J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J Mol Biol *293*, 321-331.

Wright, P.E., and Dyson, H.J. (2009). Linking folding and binding. Current Opinion in Structural Biology *19*, 31-38.

Wright, P.E., and Jane Dyson, H. (2015). Intrinsically disordered proteins in cellular signalling and regulation. Nature Publishing Group *16*.

Xie, Q., Arnold, G.E., Romero, P., Obradovic, Z., Garner, E., and Dunker, A.K. (1998). The Sequence Attribute Method for Determining Relationships Between Sequence and Protein Disorder. Genome Inform Ser Workshop Genome Inform *9*, 193-200.

Yamada, J., Phillips, J.L., Patel, S., Goldfien, G., Calestagne-Morelli, A., Huang, H., Reza, R., Acheson, J., Krishnan, V.V., Newsam, S., *et al.* (2010). A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins. Mol Cell Proteomics *9*, 2205-2224.

Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., and Yokoyama, S. (2008). Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode N(epsilon)-(o-azidobenzyloxycarbonyl) lysine for site-specific protein modification. Chemistry & biology *15*, 1187-1197.

Yang, C.H., Kuo, W.T., Chuang, Y.T., Chen, C.Y., and Lin, C.C. (2013). Cyclin B1 destruction box-mediated protein instability: the enhanced sensitivity of fluorescent-protein-based reporter gene system. BioMed research international *2013*, 732307.

Yang, W., Gelles, J., and Musser, S.M. (2004). Imaging of single-molecule translocation through nuclear pore complexes. Proceedings of the National Academy of Sciences of the United States of America *101*, 12887-12892.

Yaseen, N.R., and Blobel, G. (1999). GTP hydrolysis links initiation and termination of nuclear import on the nucleoporin nup358. J Biol Chem *274*, 26493-26502.

Yoshimura, S.H., and Hirano, T. (2016). HEAT repeats – versatile arrays of amphiphilic helices working in crowded environments? Journal of cell science *129*, 3963-3970.

Yoshimura, Shige H., Kumeta, M., and Takeyasu, K. (2014). Structural Mechanism of Nuclear Transport Mediated by Importin β and Flexible Amphiphilic Proteins. Structure *22*, 1699-1710.

Yudin, D., and Fainzilber, M. (2009). Ran on tracks--cytoplasmic roles for a nuclear regulator. Journal of cell science *122*, 587-593.

Zhou, H.X. (2010). From induced fit to conformational selection: A continuum of binding mechanism controlled by the timescale of conformational transitions. Biophysical Journal *98*, L15-L17.

Zhu, Y., Liu, T.W., Madden, Z., Yuzwa, S.A., Murray, K., Cecioni, S., Zachara, N., and Vocadlo, D.J. (2016). Post-translational O-GlcNAcylation is essential for nuclear pore integrity and maintenance of the pore selectivity filter. Journal of molecular cell biology *8*, 2-16.

Zlatkine, P., Mehul, B., and Magee, A.I. (1997). Retargeting of cytosolic proteins to the plasma membrane by the Lck protein tyrosine kinase dual acylation motif. Journal of cell science *110 (Pt 5)*, 673-679.