A three-dimensional view of the interface between nuclear envelope and chromatin

Camille Samson

université PARIS-SACLAY

ÉCOLE DOCTORALE
Innovation thérapeutique : du fondamental à l'appliqué

Freie Universität Berlin

NNT : 2018SACLS048

Une représentation en trois dimensions de l'interface entre l'enveloppe nucléaire et la chromatine

Thèse de doctorat de l'Université Paris-Saclay et de la Freie Universität de Berlin
Préparée à l'Université Paris-Sud

École doctorale n°569 :
Innovation thérapeutique : du fondamental à l'appliqué (ITFA)
Spécialité de doctorat: biochimie et biologie structurale

Thèse présentée et soutenue au CEA Saclay, 04/04/2018, par :

Camille Samson

Composition du Jury:

Prof Herman Van Tilbeurgh, Université Paris Sud	Président
Prof Christian Freund, FU	Rapporteur
Dr Nicolas Wolff, Institut Pasteur	Rapporteur
Dr Eric Schirmer, Université d'Edimbourg	Rapporteur
Dr Louis Renault, I2BC	Examinateur
Dr Chris Weise, FU	Examinateur
Prof Hartmut Oschkinat, FU	Co-directeur de thèse
Dr Sophie Zinn Justin, I2BC	Directeur de thèse

Thèse de doctorat

ACKNOWLEDGMENTS

The work presented in this thesis was mainly performed at the Laboratory of Structural Biology and Radiobiology (LBSR) of CEA Saclay. First, I would like to thank Dr. Sophie Zinn-Justin, my supervisor, for giving me the opportunity to do this PhD, for all her help and advices throughout my studies, and for all the support. Thank you for your listening, your availability and above all, thank you for your great enthusiasm and your positivity. It was a real pleasure to work with you during these three years. I would like to thank Prof Hartmut Oschkinat for his kindness and his help when I was in Berlin.

Thanks to Drs. Christian Freund, Nicolas Wolff and Eric Schirmer who took the time to read this thesis and have accepted to be part of my thesis committee as reviewers. My thanks also to Drs. Louis Renault, Chris Weise and Prof. Herman Van-Tilbeurgh who have kindly accepted to judge this work.

During my thesis, I had the opportunity to spend 9 months at the FMP (Leibniz-Institut Für Molekulare Pharmakologie) in Berlin in the team of Dr. Philipp Selenko. I thank Philipp for his hospitality and advice as well as the members of his team and more particularly Dr. François-Xavier Theillet for his precious help and availability.

I sincerely thank every people who participated in the results presented in this thesis. Thanks to all the members of the LBSR for their advices, their help but also for all the moments spent together outside the laboratory. Thanks to Danni Liu, Virginie Ropars, Simona Miron, Gaëlle Garet, Ambre Petitalot, Aurélien Pitois, Clément Némoz, Florian Celli, Nada Essawy, Jingqi Dai, Amandine Gontier, Audrey Comte and all the others. You supported me a lot during this thesis and you were always there when I needed it.

A huge thank you to Dr. Brigitte Buendia (Paris-Diderot University) for having realized all the immunofluorescence and PLA experiments in cells. For the electron microscopy work, thanks to Dmytro Puchkov (Core Facility Cellular Imaging du FMP, Berlin) and a big thanks to Ana Arteni (Institut de Biologie Intégrative de la Cellule, Gif-Sur-Yvette) for all these hours spent with me and for having allowed me to obtain these beautiful images of filaments.

I would also like to thank Dr. Christophe Velours, from the LEBS, for carrying out the experiments of UCA. Finally I would like to thank Max Zinke, Kitty Hendriks and Prof. Adam Lange at the FMP Berlin, for having realized and analyzed the experiments of solid-state NMR and for their help and kindness.

Thanks to the group “Flow-Cytech”, especially to Ernesto Marcos, Brice Targat, Julien Lemaitre, André Pozo Rodriguez, Mario Gomez, Lamine Alaoui, Sabrina Guenounou, Pierre Rosembaum, and Nicolas Tchitchek, for all these evenings, weekends and holidays spent together to unwind during this thesis.

And last but not least, I would like to say a big thank you to my boyfriend Adrien, to my parents Irène and Benoit and to my brothers Orso, Elie and Nathan who supported me during these three years, and more, in all the good or even bad times. Thanks in particular to my family for helping me to draft scientific papers.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

ABBREVIATIONS

Aa	Amino Acid
AD-EDMD	Autosomal Dominant-Emery Dreifuss Muscular Dystrophy
APS	Atypical progeria syndromes
AR-EDMD	Autosomal Recessive-Emery Dreifuss Muscular Dystrophy
ATP	Adenosine triphosphate
BAF	Barrier-to-autointegration factor
CEA	French Alternative Energies and Atomic Energy Commission
C. elegans	*Caenorhabditis elegans*
CMC	Cardiomyopathy with conduction defects
CMT	Charcot-Marie-Tooth disease
CTSL	Cathepsin L
DSBs	Double strand breaks
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DTT	Dithiothréitol
E. Coli	*Escherichia coli*
EDMD	Emery-Dreifuss muscular dystrophy
EDTA	Acide éthylène diamine tétra-acétique
EM	Electron Microscopy
ER	Endoplasmic Reticulum
F-actin	Filamentous actin
FBTSA	Fluorescence-based Thermal Shift Assay
FLIP	Fluorescence loss in photobleaching
FMP	Leibniz-institut für molekulare pharmakologie

FPLD	Dunnigan-type familial partial lipodystrophy
FRET	Fluorescence resonance energy transfer
G-actin	Globular actin
GCL	Germ cell less
GF	Gel Filtration
GFP	Green Fluorescent Protein
GST	Glutathione-S-transferase
HDAC3	Histone deacetylase 3
HGPS	Hutchinson-Gilford progeria syndrome
HhH motif	Helix-hairpin-Helix motif
HSQC	Heteronuclear single quantum coherence spectroscopy
Icmt	Isoprenylcysteine carboxyl methyltransferase
IgFold domain	Immunoglobulin-like domain
INM	Inner Nuclear Membrane
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
KASH domain	Klarsicht, ANC-1 and SYNE/Nesprin-1 and 2 Homology domain
LADs	Lamina-associated domains
LAS	Lamina-associated sequences
LB	Lysogeny Broth
LBR	Lamin-B receptor
LBSR	Structural Biology and Radiobiology Laboratory
LEBS	Enzymology and Structural Biochemistry Laboratory
LEM domain	LAP2-Emerin-Man1 domain
LGMD	Limb-girdle muscular dystrophy type 1B

LINC complex	Linker of Nucleoskeleton and Cytoskeleton complex
Lmo7	Lim-domain-only-7
MAD	Mandibuloacral dysplasia
MSC motif	MAN1/Src1p/C-terminal motif
NARF	Nuclear Prelamin A Recognition Factor
NE	Nuclear Envelope
NGPS	Nestor Guillermo Progeria Syndrome
NHEJ	Nonhomologous end-joining
Ni-NTA	Nickel-nitrilotriacetic acid
NLS	Nuclear Localization Signal
NMR	Nuclear Magnetic Resonance
NPC	Nuclear Pore Complex
NRF2	Nuclear Factor (erythroid-derived 2)-like 2
OD	Optical Density
O-GlcNA	O-linked-Nacetylglucosamine
ONM	Outer Nuclear Membrane
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylenglycol
PIC	Pre-integration complex
PKA	Protein Kinase A
PLA	Proximity Ligation Assay
PMSF	Phenylmethyl Sulfonyl Fluoride
PS	Progeroid syndromes
Ran-GTP	RAs-related nuclear protein-Guanosine Triphosphate

Rce1	Ras Converting CAAX Endopeptidase 1
RD	Restrictive dermopathy
RNP	Ribonucleoprotein
ROS	Reactive oxygen species
S.cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodium dodecyl sulfate
SDS-PAGE	Polyacrylamide gel electrophoresis with SDS
Src	Tyrosine kinase Src
SUN domain	Sad1 and UNC-84 domain
TAN lines	Transmembrane actin-associated nuclear lines
TEV	Tobacco etch virus protease
ThT	Thioflavin
TM	Transmembrane domain
Trp	Translocated promoter region protein
UHM motif	U2AF homology motif
WT	Wild-Type
XL-EDMD	X-linked-Emery Dreifuss Muscular Dystrophy
53BP1	P53-binding protein 1

INTRODUCTION

The defining feature of eukaryotic cells is that they have membrane-bound organelles such as mitochondria and the Golgi apparatus (figure 1). The bigger organelle specific to eukaryotic cells is the nucleus, with a diameter of 10 to 20 µm. It has two major functions: it stores cell hereditary material (DNA), and it coordinates cell activities, which include growth, intermediary metabolism, protein synthesis, and cell division. Eukaryotes usually have a single nucleus, but a few cell types, such as mammalian red blood cells, have no nucleus, and a few others, as some muscle cells, have many nuclei[1,2].

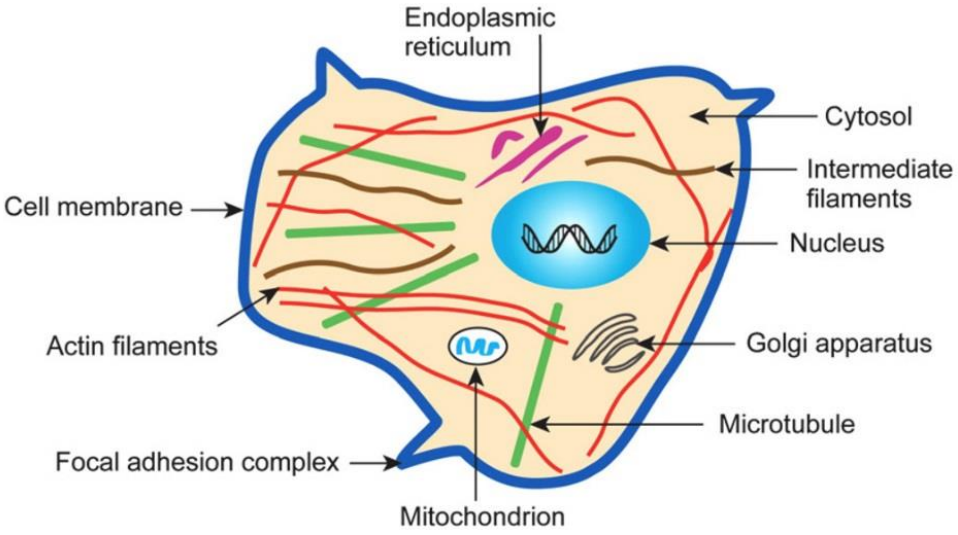


Figure 1 : Scheme of a eukaryotic cell with its different components[3].

A nuclear envelope separates the cytoplasm from the nucleoplasm (figure 2). This envelope contains the outer nuclear membrane, the inner nuclear membrane and a lot of proteins that can be gathered into three groups. The first group contains the nuclear pore complexes, which are large protein complexes crossing the nuclear envelope. They mediate the selective exchange of components, including RNA, ribosomal proteins, signaling molecules and lipids between the cytoplasm and the nucleus. The second group is composed of integral membrane proteins, as for example the LEM domain proteins Emerin, MAN1 and LAP2 (for Lamina-Associated Protein 2)[4]. The last group corresponds to the nuclear lamina, which is a meshwork of lamin filaments underlying the nuclear envelope (figure 2).

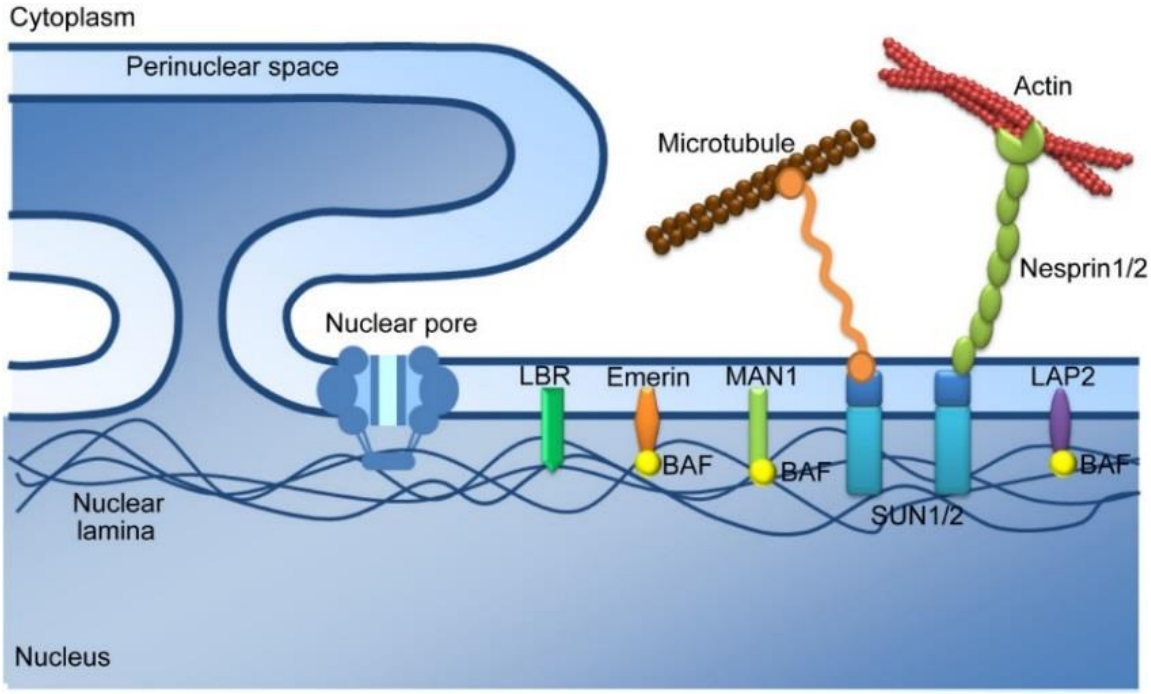


Figure 2 : Architecture of the nuclear envelope[5].

The nuclear envelope (NE) is the focus of an increasing number of studies because mutations in genes encoding NE proteins cause a panel of human diseases called nuclear envelopathies. A large number of NE-associated diseases are caused by inherited mutations in the gene *LMNA* coding for A-type lamins[6]. These laminopathies include Emery-Dreifuss muscular dystrophy, dilated cardiomyopathy with variable muscular dystrophy, Dunnigan-type familial partial lipodystrophy, Charcot-Marie-Tooth type 2 disease, mandibuloacral dysplasia, and the Hutchinson-Gilford progeria syndrome. Research on laminopathies and more generally on nuclear envelopathies has provided novel clues about nuclear envelope function.

I. NUCLEAR ENVELOPE: STRUCTURE

Each component of the nuclear envelope (outer and inner nuclear membranes, nuclear pore complexes, membrane proteins and nuclear lamina) forms distinct domains but all are interconnected. The architecture of these domains is poorly characterized and their structural study remains nowadays challenging.

1. THE NUCLEAR PORE COMPLEX

a. Structure

The vertebrate nuclear pore complex (NPC), which is one of the largest macromolecular complexes in the cell with an estimated molecular mass of 120 kDa, is embedded into the outer and the inner nuclear membranes and mediates macromolecular transport across the NE[7]. In mammalian cells, the number of nuclear pore complexes varies between 4000 and 6000 and they appear like a ring of a 100 nm of diameter[8].

Several 3D structures of these nuclear pore complexes were recently reported. For example, the groups of O. Medalia at University of Zürich and M. Beck at EMBL Heidelberg solved the 3D structure of a human nuclear pore complex, using cryo-electron tomography[9] and molecular modeling combined with cross-linking mass spectrometry and cryo-electron tomography, respectively[10]. The group of A. Hoelz at Caltech obtained the 3D structure of the inner ring complex using X-ray crystallography and cryo-electron tomography[11].

The nuclear pore exhibits an octagonal cylindrical structure. More precisely, it is composed of two rings, cytoplasmic and nucleoplasmic, with 8 spokes to connect both rings, one central region, 16 filaments (8 cytoplasmic and 8 nucleoplasmic) and one terminal ring (figure 3). The 8 nucleoplasmic filaments adopt a basket arrangement (the nuclear basket) stabilized by the presence of the terminal ring.

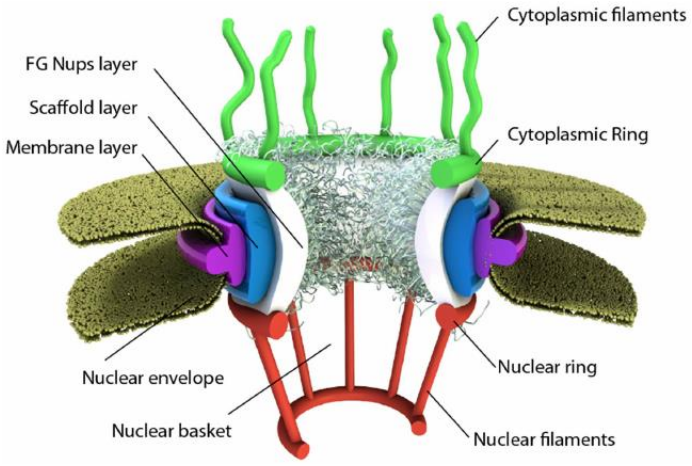


Figure 3 : Model of nuclear pore complex substructures.

The central channel exhibits eight-fold rotational symmetry and has eight cytoplasmic filaments as well as eight nuclear filaments protruding into the cytoplasm and nucleoplasm respectively. The nuclear filaments are bound via a ring, resulting in a basket structure[11].

This protein complex comprises approximatively 50 to 100 proteins termed nucleoporins. All nucleoporins share basic structural components and can be classified in three classes[12]. The first group contains transmembrane α-helix proteins, which mostly constitute the outside of NPC central region, whereas the second group, which contains β-propeller and α-solenoid folds, constitutes the inside of the NPC. The last group is composed of the FG-nucleoporins. The FG-nucleoporins contain a structured domain that serves as an NPC anchor point and a largely unstructured domain composed of phenylalanine and glycine repeat sequences that forms some kind of entropic barrier to diffusion. This group may contribute to the formation of the NPC inner central framework and peripheral structures.

b. Nuclear import/export in cells

Small molecules (less than 40kDa), ions and metabolites are able to diffuse freely inside the nucleus but it is not the case of proteins, various types of RNA and ribonucleoproteins (RNPs), which are actively translocated through the NPCs. These molecules can be thought of as cargoes[13,14].

Concerning the import, the first step is the recognition of a specific sequence on the cargoes, the nuclear localization signal (NLS), which can be complex but is predominantly a stretch of basic residues. This NLS is recognized by importin-α, which will, after binding to the NLS, become an adapter for importin-β and then, the complex is translocated through the NPC into the nucleus, where importins are dissociated from the complex by Ran-GTP.

Ran (RAs-related Nuclear protein) belongs to the family of small proteins G with a GTPase activity and is able to hydrolyze GTP in GDP when activated by a Ran GTPase activating protein[13].

Afterwards, importins can diffuse to the cytoplasm, where the GTP molecules are hydrolyzed into GDP, which leads to the release of α and β importins then available for a new protein-NLS import cycle (figure 4).

The export complex is formed inside the nucleus based on the recognition of a nuclear export signal (typically a leucine rich sequence) by exportins, in the presence of Ran-GTP (figure 4). After passing through the nuclear pores, exportins release the cargo molecule due to activation of the Ran GTPase activity.

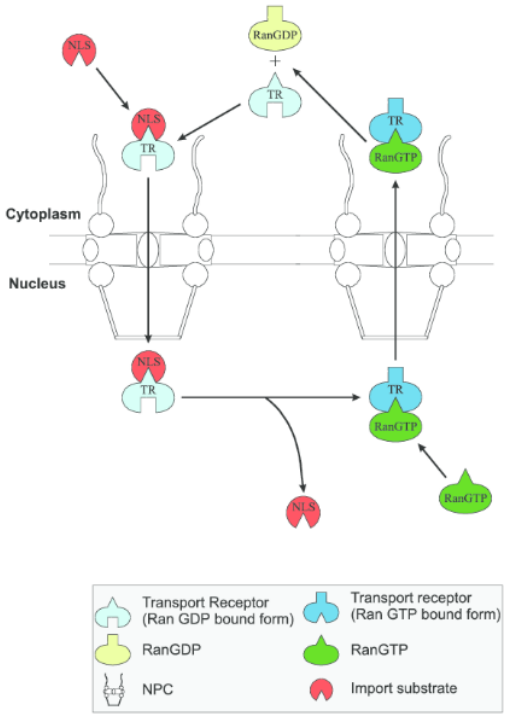


Figure 4 : Import and export through the NPC[13].

c. *Nuclear lamina acts as anchor for NPCs*

An interaction between NPCs and the lamina was first proposed after the observation of filamentous structures associated with the baskets of NPCs in amphibian oocytes[15,16]. Afterwards, it was shown that NPCs are motionless after assembly, and, from the study of nuclear membrane disintegration using several detergents, it was proposed that lamina could be able to cradle the nuclear pores[17].

Recently, using a BioID approach, A-type lamin C (LaC) was find to be a specific partner of NPCs, more than A-type lamin A (LaA) or B-type lamin (LaB)[18]. Indeed, LaC interacts prominently compared to LaA or LaB with the Trp protein (Translocated promoter region protein), a component of the nuclear basket, and with protein Nup214, a FG-Nup of the NPCs. In more details, it was shown that interaction between NPCs and lamins is regulated by the status of the lamin post-IgFold tail and that A-type lamins may contribute directly to NPCs distribution.

2. THE NUCLEAR LAMINA

a. *Expression and localization*

The nuclear lamina, a filamentous protein network underlying the inner nuclear membrane, has emerged as a critical player from basic cell biology to human health[19]. Lamin genes are found in all metazoa examined to date, but are absent in plants and unicellular organisms[20]. Three lamin genes are present in mammalian genomes[21]. On one hand, *LMNB1* and *LMNB2* genes encode lamin B1 and B2 (B-type lamins), proteins expressed at all stages of development and essential for cell survival[22,23]. On the other hand, the *LMNA* gene encodes prelamin A and lamin C (A-type lamins), which are expressed later in organism development. In addition to these 4 major isoforms, 3 minor isoforms are also produced: AΔ10, C2 and B3[24,25].

A-type and B-type lamins show different behaviors and spatial distributions. To observe these different properties, the most evident is to focus on lamin behavior during mitosis. At the onset of prophase, nuclear lamin is phosphorylated in order to be disassembled. A-type lamins are first phosphorylated by Cdk1 and become soluble whereas B-type lamins are phosphorylated by protein kinase C and stay associated to the nuclear envelope (likely due to the presence of the farnesyl moiety) as it is dispersed into the endoplasmic reticulum[26–28]. At this stage, both lamins are distributed throughout the cytoplasm. Afterwards, A and B-type lamins also show different temporal orders of re-assembly at the nuclear envelope. After lamin dephosphorylation by PP1A, B-type lamins bound to nuclear membranes accumulate around chromosomes[29,30]. Although A-type lamins are thought to bind chromatin at a later stage, after the major components of the NE including NPC are assembled[30], live-cell imaging analyses allowed the visualization in early anaphase of a small pool of A-type lamins at specific chromosome regions, together with BAF, emerin and LAP2alpha, from where it extends at later stages[29,31].

b. Maturation

Like for all proteins, the *LMNA* and *LMNB1/2* transcription step takes place inside the nucleus. Then, the produced mRNA is exported to the cytoplasm where it is translated into lamin C, lamin B1, lamin B2 or prelamin A, the precursor form of lamin A, these 3 last lamins all containing a prenylation motif, the CaaX box (figure 5)[28,32]. Lamin C lacks a –CAAX sequence and its carboxyl terminus is not further modified after synthesis, whereas other lamins are modified in order to become mature. First maturation steps are common to lamins A, B1 and B2.

Initiation of the -CaaX sequential processing consists of the addition of a farnesyl group onto the cysteine residue by a farnesyltransferase. This modification leads to the removal of the three terminal amino acids by an endopeptidase. The cleavage of the –aaX motif is done by different enzymes according to the lamin type. Concerning prelamin A, the Zinc metalloprotease ZMPSTE24 is used for this step whereas for lamins B1 and B2, it is a different protease, Rce1 (Ras-converting enzyme 1).

A final common step for all farnesylated lamins is the carboxymethylation of the terminal cysteine by the isoprenylcysteine carboxyl methyltransferase (Icmt). After this step, lamin B1 and lamin B2 have reached their mature states. This is not the case of prelamin A, which possesses a second cleavage site for ZMPSTE24. This last step leads to the removal of 15 amino acids from the C-terminus of lamin A, including the farnesylated/carboxymethylated cysteine residue. The presence of a farnesyl residue attached to B-type lamins suggests that these lamins are more closely associated to the inner nuclear membrane than A-type lamins[33].

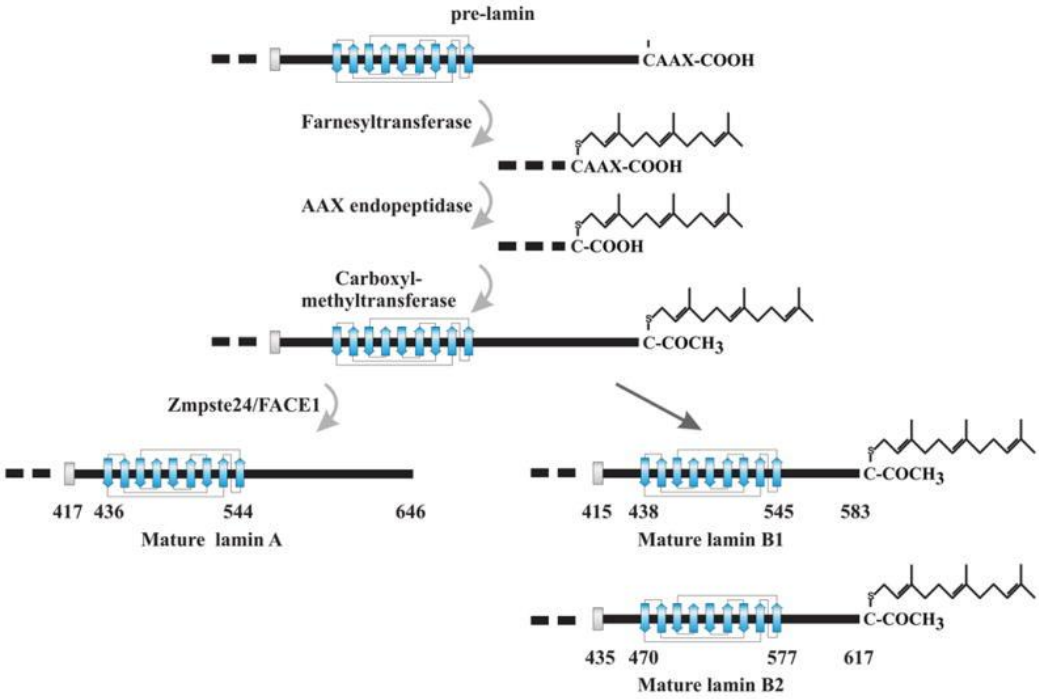


Figure 5 : The maturation process of lamins A, B1 and B2[28]. The common C-terminal CAAX motif is composed of a cysteine (C), followed by two aliphatic amino acids (aa) and a random amino acid (X). Its farnesylation is the first step of the lamin maturation process.

The prelamin A maturation is a complex process, and some genetic mutations driving to failures to generate mature lamin A are responsible for diseases. In particular, it was observed that a mutation modifying the second cleavage site for ZMPSTE24 (Leu647Arg)[34] or mutations modifying the enzyme ZMPSTE24 cause progeroid syndromes in humans[35,36].

A *de novo* dominant splice-donor site mutation within exon 11 (G608G) was observed in several patients[37]. It results in the synthesis of a mutant prelamin A, called progerin (figure 6). This protein is characterized by a deletion of 50 amino acids near the C-terminal tail of prelamin A, but its –CaaX motif is intact. The deletion eliminates the site for the second endoproteolytic cleavage by ZMPSTE24.

The mutated protein can undergo farnesylation, endoproteolytic release of the –aaX, and carboxyl methylation but stays, at the end of the maturation process, in a farnesylated form. It remains associated with the nuclear membrane, eliciting nuclear blebbings and other nuclear abnormalities in cells, including disrupted lamin-heterochromatin interactions and alterations in gene transcription. All these defects cause a known syndrome, the Hutchinson–Gilford progeria syndrome (HGPS), characterized by the rapid appearance of aging at the beginning of childhood[37,38]. Affected children look normal at birth and in early infancy, but then grow more slowly than other children and show different abnormalities as stiffness of joints, cardiovascular disease, atherosclerosis and loss of body fat and hair.

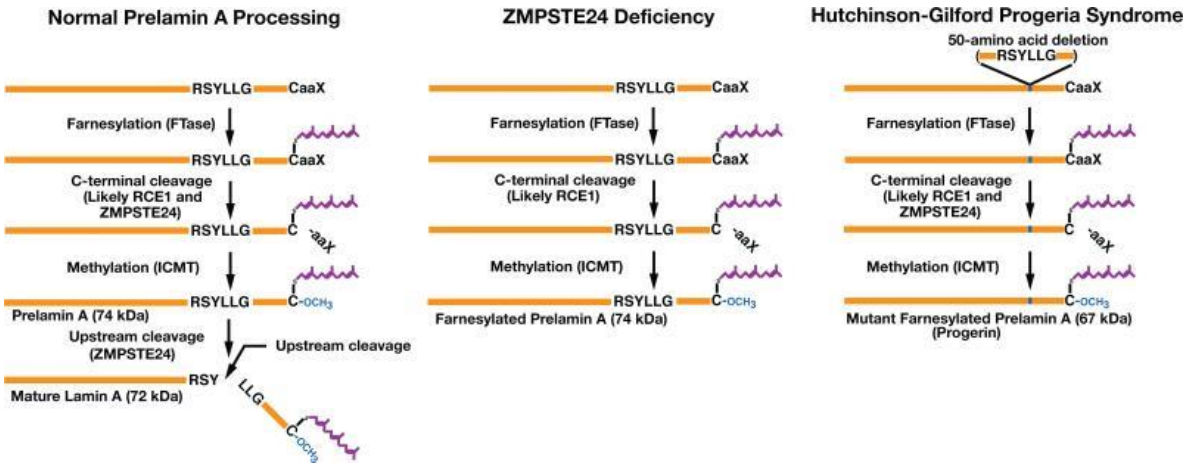


Figure 6 : Biogenesis of lamin A in healthy humans as well as in patients with either a ZMPSTE24 deficiency or the silent mutation causing HPGS[39].

c. Structure

The structure of lamins is composed of three units that are common among intermediate filament proteins. These are a long central-helical rod domain of 40 kDa, an N-terminal (head) domain and a C-terminal (tail) domain (figure 7a)[32,40].

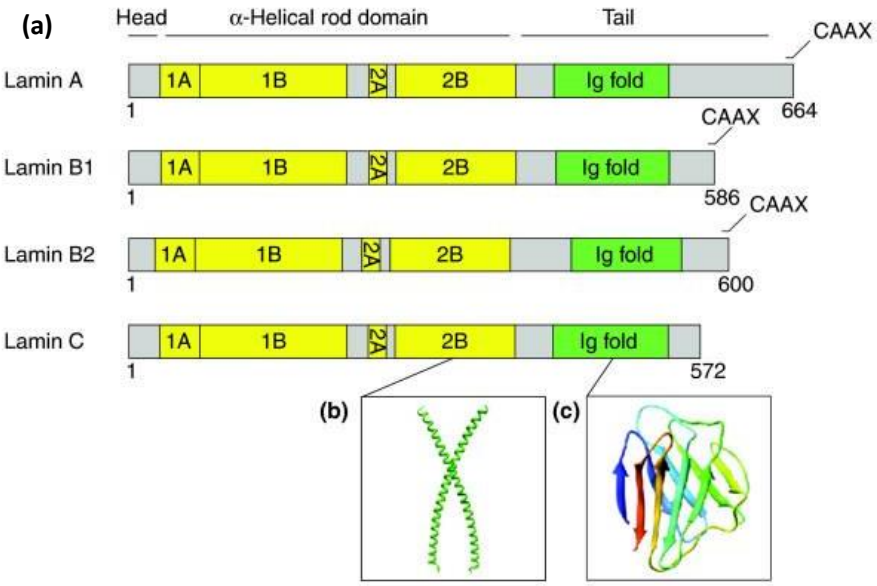


Figure 7: Structural organization of nuclear lamins.

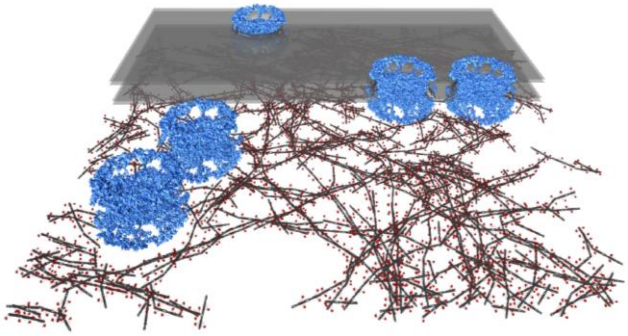
(a) Architecture of major lamin variants. (b) X-ray structure of the human lamin A/C coiled-coil 2 fragment (PDB code: 1X8Y)[41]. (c) NMR structure of the human lamin A/C IgFold domain (PDB code: 1IVT)[42].

The central domain contains α-helical segments, called 1A, 1B, 2A and 2B, showing the heptad repeat periodicity of coiled-coil proteins, and connected by short linkers. A 3D structure of the coiled-coil 2B fragment corresponding to human lamin A was solved by crystallography[41] (figure 7, b). The amino-terminal head domain is variable in size but generally shorter than 40 residues long. It is predicted as unstructured. The carboxy-terminal tail domain contains a NLS and a globular immunoglobulin-like domain (IgFold). In B-type lamins, this globular domain is followed by a CAAX box, whereas this box is absent in lamin C and is cleaved during maturation in lamin A. The 3D structure of the human A-type lamin Igfold was solved by NMR[42] and crystallography[43] (figure 7, c). It shows a compact β-sandwich formed by two β-sheets and nine β-strands connected by short loops. This Igfold is reported to mediate diverse protein-protein and protein-ligand interactions. However no 3D structure of a complex involving this Igfold was reported until now.

Then, first high resolution pictures of lamin filaments were recently obtained using cry-electron tomography by the groups of W. Baumeister at Martinsried[44] and the group of O. Medalia at University of Zürich[45]. The first group obtained images of vitrified interphase HeLa cells, and identified entangled, 4-nm-diameter filaments in a volume of 30 nm in thickness underlying the NE, which they assigned to lamins.

The second group obtained images of vimentin-null mouse embryonic fibroblasts, which were treated as following: the plasma membrane and cytoplasmic content were removed by a short exposure to a mild detergent, and most of the chromatin was removed by treating the nuclei with nuclease. The lamin filaments were identified using gold-labelled antibodies against lamins. Thus this group reported the presence of ~3.5 nm thick lamin filaments forming a complex meshwork within a 14±2 nm thick layer at the nuclear periphery. Statistical analysis of these filaments indicated a length of 380±122 nm. In figure 9, I have shown 2 views of lamins presented in the paper of Medalia and coworkers, showing the filament network at the nuclear periphery (figure 8a) and a model of the lamin oligomer consistent with the cryoelectron tomography data (figure 8b).

(a)



(b)

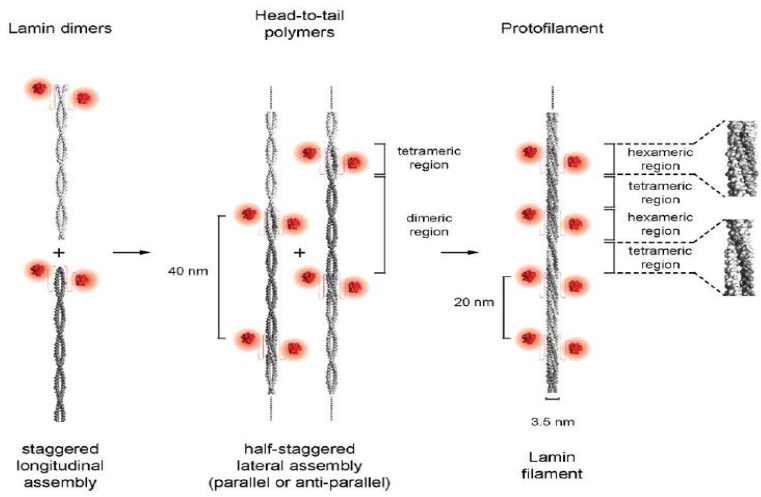


Figure 8 : Lamin models as deduced from the analysis of cryo-electron tomography images.

(a) The lamin network (rod domain in dark grey and Igfold in red), NPCs (in blue) and the lipid bilayers (in light grey) were modeled into a region of a cryo-tomogram containing the lamin meshwork. (b) Averaging and classification of lamin filament images obtained *in situ* confirmed that lamin filaments are composed of at least two half-staggered head-to-tail polymers. Igfold domains are displayed in red[46].

Both A and B-type lamins are able to oligomerize at the nuclear envelope. As it is well described in figure 8b, lamins form dimers through their α-helical rod domain, oriented in a parallel manner, and these dimers form the basic building block of higher-order lamin assemblies. Indeed, each dimer can further associate with another dimer through head to tail interactions, to form a polymer. At least two polymers finally associate laterally in a staggered conformation to form a filament showing a diameter of about 3-4 nm[45].

A- and B-type lamin meshworks are similar in appearance, but are mostly separated from each other, with some points of colocalization due to interactions between lamins[46]. Recently, A and B-type lamins were visualized directly in cells using super resolution microscopy, in order to confirm and extend these previous data (figure 9)[18]. It was observed that lamins are assembled into structurally and functionally separated meshworks or microdomains, with no preferential alignment.

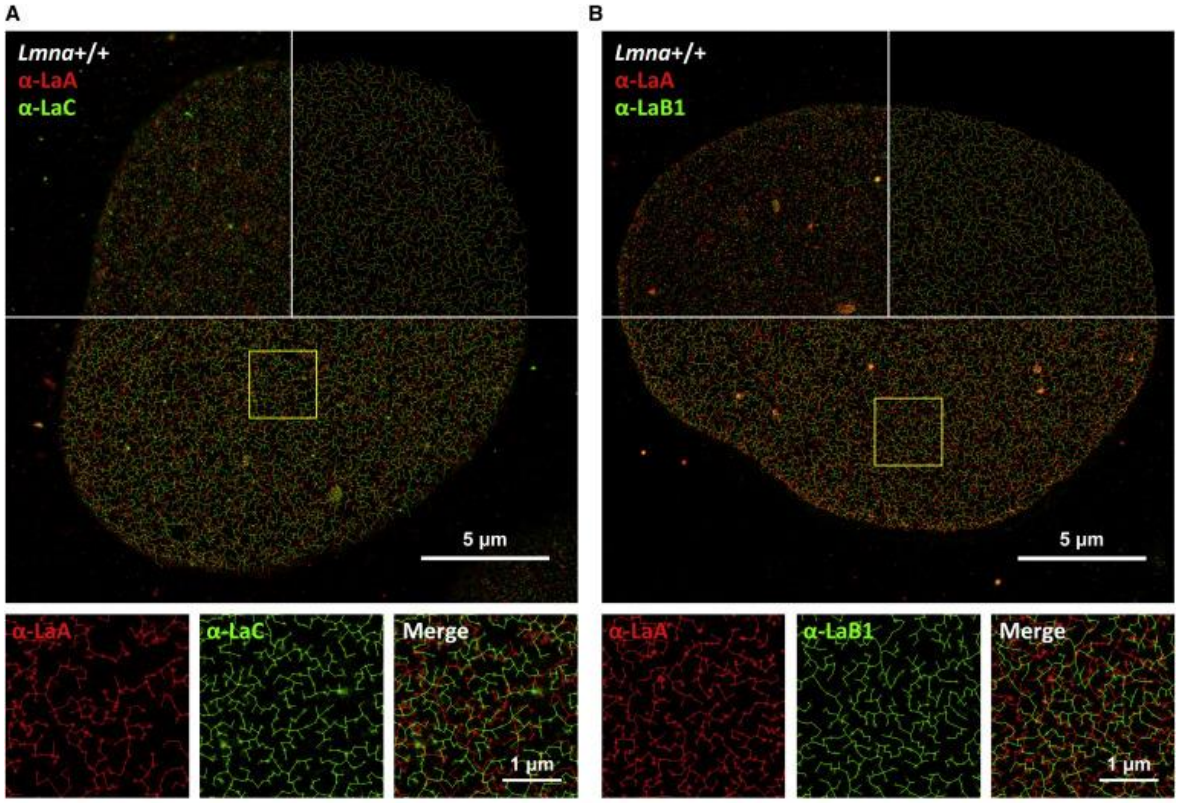


Figure 9 : Endogenous lamins observed using super resolution microscopy (d-STORM analysis), based on conventional indirect immunofluorescence labeling employing primary antibodies against lamins A, C, and B1[18].

3. THE INNER NUCLEAR MEMBRANE PROTEINS

The outer nuclear membrane (ONM) is continuous with the endoplasmic reticulum and has ribosomes bound to its cytoplasmic surface. However, the ONM contains specific proteins, which are not found in the endoplasmic reticulum membrane[47]. The inner nuclear membrane adheres to the nuclear lamina matrix and proteomic studies revealed that this membrane is enriched in at least 50 integral membrane proteins[48]. I will focus below on the most studied of these inner nuclear membrane proteins.

a. LEM domain proteins

Among the inner nuclear membrane proteins, a prominent family shares the ability to directly bind to the nuclear lamina and to tether the nuclear envelope to chromatin, through interactions with a conserved chromatin-associated protein named barrier-to-autointegration factor (BAF). This family is formed by the LAP2-emerin-MAN1 (LEM) domain proteins, which all exhibit a LEM domain, whose 3D structure is known[49–52] (figure 10). This domain is composed of approximatively 40 amino acids, which form a three residue N-terminal helical turn (in blue) and two parallel α-helices (in red and green) connected by a long extended loop (in yellow).

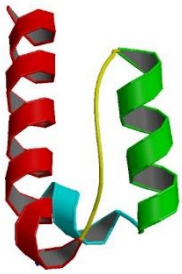


Figure 10 : Emerin LEM domain 3D structure, as solved by NMR *(PDB 2ODC)*[52].

In mammals, LEM domain proteins can be categorized into three groups (figure 11)[53]. Group I is composed of proteins with an N-terminal LEM domain (or a LEM-like domain, i.e. a domain structurally homologous to the LEM domain)[50,51,54], followed by a large and disordered region. Some of these proteins possess a single transmembrane domain at the C-terminus that anchors the protein to the inner nuclear membrane, whereas some are soluble nuclear proteins. In this group, we find proteins emerin and LAP2. The 3D structure of emerin LEM domain in complex with BAF was solved on the basis of NMR data[52]. This LEM domain binds to BAF with an affinity of 200-600 nM[52,55]. LAP2 can be expressed as six isoforms. Two of these isoforms, LAP2α and LAP2ζ, lack a transmembrane domain and localize in the nucleoplasm.
All these isoforms possess an N-terminal LEM-like domain (residues 5-48) followed by a LEM domain (residues 109-153)[50,52]. Two-hybrid assays revealed that the LAP2 N-terminal region common to all isoforms was sufficient to bind BAF[56]. More precisely, region 67-195, but not region 1-67 nor region 138-373, was capable of BAF binding. As the LEM domain comprises residues 109 to 153, the linker between the LEM-like and the LEM domain together with the LEM domain seem sufficient for BAF binding. Cai et al.[51] further confirmed using NMR that the LEM domain is the main LAP2 binding site for BAF. However they observed that the complex between LAP2 LEM and BAF has a short lifetime of about 1.5-2 ms. Finally, Huang et al.[57] in the same team also confirmed later using NMR that the LEM domain of MAN1, another protein belonging to group I, binds to BAF.

Group II is composed of proteins with an N-terminal LEM domain, followed by an unfolded nucleoplasmic region, two transmembrane domains and a carboxyl-terminal winged-helix MSC (for MAN1/Src1p/C-terminal motif) domain, that directly binds to DNA. Among proteins of this group, there are LEMD2 and MAN1, which possesses an additional C-terminal UHM (U2AF homology motif) domain[58]. The MSC domain of LEMD2 was recently shown to directly interact with the ESCRT-II/ESCRT-III hybrid protein Cmp7p (CHMP7), in order to recruit additional ESCRT-III proteins to holes in the nuclear membrane and thus contribute to seal the membrane holes[59]. The UHM domain of MAN1 is essential for binding to SMAD2/3, which are important actors of the TGF-β pathway[60]. Finally, group III is composed of proteins that carry internal LEM domains and multiple ankyrin repeats. Representatives of this group include LEM3/Ankle1 and LEM4/Ankle2.

LEM3 shuttles between the cytoplasm and the nucleoplasm[61] whereas LEM4 has a transmembrane domain and localizes either at the endoplasmic reticulum in human cells, or at the inner nuclear envelope in worms[62]. LEM4 is essential for BAF dephosphorylation at mitotic exit. It acts in part by inhibiting BAF's mitotic kinase VRK-1 and by recruiting PP2A. It is required for PP2A to dephosphorylate BAF.

The composition of the LEM domain protein family differs between organisms. In metazoa, at least one member from each group is present, whereas in yeast only group II proteins were identified[53].

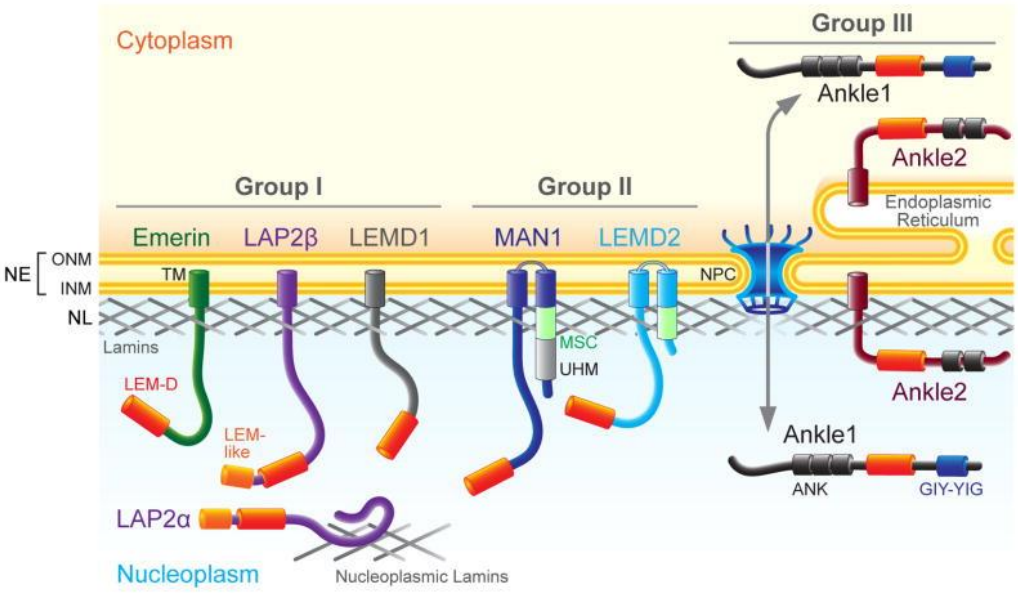


Figure 11 : Human LEM domain protein family[53].

I inserted in the following pages of this thesis, the accepted version of a review on LEM proteins written for Methods in Enzymology and to which I contributed.

Purification and structural analysis of LEM-domain proteins

Isaline Herrada[*], Benjamin Bourgeois[*2], Camille Samson[*], Brigitte Buendia[†], Howard J Worman[‡] & Sophie Zinn-Justin[*,1]

* Institute for Integrative Biology of the Cell, CEA, CNRS, Université Paris-Sud, Batiment 144 – CEA SACLAY, Gif-sur-Yvette, F-91191, France

† Institut de Biologie Fonctionnelle et Adaptative, Université Paris Diderot-Paris 7, CNRS, UMR8251, 4 rue M.A. Lagroua Weill Halle, 75205 Paris cedex 13, France

‡ Department of Medicine and Department of Pathology and Cell Biology, College of Physicians and Surgeons, Columbia University, 630 West 168th Street, New York, NY 10032

[1] Corresponding author e-mail: sophie.zinn@cea.fr

[2] Present address: Center for Integrated Protein Science Munich, Department Chemie, Technische Universität München, 85478 Garching, Germany

Contents

Abstract

LAP2-emerin-MAN1 (LEM)-domain proteins are modular proteins characterized by the presence of a conserved motif of about 50 residues. Most LEM domain proteins localize at the inner nuclear membrane, but some are also found in the endoplasmic reticulum or nuclear interior. Their architecture has been analyzed by predicting the limits of their globular domains, determining the 3D structure of these domains and in a few cases calculating the 3D structure of specific domains bound to biological targets. The LEM domain adopts an α-helical fold also found in SAP and HeH domains of prokaryotes and unicellular eukaryotes The LEM domain binds to BAF (barrier-to-autointegration factor; *BANF1*), which interacts with DNA and tethers chromatin to the nuclear envelope. LAP2 isoforms also share an N-terminal LEM-like domain, which binds DNA. The structure and function of other globular domains that distinguish LEM domain proteins from each other have been characterized, including the C-terminal dimerization domain of LAP2α and C-terminal WH and UHM domains of MAN1. LEM domain proteins also have large intrinsically disordered regions that are involved in intra- and inter-molecular interactions and are highly regulated by posttranslational modifications in vivo.

1. Introduction

The nuclear envelope (NE) forms the boundary of the nuclear compartment in eukaryotic cells. The NE includes two nuclear membranes (inner and outer), nuclear pore complexes and the nuclear lamina. The inner nuclear membrane (INM) has numerous integral membrane proteins and closely contacts nuclear intermediate filament proteins, lamins, to form 'nuclear lamina' networks that line the membrane. It is poorly understood how these proteins are organized at the structural level: the nuclear periphery is densely packed with proteins and chromatin, and structural investigation of networks of INM proteins and lamins *in situ* remains challenging (Gruenbaum and Medalia, 2015). Current knowledge of the inner aspect of the NE is therefore based on the identification of direct or indirect interactions between integral INM proteins and their partners including lamins, chromatin and transcriptional regulators.

The most extensively studied INM proteins and functional complexes involve LAP2-emerin-MAN1 (LEM)-domain proteins. This protein family is characterized by a globular LEM domain, initially identified in Lamina-Associated Polypeptide 2 (LAP2), Emerin and MAN1 (Lin et al., 2000). LEM-domain proteins are often anchored at the INM by binding to lamins. However, INM anchoring is not a feature of all isoforms, and no consensus lamin binding motif has been identified in LEM-domain proteins. For example, the LAP2 gene is alternatively spliced to produce six protein isoforms; LAP2β, γ, δ and ε each have a transmembrane domain, whereas LAP2α and LAP2ξ do not (Berger et al., 1996; Furukawa et al., 1995). However the LAP2α isoform-specific region from amino acid 616 to amino acid 693 binds the tail domain of A-type lamins (Dechat et al., 2000). LAP2α is essential for solubilizing a small fraction of lamin A in the nucleoplasm (Naetar et al., 2008).

Yeast two-hybrid experiments showed that LAP2β, the largest INM-localized isoform, uses residues 298 to 373 to interact with lamins B1 and B2 (Furukawa et al., 1998). This lamin-binding region is conserved in most LAP2 isoforms with a transmembrane domain.

The LEM domains of LAP2, emerin and MAN1 directly bind to barrier-to-autointegration factor (BAF) (Cai et al., 2001; Lee et al., 2001; Mansharamani and Wilson, 2005; Shumaker et al., 2001). BAF is highly conserved among multicellular eukaryotes (Cai et al., 1998) and is essential in *C. elegans* (Zheng et al., 2000). BAF co-localizes with chromosomal DNA during both interphase and mitosis in *Drosophila* (Furukawa, 1999), but localizes dynamically at the NE, nucleoplasm and cytoplasm in *C. elegans* and mammalian cells (Jamin and Wiebe, 2015). Through mechanisms that are not understood, BAF helps tether chromatin to the nuclear envelope (Asencio et al., 2012; Margalit et al., 2005) and functions as an epigenetic regulator (Montes de Oca et al., 2011).

2. "Divide and conquer" approach to study LEM-domain proteins

In mammals, three major types of LEM proteins can be distinguished based on their domain organization as shown in Figure 1 (Brachner and Foisner, 2011). Emerin and LAP2 are INM proteins characterized by one transmembrane segment (Figure 1A). They have an N-terminal nucleoplasmic LEM domain (shown for LAP2 in Figure 1A) and a large region predicted as unstructured, followed by the transmembrane segment and short lumenal domain. MAN1 and Lem2 are anchored at the INM by two transmembrane segments, with their N- and C-terminal regions exposed in the nucleoplasm (Figure 1E). Their LEM domains (Figure 1F) are each followed by a large region predicted to be unstructured (like emerin and LAP2β), and their C-terminal nucleoplasmic regions include either one (Lem2; not shown) or two globular domains (MAN1; shown in Figure 1G,H).

Ankle1 and Ankle 2 (also called Lem3 and Lem4) are characterized by ankyrin repeats, which are common protein-protein interaction motifs mainly found in eukaryotes (Figure 1). Ankle 1 has no transmembrane segment; it is an endonuclease that cleaves DNA and induces a DNA damage response (Brachner et al., 2012). Ankle 2 by contrast has an N-terminal transmembrane segment and localizes with the endoplasmic reticulum (Asencio et al., 2012). Ankle 2 is essential for postmitotic NE formation, and controls BAF phosphorylation by binding and coordinating the activities of kinase VRK-1 and phosphatase PP2A.

Structural analysis of LEM domain proteins has been hindered by difficulty expressing and purifying these large modular proteins, especially when they have transmembrane domains. Therefore, a “divide-to-conquer” approach has been chosen by several groups, based on bioinformatics prediction of the limits of the folded regions. Using this approach, globular domains of LEM domain proteins were systematically produced, purified and structurally characterized.

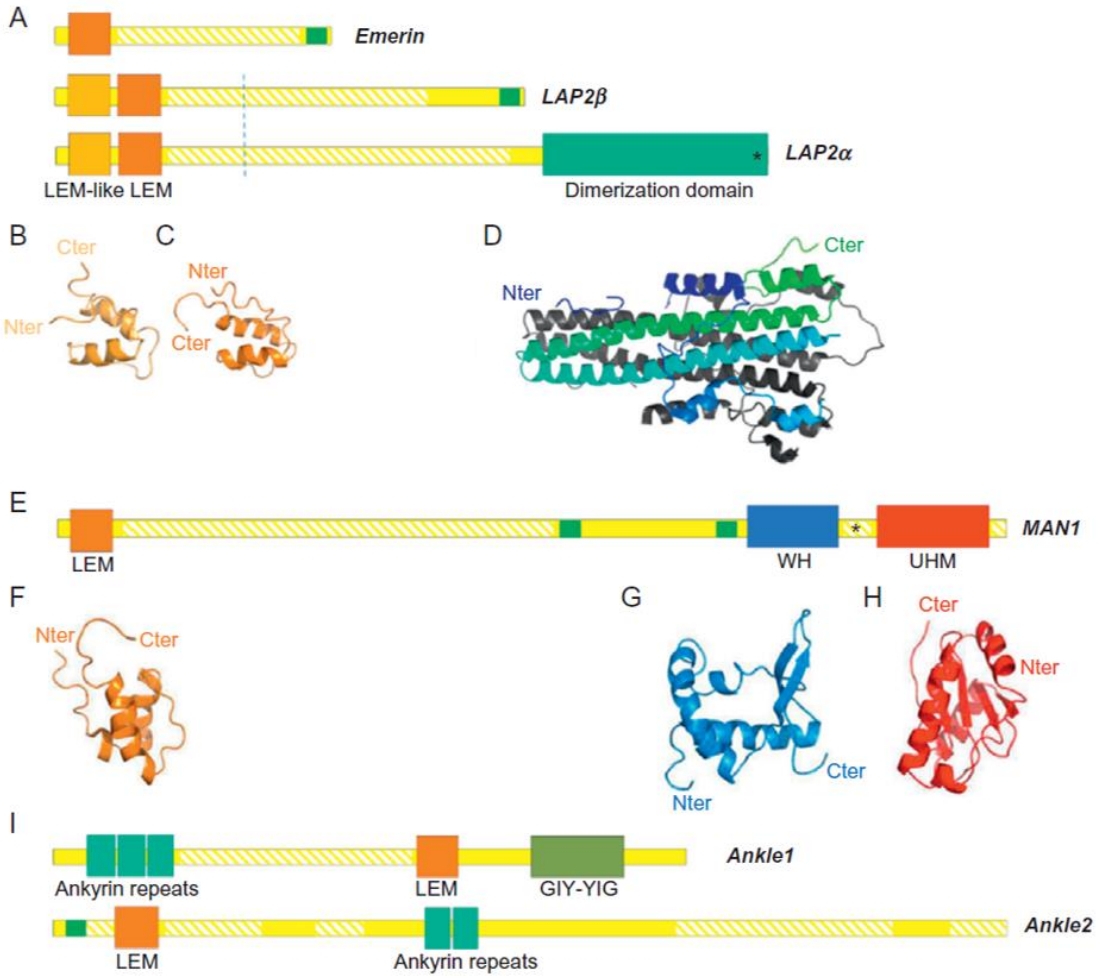


Figure 1. Schematic of LEM-domain protein architecture. Three families of LEM-domain proteins, illustrated by 6 human proteins. Hatched yellow regions are predicted as intrinsically disordered by Disopred3. Thin green boxes indicate transmembrane domains. Black dotted line indicates that LAP2 isoforms share residues 1-186 but have different C-terminal regions (after the line). **(A)** LEM-domain proteins with one transmembrane domain (emerin, LAP2β) localize at the INM; alternatively-spliced isoform LAP2α has no transmembrane domain. The star in LAP2α indicates the Arg690Cys mutation linked to dilated cardiomyopathy. **(B,C,D)** 3D structures of the LEM-like domain in light orange (PDB code 1H9E), the LEM domain in strong orange (PDB code 1H9F) and LAP2α–specific dimerization domain in green (PDB code 2V0X). **(E)** INM-localized LEM-domain proteins with two transmembrane segments are Lem2 (not shown) and MAN1. The star in MAN1 indicates the UHM Ligand Motif Lys/Arg_{763}-X-Trp_{765}-Gln_{766}-X-X-Ala_{769}-Phe_{770}. **(F,G,H)** 3D models of MAN1 globular domains, calculated by homology for the LEM domain in orange, from NMR data for the WH domain in blue (PDB code 2CH0; (Caputo et al., 2006)) and from NMR chemical shift data and molecular modeling for the UHM domain in red (Konde et al., 2010). **(I)** LEM-domain proteins with ankyrin repeats: Ankle1 is soluble; Ankle2 is ER-localized.

Structural analysis of LEM domain proteins has been hindered by difficulty expressing and purifying these large modular proteins, especially when they have transmembrane domains. Therefore, a “divide-to-conquer” approach has been chosen by several groups, based on bioinformatics prediction of the limits of the folded regions. Using this approach, globular domains of LEM domain proteins were systematically produced, purified and structurally characterized.

2.1. The LEM domain common to LAP2, emerin and MAN1

Biochemical characterization of the LEM domain started with the chemical synthesis of the LEM domains of human emerin (residues 2 to 54) and LAP2 (residues 103 to 159) and the LEM-like domain of LAP2 (residues 2 to 56) (Laguri et al., 2001; Wolff et al., 2001). The LEM-like domain is a highly divergent version of the LEM domain. Whereas the identity between the LEM domains of emerin and LAP2 is 33%, the LEM and LEM-like domains of LAP2 are only 18% identical. Milligrams of each peptide were produced, purified by HPLC and dissolved in a sodium phosphate buffer (pH 6.3). They were first analyzed by analytic ultracentrifugation, which showed that all these peptides are monomeric at concentrations less than 50 μM. The three peptides were then analyzed using proton Nuclear Magnetic Resonance (NMR; 500 and 600 MHz spectrometers), at a concentration of about 1 mM at 298K, to calculate their three-dimensional structures in solution.

Superimposition of the resulting structures revealed that they all adopted the same fold, mainly composed of two large parallel alpha helices stabilized by intramolecular electrostatic interactions and hydrophobic contacts (Figure 2A).

The N-terminal region of LAP2 containing both the LEM-like and LEM domains was also studied by NMR (Cai et al., 2001). A recombinant His_6-tagged LAP2 fragment (residues 1 to 168) was produced in *E. coli* as a ^{15}N and ^{13}C labeled protein. It was purified by affinity and gel filtration chromatography in phosphate buffer pH 7.2. Multidimensional ^{1}H, ^{15}N, ^{13}C NMR experiments were performed to solve its 3D structure. This fragment contained two structurally independent, non-interacting domains that are connected by a flexible linker and adopt the same α-helical fold. NMR chemical shift mapping demonstrated that the LEM domain binds BAF, whereas the LEM-like domain binds DNA. These interactions involve similar regions of the LEM fold, specifically helix 1, the loop connecting the two large helices and the N-terminus of helix 2. The distinct binding properties of the LEM-like and LEM domains are determined by their surface residues in these locations: predominantly positively charged in the case of the LEM-like domain, and mainly hydrophobic for the LEM domain.

There are few structural domains of less than 50 residues in the Protein Data Bank. A DALI search highlighted that the LEM and LEM-like domains are structurally highly related to the SAF-Acinus-PIAS (SAP) and the Helix-extended loop-Helix (HeH) motifs. For example, the emerin LEM structure can be superimposed onto the SAP structures of human nuclear protein Hcc-1 and yeast protein Tho1, and the HeH domain of *E. coli* transcriptional terminator protein Rho (Figure 2B). The SAP domain is a eukaryotic putative DNA-binding module found in chromatin-associated proteins that likely targets these proteins to specific chromosomal locations (Aravind and Koonin, 2000).

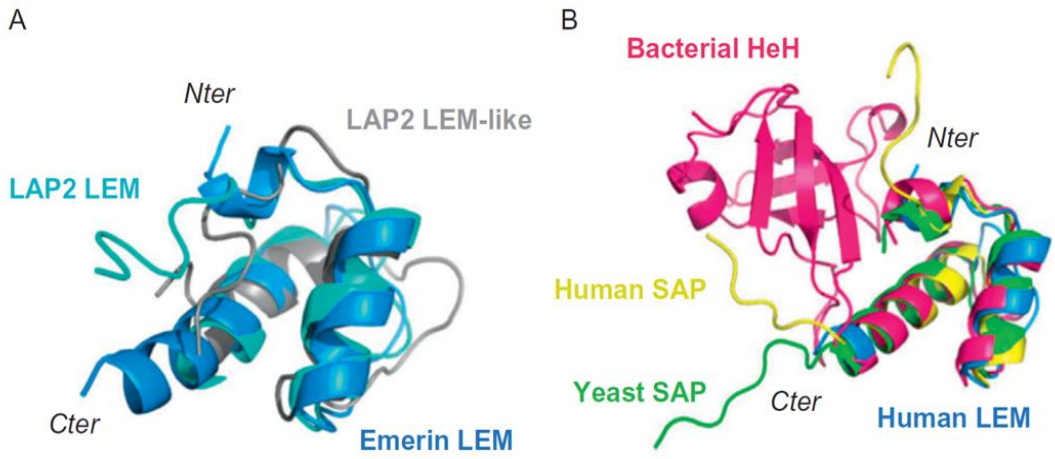


Figure 2. The LEM domain fold. (A) Superimposition of the LEM structure of emerin (PDB 2ODC in marine blue) with the LEM (PDB 1H9F in cyan; DALI rmsd 2.2 Å) and LEM-like (PDB 1H9E in grey; DALI rmsd 3.0 Å) structures of LAP2. **(B)** Superimposition of the LEM fold with the bacterial HeH structure from the RNA binding domain of the transcription termination factor rho (PDB 1A62 in magenta; DALI rmsd 1.7 Å), the yeast SAP structure of the RNA binding protein Tho1 (PDB 4UZW; DALI rmsd 1.4 Å) and the human SAP structure of the ribonucleoprotein Hcc-1 (PDB 2DO1; DALI rmsd 2.2 Å).

HeH domains from bacteria, bacteriophages and plants are known (or predicted) to bind nucleic acids, suggesting that this was the ancestral function of the HeH fold (Aravind and Koonin, 2001; Aravind et al., 2002). The SCOP database gathers all these domains into a category named "LEM/SAP Helix-extended loop-Helix (HeH) motif". Brachner & Foisner showed in 2011 that orthologs of metazoan LEM-domain proteins in unicellular eukaryotes, which lack BAF, contain a SAP or HeH instead of a LEM motif, suggesting the LEM domain evolved from an ancestral SAP/HeH domain found in chromosome tethering proteins, concomitant with the emergence of BAF (Brachner and Foisner, 2011).

2.2. The C-terminal domain specific to LAP2α

LAP2α contains an N-terminal region from amino acid 1 to amino acid 186 common to all LAP2 isoforms (including LEM-like and LEM domains, see Figure 1A,B,C) and a specific C-terminal region from amino acid 187 to amino acid 693 responsible for many of its known functions (see Figure 1A,D). This last region is specifically involved in LAP2α retention in postmitotic nuclei and essential for A-type lamin binding (Dechat et al., 1998; Dechat et al., 2000; Vlcek et al., 1999). It is mutated at position 690 (Arg to Cys) in two brothers with dilated cardiomyopathy, but segregation of the disease-causing allele within the family is unclear (Taylor et al., 2005). LAP2α variant $Arg_{690}Cys$ is reported to be impaired in its ability to bind lamin A. To identify the molecular basis of the LAP2α / lamin A interaction, full-length murine LAP2α was digested by chymotrypsin and fragments were analyzed by mass spectrometry and N-terminal sequencing. A stable fragment was identified that comprised residues 459 to 693. This fragment was expressed in *E. coli* and purified by metal affinity chromatography followed by size exclusion chromatography. Its three-dimensional structure was determined by X-ray crystallography (Figure 1D; (Bradley et al., 2007)). The C-terminal region of LAP2α is an elongated α-helical dimer. Each monomer exhibits six helices with the fourth and fifth arranged into a four-stranded coiled coil. No complex between this C-terminal dimerization domain and A-type lamin tails could be observed using chromatography, calorimetry or ultracentrifugation experiments. Only a solid-phase overlay assay suggested transient binding between the two molecules and pointed to six LAP2α residues involved in lamin recognition. The $Arg_{690}Cys$ substitution was introduced into the murine LAP2α fragment, but did not impair lamin binding in the solid-phase overlay assay. This illustrates the difficulty encountered in understanding the molecular basis of lamin recognition by LEM-domain proteins.

2.3. The WH and UHM domains in MAN1

MAN1 and LEM2 are LEM-domain proteins anchored in the INM by two transmembrane segments. They exhibit two nucleoplasmic regions: an N-terminal region including the LEM domain and a C-terminal region (see Figure 1E,F). This last region contains two globular domains in MAN1 (Figure 1G,H) and one globular domain in Lem2. Both globular domains in the C-terminal region of human MAN1 were analyzed by NMR to determine their three-dimensional structure.

A fragment comprising MAN1 C-terminal residues 655-775, which are shared with Lem2, was produced as a ^{15}N, ^{13}C labeled GST fusion protein and purified by affinity chromatography. The GST fragment was cleaved using thrombin and retained on an affinity column. The MAN1 fragment was eluted in the flow-through and then dialyzed to a concentration of about 0.5 mM in phosphate buffer (pH 6.0) containing 150 mM NaCl. It was analyzed by multidimensional ^{1}H, ^{15}N and ^{13}C NMR, and the 3D structure of the region between residues 666 and 750 was solved from data acquired at 303K on 600 MHz and 900 MHz spectrometers (Caputo et al., 2006). The N-terminal half of this region is mainly α-helical, whereas the C-terminal half is composed of two large β-strands arranged in a twisted anti-parallel β-sheet (Figure 1G). The helices form a three-helix bundle. A three-stranded β-sheet is packed onto the three-helix bundle. The α/β interface is mainly hydrophobic. The DALI server clearly identified the MAN1 region from residue 666 to residue 750 as adopting a Winged Helix (WH) fold also found in several DNA binding domains belonging to transcription factors. The helix that contacts DNA in these transcription factors is positively charged in MAN1 and, consistently, the MAN1 fragment binds DNA (Caputo et al., 2006).

The MAN1 fragment from amino acid 775 to amino acid 911 is not found in Lem2. A point mutation in this region abolishes the capacity of MAN1 to bind R-Smads and repress transcription of the R-Smad-mediated signaling pathway (Pan et al., 2005). Consistently Lem2 is not capable of antagonizing R-Smad-mediated signaling activity (Brachner et al., 2005). The MAN1 fragment (residues 755 to 911) was produced as a ^{15}N, ^{13}C labelled His_6-tagged protein to characterize its 3D structure by NMR (Konde et al., 2010). It was purified by affinity chromatography and gel filtration in a Tris buffer (pH 6.7) containing 150 mM NaCl. NMR analysis provided the secondary structure of the MAN1 fragment. However, conformational exchange precluded the determination of its 3D structure. Sequence alignment strongly suggested that residues 785 to 880 adopt a U2AF-Homology Motif (UHM) fold (Kielkopf et al., 2004). NMR chemical shift data were consistent with this prediction. From NMR and bioinformatics analyses, it was possible to calculate a model for the UHM domain of MAN1 (Figure 1H). Broad NMR signals suggesting conformational exchange were observed in the linker region between the WH and UHM domains (residues 755 to 785) as well as around the hydrophobic cavity of the UHM domain. UHM domains interact with peptides exhibiting a UHM Ligand Motif (ULM). This motif consists of a patch of positively charged amino acids followed by a conserved tryptophan residue (Kielkopf et al., 2004). In the MAN1 linker region, the motif Lys/Arg_{763}-X-Trp_{765}-Gln_{766}-X-X-Ala_{769}-Phe_{770} is highly conserved through metazoans (this motif is represented by a star in Figure 1E). When Trp_{765} and Gln_{766} are mutated to alanines, the radius of gyration of the MAN1 fragment significantly increases, as shown by Small Angle X-ray Scattering (SAXS), and the NMR signals of the residues of the UHM hydrophobic cavity are changed. From these data, a model of the MAN1 fragment was proposed in which Trp_{765} anchors the linker onto the UHM domain through a conserved intramolecular interaction.

The biological significance of this interaction is currently unknown. However, mutating Trp_{765} and Gln_{766} to alanines significantly weakens the interaction between the MAN1 fragment and the MH2 domain of Smad2.

3. Analysis of predicted unstructured regions in LEM-domain proteins

Large Intrinsically Disordered Regions (IDRs) have been identified in LEM-domain proteins, either between two globular domains or between a globular domain and a transmembrane segment (Figure 1A,E,I). The functional role of these regions is not yet understood. They might contain MoRFs, small (10-70 residues) fragments that undergo a disorder-to-order transition upon binding to their partners. However, no 3D structure of a LEM-domain protein IDR bound to its partner has been solved. They are largely post-translationally modified and phosphorylated in a cell-cycle dependent manner (Ellis et al., 1998; Hirano et al., 2009; Hirano et al., 2005; Yip et al., 2012). Phosphorylation of the emerin IDR is also triggered by the application of mechanical force on the nucleus (Guilluy et al., 2014).

The IDR located between the WH and UHM domains of MAN1 (residues 750 to 785) has been studied while analyzing the three-dimensional structure of the whole C-terminal nucleoplasmic region of MAN1. This IDR binds to MAN1 UHM domain, thus regulating its recognition properties (Konde et al., 2010). Mutation of the IDR sequence Trp_{765}-Gln_{766} weakens both the intramolecular IDR-UHM interaction and the intermolecular MAN1-Smad2 interaction (Figure 1E). This suggests that the IDR position onto the UHM is critical for Smad2 recognition. Three residues were identified as phosphorylated in this IDR, namely Ser_{777}, Ser_{781} and Thr_{783} (http:// www.phosphosite.org/ proteinAction.do?id=7784). Ser_{777} and Thr_{783} are potential targets of Pro-directed kinases such as P38 (mitogen-activated protein kinase); Ser_{777} is also a potential target of glycogen synthase kinase-3β (GSK-3β) (predictions: elm.eu.org).

These kinases might regulate the intramolecular interaction between MAN1 IDR and UHM regions, and thus the binding properties of MAN1. Other larger IDRs have been described in emerin and MAN1 N-terminal nucleoplasmic regions. Production and purification of these regions should enable a biophysical study of their structural and binding properties, either before or after post-translational modification.

3.1. Production and purification of predicted unstructured regions of emerin and MAN1

3.1.1. Emerin nucleoplasmic region

The emerin nucleoplasmic region contains a LEM domain and a large IDR (Figure 1A). We attempted to produce in *E. coli* and purify by chromatography the whole nucleoplasmic fragment of emerin from amino acid 1 to amino acid 221 and a smaller fragment from amino acid 1 to amino acid 187 (emerin 1-187), but only the latter was soluble in our conditions. We expressed it as a 8His-tag, followed by a TEV site and the emerin 1-187 sequence. The cDNA coding for this construct was optimized for *E. coli,* synthesized by Genscript and cloned into a pETM13-LIC vector (pETM13 expression vector modified in our laboratory for LIC cloning system). The emerin 1-187 sequence contains four positions with amino acid substitutions or deletions in patients with X-linked Emery-Dreifuss muscular dystrophy (EDMD; positions 54, 95-99, 133 and 183). Studying the structural consequences of these mutations is of particular interest knowing that most cases of X-linked EDMD are caused by mutations leading to loss of emerin expression.

Detailed protocol for emerin 1-187 production and purification:

1. *E. coli* cells are grown at 310K in LB media (or M9 media for NMR samples). When the OD_{600} is between 1 and 1.5, add 0.5 mM IPTG and induce expression overnight at 293K.

2. Lyse cells by sonication in lysis buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 5% glycerol and 1% Triton X-100).

3. As the protein is produced in inclusion bodies, the pellet is resuspended in Buffer A8 (50 mM Tris-HCl pH 8, 150 mM NaCl, 20 mM imidazole, 8 M urea) and centrifuged 20 min at 20,000g at room temperature.

4. Load the supernatant on a Ni-NTA column (GE Healthcare) equilibrated in Buffer A8.

5. Wash the column with Buffer A8 and elute protein using 100% Buffer B8 (50 mM Tris-HCl pH 8, 150 mM NaCl, 1 M Imidazole, 8 M urea).

6. Fractions containing emerin 1-187 are pooled and dialysed against 20 mM Tris-HCl pH 8, 30 mM NaCl at room temperature in three steps (one night and twice for two hours), and stored at 277K.

3.1.2 The N-terminal nucleoplasmic region of MAN1

The MAN1 nucleoplasmic region contains a LEM domain and a large IDR (Figure 1E). We first attempted to produce recombinant MAN1 (amino acid 1 to amino acid 471; MAN1 1-471) in *E. coli* as a 6His-tagged fusion protein, with a TEV site between the tag and the MAN1 sequence. However due to weak expression yield and the production of several truncated MAN1 peptides, we optimized the cDNA sequence for expression in *E. coli* and the synthetic gene (ProteoGenix) was cloned into our pETM13-LIC vector.

Detailed protocol for MAN1 1-471 production and purification:

1. *E. coli* cells are grown at 310K in LB media (or M9 media for NMR samples) and induced 4 h at 310K, with 0.5 mM IPTG when the OD_{600} is between 0.8 and 1.

2. Lyse cells by sonication: 30 pulses of 1 sec each (power from 20 to 70%) separated by intervals of 1 sec in lysis buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 5% glycerol, 1% Triton X-100).

3. Incubate cell lysates with benzonase for 20 min at room temperature and centrifuge 20 min at 20 000g.

4. As the protein is soluble, load the supernatant on a Ni-NTA column (GE Healthcare) equilibrated in Buffer A (50 mM Tris-HCl pH 8, 150 mM NaCl, 20 mM Imidazole).

5. Wash the column with Buffer A, and elute using a 20 min gradient from 0% to 100% Buffer B (50 mM Tris-HCl pH 8, 150 mM NaCl, 1 M imidazole).

6. Fractions containing MAN1 1-471 are pooled, concentrated by centrifugation using Vivapsin sample concentrators (GE Healthcare), injected on a gel filtration column (Superdex 200, Hi-load, 120 mL, GE Healthcare) equilibrated in a buffer containing 150 mM NaCl and 1 mM tris(2-carboxyethyl)phosphine.

7. Store the purified MAN1 1-471 polypeptide at 277K.

3.2. Association between globular and predicted unstructured regions in MAN1

In the MAN1 C-terminal nucleoplasmic region, a ULM motif was detected in the linker region that interacts with the UHM domain (Konde et al., 2010). Similarly, within MAN1 residues 1-471 a conserved motif Arg_{190}-Arg_{191}-Lys_{192}-Pro_{193}-His_{194}-Ser_{195}-Trp_{196}-Trp_{197}-Gly_{198} could play the role of a ULM motif. We therefore tested the entire purified N-terminal nucleoplasmic region of MAN1 for potential binding to the C-terminal nucleoplasmic region using 2D ^{1}H, ^{15}N NMR.

We prepared two NMR samples: both contained 200 µl ^{15}N labeled MAN1 1-471 at 100 µM in phosphate buffer (pH 6.7) with 150 mM NaCl; one tube also contained unlabeled MAN1 755-911 at 100 µM. 2D ^{1}H-^{15}N HSQC spectra were acquired on both samples at 293K on a 600 MHz spectrometer. Superimposition of these two spectra revealed an interaction between the two MAN1 nucleoplasmic regions (unpublished results). Indeed, several MAN1 1-471 NMR signals decreased in intensity upon addition of MAN1 755-911. These signals were assigned to the MAN1 sequence Trp_{196}-Trp_{197}-Gly_{198} thus revealing that this motif behaves like a new ULM. We suggest that the two MAN1 ULM motifs, Trp_{196}-Trp_{197}-Gly_{198} and Trp_{765}-Gln_{766}, can compete with each other for binding to the UHM domain leading to regulation of MAN1-SMAD2/3 complex formation.

3.3. Self-association of emerin via predicted unstructured region

A role for the emerin IDR in promoting self-association events was suggested by two studies. First, yeast 2-hybrid experiments revealed that a fragment comprising emerin residues 1-225 binds to itself, suggesting C-terminally-truncated emerin could form homodimers and/or multimers (Sakaki et al., 2001). Second, *in vitro* GST-pulldown experiments using recombinant emerin residues 1-221, and co-immunoprecipitation of full-length emerin from HEK293T cell extracts, also strongly suggested that emerin could oligomerize both *in vitro* and in cells (Berk et al., 2014). To identify which emerin region was responsible for self-association, N-terminally His-tagged emerin fragment was bound to Ni^{2+}-NTA agarose under denaturing conditions and the bead-bound polypeptide was incubated with purified GST-tagged emerin fragments in an appropriate binding buffer. After washing the beads, proteins were eluted using SDS-sample buffer and the binding reactions were resolved on gels. This experiment showed that GST-tagged emerin 170-220 is sufficient to bind emerin 1-221, whereas emerin 1-160 does not bind. Further pull-down assays suggested that emerin 170-220 recognized both fragments 1-132 and 170-220.

In Hela cells co-expressing GFP-tagged and Flag-tagged emerin with different deletions, fragment 170-220 was also essential for emerin-emerin co-immunoprecipitation (Berk et al., 2014). Thus self-association is due to sequences that are predicted to be unstructured within the emerin nucleoplasmic region.

4. LEM-domain proteins with their partners

LEM-domain proteins interact with lamins and with other INM proteins such as SUN1, SUN2, LAP1 and contribute to NE architecture (Haque et al., 2010; Patel et al., 2014; Shin et al., 2013). These interactions play critical roles in nuclear structure and chromatin organization. However, the difficulty in obtaining pure and functional fragments of LEM-domain proteins in their proper oligomerization and modification states makes it difficult to reconstitute *in vitro* biologically relevant complexes. Thus current knowledge is based on studies that focused on binding between specific functional domains and specific partners.

4.1. 3D structure of the LEM-BAF complex

The first LEM-domain complex successfully studied consisted of the LEM domain of emerin and the small (10 kDa) protein BAF. BAF is a centrosymmetric homodimer that binds DNA and thereby bridges two DNA molecules, thus contributing to DNA compaction (Cai et al., 1998). Various studies including NMR spectroscopy, X-ray crystallography and site-directed mutagenesis have shown that BAF is a α-helical protein that recognizes DNA through a pair of helix-hairpin-helix (HhH) motifs (Bradley et al., 2005; Umland et al., 2000). BAF binding to emerin 1-187 can be observed in solution by NMR (unpublished results; Figure 3A). Clore and colleagues determined the three-dimensional structure of the complex between BAF and the LEM domain of emerin (Figure 3B; (Cai et al., 2007)).

In this structure, one LEM domain interacts with two BAF monomers, as observed by NMR, light scattering and analytic centrifugation, with a dissociation constant of 0.6 μM as shown by Isothermal Titration Calorimetry (ITC). However, further studies suggested that post-translational modifications of emerin's IDR also regulate BAF recognition (Berk et al., 2013; Tifft et al., 2009). Information obtained from the structure of the complex showed that BAF binding sites to the LEM domain and DNA do not overlap. Thus, BAF could bind simultaneously to emerin and DNA. However, additional experiments suggested that, in cells, expression of BAF variant $Gly_{25}Glu$, with Gly_{25} being located at the interface between BAF and DNA, affected emerin localization during telophase (Haraguchi et al., 2001).

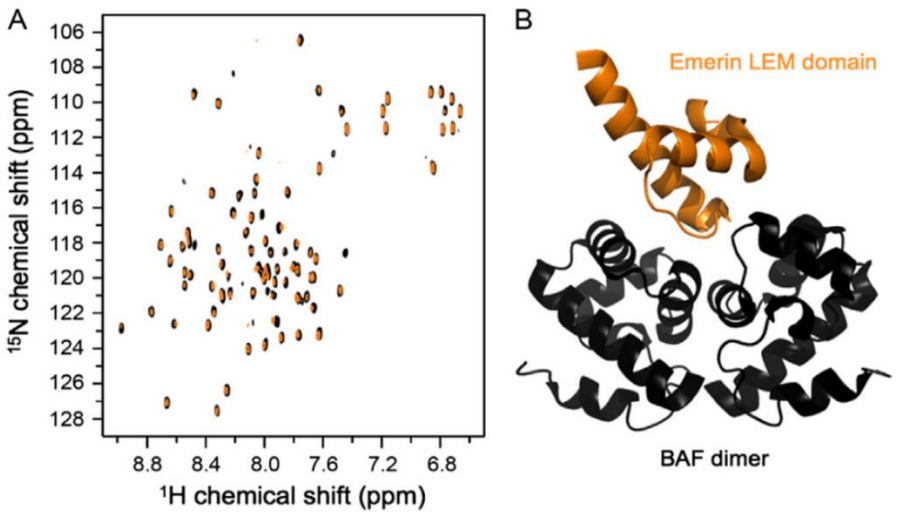


Figure 3. The emerin-BAF interaction. (A) Superimposition of two NMR ^{1}H-^{15}N HSQC spectra revealing the interaction between emerin 1-187 and BAF: the red and black spectra were recorded on ^{15}N labelled BAF alone and in complex with emerin, respectively (in 40 mM phosphate buffer pH 6.7, 150 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine, 1 mM EDTA at 293K on a 600 MHz spectrometer). Disappearance of ^{1}H-^{15}N signals indicates binding to the unlabelled partner. **(B)** 3D structure of the LEM-BAF complex: the emerin LEM-domain is blue, the BAF dimer is red (PDB code 2ODG).

4.2. Interaction of the MAN1 WH domain with DNA

The MAN1 residues 655-775 interact with a 211-base pair linear DNA molecule as observed by electrophoretic mobility shift assay (Caputo et al., 2006). The apparent affinity of the MAN1 fragment for this DNA molecule is 50 nM. Mutations of positively charged residues of the helix predicted to contact DNA, abolish DNA binding. Thus, the recognition helix of the MAN1 WH domain is involved in the binding of the carboxyl-terminal region of MAN1 to DNA.

4.3. MAN1 interactions with SMAD2/3 and PPM1A

Three groups reported that the MAN1 C-terminal region physically interacts with R-Smad proteins to repress the transforming growth factor-β (TGF-β) signaling pathway (Hellemans et al., 2004; Lin et al., 2005; Pan et al., 2005). Mutations in the gene encoding MAN1 cause sclerosing bone dysplasias, which sometimes have associated skin abnormalities (Hellemans et al., 2004). Fibroblasts from affected individuals are haploinsufficient with respect to MAN1 expression. These perturbations enhanced TGF-β signaling, since downstream genes targeted by this pathway were upregulated (Hellemans et al., 2004).

The MAN1 fragment from amino acid 755 to amino acid 911 recognizes the MH2 domain of Smad2 with micromolar affinity. From biochemical, NMR and SAXS analyses, it was possible to calculate a model of the complex between this MAN1 fragment and the Smad2 MH2 domain (Bourgeois et al., 2013). As predicted by this model, experiments *in vitro* showed that MAN1 binds to Smad2 alone, and to the activated Smad2-Smad4 complex. However, in cells, MAN1 does not bind Smad4-containing complexes. Overexpression of MAN1 leads to Smad2 dephosphorylation, thus hindering Smad2 binding to Smad4. *In vitro*, MAN1 binds directly to the phosphatase PPM1A, which catalyzes dephosphorylation of Smad2.

These results suggest a mechanism through which the MAN1-specific C-terminal region inhibits TGF-β signaling (Bourgeois et al., 2013).

They demonstrate that this MAN1 region recognizes different forms of Smad2 (monomers, homo and heterotrimers) and show that MAN1, by recruiting Smad2 to the nuclear envelope, facilitates its dephosphorylation by PPM1A and thus its inactivation.

5. Concluding remarks

LEM-domain proteins are involved in diverse cellular processes including DNA replication and cell cycle control, chromatin organization, nuclear assembly, regulation of gene expression and signaling pathways, as well as retroviral infection (Barton et al., 2015; Brachner and Foisner, 2011; Li and Craigie, 2006). The INM LEM-domain proteins LAP2, emerin and MAN1 have been intensively studied using biophysical and cellular approaches. Their LEM domain directly recognizes the DNA binding protein BAF (Cai et al., 2001). However, these proteins are membrane proteins and their soluble regions are largely predicted as intrinsically disordered. LAP2α dimerizes through its C-terminal α-helical domain (Bradley et al., 2007). Emerin also self-assembles into poorly characterized oligomers (Berk et al., 2014). It is thus difficult to identify their functional states *in vitro* and to reconstitute complexes between these proteins and their biological partners. First attempts to stabilize nuclear envelope complexes has recently led to the resolution of the X-ray structure of the KASH-SUN complex, located between the outer and inner nuclear membranes (Sosa et al., 2012). Moreover, development of mass spectrometry and NMR techniques now enables description of post-translational modification events in LEM-domain proteins, and their consequences on protein structure and binding properties.

Integrative approaches, including structural characterization of protein subcomplexes based on a panel of biophysical techniques and development of new tools in biochemistry and imaging techniques for identification of protein-protein interfaces might ultimately provide a better view of inner NE architecture.

Acknowledgments

This work was supported by AFM grants 17243 (research grant to SZJ) and 18159 (full PhD fellowship to CS), by FRM grant FDT20140931008 (one year PhD fellowship to IH), and by the French national infrastructure FRISBI program.

References

Aravind, L., and Koonin, E.V. (2000). SAP - a putative DNA-binding motif involved in chromosomal organization. Trends Biochem. Sci. *25*, 112-114.

Aravind, L., and Koonin, E.V. (2001). Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. Genome Research *11*, 1365-1374.

Aravind, L., Mazumder, R., Vasudevan, S., and Koonin, E.V. (2002). Trends in protein evolution inferred from sequence and structure analysis. Curr. Opin. Struct. Biol. *12*, 392-399.

Asencio, C., Davidson, I.F., Santarella-Mellwig, R., Ly-Hartig, T.B., Mall, M., Wallenfang, M.R., Mattaj, I.W., and Gorjanacz, M. (2012). Coordination of kinase and phosphatase activities by Lem4 enables nuclear envelope reassembly during mitosis. Cell *150*, 122-135.

Barton, L.J., Soshnev, A.A., and Geyer, P.K. (2015). Networking in the nucleus: a spotlight on LEM-domain proteins. Curr. Opin Cell Biol. *34*, 1-8.

Berger, R., Theodor, L., Shoham, J., Gokkel, E., Brok-Simoni, F., Avraham, K.B., Copeland, N.G., Jenkins, N.A., Rechavi, G., and Simon, A.J. (1996). The characterization and localization of the mouse thymopoietin/lamina-associated polypeptide 2 gene and its alternatively spliced products. Genome Research *6*, 361-370.

Berk, J.M., Maitra, S., Dawdy, A.W., Shabanowitz, J., Hunt, D.F., and Wilson, K.L. (2013). O-Linked beta-N-acetylglucosamine (O-GlcNAc) regulates emerin binding to barrier to autointegration factor (BAF) in a chromatin- and lamin B-enriched "niche". J. Biol. Chem. *288*, 30192-30209.

Berk, J.M., Simon, D.N., Jenkins-Houk, C.R., Westerbeck, J.W., Gronning-Wang, L.M., Carlson, C.R., and Wilson, K.L. (2014). The molecular basis of emerin-emerin and emerin-BAF interactions. J. Cell Science *127*, 3956-3969.

Bourgeois, B., Gilquin, B., Tellier-Lebegue, C., Ostlund, C., Wu, W., Perez, J., El Hage, P., Lallemand, F., Worman, H.J., and Zinn-Justin, S. (2013). Inhibition of TGF-beta signaling at the nuclear envelope: characterization of interactions between MAN1, Smad2 and Smad3, and PPM1A. Science Signaling *6*, ra49.

Brachner, A., Braun, J., Ghodgaonkar, M., Castor, D., Zlopasa, L., Ehrlich, V., Jiricny, J., Gotzmann, J., Knasmuller, S., and Foisner, R. (2012). The endonuclease Ankle1 requires its LEM and GIY-YIG motifs for DNA cleavage in vivo. J. Cell Science *125*, 1048-1057.

Brachner, A., and Foisner, R. (2011). Evolvement of LEM proteins as chromatin tethers at the nuclear periphery. Biochemical Society Transactions *39*, 1735-1741.

Brachner, A., Reipert, S., Foisner, R., and Gotzmann, J. (2005). LEM2 is a novel MAN1-related inner nuclear membrane protein associated with A-type lamins. J. Cell Science *118*, 5797-5810.

Bradley, C.M., Jones, S., Huang, Y., Suzuki, Y., Kvaratskhelia, M., Hickman, A.B., Craigie, R., and Dyda, F. (2007). Structural basis for dimerization of LAP2alpha, a component of the nuclear lamina. Structure *15*, 643-653.

Bradley, C.M., Ronning, D.R., Ghirlando, R., Craigie, R., and Dyda, F. (2005). Structural basis for DNA bridging by barrier-to-autointegration factor. Nature Struct. & Molec. Biol. *12*, 935-936.

Cai, M., Huang, Y., Ghirlando, R., Wilson, K.L., Craigie, R., and Clore, G.M. (2001). Solution structure of the constant region of nuclear envelope protein LAP2 reveals two LEM-domain structures: one binds BAF and the other binds DNA. EMBO J. *20*, 4399-4407.

Cai, M., Huang, Y., Suh, J.Y., Louis, J.M., Ghirlando, R., Craigie, R., and Clore, G.M. (2007). Solution NMR structure of the barrier-to-autointegration factor-Emerin complex. J. Biol. Chem. *282*, 14525-14535.

Cai, M., Huang, Y., Zheng, R., Wei, S.Q., Ghirlando, R., Lee, M.S., Craigie, R., Gronenborn, A.M., and Clore, G.M. (1998). Solution structure of the cellular factor BAF responsible for protecting retroviral DNA from autointegration. Nature Structural Biology *5*, 903-909.

Caputo, S., Couprie, J., Duband-Goulet, I., Konde, E., Lin, F., Braud, S., Gondry, M., Gilquin, B., Worman, H.J., and Zinn-Justin, S. (2006). The carboxyl-terminal nucleoplasmic region of MAN1 exhibits a DNA binding winged helix domain. J. Biol. Chem. *281*, 18208-18215.

Dechat, T., Gotzmann, J., Stockinger, A., Harris, C.A., Talle, M.A., Siekierka, J.J., and Foisner, R. (1998). Detergent-salt resistance of LAP2alpha in interphase nuclei and phosphorylation-dependent association with chromosomes early in nuclear assembly implies functions in nuclear structure dynamics. EMBO J. *17*, 4887-4902.

Dechat, T., Korbei, B., Vaughan, O.A., Vlcek, S., Hutchison, C.J., and Foisner, R. (2000). Lamina-associated polypeptide 2alpha binds intranuclear A-type lamins. J. Cell Science *113*, 3473-3484.

Ellis, J.A., Craxton, M., Yates, J.R., and Kendrick-Jones, J. (1998). Aberrant intracellular targeting and cell cycle-dependent phosphorylation of emerin contribute to the Emery-Dreifuss muscular dystrophy phenotype. J. Cell Science *111*, 781-792.

Furukawa, K. (1999). LAP2 binding protein 1 (L2BP1/BAF) is a candidate mediator of LAP2-chromatin interaction. J. Cell Science *112*, 2485-2492.

Furukawa, K., Fritze, C.E., and Gerace, L. (1998). The major nuclear envelope targeting domain of LAP2 coincides with its lamin binding region but is distinct from its chromatin interaction domain. J. Biol. Chem. *273*, 4213-4219.

Furukawa, K., Pante, N., Aebi, U., and Gerace, L. (1995). Cloning of a cDNA for lamina-associated polypeptide 2 (LAP2) and identification of regions that specify targeting to the nuclear envelope. EMBO J. *14*, 1626-1636.

Gruenbaum, Y., and Medalia, O. (2015). Lamins: the structure and protein complexes. Curr. Opin. Cell Biol. *32C*, 7-12.

Guilluy, C., Osborne, L.D., Van Landeghem, L., Sharek, L., Superfine, R., Garcia-Mata, R., and Burridge, K. (2014). Isolated nuclei adapt to force and reveal a mechanotransduction pathway in the nucleus. Nature Cell Biology *16*, 376-381.

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

Haraguchi, T., Koujin, T., Segura-Totten, M., Lee, K.K., Matsuoka, Y., Yoneda, Y., Wilson, K.L., and Hiraoka, Y. (2001). BAF is required for emerin assembly into the reforming nuclear envelope. J. Cell Science *114*, 4575-4585.

Hellemans, J., Preobrazhenska, O., Willaert, A., Debeer, P., Verdonk, P.C., Costa, T., Janssens, K., Menten, B., Van Roy, N., Vermeulen, S.J., *et al.* (2004). Loss-of-function mutations in LEMD3 result in osteopoikilosis, Buschke-Ollendorff syndrome and melorheostosis. Nature Genetics *36*, 1213-1218.

Hirano, Y., Iwase, Y., Ishii, K., Kumeta, M., Horigome, T., and Takeyasu, K. (2009). Cell cycle-dependent phosphorylation of MAN1. Biochemistry *48*, 1636-1643.

Hirano, Y., Segawa, M., Ouchi, F.S., Yamakawa, Y., Furukawa, K., Takeyasu, K., and Horigome, T. (2005). Dissociation of emerin from barrier-to-autointegration factor is regulated through mitotic phosphorylation of emerin in a xenopus egg cell-free system. J. Biol. Chem. *280*, 39925-39933.

Jamin, A., and Wiebe, M.S. (2015). Barrier to Autointegration Factor (BANF1): interwoven roles in nuclear structure, genome integrity, innate immunity, stress responses and progeria. Curr. Opin. Cell Biol. *34*, 61-68.

Kielkopf, C.L., Lucke, S., and Green, M.R. (2004). U2AF homology motifs: protein recognition in the RRM world. Genes & Development *18*, 1513-1526.

Konde, E., Bourgeois, B., Tellier-Lebegue, C., Wu, W., Perez, J., Caputo, S., Attanda, W., Gasparini, S., Charbonnier, J.B., Gilquin, B., *et al.* (2010). Structural analysis of the Smad2-MAN1 interaction that regulates transforming growth factor-beta signaling at the inner nuclear membrane. Biochemistry *49*, 8020-8032.

Laguri, C., Gilquin, B., Wolff, N., Romi-Lebrun, R., Courchay, K., Callebaut, I., Worman, H.J., and Zinn-Justin, S. (2001). Structural characterization of the LEM motif common to three human inner nuclear membrane proteins. Structure *9*, 503-511.

Lee, K.K., Haraguchi, T., Lee, R.S., Koujin, T., Hiraoka, Y., and Wilson, K.L. (2001). Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF. J. Cell Science *114*, 4567-4573.

Li, M., and Craigie, R. (2006). Virology: HIV goes nuclear. Nature *441*, 581-582.

Lin, F., Blake, D.L., Callebaut, I., Skerjanc, I.S., Holmer, L., McBurney, M.W., Paulin-Levasseur, M., and Worman, H.J. (2000). MAN1, an inner nuclear membrane protein that shares the LEM domain with lamina-associated polypeptide 2 and emerin. J. Biol. Chem. *275*, 4840-4847.

Lin, F., Morrison, J.M., Wu, W., and Worman, H.J. (2005). MAN1, an integral protein of the inner nuclear membrane, binds Smad2 and Smad3 and antagonizes transforming growth factor-beta signaling. Human Molecular Genetics *14*, 437-445.

Mansharamani, M., and Wilson, K.L. (2005). Direct binding of nuclear membrane protein MAN1 to emerin in vitro and two modes of binding to barrier-to-autointegration factor. J. Biol. Chem. *280*, 13863-13870.

Margalit, A., Segura-Totten, M., Gruenbaum, Y., and Wilson, K.L. (2005). Barrier-to-autointegration factor is required to segregate and enclose chromosomes within the nuclear envelope and assemble the nuclear lamina. Proc. Natl. Acad. Sci. U.S.A. *102*, 3290-3295.

Montes de Oca, R., Andreassen, P.R., and Wilson, K.L. (2011). Barrier-to-Autointegration Factor influences specific histone modifications. Nucleus *2*, 580-590.

Naetar, N., Korbei, B., Kozlov, S., Kerenyi, M.A., Dorner, D., Kral, R., Gotic, I., Fuchs, P., Cohen, T.V., Bittner, R., *et al.* (2008). Loss of nucleoplasmic LAP2alpha-lamin A complexes causes erythroid and epidermal progenitor hyperproliferation. Nature Cell Biol. *10*, 1341-1348.

Pan, D., Estevez-Salmeron, L.D., Stroschein, S.L., Zhu, X., He, J., Zhou, S., and Luo, K. (2005). The integral inner nuclear membrane protein MAN1 physically interacts with the R-Smad proteins to repress signaling by the transforming growth factor-{beta} superfamily of cytokines. J. Biol. Chem. *280*, 15992-16001.

Patel, J.T., Bottrill, A., Prosser, S.L., Jayaraman, S., Straatman, K., Fry, A.M., and Shackleton, S. (2014). Mitotic phosphorylation of SUN1 loosens its connection with the nuclear lamina while the LINC complex remains intact. Nucleus *5*, 462-473.

Sakaki, M., Koike, H., Takahashi, N., Sasagawa, N., Tomioka, S., Arahata, K., and Ishiura, S. (2001). Interaction between emerin and nuclear lamins. J. Biochemistry *129*, 321-327.

Shin, J.Y., Mendez-Lopez, I., Wang, Y., Hays, A.P., Tanji, K., Lefkowitch, J.H., Schulze, P.C., Worman, H.J., and Dauer, W.T. (2013). Lamina-associated polypeptide-1 interacts with the muscular dystrophy protein emerin and is essential for skeletal muscle maintenance. Developmental Cell *26*, 591-603.

Shumaker, D.K., Lee, K.K., Tanhehco, Y.C., Craigie, R., and Wilson, K.L. (2001). LAP2 binds to BAF.DNA complexes: requirement for the LEM domain and modulation by variable regions. EMBO J. *20*, 1754-1764.

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

Taylor, M.R., Slavov, D., Gajewski, A., Vlcek, S., Ku, L., Fain, P.R., Carniel, E., Di Lenarda, A., Sinagra, G., Boucek, M.M., *et al.* (2005). Thymopoietin (lamina-associated polypeptide 2) gene mutation associated with dilated cardiomyopathy. Human Mutation *26*, 566-574.

Tifft, K.E., Bradbury, K.A., and Wilson, K.L. (2009). Tyrosine phosphorylation of nuclear-membrane protein emerin by Src, Abl and other kinases. J. Cell Science *122*, 3780-3790.

Umland, T.C., Wei, S.Q., Craigie, R., and Davies, D.R. (2000). Structural basis of DNA bridging by barrier-to-autointegration factor. Biochemistry *39*, 9130-9138.

Vlcek, S., Just, H., Dechat, T., and Foisner, R. (1999). Functional diversity of LAP2alpha and LAP2beta in postmitotic chromosome association is caused by an alpha-specific nuclear targeting domain. EMBO J. *18*, 6370-6384.

Wolff, N., Gilquin, B., Courchay, K., Callebaut, I., Worman, H.J., and Zinn-Justin, S. (2001). Structural analysis of emerin, an inner nuclear membrane protein mutated in X-linked Emery-Dreifuss muscular dystrophy. FEBS Lett. *501*, 171-176.

Yip, S.C., Cotteret, S., and Chernoff, J. (2012). Sumoylated protein tyrosine phosphatase 1B localizes to the inner nuclear membrane and regulates the tyrosine phosphorylation of emerin. J. Cell Science *125*, 310-316.

Zheng, R., Ghirlando, R., Lee, M.S., Mizuuchi, K., Krause, M., and Craigie, R. (2000). Barrier-to-autointegration factor (BAF) bridges DNA in a discrete, higher-order nucleoprotein complex. Proc. Natl. Acad. Sci. U.S.A. *97*, 8997-9002.

b. LBR

LBR, for Lamin-B receptor, was first described in a publication from the Blobel laboratory, in 1988[63,64]. It is an inner nuclear membrane protein with two functions: its N-terminal nucleoplasmic region, which contains a TUDOR domain (figure 12a), interacts with both chromatin and the nuclear lamina, while its C-terminal multi-pass transmembrane domain is a C14 sterol reductase (figure 12b). Mutations in LBR are associated with rare human diseases, Pelger-Huët anomaly and Greenberg skeletal dysplasia, which cause harmless anomalies of blood cells to fatal developmental defects[66–68]. The severity of these diseases depends on the nature of the specific mutation, and whether one or both copies of the gene are affected. All these diseases are characterized by a defect in cholesterol synthesis due to loss of LBR-associated sterol C14 reductase activity[68]. Several disease-associated LBR point mutations reduce sterol C14 reductase activity by decreasing the affinity of LBR for the reducing agent NADPH. Other mutations lead to LBR truncations.

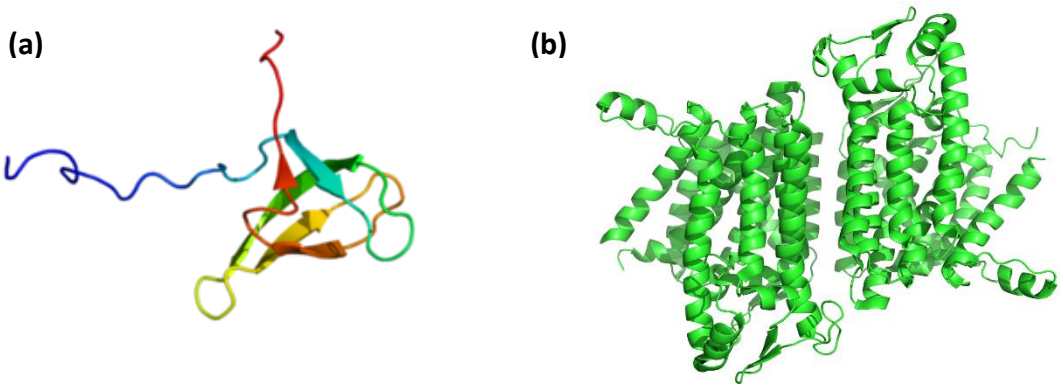


Figure 12 : Architecture of LBR.

(a) 3D structures of the TUDOR domain of LBR (PDB code: 2L8D)[70] (b) and the integral membrane region of a LBR homolog (PDB code: 4QUV)[71].

LBR was initially characterized because of its ability to bind to the nuclear lamin B, and more precisely, its capacity to mediate interaction between chromatin and lamin B. However, it was shown that the TUDOR domain of LBR is not sufficient for binding to lamins[69]. It is able to bind to free histone H3. It also recognizes H4 dimethylated on lysine 20 and is essential for chromatin compaction[70] (figure 13). A specific motif in the N-terminal nucleoplasmic region of LBR directly binds to the chromo-shadow domain of the chromatin-associated protein HP1α which interacts with histones[71,72] (figure 13).

It was reported that LBR and lamin A/C are responsible for two different mechanisms that tether heterochromatin to the NE[73]. These two mechanisms are sequentially used during cellular differentiation and development. The absence of both LBR and lamin A/C leads to loss of peripheral heterochromatin and an inverted architecture with heterochromatin localizing to the nuclear interior.

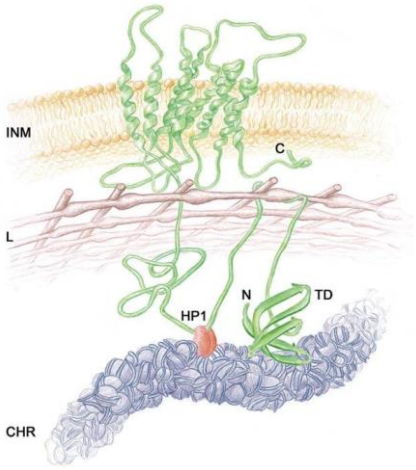


Figure 13 : Cartoon representation of LBR embedded within the nuclear envelope[74].

c. The LINC complex

Nuclear positioning is essential for many cellular processes, as cell division, migration, differentiation and polarization. In eukaryotes, this process requires the activity of actin, microtubules and microtubule-dependent motors. However, several studies revealed that another complex is fundamental to provide a structural support to the nucleus: the LInker of Nucleoskeleton and Cytoskeleton complex (LINC complex).

The LINC complex is conserved from yeast to human. It is a protein complex associated with the inner and outer membranes of the nucleus. It is composed of ONM-resident KASH-domain proteins (Klarsicht, ANC-1 and SYNE/Nesprin-1 and 2 Homology) and INM-resident SUN-domain proteins (Sad1 and UNC-84)[75,76]. Defects in these LINC complex proteins are associated with a broad spectrum of muscle pathologies (e.g., Emery-Dreifuss muscular dystrophy and dilated cardiomyopathy) and neuronal disorders[77–79].

KASH-domain proteins are membrane proteins that share several common features[76]. There are composed of large cytoplasmic domains extending into the cytoplasm and interacting with actin filaments, microtubules, intermediate filaments or centrosomes. They possess a single transmembrane helix at their N-terminus, followed by a short fragment of around 30 residues, containing a PPPX motif (X is a variable residue) and located in the perinuclear space. The transmembrane helix together with the luminal tail form the KASH domain, a characteristic signature of the family[80]. Six KASH proteins are known in humans. Four of them, called Nesprins (1-4), contain spectrin repeats in their cytoplasmic domains. Nesprin 1 and 2, in mammals, are actin-binding proteins and are characterized by the fact that they are very large, with extensive fibers that span the cytoplasm[81–83]. Nesprin-3, a shorter molecule, binds to actin too, but through binding to a giant protein, plectin[82]. The last known Nesprin, Nesprin-4, is found in lower abundance, only in a few cell types, and is a microtubule-binding protein[83].

Sad1/UNC-84 or SUN proteins were originally described in fission yeast (Sad1) and *Caenorhabditis elegans* (UNC-84)[84,85]. Afterwards, five mammalian family members (SUN1-5) have been identified[86]. Whereas SUN3, SUN4 (SPAG4), and SUN5 seem to be specifically expressed in testicles, SUN1 and SUN2 are ubiquitously expressed and can bind to all four Nesprins[87,88]. Sun proteins contain a variable nucleoplasmic region at their N-terminus, followed by a single transmembrane domain that spans the inner nuclear membrane and a region, localized into the perinuclear space, which comprises two coiled-coil repeats (CC1 and CC2) that are shown to be essential for trimerization and are also believed to act as rigid spacers to delineate the distance between the inner and the outer nuclear membranes[80,89]. Finally, the most highly conserved region between family members is the C-terminal SUN-domain (figure 14, a)[90].

The 3D structure of a part of the LINC complex has recently been solved by X-ray crystallography[90,91] (figure 14, b). This structure contains the SUN domain of SUN-2, which folds into a compact β-sandwich. It exhibits at its N-terminus a helical extension (residues 525–540) involved in the formation of a coiled-coil, followed by 20 residues (residue 567 to 587) that are assembled in order to form a flexible domain, called the KASH-lid[92]. Residues 593–601 form a well-defined loop, which surrounds and coordinates a bound cation, and is structured by an intrachain disulfide bond formed between conserved cysteines at positions 601 and 705.

It was shown that this SUN2 fragment forms a stable trimer and binds to the KASH domain of nesprin[90,91]. The SUN homotrimer serves as a platform to interact with three KASH peptides. The peptides are individually bound in three deep grooves, which are symmetrically distributed close to the trimeric interface of the SUN domain homotrimer (figure 14, c). It is interesting to observe that addition of KASH peptide does not induce large structural changes in the SUN domain. Indeed, when the structure of apo-SUN-2 is superimposed with KASH-SUN-2 structure, only the region around the KASH-lid shows significant changes.

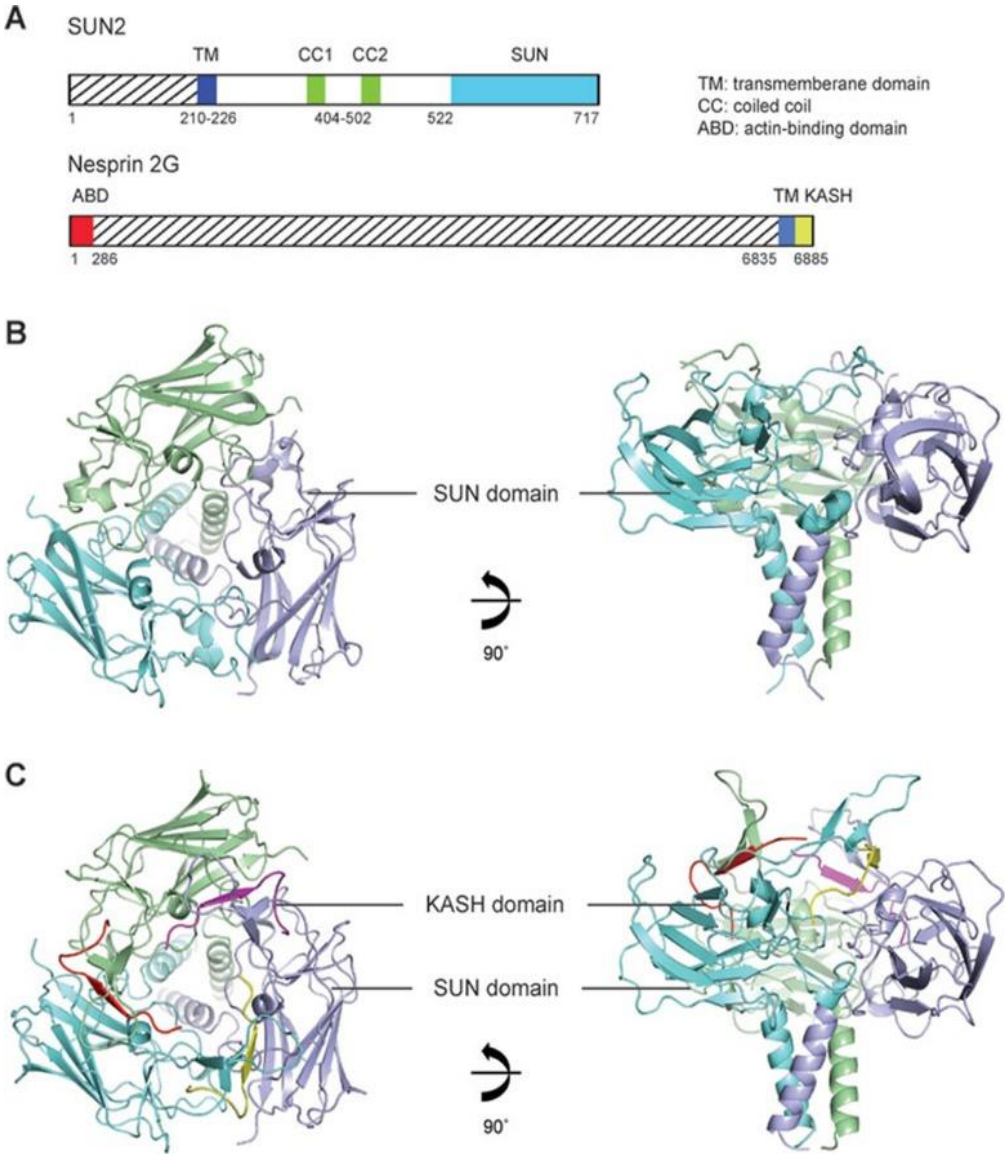


Figure 14 : Structure of free and complexed SUN2[90].

(A) Schemes of the domain organization for human SUN2 and Nesprin 2G. (B) The SUN domain homotrimer side and top views represented in cartoon (PDB: 3UNP). The three protomers of the SUN homotrimer are colored in cyan, light blue and pale green. (C) Top and side views of the SUN-KASH complex. The three KASH domain peptides are colored in yellow, magenta and red.

Unexpectedly, when the two coiled-coil regions that are located before the SUN domain were studied separately, it was showed that, while the first coiled-coil is indeed trimeric, the second coiled-coil alone is monomeric[93]. The 3D structure of the coiled-coil 2 and SUN domain revealed that this coiled coil can fold into a three helix bundle and lock the SUN domain in a monomeric and inactive conformation (figure 15). Thus, structural changes within coiled-coil 2 dictate the oligomeric state of SUN2 and thus regulate the formation of the LINC complex.

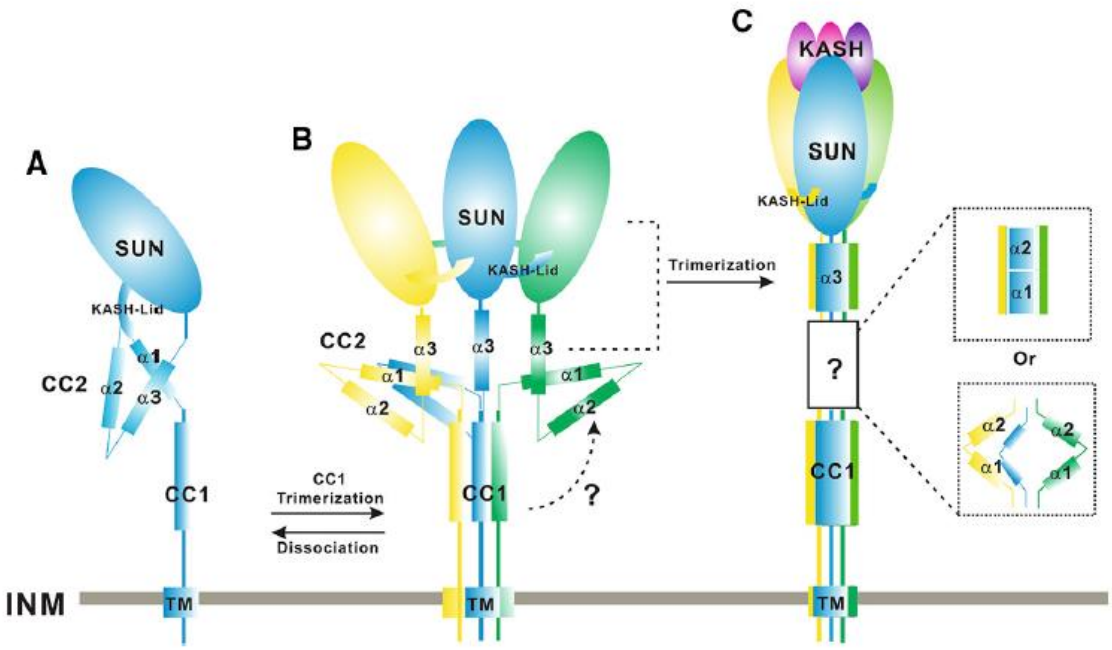


Figure 15 : Structural insights into SUN-KASH complexes across the nuclear envelope[93].

In the monomeric form of the SUN2 protein, coiled-coil 2 forms a three-helix bundle that interacts with the SUN domain and lock it in an inactive state. SUN2 trimerization mediated by coiled-coil 1 probably induces a conformational change within coiled-coil 2. This change releases helix 3 of coiled-coil 2 that is then able to trimerize, thus triggering the formation of three KASH binding sites on the SUN2 trimer.

II. NUCLEAR ENVELOPE: FUNCTIONS

1. STRUCTURAL AND MECHANICAL PROPERTIES OF THE NUCLEUS

The nuclear envelope provides sturdiness to the nucleus and this function may be more or less important depending on the tissue. Indeed, stiffness and strength of a tissue are related to the physical stress sensed in that tissue. For example, brain or fat tissues bear little stress, while adult bone tissue sustains high stress, which promotes bone growth and stiffening to match the stress[94].

In this part, I will first describe the implication of the nuclear lamina, the LINC complex, nuclear actin and emerin in the mechanical properties of the nucleus and in the cellular response to a mechanical stress. Then, I will present the implication of another protein, BAF, in nuclear shape maintenance through regulation of nuclear envelope and chromatin interaction.

a. Lamin composition determines the nucleus stiffness

Lamins contribute to nuclear stiffness[95]. In lamin-depleted Xenopus egg extracts, assembled nuclei are highly fragile[96]. Cells from mouse *LMNA*$^{-/-}$ were found to be mechanically weak[97]. Later, direct mechanical measurements on Xenopus oocyte nuclei confirmed that lamina acts as a stiff[95]. In addition, mass spectrometry analyses showed that elasticity of bulk tissue is strongly correlated to the composition of the nuclear lamina (figure 16). Indeed, the ratio of A-type lamin to B-type lamin is increasing with the stiffness of the tissue. This observation could explain the fact that only vertebrates possess A-types of lamins[94].

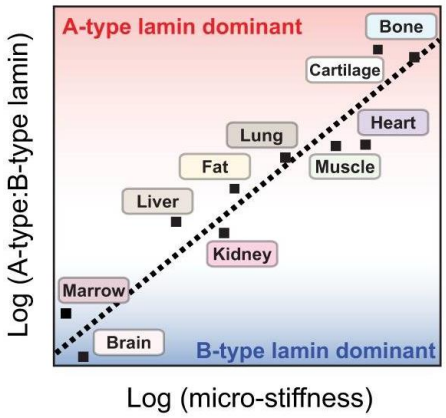


Figure 16 : Relationship between lamin composition and tissue rigidity.

A logarithmic scale representation of the ratio of A-type lamin (in red) to B-type lamin (in blue) in tissues according to their rigidity[96].

Moreover, lamins have been shown to be regulated during mechanical stress. Lamin implication in nucleus mechanical properties was first studied in nuclei with different lamina compositions, by micropipette aspiration experiments[95,98]. It was shown that B-type lamins contribute to the elastic response whereas A-type lamins contribute to the viscosity. Afterwards, lamins behavior during mechanical stress was studied in mesenchymal stem cells (MSCs) that were cultivated in a medium mimicking the extra cellular matrix of either brain or bone (figure 17). This experiment showed that in the nuclear envelope of cells grown on soft substrates, lamins are wrinkled and relaxed, whereas, on stiff substrates, lamins appear smooth. On stiff substrates, the rate of phosphorylation at four sites decreases. Phosphorylation of intermediate filament proteins modulates their solubility, conformation and organization. Lamins are known to be highly phosphorylated during mitosis, in order to be disassembled. We can do the hypothesis that decreasing phosphorylation leads to a diminution of A-type lamin solubility and an increase of lamin strengthening[94]. More than dephosphorylation, a partial unfolding of the Igfold domain could be triggered by a mechanical stress. In contrary, on soft substrates, A-type lamins are more extensively phosphorylated, more soluble and could be eventually degraded.

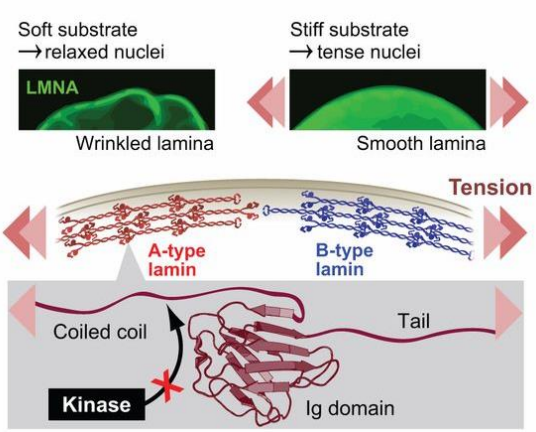


Figure 17 : Impact of mechanical stress on the Igfold domain structure[94].

Nuclei of MSCs were cultured on soft and stiff substrates. Cysteine-shotgun MS was used to identify cysteines that are exposed by stress. Cys522 of A-type lamin Igfold was identified as a stress-sensitive site.

Finally, migration of the cells in complex environments requires cellular deformability, which is limited by the cell nucleus. Indeed, lamins form a rigid shell underneath the inner nuclear membrane. To overcome this limitation, some migrating cells express a low level of lamins, whereas others secrete proteases to enlarge pores in the extracellular matrix[99]. It was recently shown that dendritic cells, when migrating through narrow channels, answer to nuclear rigidity by rapidly assembling a branched actin network around the nucleus in an Arp2/3 and Wave2 dependent manner[100]. This triggers local ruptures of the lamin A/C shell. If lamin A/C expression is suppressed, then this mechanism is no longer necessary. Formins, which are actin nucleating proteins generating unbranched actin filaments, also contribute to the formation of perinuclear actin filaments. In particular, FMN2 associates with and generates a perinuclear actin system that controls nuclear shape and positioning in cells migrating on 2D surfaces[101]. This actin-based cage protects the nucleus and its contents from damage when cells migrate through tiny spaces. It may prevent nuclear envelope rupture or facilitate nuclear envelope repair by either ESCRT III or lamin-dependent[102] or –independent pathways.

b. Interaction between the LINC complex and actin contributes to nucleus positioning

Nuclear envelope proteins together with actin control the position of the nucleus in the cell, in order to regulate cell division, differentiation and again cell migration[103]. In particular, actin is an essential actor of the rearward movement of the nucleus to orient the centrosome and polarize cells for directed migration. Actin-dependent nuclear movement in fibroblasts and myoblasts involves a LINC complex composed of nesprin-2G and SUN2. The association of these proteins with dorsal actin cables results in their assembly into linear arrays termed transmembrane actin–associated nuclear (TAN) lines[104]. Interaction between actin and the outer nuclear envelope involves the CH domains of nesprin-2G. The formin family member FHOD1 provides a second site of interaction between the actin cable and nesprin-2G[105]. The KASH domain of nesprin-2G interacts with the SUN domain of SUN2 in the perinuclear space. Finally TAN lines are positioned because of association of SUN2 with A-type lamins[106]. Emerin is also found in TAN lines and may contribute to their anchoring.

c. *Emerin mediates nuclear response to a mechanical stress*

Emerin, a protein which belongs to the group I of LEM-domain proteins, is implicated in the nucleus response to a mechanical stress, in particular through phosphorylation events, like lamins. Analysis of the role of nuclear envelope proteins in the nucleus response to a mechanical stress was done though application of a force on isolated nuclei via the LINC complex component Nesprin-1[107]. Magnetic tweezers were used to stimulate magnetic beads that were coated with anti Nesprin-1 antibodies. Successive pulses with a constant force were applied on the nucleus and bead displacement diminution was observed which means that application of a constant force leads to an increase of the nucleus resistance.

Because induction of protein phosphorylation is one of the first events that appears when mechanical force is applied to cells, tyrosine phosphorylation of nuclear proteins from isolated nuclei subjected to force was observed. Several nuclear proteins were phosphorylated after mechanical stress, but one of these proteins was more strongly phosphorylated on its tyrosines: it is the emerin protein. More precisely, emerin phosphorylation on two tyrosines (Y74 and Y95) by the kinase Src was essential for the response of the nucleus to a mechanical stress (figure 18). When these tyrosine were mutated, the nucleus could not answer anymore to force application. However the molecular consequences of these tyrosine phosphorylations are still unknown.

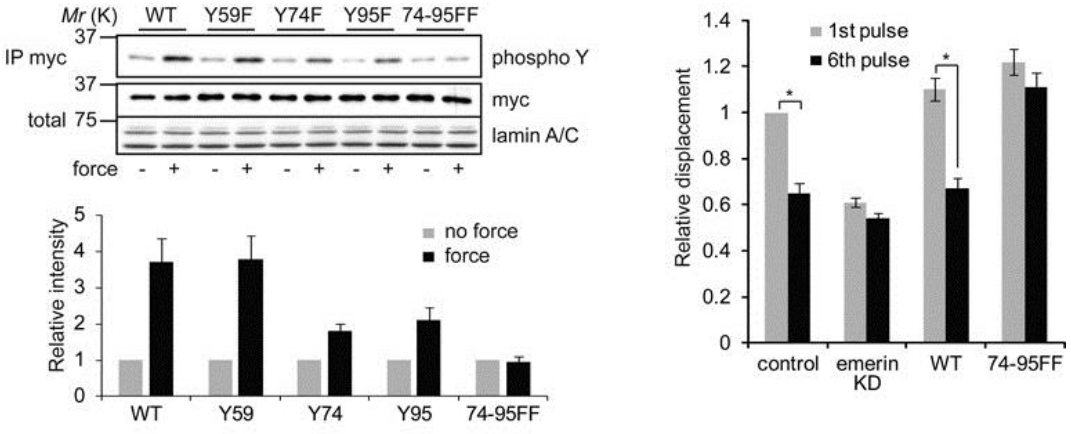


Figure 18 : Emerin is phosphorylated by Src on its tyrosines 74 and 95 in response to a mechanical stress[107].

d. BAF mediates interaction between nuclear envelope and chromatin

The small protein barrier-to-autointegration-factor (BAF) was identified in a yeast two-hybrid screen designed to find partners of the inner nuclear membrane protein LAP2[56]. BAF is a highly conserved 89-amino acid protein that is localized in the cytoplasm and the nucleoplasm of metazoan cells[108,109]. It forms a dimer in solution. The 3D structure of this dimer was solved by NMR[110] (PDB: 2EZX) and by X-ray crystallography[111] (PDB: 2BZF) (figure 19).

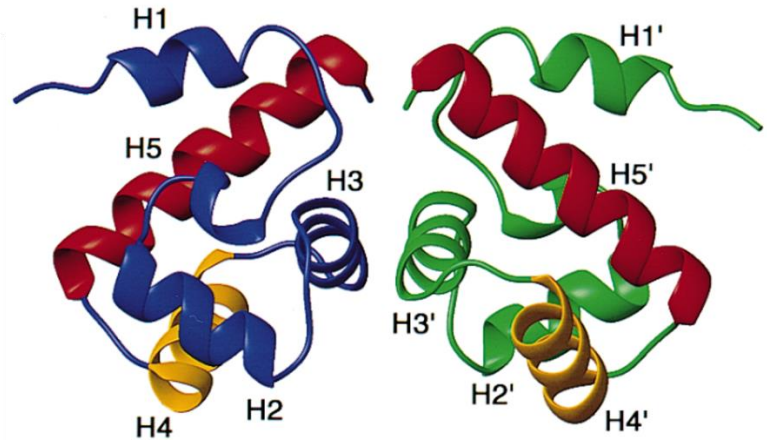


Figure 19 : Solution structure of BAF dimer, as solved by NMR (PDB: 2EZX)[110].

Ribbon diagram of the BAF dimer, with first and second helices of the DNA recognition helix-turn-helix motif in orange and red, respectively and other helices in blue for one subunit and green for the other.

Each monomer of BAF is mainly composed of five helices (H1-5). The presence of one copy and one pseudo copy of a nonspecific DNA-binding motif (helix-hairpin-helix, HhH) was observed for each monomer. This motif is characterized by two helices connected by a short turn and is formed by helices H2 and H3 and helices H4 and H5. The 3D structure of the BAF dimer in interaction with a 7-bp-DNA was then solved by X-ray crystallography, at a resolution of 2.87 Å[111]. This structure shows that the interaction is mediated by both HhH motifs (figure 20) and the N-terminal of the helix 1 (in yellow in figure 20). In the presence of a short DNA, each BAF monomer is able to bind to one DNA molecule, whereas it was shown before that BAF is able to form higher-order complexes in presence of longer DNA molecules[112].

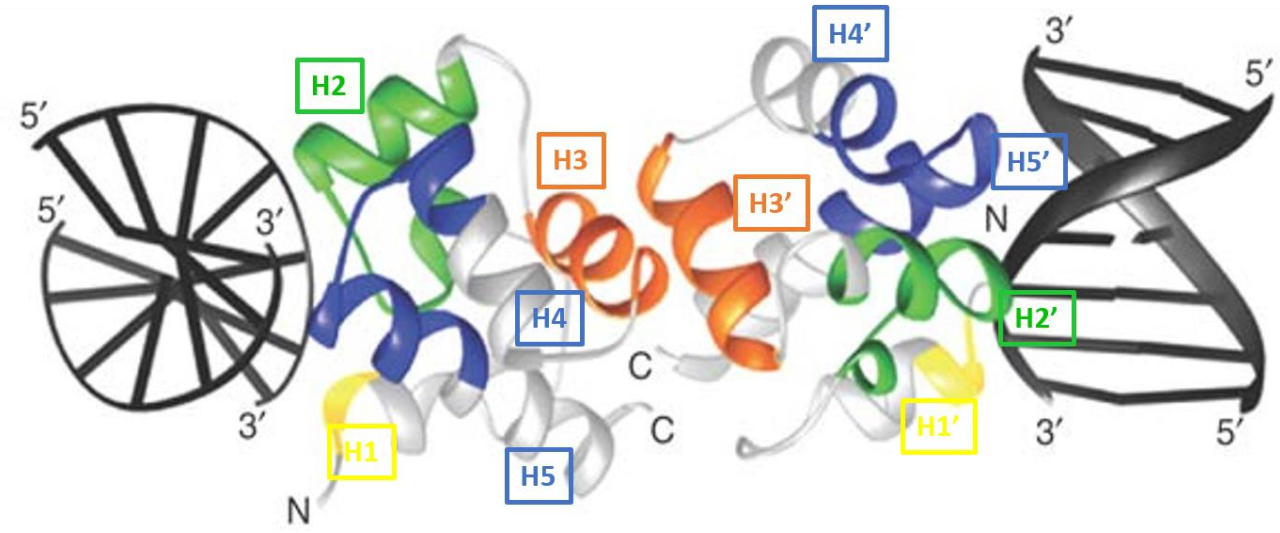


Figure 20 : Structure of the BAF dimer in interaction with a 7-bp-DNA, solved by X-ray crystallography (PDB: 2BZF)[111].

The HhH and pseudo HhH motifs are formed by helices H2 (in green) and helices H4 and H5 (in blue), respectively, with the connecting helix H3 (in orange). In yellow is represented the interacting region of α1; black, DNA; N and C, protein termini.

A first link between BAF function and chromatin was found several years ago by a group who did an assay to identify and purify the host factor conferring protection to the retroviral DNA against auto-integration[113]. More precisely, BAF was found as an essential actor for the biogenesis of MoMLV and HIV-1 retroviruses. For example, during biogenesis of MoMLV, BAF is recruited to cytoplasmic pre-integration complexes (PICs), in order to block autointegration of the viral DNA, whereas in absence of BAF, self-destruction of viral DNA is observed. BAF seems to protect DNA against integration by bridging DNA.

The most important role of BAF is its capacity to interact with chromatin in order to regulate nuclear assembly and organization. Indeed, *in vivo* studies have demonstrated that depletion of BAF from *C. elegans* embryos[112,114] or *D. melanogaster* larvae[115] induce strong effects on chromosome condensation and segregation, notably the retention of mitotic chromatin modification during late anaphase and telophase. In addition, a direct effect of BAF on nuclear envelope assembly and function was observed. Indeed, depletion of BAF using short RNAi resulted in nuclear shape and size defects[116] and moreover, this depletion compromised severely chromatin segregation. Furthermore, protein kinase VRK-1 phosphorylates BAF *in vivo* to regulate its localization and function[116].

Phosphorylation of BAF not only abrogates its DNA binding activity by introducing negative charges on the DNA interacting surface, but reduces the binding to LEM domain proteins and leads to mislocalization from the newly formed NE[117]. Co-expression of VRK1 and GFP-BAF greatly diminishes the association of BAF with the nuclear chromatin/matrix and leads to its dispersal throughout the cell.

On the opposite, depletion of VRK-1 by RNAi for 13h causes strong accumulation of BAF on chromatin throughout mitosis, which clearly indicates that VRK-1 activity is required for BAF to dissociate from mitotic chromatin. In addition, an increase in chromatin association of BAF in anaphase was observed in absence of VRK-1 in embryo cells. Finally, more than dissociation from mitotic chromatin, depletion of VRK1 and the consequent reduction in BAF phosphorylation caused abnormalities in the nuclear envelope of interphase cells. One explanation could be that an increased fraction of unphosphorylated BAF could extend the duration of interactions between BAF and its DNA or protein partners which could be responsible for the distorted structure of the nuclear envelope[118].

Analysis of BAF localization during the cell cycle in HEK293 cells confirmed that during interphase, total BAF and phosphorylated BAF are present in both the cytoplasm and the nucleus[119]. In early telophase, phosphorylated BAF accumulates at the central region of chromosomes, and in late telophase, it already dispersed to the cytoplasm, while total BAF translocates to the periphery of chromosomes. Thus, unphosphorylated BAF is mainly localized at the periphery of chromosomes through telophase, where the new NE is formed. Emerin localizes to the central region of chromosomes in a BAF-dependent manner[31]. BAF phosphorylation controls the recruitment of emerin to the central region of telophase chromosomes, which is essential for new NE assembly.

In addition to its role in nuclear envelope association around chromosomes, BAF was recently found as a spindle factor for shaping a single nucleus instead of an ensemble of individualized mitotic chromosomes[120]. Indeed, a micronucleation phenotype was caused by a siRNA targeting BAF. Moreover, BAF mediated DNA cross-bridging in order to form a network at the chromatin surface and this increased the mechanical resistance of chromatin. Finally, they showed that this chromatin compaction could be reversed by addition of recombinant VRK1. In conclusion, as it is represented on the figure 21, BAF-mediated DNA cross-bridging allows formation of a compact and mechanically stiff chromatin surface that shapes a single nucleus and this BAF-DNA interaction is disrupt by BAF phosphorylation by VRK1 at the early stages of mitosis[120].

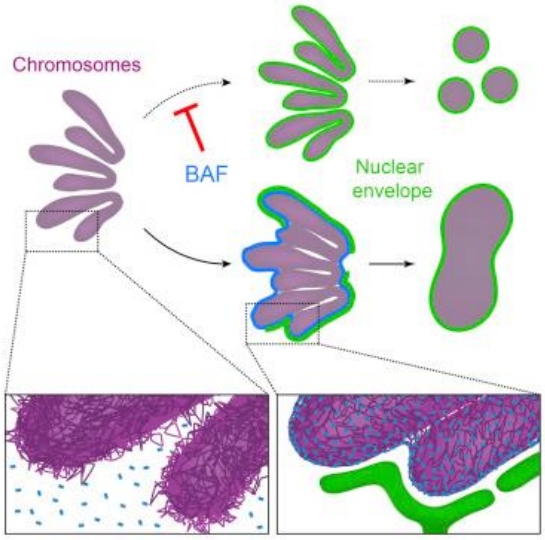


Figure 21 : Model of the BAF-DNA network forming at the surface of anaphase chromosomes prior to nuclear envelope reassembly.

BAF, in blue, increases chromatin (in purple) surface stiffness and forces formation of a single nucleus. Nuclear envelope is labelled in green[122].

2. GENE EXPRESSION

Since several years, different studies focused on the organization of chromatin at the nuclear periphery (heterochromatin) (figure 22). Its nature and its role differ from those of chromatin located in the nuclear interior (euchromatin)[121]. Indeed, chromatin at the nuclear periphery is compact and transcriptionally silent. Inversely, internal chromatin is unpacked and transcriptionally active. The association of a promoter with the inner nuclear membrane (INM) is neither necessary nor sufficient for repression. So how does subnuclear localization influence gene expression? Recent work argues that the common denominator between genome organization and gene expression is the modification of histones and in some cases of histone variants[122].

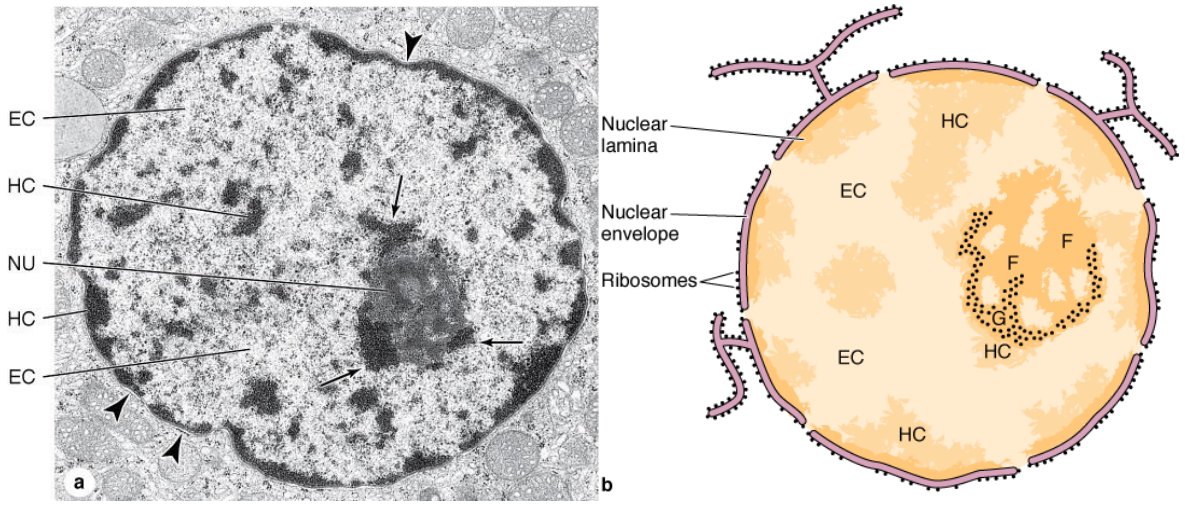


Figure 22 : Organization of the chromatin at the nuclear periphery[123].

(a) Structure of euchromatin (EC), heterochromatin (HC) and nucleolus (NU) observed by electron microscopy (b) Schematic representation of the nucleus and of the nucleolus (with the pars granulosa (G) and pars fibrosa (F)).

The nuclear envelope is a tethering point for chromatin[124,125]. Several nuclear envelope components are important for this function, as the mammalian nuclear pore complexes, composed of elements that can be separated in two groups, some which could activate gene expression whereas others could have a silencing function[126,127] and the lamina, which is close to compact heterochromatin.

a. *Nuclear pore complexes and transcription activation*

NPCs have been implicated in chromatin organization and gene regulation[128,129]. Studies in yeast revealed that Nups can associate with promoters of active genes[130] and that the expression of inducible genes is increased by interactions with nuclear pores[131]. For example, *in yeast*, the nucleoporin Nup84 of pore complexes and its subcomplex components are able to activate transcription when they are fused to a DNA-binding domain[132]. In the absence of Nup84, transcriptional activation by most nucleoporins is impaired.

However, a genome-wide study of Nup-chromatin association in human cells revealed a correlation between the binding sites of Nups and regions enriched in repressive histone modifications[133], which exhibited characteristics of sequences known to associate with the nuclear periphery[134]. In higher eukaryotes, some nucleoporins have two separate pools: one activates genes in the nucleoplasm while the other in the peripheral NPC structures has a silencing function[126,127].

b. *Lamins and gene silencing*

Nuclear periphery has been implicated in gene expression regulation. Molecular mapping of chromatin regions associated to the lamina identified large genomic regions called lamina-associated domains (LADs)[135]. Such LADs are proposed to help organize chromosomes inside the nucleus and have been associated with gene repression. Furthermore, localization of these LADs close to the nuclear lamina is facilitated by the presence of lamina-associated sequences (LAS), which contain the GAGA motif[136]. This repeated motif directs lamina association and is bound by the transcriptional repressor cKrox, in a complex with HDAC3 and Lap2β. Knockdown of cKrox or HDAC3 results in dissociation of LASs/LADs from the nuclear lamina.

In both mammals and nematodes, chromatin associated with LADs is densely packed and enriched for repressive histone modifications, most notably H3K9me2/me3[122,134,137]. In nematodes, elimination of two histone methyltransferases, MET-2 and SET-25, compromised INM binding[138].

In metazoans, LEM-domain proteins are found in regions enriched in repressive epigenetic marks and which contain all features that define LADs[53]. This suggests that LEM-domain proteins may contribute to LAD establishment or maintenance. In addition, proteins which promote the formation of repressed chromatin may contribute to LAD maintenance. Two LEM-domain proteins, emerin[139] and LAP2β[140] are able to bind the histone deacetylase HDAC3, which allows the deacetylation of histone H4, in order to inhibit transcription activity. This confirms that LEM-domain proteins can maintain LADs by recruiting partners that contribute to repressed chromatin states.

c. BAF and gene expression

BAF's roles in transcriptional regulation are major open questions, and might involve interactions with a variety of players including transcription factors, chromatin modifiers and nuclear lamina components. A first example is BAF association with Sox2 in embryonic stem cells that could contribute to the formation or regulation of Sox2 complexes to maintain pluripotency[141]. Another example is found in the photoreceptor development and function field. Indeed, BAF was identified in a yeast two-hybrid screen that was designed to identify partners of the homeodomain transcription factor Crx, implicated in regulating the expression of photoreceptor and pineal genes[142]. BAF is able to repress the Crx-dependent reporter activity *in vivo* by directly interacting with Crx[142]. This study also shows that Crx is not the only transcription factor that can bind to BAF, but four other paired-like homeodomain proteins (Chx10, Pax-6, Otx1, and Otx2) also have the ability to interact with BAF *in vitro*.

3. GENOMIC INSTABILITY AND DNA DAMAGE RESPONSE

Genomic instability is defined as a tendency of the genome to acquire mutations and epimutations as well as alterations in gene or chromosome dosage. The origin of this phenomenon is none other than failures in processes important for genome stability and replication[143]. DNA damage is the result of constant attack of the genome by endogenous and exogenous agents. It can result from side products of our normal metabolic activities as free radicals and reactive oxygen (ROS) or reactive nitrogen species; it can also result from environmental factors such as UV radiation, X-rays and chemical compounds. Furthermore, DNA lesions might be due to deficiencies in DNA replication or loss of telomere function. The most deleterious kinds of lesions, which are DNA double-strand breaks, can be repaired by nonhomologous end-joining (NHEJ) or homologous recombination (HR).

Several lines of evidence have linked laminopathies with increased genomic instability[143]. In particular, expression of several A-type lamin mutants has been associated with impairment in the ability of cells to properly repair DNA damage and maintain telomere homeostasis. Defects in DNA repair were first reported in the premature aging laminopathy HGPS and the *Zmpste24*$^{-/-}$ mouse model of progeria[144]. Later studies indicated that accumulation of unrepairable DNA damage in laminopathies such as HGPS and Restrictive Dermopathy is in part due to elevated levels of ROS and greater sensitivity to oxidative stress[145]. A whole variety of molecular mechanisms have been proposed to contribute to the defects in DNA repair and the increased genomic instability in progeria.

From the observation that progeria cells exhibit delayed recruitment of 53BP1 and RAD51 to γ-H2AX-labeled DNA repair foci after irradiation, the role of lamins in HR and NHEJ pathways was studied. The role of lamins in telomere maintenance was also examined.

a. A-type lamins and 53BP1

P53-binding protein 1 (53BP1) is an important regulator of the cellular response to DNA double strand breaks (DSBs) that can bind to chromatin close to DSBs in order to carry out several functions[146].

In NHEJ processes, 53BP1 was shown to be important because it participates in the repair of short-range DNA DSBs by inhibiting end-resection and facilitating the recruitment of the NHEJ DNA repair machinery.

A-type lamins are known to stabilize this protein and one of the mechanisms behind genomic instability in A-type lamins-deficient cells is the loss of the DNA repair factor 53BP1. This was observed in several studies and notably through experiments showing that depletion in A-type lamins by lentiviral transduction with shRNAs leads to decreased levels of the 53BP1 protein but not of its transcript levels[143].

A search for mechanisms by which A-type lamins affect DNA DSB repair revealed a role for cathepsin L (CTSL) in regulating 53BP1 protein stability[143]. CTSL is a cysteine protease ubiquitously expressed in tissues that was recently shown to be able to localize in the nucleus instead of its classical localization in the lysosomes. It was found that loss of A-type lamins induces transcriptional upregulation of CTSL or increased stability of its transcripts. Moreover, increase in CTSL is directly responsible for the decrease in 53BP1 protein levels so that A-type lamins seems to stabilize 53BP1 by inducing transcriptional downregulation of CTSL.

In addition, in progeria cells, loss of components of the NuRD complex and decreased levels of the histone acetyltransferase Mof were observed and associated with impaired recruitment of DNA repair factors and increased basal DNA damage[147,148]. This decrease in Mof levels correlates with lower global levels of H4K16 acetylation, a histone mark associated with chromatin compaction, and decreased recruitment of DNA repair factors such as 53BP1. The importance of this H4K16 acetylation is demonstrated by the restoration of 53BP1 recruitment upon ectopic expression of Mof.

b. *A-type lamins defects and alteration in telomere biology*

There is evidence in the literature that A-type lamins associate with telomeres, putatively via telomere-binding proteins. This association has been shown to be important for the proper distribution of telomeres within the 3D nuclear space. A study compared nuclear distribution of telomeres between wild-type mouse embryonic fibroblasts and fibroblasts devoid of A-type lamins, and showed a clear difference in telomere distribution between the two genotypes[149].

This confirmed that A-type lamins participate in the correct distribution of telomeres throughout the entire nuclear volume and that defects in these lamins lead to a nuclear decompartmentalization of telomeres.

A second important feature is the maintenance of a heterochromatic structure for telomere length homeostasis and interestingly, A-type lamins participate in the maintenance of histone marks found at telomeric heterochromatin. Indeed, in mice fibroblasts expressing a lamin A mutant associated with HGPS (HGPS), a down-regulation of H3K9me3 and H3K27me3 and an up-regulation of H4K20me3 were observed[150–152]. These changes found in HGPS cells are different from changes observed in mouse fibroblasts expressing no lamin A.

Indeed, in this second case, no change in H3K9me3 levels was observed whereas a significant decrease in H4K20me3 levels was noted. Altogether, these results confirm implication of A-type lamins in the maintenance of histone marks[149].

Then, A-type lamin mutants exhibited a defective ability to process dysfunctional telomeres through the NHEJ pathway[149]. First, loss of 53BP1 function was recently associated with defective processing of dysfunctional telomeres by the NHEJ repair pathway[153]. Then, chromosome fusion percentages were evaluated in mouse fibroblasts, expressing or not A-type lamins. Clearly, in presence of a dominant-negative mutant of the telomeric binding protein TRF2, which induces telomere dysfunction and chromosome end-to-end fusions, the number of fusions increased in mouse fibroblasts expressing no A-type lamins [149]. Because of these two important observations, we can speculate that A-type lamins provide a platform for the association of 53BP1 to dysfunctional telomeres.

Finally, another study confirmed the implication of lamin A/C in dysfunctional telomeres processing. First, co-immunoprecipitation of endogenous TRF2 together with lamin A/C in IMR90s revealed an interaction between both proteins[154]. Then, the specificity of this interaction was confirm by pull-down assays using endogenous lamin A/C and GFP-TRF2 (functional TRF2), GFP-TRF2ΔBΔM (non-functional TRF2) or wild-type GFP-TRF1, in IMR90s. Indeed, only interaction between lamin A/C and functional TRF2 (GFP-TRF2) was observed in these conditions. In addition, pull-down assays revealed that lamin A mutant that causes HGPS (GFP-progerin) could not interact with endogenous TRF2.

TRF2 has been shown to bind interstitial telomeric sequences[155], thus stabilizing the telomeric t-loop; the impact of a decrease in lamin A/C levels or the *LMNA* mutation that causes HGPS (progerin) was observed. Both led to a reduction of interstitial t-loop formation and to telomere loss.

c. Nuclear envelope and mobility of DNA breaks

A recent study also implicated components of the nuclear envelope, such as the LINC complex and associated microtubule motor proteins (Kinesins), in the mobility of dysfunctional telomeres as well as of DNA DSBs within the nucleus. In addition, 53BP1/LINC/microtubule-dependent movement is important for NHEJ-based repair of dysfunctional telomeres and DSBs[156]. All together, these studies indicate that damaged DNA/chromatin exhibits increased mobility within the nucleus, and that this mobility is critical for its proper repair. The DNA repair factor 53BP1, LINC-microtubule connections, and the lamin network, all seem to play a role in these processes. Characterizing how these connections are established and regulated will be important for a mechanistic understanding of how nuclear architecture impacts genome stability.

d. A-type lamins and replicative stress

Lamin A/C is present at sites of early replication in normal human fibroblasts[152,157]. Lamin A/C-depleted cells are unable to restart most replication forks after treatment with HU compared to cells with lamin A/C, indicating that lamin A/C is required for the resolution of stalled replication forks[158]. This suggests that lamin A/C provides a platform for the resolution of stalled replication fork intermediates.

III. EMERIN, A NUCLEAR ENVELOPE PROTEIN

1. EMERIN STRUCTURE

Emerin is a type II integral INM protein of 254 amino acids (aa). Its gene was discovered in 1994 because it is mutated in patients with X-linked EDMD[159]. Emerin shows ubiquitous localization with the highest expression in skeletal and cardiac muscles and is mainly present at the inner nuclear membrane[160,161]. Emerin possesses a single transmembrane (TM) domain (aa 223-243) for anchoring at the inner nuclear membrane and in addition, is composed of a small C-terminal domain in the luminal space (aa 244-254) and a large hydrophilic N-terminal region localized in the nucleoplasm (aa 1-222). This N-terminal region exhibits a small globular LEM domain and a large region predicted as unstructured. Structure of the LEM domain was solved by NMR[49] and as all LEM-domain proteins (as presented in the introduction), it is composed of one three-residue N-terminal α-helix and two large parallel α-helices separated by a loop with conserved hydrophobic residues[51].

Most mutations that cause EDMD or other myopathies yield to premature stop codons, and truncated emerin is rapidly degraded in cells[162]. However a small number of mutations are missense or short in-frame deletions. Seven of these mutations were reported in patient cells, and an eighth mutation was recently identified in the group of Dr G. Bonne (Institut de Myologie, Paris). During my PhD, I focused my work on these eight mutations. Five of them affect residues of the predicted unstructured nucleoplasmic region (S54F, Δ95-99, Q133H, P183T, and P183H) and cause a classical EDMD phenotype[163]. Two missense mutations P22L and T43I (Bonne, personal communication) and the short in-frame deletion mutation ΔK37[164] affect the LEM domain and are associated with isolated cardiac defects.

Recent studies provided biochemical and cellular evidences that emerin can interact with itself[165]. First, yeast 2-hybrid experiments revealed that region 1-225 of emerin binds to itself, indicating that C-terminus truncated emerin can form homodimers and/or multimers[166]. Second, *in vitro* GST-pulldown experiments using recombinant emerin fragment 1-221 and co-immunoprecipitation of full-length emerin from HEK293T cell extracts also strongly suggested that emerin could oligomerize both *in vitro* and in cells[165]. A model was proposed to explain emerin intermolecular association data (figure 23)[165].

The R region (aa 187 to 220) could be involved in homotypic interactions (R to R) and heterotypic interactions between itself and the S region (aa 159 to 162), whereas the LEM domain (aa 1 to 47) could interact with the AR region (aa 170 to 220). Consequently, this model proposes two emerin "backbone" configurations depending on LEM domain positions. A down backbone configuration (figure 23. A), which corresponds to BAF- or chromatin-associated positions of the LEM domain, and an up backbone configuration (figure 23. B), which allows emerin self-association.

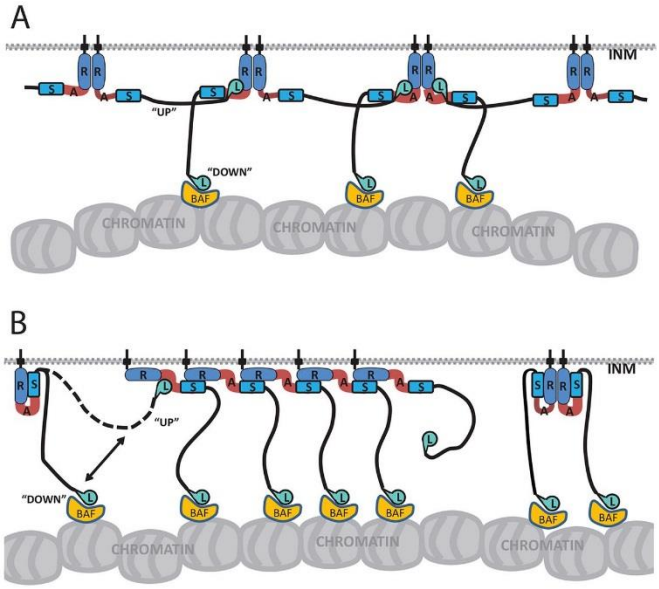


Figure 23 : Model for emerin intermolecular association[165].

(A) Down backbone configuration and (B) up backbone configuration.

2. EMERIN INTERACTING PROTEINS

Emerin interacts with several proteins essential for nuclear structure and chromatin organization. Best described emerin binding partners are A-type lamins and BAF. In addition, human emerin directly binds to at least 14 other proteins *in vitro* (figure 24)[167].

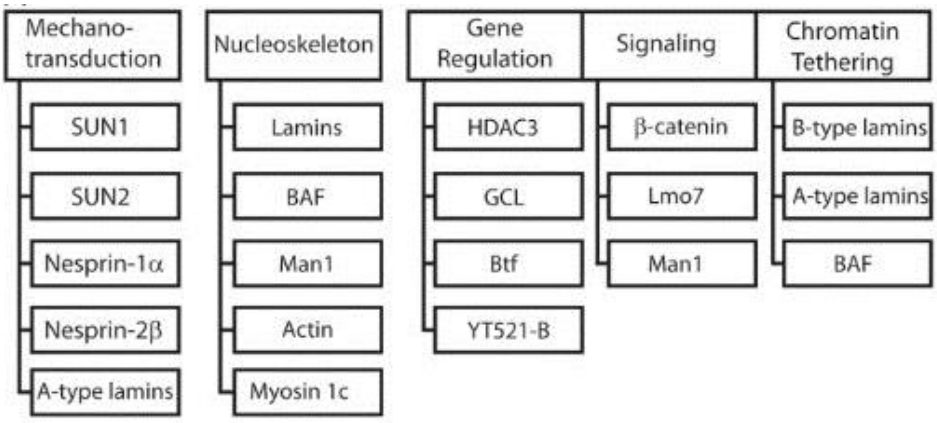


Figure 24 : Direct binding partners of emerin[167].

a. Emerin forms a ternary complex with BAF and A-type lamins

All LEM-domain proteins localized at the INM are supposed to interact with A-type lamins but interestingly, no consensus lamin-binding motif has been identified in LEM-domain proteins.

Several studies report the existence of an interaction between emerin and lamins and confirm that this interaction provides the structural backbone for proper nuclear localization and retention of emerin. A first observation of this interaction was done using immunofluorescence in Green Monkey Kidney cells: co-localization of emerin with A and B-type lamins was revealed in interphase cells[162]. Afterwards, this interaction was confirmed by several co-immunoprecipitation experiments[168,169]. The fact that the emerin and A-type lamins directly interact was demonstrated using Biomolecular Interaction Analysis (BIA)[170] and yeast two hybrid assays[166]. Finally, structural associations between emerin and lamins were investigated in four human cell lines. Absence of lamins A/C led to a mis-localization of endogenous and exogenous emerin in the endoplasmic reticulum. In addition, in cell lines which express lamin C, but not lamin A, both emerin and lamin C were mis-localized[169].

Because of these results, a hypothesis could be that lamin A mediates association of lamin C with the lamina and then, lamin C stabilizes association of emerin with the inner nuclear membrane.

Concerning molecular details of this interaction, only few data were reported. The region of lamin A responsible for emerin binding was delimitated using yeast two-hybrid experiments[166]. It consists of the lamin A tail domain, and more precisely, the region between amino acids 384 to 566. On the other hand, the region of emerin responsible for its interaction with A and B-type of lamin tails was deduced from GST-pulldown experiments. Two regions of emerin (residues 55-132 and 159-178) were found as necessary for lamin A and B1 tails binding and two regions of emerin (residues 1-132 and 159-220) were found to be sufficient for these bindings[165].

Emerin interacts with BAF through its LEM domain. The 3D structure of the complex between emerin LEM domain and BAF was solved using NMR and molecular modeling[52] (figure 25). On one hand, the LEM domain of emerin and more in details, its helix α1, its subsequent loop, and the N-terminal end of its helix α2 (in green) can bind directly to the dimer of BAF. On the other hand, BAF dimer interacts with the LEM domain through the C-terminal end of helix α2, the hairpin turn and the helix α3 of its first monomer (in red) and through the hairpin turn between α2 and α3, the C-terminal end of α3, and the central portion of α4 of its second subunit (in blue). In addition, superimposition of the two known 3D structures of BAF dimer in interaction with either DNA or the LEM domain clearly revealed that BAF binding sites for both molecules do not overlap (figure 25). Thus, BAF can probably simultaneously bind to emerin and DNA.

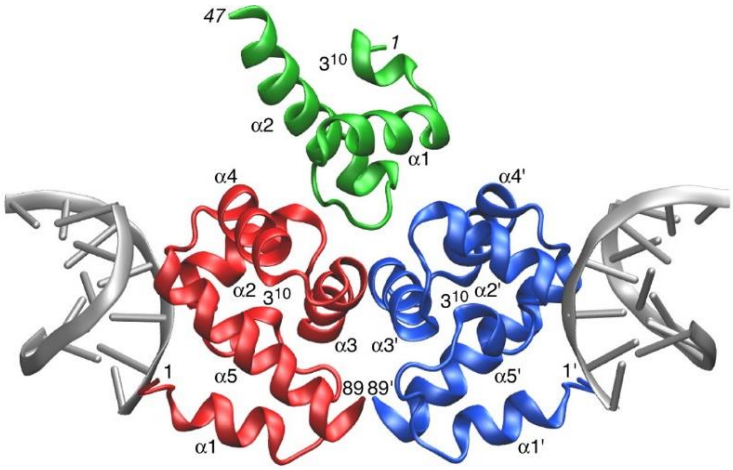


Figure 25 : The BAF_2 (in blue and red) - Em^{LEM} (in green) interface (PDB: 2ODG) does not overlap the BAF_2-DNA_2 (in grey) interface (PDB: 2BZF), as deduced from the superimposition of the 3D structures of BAF dimer bound to either emerin or DNA[52].

A direct interaction between emerin and BAF at the nuclear envelope of living cells was demonstrated by Fluorescence Resonance Energy Transfer (FRET)[171]. But interestingly, in the same study, Fluorescence Recovery After Photobleaching (FRAP) and Fluorescence Loss In Photobleaching (FLIP) experiments showed that, whereas BAF was highly mobile at the nuclear envelope, emerin was much less mobile. Thus, a "touch and go" model was proposed according to which BAF interacts frequently but transiently with emerin in interphase cells.

It was proposed that emerin interacts with A-type lamins in order to provide structural support to the nucleus and with BAF in order to regulate gene expression. Recently, studies showed that the three proteins are able to interact together to form a ternary complex. First, autoradiography experiments using 150 nM ^{35}S-lamin A, 250 nM ^{35}S-BAF, or both proteins, which were incubated with immobilized emerin, showed that A-type lamins and BAF together form stable complexes with emerin[55]. Moreover, K. Wilson and co-workers published that binding of ^{35}S-emerin to prelamin A tail-containing strips is enhanced fourfold by wild-type BAF[172]. Then, it was shown, *in vitro* and in HeLa cells, that phosphorylation of BAF on serine 4 leads to mis-localization of emerin, which failed to be at the inner nuclear envelope. Binding of prelamin A tail to BAF was also significantly reduced in the case of BAF mutant S4E. In contrast, BAF S4E had no effect on emerin binding to lamin A[172].

b. *Structural binding partners*

In addition to lamins, emerin interacts with other proteins that provide structural support to the nucleus. First, emerin was report to directly interact with actin. Cytoplasmic actin was well studied during several years but afterwards actin was also found in the nucleus[173]. It was described as a component of chromatin remodeling complexes. It was also suggested as involved in nuclear architecture. Globular actin (G-actin) was shown to co-immunoprecipitate with emerin in myoblast extracts[168]. Another study demonstrated that emerin can form a complex with lamin A and actin, but only in differentiated muscle cells[174]. This binding was strongly increased in the presence of an active serine-threonine phosphatase. Thus dephosphorylation is required to allow emerin-actin binding. Afterwards, it was demonstrated that emerin can also bind and stabilize filamentous actin (F-actin) by capping its pointed end[175]. Recently, emerin polypeptides bearing various mutations have been tested *in vitro* for binding to different emerin partners and interaction with actin was confirmed.

It was shown that the LEM domain as well as the central region of emerin are implicated in this binding[167]. Finally, emerin is able to bind nuclear myosin I, in cells and directly *in vitro*[176].

Other important emerin partners are SUN-domain and Nesprin proteins. This group of proteins clearly plays a structural role in the nucleus because they form the LINC complex (first part of this introduction). Nesprin-1α binds to emerin *in vitro*, as shown by blot overlay[177]. Interestingly, this protein was found as binding also lamin A, suggesting that it might enhance a crosslink between emerin and A-type lamins at the nuclear envelope. More recently, another Nesprin isoform, Nesprin-2 was found to bind both emerin and lamin A/C tail[178]. Moreover, it was shown that absence of Nesprin-2 in cells leads to mis-localization of emerin. Finally, concerning SUN-domain proteins, it was demonstrated, by immunoprecipitation, that emerin binds to the nucleoplasmic domains of SUN1 (aa 223–302) and SUN2 (aa 1-174). Although a weak emerin interaction was also observed with the proximal end (residues 1–208) of SUN1[179].

c. *Gene regulatory binding partners*

Supporting a role for emerin in gene expression, several gene regulatory proteins have been reported to interact with emerin. First, emerin associates directly with the histone deacetylase 3 (HDAC3)[139] and is able to enhance its enzymatic activity by 2.5 fold *in vitro*.

Second, co-immunoprecipitation assays from HeLa cell nuclear extracts, and a microtiter well binding assay, demonstrated that three different regions of emerin (residues 34-83, 175-196 and 207-217) interact with the protein germ cell less (GCL)[55]. GCL is a conserved protein that binds to the transcription factor E2F, which activates genes required for entry into S-phase and it was shown that emerin interacts with GCL in order to co-repress E2F. Interestingly, competition assays demonstrated that BAF and GCL compete for binding to emerin *in vitro*[55].

Adding to the list of gene regulatory binding emerin partners, a yeast-two-hybrid screening of a human heart cDNA library identified a nuclear splicing associated factor called YT521-B[180]. More in details, it was demonstrated that the C-terminus of YT521-B is sufficient to bind emerin *in vitro* and influence splice-site selection *in vivo*. On the other hand, it was shown that this protein binds emerin on the two regions flanking the lamin A binding domain, partially overlapping the BAF and lamin A binding regions.

In a similar yeast-two-hybrid assay, the pro-apoptotic transcription repressor called Btf was identified as an emerin interacting protein[181]. On one hand, this interaction was shown to require emerin residue 45-83 and 175-217 and on the other hand, it does require residues 377–646 of Btf. It is important to note that Btf can relocalize from the nuclear interior to the nuclear envelope during apoptosis induction, which suggests that emerin could be required to signal apoptosis within the nucleus.

The latest indication of emerin implication in gene regulation involves β-catenin and Lim-domain-only-7 (Lmo7), two signaling transcription factors that shuttle between the cell surface and the nucleus. The first one, β-catenin, is known to mediate Wnt signaling, and absence of emerin in cells seems to increase Wnt signaling, which suggests that emerin normally attenuates this pathway[182]. The second one, Lmo7, is a transcription factor that activates many genes including the emerin gene[176]. This suggests that Lmo7 binding to emerin protein can feedback-regulate emerin gene expression.

3. EMERIN POST-TRANSLATIONAL MODIFICATIONS

Emerin plays several functions at the nuclear envelope through binding to a lot of different partners. These interactions are differentially regulated by post-translational modifications. Major modifications in cells are phosphorylations and for example, emerin can be phosphorylated on several sites, like serine and threonine residues including S49, S66, T67, S120, S163 and S175[183–186]. Then, five residues in human emerin were identified as undergoing cell cycle-dependent phosphorylation using a Xenopus egg mitotic cytosol model system[184]. In addition, emerin is hyperphosphorylated in four different forms in metaphase and early S-phase cells[163]. Finally, during infection by Herpes simplex virus type 1, emerin is phosphorylated by several kinases[187,188].

a. Emerin phosphorylation

Human emerin possesses at least 42 published sites of phosphorylation *in vivo* (25 serines, 4 threonines and 13 tyrosines)[167,189] (figure 26). It was reported that emerin is a major target of phosphorylation during mitosis and interphase, but the proteins and pathways that mediate phosphorylation on these sites, and the functional consequences of these modifications, remain poorly understood.

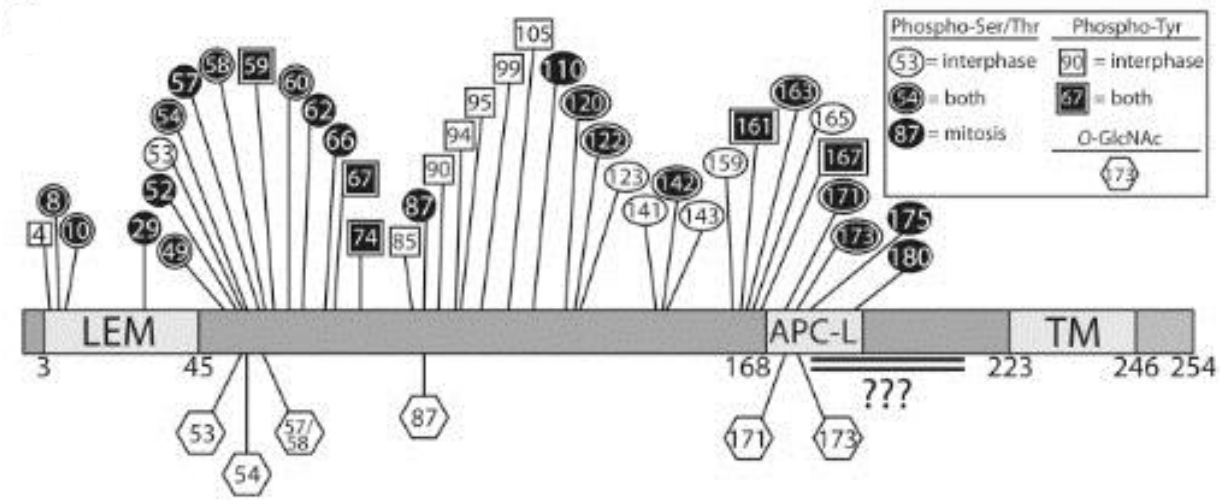


Figure 26 : Emerin phosphorylation sites identified by mass spectrometry in cells[167].

Hexagons, *O*-GlcNAc sites; circles, phospho-Ser/Thr; squares, phospho-Tyr; white, asynchronous cultures; black, mitotic cultures and conditions; black with outline, sites identified in both asynchronous and mitotic cells. The double-underlined region has at least two *O*-GlcNAc sites and potentially other modifications that are uncharacterized due to the large size of the corresponding trypic peptide and poor recovery by mass spectrometry.

1) Tyrosine phosphorylation

Several independent proteomic studies were achieved in order to identify tyrosine phosphorylation sites on emerin and 13 sites were found (figure 27). These sites are indicated with black dots on figure 27. All these sites found in human emerin are conserved in mouse and only six are conserved in *Xenopus*. In cells, tyrosine phosphorylation of emerin was validated by immunoprecipitation assays[189].

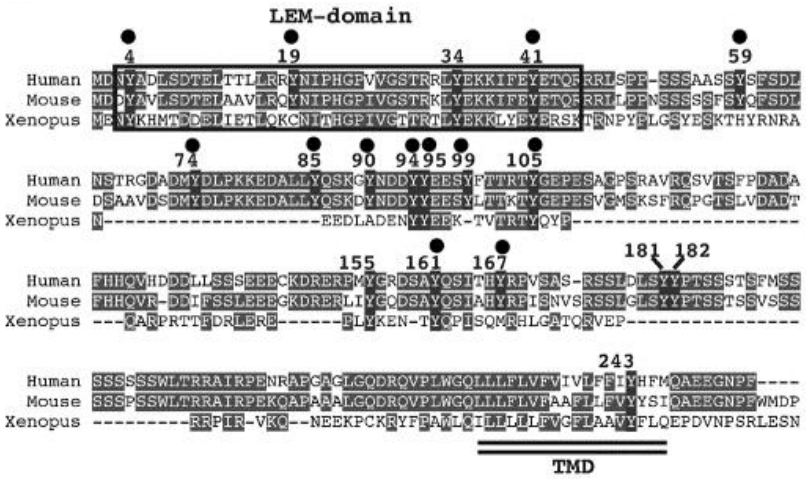


Figure 27 : Tyrosine phosphorylation sites of the emerin and amino acid sequences of emerin from human, mouse and *Xenopus laevis*[190].

At least two non-receptor Tyrosine kinases target emerin directly: Src and Abl. LC-MS/MS analysis revealed three major tyrosines that are phosphorylated by Src: Y59, Y74 and Y95[190]. Concerning Abl, one site was recovered and identified as Y167.

Tyrosine phosphorylation might regulate emerin interactions with specific partners and in particular, it was found as important for emerin binding to BAF[184,190]. First, BAF interacts with the LEM domain of emerin, yet this domain has four tyrosines in human, with at least three that are phosphorylated by unknown kinases *in vivo*. In addition, five tyrosines were found as impacting the interaction between BAF and emerin: Y4, Y19, Y34 and Y41 and interestingly, Y161, which is not localized closed to the LEM domain. Substitutions that removed a single hydroxyl moiety decreased (Y19F, Y34F, Y161F) emerin binding to BAF *in vitro* and *in cells*[190]. Concerning Y4F substitution, it was shown to reduce emerin 1-176 binding to BAF *in vitro*, but conferred enhanced binding to BAF when expressed in cells as full-length GFP-emerin[190].

2) Serine and threonine phosphorylation

Concerning emerin phosphorylation on serines and threonines, several known modifiers were suggested as protein kinase A (PKA), GSK3β, PKCδ and ERK2/MAPK.

For example, ERK2 was shown to directly interact with emerin *in vivo* during lentiviral infection[191] and this led to a hyperphosphorylation of emerin that caused an altered distribution of emerin at the inner nuclear envelope and a mislocalization of emerin to the outer nuclear membrane[188].

b. Emerin O-glycosylation

Emerin is highly modified through serine, threonine and tyrosine phosphorylations but it also possesses all features characteristic of O-linked-Nacetylglucosamine-modified proteins by, which are the presence of large disordered regions containing a lot of serine and threonine residues. O-linked-Nacetylglucosamine is a single sugar modification of serine or threonine residues that is abundant, reversible and dynamic, and this modification can compete or augment phosphorylation to regulate signaling, transcription, mitosis, and stress responses.

Immunoprecipitation assay with HeLa cells confirmed that emerin is O-GlcNAcylated *in vivo*[167]. Afterwards, different emerin polypeptides were O-GlcNAcylated *in vitro*, in order to identify which emerin residues are modified. At least two regions, comprising residues 1–70 and 170–220, were targeted by the OGT enzyme *in vitro*. More in details, eight emerin residues were O-GlcNAcylated *in vitro* and five residues were identified: Ser53, Ser54, Ser87, Ser171 and Ser173. In cells, only serine 173 was identified as significantly O-GlCNAcylated.

Thereafter, interaction between O-GlcNAcylated emerin and BAF was studied. Interestingly, it was shown that only modification on Ser173 could impact BAF binding to emerin and more in details, phosphorylation instead of O-GlcNAcylated at Ser173 reduced significantly BAF binding. This result could be explained by different mechanisms. First, a phosphorylation versus an O-GlcNAcylation at Ser173 could change emerin conformation. Another hypothesis is that modifications on Ser173 could impact binding to a third partner, which could impact binding between emerin and BAF afterwards.

IV. THE NUCLEAR ENVELOPE AND DISEASES

Many human diseases, called envelopathies, are associated with NE protein defects. A large proportion of them are associated with mutations in lamin A/C and emerin.

1. EMERY-DREIFUSS MUSCULAR DYSTROPHY (EDMD)

EDMD is a group of genetic and degenerative diseases that primarily affects voluntary muscles. Its name comes from two physicians, Alan Emery and Fritz Dreifuss, who first described this disorder among a large Virginian family in 1966[192]. EDMD is characterized by contractures of major tendons, slowly progressive skeletal muscle wasting and weakness, and dilated cardiomyopathy with potentially lethal ventricular conduction system defects that can cause sudden cardiac arrest.

Mutations in *EMD* (encoding emerin) and *FHL1* (encoding FHL1) cause X-linked EDMD (XL-EDMD), whereas mutations in *LMNA* (encoding lamin A and C) cause an autosomal dominant (AD-EDMD) or autosomal recessive EDMD (AR-EDMD). For all forms of EDMD the diagnosis is based on clinical findings and family history. The diagnosis of X-linked EDMD also relies on failure to detect emerin or FHL1 protein in various tissues and molecular genetic testing of *EMD* or *FHL1*. The diagnosis of AD-EDMD and AR-EDMD relies on molecular genetic testing of *LMNA*[193].

a. X-Linked EDMD

The first gene responsible for the X-linked form was identified in 1994[159]. It corresponds to the *STA* (or *EMD*) gene and mutations in this gene were found in several patients. Moreover, this gene was found as coding for the emerin protein.

DNA extraction from peripheral blood lymphocytes was done to detect mutations on the emerin gene which cause X-linked EDMD. Almost 100 mutations were found in this *STA* gene and these are approximately composed of 39.5% of small deletions, 31% of non-sense mutations, 15.5% of mutations in splice sites, 4% of large deletion, 8.5% of missense mutations and 1.5% of mutations in the promoter.

Although, the majority of mutations results in complete loss of emerin, some mutations have been identified as reducing expression level and protein amount[194].

Interestingly, despite the different mutations in EDMD producing varying effects on emerin expression, the clinical phenotype of all the patients is similar.

Several studies focused on the emerin mutants that are still present in cells. Five emerin variants were observed at the nuclear membrane: S54F, Δ95-99, Q133H and P183T/H. Interestingly, all the corresponding mutations are detected in the predicted unfolded part of emerin. The behavior of four of these mutants (S54F, Δ95-99 and P183H/T) was studied by transfection of GFP-mutants in C2C12 myoblasts[168]. All mutants displayed reduced targeting and retention at the nuclear envelope in comparison to wild-type emerin and this was particularly significant in the case of mutant Δ95-99. The observed immunofluorescence intensity decrease might correspond to a decrease in expression, in stability or to a wrong localization.

Concerning mutation Q133H, it was shown to reduce the amount of emerin compared to control[195] and study of the corresponding GFP-mutant in COS-7 cells showed a correct localization.

Furthermore, emerin polypeptides bearing four of these mutations (S54F, Δ95-99, Q133H and P183H) were tested *in vitro* for binding to different partners: BAF, lamin A, GCL, Btf, YT521-b, Lmo7, HDAC3, F-actin and MAN1[196] (figure 28). On one hand, all tested mutants are still able to bind BAF. On the other hand, interaction between all emerin mutants and HDAC3 is disrupted. Interestingly, mutant Δ95-99 is the most functionally affected. Indeed, this mutant cannot bind anymore, in addition to HDAC3, A-type lamins, GCL, Btf, actin and MAN1. Like mutant Δ95-99, mutant Q133H cannot bind actin and MAN1 whereas mutant S54F cannot bind Btf. Finally, mutant P183H is the only one that cannot bind to Lmo7.

Recently a new in frame *EMD* deletion was found: ΔK37. This mutation affects the LEM domain and is associated with isolated cardiac defects in patients[164].

	S54F	Δ95-99	Q133H	P183H
BAF	+	+	nt	+
Lamin A	+		nt	+
GCL	+		+	+
Btf			+	+
YT521-B	nt	nt	+	+
Lmo7	+	+	+	
HDAC3				
Actin	+			+
MAN1	+			+

Figure 28 : X-linked EDMD mutation impacts on emerin binding to its partners[196].

Scoring: normal binding (+), weakened binding (± and gray), and undetectable binding (black box). nt, not tested.

b. Autosomal EDMD

Although most cases of EDMD are X-linked, a rare autosomal dominant form (AD-EDMD) was also reported. The gene responsible for AD-EDMD was identified in 1999. It is localized on chromosome 1q11-q23 and corresponds to the *LMNA* gene, which encodes A-type lamins[197].

Unlike X-linked EDMD, the majority of mutations which cause AD-EDMD are missense mutations and lead to the production of an equimolar mixture between normal and mutated lamins[198,199].

Three mechanisms are possible to cause autosomal dominance. The first one is a mechanism in which the reduced level of functional lamin A/C is not sufficient for normal lamina function. The second proposed mechanism is that lamin A/C mutants play a deleterious function. The last one is that, knowing that lamins form dimers in order to form filaments, a functional multimer exists, which cannot tolerate a defective subunit.

2. OTHER LAMINOPATHIES

More than EDMD, mutations in the *LMNA* gene have been shown to cause a large spectrum of diseases and the term of laminopathies has been adopted to describe them[200–203]. So far, at least eight different laminopathies have been associated to mutations in *LMNA*, including dilated cardiomyopathy with conduction defects (CMC), Limb-girdle muscular dystrophy type 1B (LGMD), Dunnigan-type familial partial lipodystrophy (FPLD), Charcot-Marie-Tooth disease (CMT), mandibuloacral dysplasia (MD), restrictive dermopathy (RD) and Hutchinson-Gilford progeria syndrome (HGPS). A representative set of mutations causing these different pathologies, including AD-EDMD, is represented on figure 29[199].

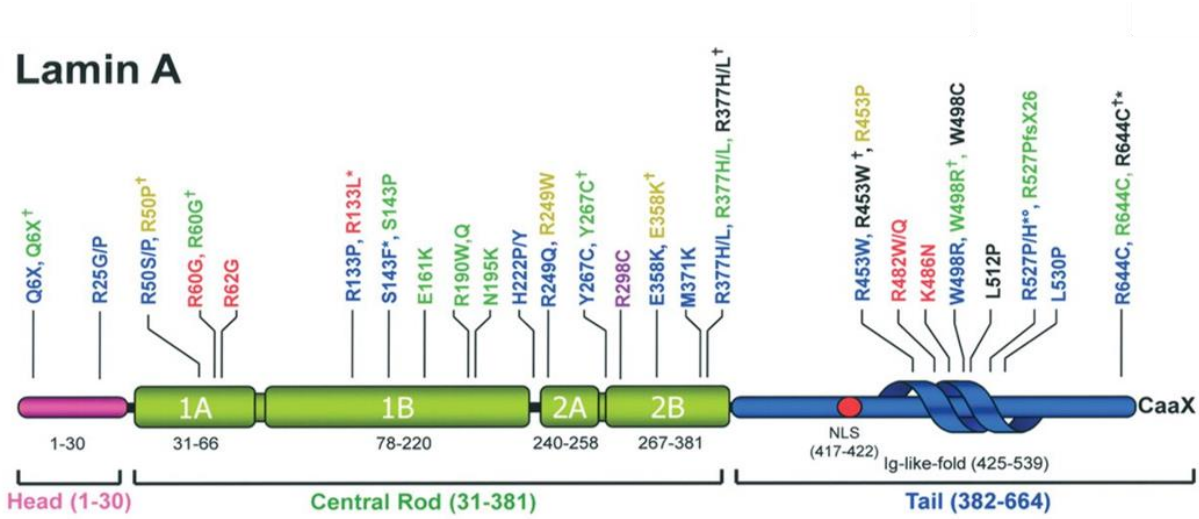


Figure 29 : Distribution of lamin A mutations and their related laminopathies, observed after clinical diagnosis[204,205] with A-EDMD in blue, DCM in green, LGMD1B in black, CMD in gold, FPLD2 in red and CMT2B in purple.

'†' indicates that the same amino acid change causes different laminopathies. *R133L also causes WRN, *S143F additionally results in HGPS, *R527H causes MAD too and *R644C also gives rise to a range of other disorders. °Patients carrying R527C develop either a severe form of MAD and/or progeria[199].

a. *Cardiomyopathy with conduction defects (DCM)*

Cardiomyopathies are characterized by cardiomyocyte dysfunction and tissue-wide remodeling of the myocardium leading to functional decline. *LMNA* was found as implicated in this kind of disease in 1999[206], by Fatkin et al., and accounts for approximatively 6-8% of CMD cases in humans. The number of mutations causing CMD increased and at least eight mutations are known[198] (figure 29).

This kind of disease can occur as an isolated phenotype or more frequently in combination with a skeletal muscle dystrophy such as EDMD or Limb girdle muscular dystrophy[207,208].

b. *Limb-girdle muscular dystrophy type 1B (LGMD-1B)*

LGMD-1B, like other muscular dystrophies, is primarily a disorder of voluntary muscles. Indeed, it affects mainly the proximal limb-girdle musculature[209]. This group of diseases comprises 15 different types inherited as both autosomal dominant and recessive forms.

As it is represented on figure 29, LGMD1B mutations (black on figure 29) tend to cluster in both the Ig-like fold and coil 2.

c. *Dunnigan-type familial partial lipodystrophy (FPLD)*

Lipodystrophies are a group of diseases characterized by the absence or reduction of subcutaneous adipose tissue. Dunnigan-type FPLD is a rare autosomal dominant disease characterized by loss of subcutaneous adipose tissue from extremities and trunk and accumulation of fat in the head and neck areas[210].

A first mutation was identified, R482Q[211] and afterwards, five other mutations were found. As it is shown on figure 29, at least half of the known mutations that cause this pathology are localized in the lamin tail region[147].

d. *Charcot-Marie-Tooth disease (CMT)*

Charcot-Marie-Tooth diseases form a heterogeneous group of hereditary motor and sensory neuropathies. This group can be classify in two subgroups: the first one is composed of demyelinating type diseases (type 1) and the second one of axonal type diseases (type 2)[38].

Mutations in the *LMNA* gene are associated with axonal CMT. The first mutation was discovered in three consanguineous Algerian families. Genome analysis identified a missense mutation causing an R298C substitution.

e. Progeroid syndromes (PS)

Mutations in lamins cause muscle diseases, adipose tissue diseases and neuropathies, as shown in Figure 29. They also cause accelerated aging syndromes also called progeroid syndromes, as presented in the following paragraphs.

1) Hutchinson-Gilford progeria syndrome (HGPS)

Hutchinson-Gilford progeria (HGPS) is a rare genetic disorder that causes an average age of death at 13.4 years due to coronary artery disease[37]. It characterized by abnormalities in skin, bone, fat tissue, hair and blood vessels. The most common HGPS mutation is located at codon 608 and like I already explained in the first part of this introduction, it is a silent mutation G608G. However, this mutation induces a cryptic splicing donor site that generates a 150-nucleotide deletion in the mRNA sequence. In consequence, the resulting prelamin A bears a 50-amino acid in-frame deletion and lacks the proteolytic cleavage site for ZMPSTE24, which is essential for C-terminal farnesyl group release[37,212]. The incompletely processed prelamin A is commonly called progerin. In cells, progerin accumulation correlates with nuclei that exhibit altered size and shape with nuclear envelope interruptions. In addition, the presence of progerin induces a loss of heterochromatin and H3K9 trimethylation, whereas H4K20 trimethylation level increases[151,152,213,214]. Progerin also impairs the response to oxidative stress through, for example, NRF2 (Nuclear Factor (erythroid-derived 2)-like 2) transcription factor sequestration, which leads to an accumulation of ROS (reactive oxygen species) and an impairment of the DNA damage response.

Because of its farnesylation, progerin enhances some interactions that might depend on its farnesyl moiety. For example, it was shown that farnesylated prelamin A interacts with NARF (Nuclear Prelamin A Recognition Factor)[215] and SUN1 proteins[216] whereas interaction with HP1α is disrupted[217].

Because it was demonstrated *in vivo* that BAF could interact with different prelamin A forms[218], and because accumulation of BAF in the nucleus of HGPS cells was observed, it was proposed that some chromatin defects could be triggered by the progerin-BAF interaction. To confirm this hypothesis, a recent study demonstrated first that progerin could bind BAF in HEK293 cells. Indeed, after expression of FLAG-tagged progerin in HEK293 cells, recruitment of endogenous BAF at the nuclear rim was observed[219]. In addition, co-immunoprecipitation studies confirmed that progerin is able to bind BAF WT.

Interestingly, in this same study, it was shown that BAF is necessary to mediate both prelamin A and progerin effects on nuclear distribution of LAP2-alpha and HP1α, which are proteins involved in chromatin-related processes (figure 30). These observations could explain how the progerin-BAF interaction plays a role on chromatin organization: accumulation of progerin induces BAF translocation from the cytoplasm to the nuclear lamina where both proteins interact and this could help progerin to interact with DNA organizing proteins to modify chromatin organization through BAF.

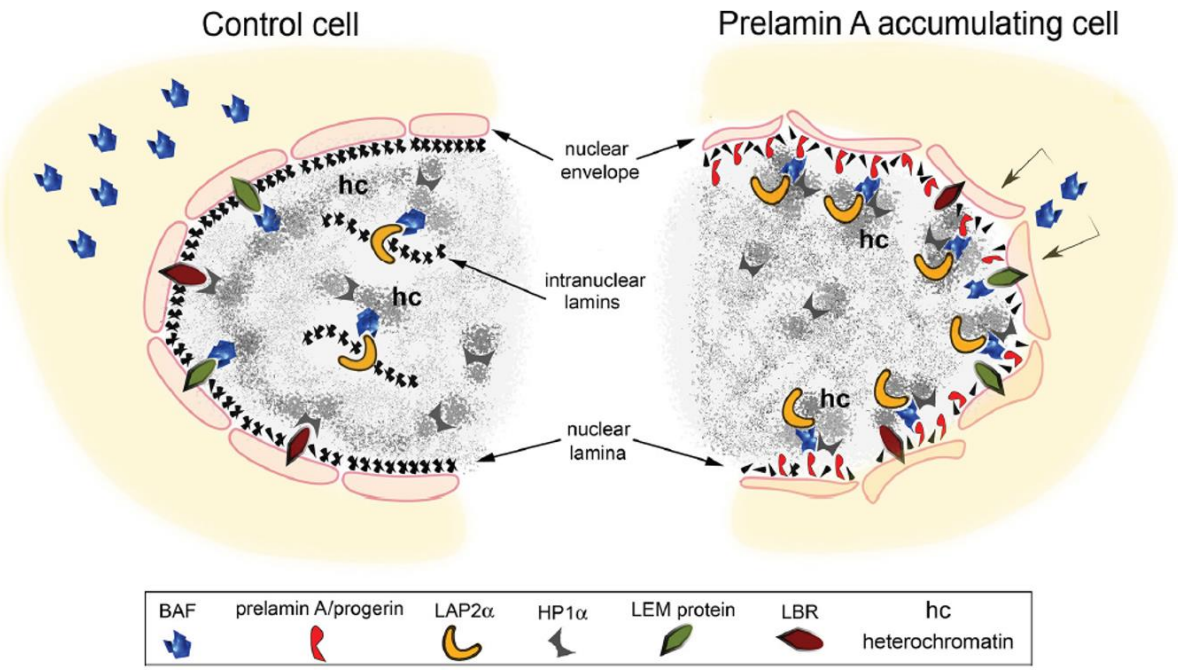


Figure 30 : Speculative cartoon of the mechanism involving prelamin A-BAF interaction in chromatin organization changes associated to prelamin A accumulation[219].

Another protein was found as interacting with progerin and could contribute to the development of the senescence phenotype of HGPS and aged cells. Indeed, different techniques as proteomic studies, exploiting RNA interference technology to target the mutated pre-spliced or mature LMNA mRNAs[220] or two-hybrid approaches using a cDNA library to search for progerin-interactive partner proteins[221], have identified lamins A/C as progerin-binding partners[222]. More in details, a direct interaction was found between progerin and A-type lamins whereas no interaction was detected with B-type lamins.

Interestingly, an ELISA assay in which His-tagged lamin A middle region (His-LMNA-M; from residues 301 to 564) was fixed in a 96-well plate and incubated with GST-fused recombinant lamin A C terminal region (GST-LMNA) or with progerin C-terminal region (GST-progerin) clearly revealed that progerin–lamin A binding affinity is more than 2-fold stronger than that of lamin A–lamin A binding (figure 31).

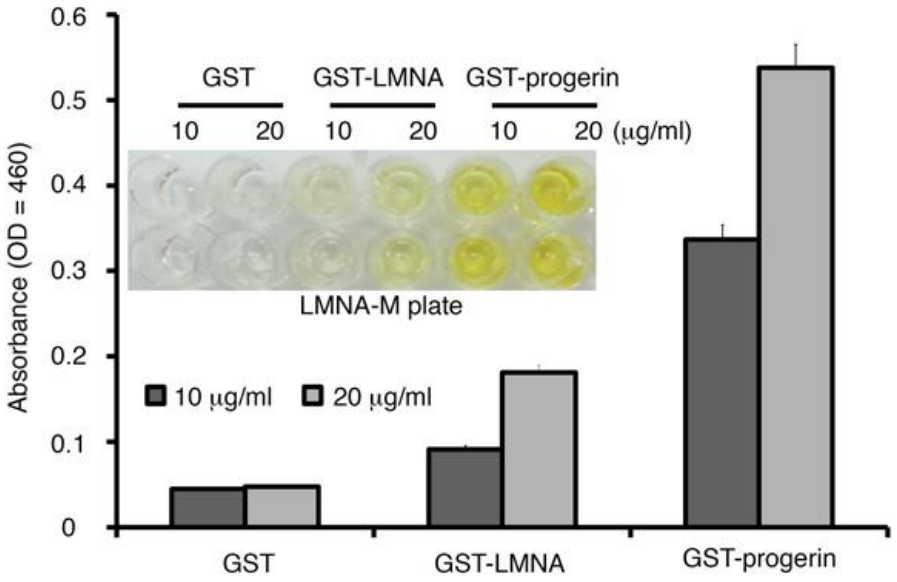


Figure 31 : ELISA results showing His-LMNA-M/ GST-LMNA and His-LMNA-M/GST-progerin interactions[222].

This result is in agreement with the HGPS cell phenotype in which A-type lamins form an irregular fiber arrangement[222]. Indeed, stronger affinity of progerin for lamin A could induce a tighter but irregular bundle formation compared to intermediate filaments formed by interaction between several lamin A/C proteins.

Another hypothesis is that interaction between progerin and lamin A may induce nuclear deformation, because it was shown that progerin alone did not induce nuclear deformation in lamin A/C-deficient cells whereas co-transfection with lamin A/C could induce it[222].

All these results are consistently leading to the hypothesis that the strong affinity of progerin for lamin A/C could contribute to nuclear lamina alteration in HGPS cells.

2) Atypical progeroid syndromes (APS)

Besides HGPS, several atypical progeria syndromes (APS) caused by *LMNA* gene mutations have been described, such as atypical Restrictive Dermopathy (the most severe), Werner syndrome and mandibuloacral dysplasia. Some mutations causing atypical progeroid syndromes were identified. First, heterozygous mutations localized in lamin A/C C-terminus like T623S[223] were found. Secondly, some mutations were identified in the coiled-coils 1A and 1B of lamin A/C, which are required for the polymerization of the nuclear lamins into higher order structures. Indeed, concerning coiled-coil 1A, the heterozygous mutants A57P[224] and both E55K and E55G[225] were observed. Similarly, several heterozygous *LMNA* mutations in the coiled-coil 1B were identified as R133L[226], L140R[224], S143F[227] and E145K[37]. Finally, a large number of mutations were found in the globular domain of lamin A/C, the Igfold: heterozygous mutations as R527C[228] and T528M associated to M540T[229] or homozygous mutations like R435C[230], R471C[203], R527C[231], R527H[232–235], A529T[236], A529V[237] and K542N[238]. All these mutations are reported on figure 32.

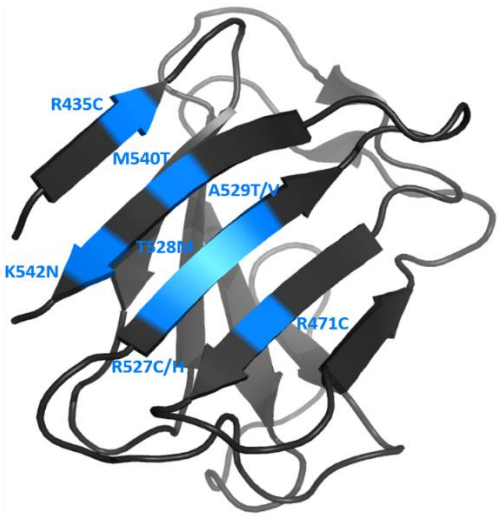


Figure 32 : Localization of Igfold residues that are mutated in APS (PDB: 1IFR[43]).

a) Restrictive dermopathy (RD)

Restrictive dermopathy is the most severe of these progeroid syndromes and was first described by Witt et al. in 1986[239]. It is a rare lethal genetic disorder that causes congenital tautness of the skin with characteristics of facial features (small mouth, small pinched nose, and micrognathia), bone mineralization defects including thin, dysplastic clavicles, and pulmonary hypoplasia[240].

A first group screened *LMNA* coding sequence and exon-intron boundaries in nine children affected by RD and did the same for the endoprotease ZMPSTE24 of seven patients. As I already mentioned, this enzyme is essential for post-translational processing of prelamin A and allows a correct maturation of lamin A to be inserted into nuclear lamina. In one case, they identified the most common heterozygous mutation responsible of HGPS (G608G) and in another case, they identified a novel splicing mutation specifically affecting lamin A. Both mutations were leading to the production and accumulation of truncated prelamin A. Concerning *ZMPSTE24*, they found a unique heterozygous insertion leading to the creation of a premature termination codon[240].

Then, the same group did another analysis on ten patients affected by RD and found three novel null mutations in *ZMPSTE24* and in all cases, a complete absence of both ZMPSTE24 and mature lamin A was observed[36].

Another study analyzed genomic DNA extracted from peripheral blood lymphocytes of a 2-year girl affected by RD. This analyze revealed the presence of a homozygous mutation, R435C[230]. Thereafter, presence of this mutation R435C was recovered in other patients[241].

b) Werner syndrome

The Werner syndrome is characterized by scleroderma-like skin changes, short stature, thinning of the hair, diabetes mellitus, soft tissue calcification and premature atherosclerosis. It was first described as a progeroid syndrome caused by mutations in the *WRN* gene, which codes for a DNA helicase but afterwards, it was shown that some features of this disorder are also present in laminopathies caused by mutations in *LMNA*. In addition, it was found that individuals with atypical Werner syndrome with mutations in *LMNA* had a more severe phenotype than those affected by the disorder due to mutated *WRN*.

Three *LMNA* missense mutations, localized in lamin A/C coiled-coil domain 1, were identified: A57P, R133L and L140R. Immunofluorescence studies on patient fibroblasts were done and revealed that these mutants cause abnormal nuclear shape and apparition of chromatin in the cytoplasm[224].

c) Mandibuloacral dysplasia (MAD)

Mandibuloacral dysplasia is a rare autosomal recessive disorder characterized by postnatal growth retardation, skeletal malformations, craniofacial anomalies, joint contractures and types A and B patterns of lipodystrophy. Owing to its slowly progressive course, the syndrome was observed in adults, and pediatric case reports are scarce[232].

A homozygous mutation, R471C, was identified in a 7-year girl and it was report that this mutation causes a severe early MAD combined with progeroid features[203].

Another homozygous mutation, R527C, was found as causing severe MAD during infancy[231]. Then, this mutation was identified in the *LMNA* gene of two children who were suffering from atypical HGPS[242].

Furthermore, one study analyzed five consanguineous Italian families, containing patients affected by typical MAD and after sequencing of the *LMNA* gene, identified a homozygous missense mutation, R527H[232]. Patient cells showed nuclei with abnormal lamin A/C distribution and abnormal nuclear shape. Several studies supported the implication of this mutation in MAD[233–235].

Finally, two homozygous missense *LMNA* mutations A529V[237] and A529T[236] were identified. A529V was found in a male and a female patients with MAD[237] and mutation analysis of a 56-year-old Japanese woman with MAD and type A lipodystrophy revealed the second one, A529T[236].

3) Nestor Guillermo Progeria Syndrome (NGPS)

Recently, two unrelated patients who exhibited Hutchinson-Gilford Progeria syndrome-like phenotypes were described[243]. These patients did not present all symptoms of HGPS syndromes. For example, they did not show signs of ischemia or atherosclerosis and these patients were much older than the average life span of progeroid patients.

The new syndrome was called Nestor Guillermo Progeria Syndrome (NGPS) and after genome sequencing, a homozygous mutation in the *BAF* gene was discovered. This mutation, A12T, does not disrupt the dimerization of the protein *in vitro* but affects BAF protein level[219,244]. Indeed, BAF was detectable in NGPS nuclei but hardly visible in the cytoplasm. Finally, this mutant also affects nuclear morphology.

Unlike in other progeroid syndromes, in NGPS cells, prelamin A processing is normal. But a recent study demonstrated that A12T mutation affects the ability of prelamin A to modify chromatin organization[219]. To prove this, an electron microscopy study was performed on HEK293 cells expressing FLAG-tagged proteins in combination with BAF. Normal chromatin organization was observed in cells expressing WT A-type lamins (LA-WT) in combination with WT BAF, whereas when WT A-type lamins were co-expressed with BAF A12T, chromatin architecture was affected (figure 33). Indeed, peripheral heterochromatin did not adhere anymore to the nuclear lamina. Interestingly, co-expression of prelamin A (LA-C661M) with WT BAF induced a specific recruitment of the heterochromatin with impressive beads-on-a-string appearance whereas in presence of BAF A12T, this beads-on-a-string appearance was not visible anymore, but only some heterochromatin clustering was observed.

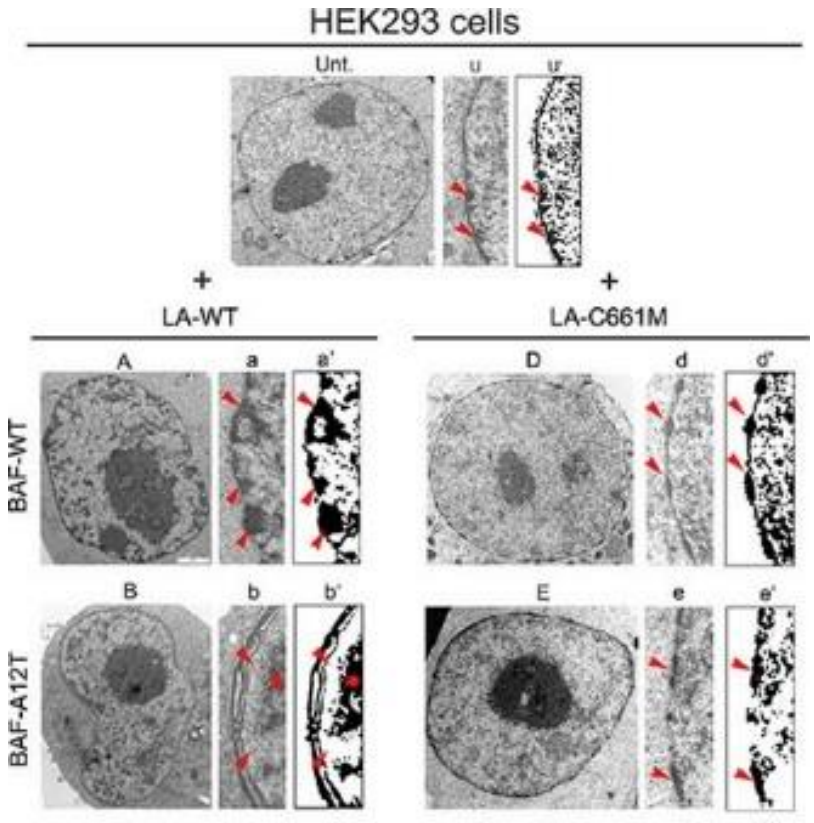


Figure 33 : Electron microscopy evaluation of the impact of prelamin A and BAF mutations on chromatin organization[219].

This observation led to new hypotheses concerning the role of BAF interaction with A-type lamins. Indeed, the authors showed that in NGPS, BAF mutation compromises prelamin A-BAF interaction but also impairs prelamin A-mediated H3K9m3 intranuclear recruitment; it could perturb the epigenetic function of BAF, affecting its ability to interact with histone H3 or involved in histone H3 modification. In HGPS, the persistence of BAF-progerin binding could impair different BAF interactions or slow down the dynamics of BAF protein complexes.

V. HYPOTHESIS AND THESIS OBJECTIVES

The nuclear envelope is composed of a double lipid bilayer, corresponding to the inner and outer nuclear membranes, and a large number of proteins. It contributes to the shape and the position of the nucleus, to the cell structure and mobility, to the organization of the genome and to the regulation of signaling pathways. Mutations in genes coding for nuclear envelope proteins, like mutations in the lamin genes, cause a large number of human diseases, called envelopathies. Structural information that describes the 3D arrangement of these proteins at the nuclear envelope is lacking, and this hampers the description of their normal and pathogenic functional mechanisms.

The LBSR (Structural Biology Laboratory and Radiobiology) works since several years on this structural characterization of protein complexes that are mutated in human diseases. The aim of my thesis was to characterize the interaction between lamin A/C and emerin, in order to obtain molecular details about the complex *in vitro*, and to understand the impact of disease-causing mutations on this interaction *in vitro* and in cells.

To achieve this goal, we decided to first describe the structure of emerin. During the beginning of my thesis, Isaline Herrada, a PhD student in the lab, showed that emerin fragment 1 to 187 oligomerizes *in vitro*. Simultaneously, the group of K. Wilson in Baltimore published that emerin could interact with itself through several regions *in vitro* and in cells[165]. Moreover Guilluy et al.[107] published that emerin is phosphorylated by Src during a mechanical stress on tyrosines 74 and 95, and this is a critical event to trigger the mechanical response of the nucleus[107]. Therefore we decided to study the impact of oligomerization and phosphorylation on emerin binding to lamin A/C.

In this manuscript I will report:

(1) the structural characterization of emerin (partly performed in collaboration with Dr Isaline Herrada).
(2) the structural description of two different complexes involving emerin and lamin.
(3) the impact of disease-causing mutations on these interactions, as observed *in vitro* but also in cells (collab. with Dr Brigitte Buendia, Univ. Paris Diderot).
(4) the role of phosphorylation by Src on emerin structure and binding properties.

RESULTS

I. STRUCTURE OF THE INNER NUCLEAR ENVELOPE PROTEIN EMERIN

The function of emerin at the nuclear envelope depends on its oligomerization state, 3D arrangement and modifications. At the beginning of my PhD, I worked together with another PhD student, Dr Isaline Herrada, in order to describe the oligomerization states of emerin fragment 1 to 187, also called EmN. This work was triggered by two observations:

- Emerin nucleoplasmic region binds to lamin A tail. Indeed, emerin nucleoplasmic region (residues 1 to 225) binds to lamin A/C (residues 1 to 566) as well as the tail region of lamin A (residues 385 to 664) as observed by yeast 2-hybrid experiments[166] and this interaction is impaired by mutations in emerin region 70 to 178 as seen by blot overlay assays[245]. However we did not observe any binding between monomeric EmN and lamin A tail by NMR.

- Emerin oligomerizes. Indeed, emerin nucleoplasmic region (residues 1 to 225) interacts with itself as observed by yeast 2-hybrid experiments[166]. Later during my PhD it was also reported that emerin self-associates *in vitro* and in cells[165].

Our hypothesis was that oligomerization of emerin could regulate lamin A tail binding. I have here inserted the 2 papers in which we describe the oligomerization of EmN from a biophysical point of view. Then I describe additional data that I obtained on the phosphorylation of EmN by a set of kinases involved in cell cycle regulation and mechanotransduction.

1. STRUCTURAL CHARACTERIZATION OF EMERIN NUCLEOPLASMIC REGION

The structure of emerin nucleoplasmic region (residues 1 to 221) was analyzed by solution NMR (figure 34). The LEM domain presents a ^{1}H-^{15}N spectrum typical of a α/β fold[49]. The narrow distribution of the peaks present on the ^{1}H-^{15}N spectra of fragments 67-170, 67-187 and 67-221 suggested that these fragments are unstructured. Assignment of the NMR signals of fragment 67-170 confirmed that this fragment is intrinsically disordered (see in Appendix 1: Samson et al., *Biomol NMR Assign* 2016). Thus emerin nucleoplasmic region is composed of a small globular LEM domain and a large disordered fragment.

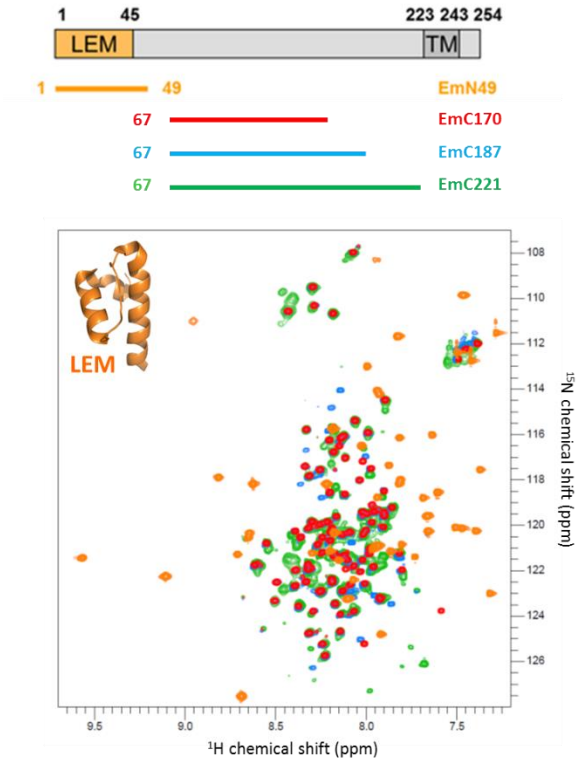


Figure 34 : Architecture of emerin and superimposition of the ^{1}H-^{15}N spectra of fragments 1-49, 67-170, 67-187 and 67-221 of emerin, recorded in 20mM phosphate buffer pH 6.5, 30mM NaCl, at 700 MHz and 303K.

In the papers inserted below, I focus on the description of the oligomeric states of emerin. The whole nucleoplasmic region of emerin, from residue 1 to residue 221, was systematically aggregating during the dialysis against a urea-free buffer. Therefore I studied oligomerization of both EmN (from residue 1 to residue 187) and EmC221 (from residue 67 to residue 221). I first characterized EmN oligomers because EmN contains the region proposed to bind to lamin A/C.

The first paper reports that EmN can self-assemble into long curvilinear filaments. These oligomers were observed by negative-staining Transmission Electron Microscopy (EM) thanks to the facilities of Dr Maïté Paternostre in our laboratory in Saclay and Dr Dmytro Puchkov at the FMP Berlin. Emerin-emerin proximities in cells were consistently detected by *in situ* Proximity Ligation Assays (collab. with Dr Brigitte Buendia, Uni Paris Diderot). Defects in self-assembly and cellular distribution were observed for several emerin variants causing Emery-Dreifuss Muscular Dystrophy. In particular, for variant del95-99, both a lack of self-assembly *in vitro* and a decrease in emerin-emerin and emerin-lamin A/C proximities in cells were observed. These first results were consistent with our first hypothesis that emerin self-assembly is essential for emerin function and in particular for lamin recognition.

Muscular dystrophy mutations impair the nuclear envelope emerin self-assembly properties

Isaline Herrada[1], Camille Samson[1], Christophe Velours[2], Louis Renault[2], Cecilia Östlund[3], Pierre Chervy[1], Dmytro Puchkov[4], Howard J Worman[3], Brigitte Buendia[5] and Sophie Zinn-Justin [1*]

[1] Laboratoire de Biologie Structurale et Radiobiologie, Institute for Integrative Biology of the Cell (I2BC), CEA Saclay Bât. 144, 91191 Gif-sur-Yvette Cedex, France

[2]Laboratoire d'Enzymologie et Biochimie Structurales, Institute for Integrative Biology of the Cell (I2BC), CNRS Bât.34, 1 avenue de la terrasse, 91190 Gif-sur-Yvette, France

[3]Department of Medicine and Department of Pathology and Cell Biology, College of Physicians and Surgeons, Columbia University, 630 West 168th Street, New York, NY, 10032, USA

[4]Department of Molecular Pharmacology and Cell Biology, Leibniz-Institut für Molecular Pharmakologie (FMP), 13125 Berlin, Germany

[5]Laboratoire de Physiologie du Muscle Strié, Université Paris Diderot-Paris 7, CNRS, UMR 8251, Institut de Biologie Fonctionnelle et Adaptative, 4 rue M.A. Lagroua Weill Halle, 75205 Paris Cedex 13, France

* To whom correspondence should be addressed: Sophie Zinn-Justin, I2BC CNRS / Univ. Paris South & IBITECS CEA, CEA Saclay Bât 144, 91191 Gif-sur-Yvette Cedex, France, Tel.: +33(0)169083026; Fax: +33(0)169084712; Email: sophie.zinn@cea.fr

Keywords: nuclear envelope, lamin, protein filaments, nuclear structure, protein oligomerization, electron microscopy

Abstract

More than 100 genetic mutations causing X-linked Emery-Dreifuss muscular dystrophy have been identified in the gene encoding the integral inner nuclear membrane protein emerin. Most mutations are nonsense or frameshift mutations that lead to the absence of emerin in cells. Only very few cases are due to missense or short in-frame deletions. Molecular mechanisms explaining the corresponding emerin variants loss of function are particularly difficult to identify because of the mostly intrinsically disordered state of emerin nucleoplasmic region. We now demonstrate that this EmN region can be produced as a disordered monomer, as revealed by Nuclear Magnetic Resonance, but rapidly self-assembles *in vitro.* Increases in concentration and temperature favor the formation of long curvilinear filaments with diameters of approximately 10 nm, as observed by electron microscopy. Assembly of these filaments can be followed by fluorescence through Thioflavin-T binding and by Fourier-transform Infrared spectrometry through formation of β-structures. Analysis of the assembly properties of five EmN variants reveals that del95-99 and Q133H impact filament assembly capacities. In cells, these variants are located at the nuclear envelope but the corresponding quantities of emerin-emerin and emerin-lamin proximities are decreased compared to wild-type protein. Furthermore, variant P183H favors EmN aggregation *in vitro* and variant P183T provokes emerin accumulation in cytoplasmic foci in cells. Substitution of residue Pro183 might systematically favor oligomerization, leading to emerin aggregation and mislocalization in cells. Our results suggest that emerin self-assembly is necessary for its proper function and that loss of either the protein itself or its ability to self-assemble causes muscular dystrophy.

Introduction

In the metazoan nucleus, the genome is surrounded by the nuclear envelope, containing a double lipid membrane and a large number of proteins. A filamentous protein meshwork called the nuclear lamina lines the inner nuclear membrane. It is primarily composed of intermediate filament proteins called A-type (lamin A/C) and B-type lamins. It is anchored to the nuclear membrane by an extensive array of inner nuclear membrane proteins. Lamins and inner nuclear membrane proteins contribute to the architecture of the nuclear envelope. Emerin is one of these inner nuclear membrane proteins. Its gene was discovered in 1994 because it is mutated in patients with an X-linked Emery-Dreifuss muscular dystrophy (EDMD; [1]). Its mRNA shows ubiquitous tissue distribution with the highest expression in skeletal and cardiac muscles. In 1996, availability of anti-emerin antibodies enabled localization of the protein at the nuclear membrane [2, 3]. More precisely, emerin is anchored at the inner nuclear membrane via its hydrophobic C-terminal domain from amino acid 223 to amino acid 243 [4]. The fragment of emerin from amino acid 117 to 170 is sufficient for nuclear accumulation but only the fragment from amino acid 3 to 170 is capable of targeting an integral membrane protein to the inner nuclear membrane [5]. In the heart and cultured cardiomyocytes, emerin might additionally be associated with intercalated discs [6]. Thus, it could play a ubiquitous role in association of the nuclear membrane with the lamins and a tissue-specific role in heart in membrane anchorage to the cytoskeleton.

Emerin interacts with several proteins essential for cytoplasmic and nuclear structure and chromatin organization. Binding partners include lamin A/C, B-type lamins, actin, BAF and lamina-associated polypeptide 1 [7-11]. Emerin interaction with lamin A/C was revealed by several teams and is supported by the fact that a large number of mutations in lamin A/C also cause EDMD [12]. Emerin generally colocalizes with lamins, being mainly at the nuclear periphery in interphase [13]. Furthermore, lamin A is essential for anchorage of emerin to the inner nuclear membrane [14]. However, emerin and lamins do not remain colocalized during mitosis, demonstrating that their interaction is regulated in a cell cycle-dependent manner. Emerin also binds the integral inner nuclear membrane protein lamina-associated polypeptide 1, which similarly associates with lamins and causes muscular dystrophy and cardiomyopathy when deleted from striated muscle [11]. Emerin and lamin A/C further bind actin at the late stages of myotube differentiation and in mature muscle and these interactions are regulated by phosphorylation [15]. Wilson and co-workers proposed that emerin region between residues 1 and 221 caps the pointed end of actin filaments [16]. Lammerding and co-workers showed that emerin regulates the activity of the mechanosensitive transcription factor megakaryoblastic leukemia 1, a myocardin family member that is pivotal in cardiac development and function, through modulation of nuclear and cytoskeletal actin polymerization [17]. Furthermore, emerin organizes actin flow for centrosome orientation and nuclear movement in migrating cells [18]. Emerin is also tyrosine phosphorylated after force is applied on isolated nuclei and this event mediates the nuclear mechanical response to tension [19].

As emerin-deficient mouse embryo fibroblasts have apparently normal nuclear mechanics but impaired expression of mechanosensitive genes in response to strain, it has been suggested that disease-causing loss of function mutations do not act by increasing nuclear fragility but through altered transcriptional regulation[20]. Interaction of emerin with the DNA-bridging protein BAF is important for nuclear membrane targeting to chromatin and chromatin decondensation during nuclear assembly [21] and is regulated through mitotic phosphorylation [22, 23].

Most mutations in the gene encoding emerin in patients with X-linked EDMD are nonsense or out-of-frame mutations that lead to the absence of emerin in patient cells [24]. Only very few EDMD cases are due to missense or short in-frame deletions. Attempts to describe the impact of these mutations on emerin structure and binding properties have been hampered by difficulties in obtaining pure and soluble fragments of the emerin nucleoplasmic region from amino acid 1 to 221 [25]. The three-dimensional structure of the emerin N-terminal LEM domain, from amino acid 1 to 45, has been solved [26] and a model of its structure in complex with a BAF dimer has been proposed [27]. However, it was demonstrated in HeLa cells that the LEM domain and distal tyrosines (a.a. 59, 74, 95, 161) contribute to BAF binding [28]. The region of emerin from amino acid 46 to 221 is predicted mainly intrinsically disordered [29]. It contains the lamin and actin binding regions of emerin. It also comprises four positions (a.a 54, 95-99, 133, 183) whose substitution or deletion do not hinder emerin expression but cause EDMD [7, 30-32]. The structural consequences of these changes in the protein are yet unknown. They could preclude specific emerin modification events or locally modify the three dimensional structure and conformational plasticity of the protein. They could also affect assembly of protein complexes involving emerin, either by altering emerin oligomerization and/or hindering partner recognition.

We now present a combined biochemical and cellular analysis of the nucleoplasmic region of emerin, describing its self-assembly properties and providing novel data on the architecture of emerin oligomers. Wilson and co-workers recently proposed that residues 187-220, at the C-terminus of the nucleoplasmic region, are essential for emerin intermolecular association both *in vitro* and at the nuclear envelope [29]. We further show that emerin fragment EmN, from residue 1 to residue 187, forms filaments *in vitro*, and that two variants with alterations in this fragment modify the organization of emerin at the nuclear envelope. From the comparison between results obtained *in vitro* and in cells on the full-length emerin protein, we discuss mechanisms that may lead to self-assembly defect and loss of function of emerin variants with amino acid substitutions or a small deletion associated to X-linked EDMD.

Results and Discussion

The region of emerin from amino acid 1 to 187 is able to self-associate either through disulfide bridge formation or via hydrophobic interactions – We produced two recombinant fragments of emerin in *Escherichia coli*: the whole nucleoplasmic region preceding the transmembrane domain from amino acid 1 to 221 (EmN0) and a smaller nucleoplasmic fragment from amino acid 1 to 187 (EmN) (Figure 1A,B). Bioinformatics analysis using the Disopred3 webserver predicted that, outside of the LEM domain, these fragments, which include amino acid substitutions and deletions in expressed disease-associated emerin variants, are essentially intrinsically disordered (Figure 1C). However, three regions show a tendency for local order (Metadisorder score lower than 0.7); they comprise residues 74 to 106, residues 159 to 186 and residues 203 to 208, respectively. The two first regions are rich in conserved hydrophobic residues (red squared in Suppl. Figure 1). One of these regions contains the residues deleted in the EDMD-associated variant del95-99 (blue squared in Suppl. Figure 1). The capacities of these regions to locally adopt preferred 3D structures are currently unknown.

As EmN0 and EmN are mainly present in inclusion bodies when expressed in bacterial cells, we purified these fragments by affinity chromatography in 8 M urea. The two fragments were then refolded when dialyzed in Tris 50 mM (pH 6.8), NaCl 150 mM. EmN0 was essentially only obtained as aggregates in all tested conditions (including pH from 5.0 to 9.0 and NaCl concentrations from 30 to 500 mM), as previously observed [25]. Therefore, we decided to use the more soluble EmN for the rest of our study. Analysis of the NMR ^{1}H - ^{15}N HSQC spectrum of refolded EmN showed that this emerin fragment is only partially structured. Indeed, about 50 HSQC signals show non-random coil ^{1}H chemical shift and/or correspond to positive ^{1}H -> ^{15}N nOe values (Figure 1D). Most of these signals can be assigned to the emerin LEM domain based on previous ^{1}H chemical shift assignments (BMRB entry 5074). The remaining HSQC signals have a ^{1}H chemical shift comprised between 7.8 and 8.4 ppm (Figure 1D). Such low dispersion of the backbone ^{1}H resonances suggests that they correspond to EmN unstructured residues. Moreover, analysis of the NMR ^{1}H -> ^{15}N nOe experiment shows that the corresponding ^{1}H-^{15}N correlations are absent from the saturated experiment spectrum when compared to the ^{1}H - ^{15}N HSQC spectrum, which confirmed that these EmN residues are highly flexible in solution (Figure 1D).

We next tested the impact of the refolding buffer on emerin solubility and oligomerization by analytical ultracentrifugation (Figure 2A). In Tris 20 mM (pH 8.0), EmN at 20 μM exists mainly as two different species in solution, as observed in sedimentation velocity experiments. These species are characterized by sedimentation coefficients of 2.2 and 3.3 S, as well as form factors f/f_0 of 1.52 and 1.58, respectively. They correspond to elongated monomeric and dimeric EmN, as confirmed by sedimentation equilibrium experiments (Suppl. Figure 2). Non-reducing SDS-PAGE analysis of the EmN samples in the presence of increasing β-mercaptoethanol concentrations showed that the unique cysteine of EmN (Cys147) is responsible for dimerization through disulfide bridge formation (Figure 2B).

Increasing the salt concentration close to physiological ionic strength favored the formation of larger molecular mass oligomers and/or aggregates of emerin (Table in Figure 2A) that were too large to be observed by NMR. The presence of these large species at high salt concentration suggested that hydrophobic interactions play a role in emerin self-assembly and/or aggregation. Similar experiments carried out after dialysis at lower pHs (5.0 and 6.8), and thus closer to the isoelectric point (5.0) of EmN, led to a massive precipitation of the samples (data not shown). This further illustrated that increasing the weight of hydrophobic interactions compared to electrostatic interactions favored emerin self-assembly and/or aggregation. Finally gel filtration experiments were performed in Tris 20 mM (pH 8.0) at two different protein concentrations (40 and 80 μM, corresponding to 1 and 3 mg/ml, respectively). Increasing EmN concentration led to the formation of high molecular oligomers in these conditions (Figure 2C).

We concluded that EmN was capable of forming different oligomers at physiological ionic strength and that addition of reducing agent and/or salt regulated EmN oligomerization / aggregation state. As the nucleus is characterized by a relatively reducing environment [33] and as the emerin unique cysteine is not conserved between species (residue 147; Suppl. Figure 1), the biological relevance of the disulfide-mediated dimer seemed limited. We choose to control EmN oligomeric state by adding a reducing agent (10 mM β-mercaptoethanol) to the urea buffer and analyze EmN after an overnight dialysis in Tris 20 mM (pH 8.0), NaCl 30 mM. In these reducing and low salt conditions, EmN is initially mainly monomeric (Suppl. Figure 3).

Moreover, it forms high molecular weight oligomers / aggregates with time, as revealed for example by following using NMR a sample of EmN at 600 μM in Tris 20 mM (pH 7.2), NaCl 30 mM, 50 mM β-mercaptoethanol: the EmN NMR signals disappeared within a few days at 10°C because of the formation of species too large to be observed by NMR.

EmN forms filaments rich in β-structure *in vitro* – We studied oligomerization of EmN over time using negative staining electron microscopy (EM). After dialysis, the protein sample was split into four aliquots which were concentrated to either 80 or 400 μM, and either kept at room temperature or heated at 338K during 1 hour; all aliquots were finally kept at room temperature during 3 days. Observation of these aliquots using EM (Figure 3) revealed that at 293K EmN forms either amorphous aggregates and few irregular filament-like structures (80 μM; Figure 3A) or spherical particles of a diameter of about 10 nm as well as filament-like structures (400 μM; Figure 3B). After heating the sample at 338K for 1 hour, EmN forms spherical particles as well as short filaments already at 80 μM (Figure 3C). At high concentration (400 μM) heating promotes the formation of long and curvilinear filaments with diameters of about 10 nm (Figure 3D). These data show that increase in both the temperature (293 to 338K) and the concentration (80 and 400 μM) favors the *in vitro* formation of long and well-separated filaments. In order to test the impact of oxidation on filament assembly, either H_2O_2 or β-mercaptoethanol was added to the EmN sample just before heating during 1 hour at 338K.

EM analyses showed that only amorphous aggregates form in the presence of H_2O_2 (Figure 3E). Thus, although oxidative conditions favor dimer formation (as shown in Figure 2B), they inhibit filament assembly. Instead, reducing conditions allow filament assembly (Figure 3F).

We monitored first structural events associated to oligomerization of EmN by the fluorescent dye thioflavin T (ThT). This benzothiol dye has been used for decades in the diagnosis of protein-misfolding diseases and in kinetic studies of self-assembly. It interacts with β-sheet structures, in particular by docking onto surfaces formed by a single tyrosine ladder [34]. This interaction is revealed by a significant increase in fluorescence intensity at 480 nm. We observed a ten-fold increase in fluorescence intensity at 480 nm when EmN at 300 μM is incubated during either 1 hour at 338K or 3 hours at 310K (Figure 4A). This suggested that EmN with time forms species characterized by the presence of amyloid-like β-structure. Furthermore, the emerin samples were characterized by ATR-FTIR spectroscopy. The amide I spectral region (1600 – 1700 cm^{-1}) is informative about the stretching mode vibrations of the backbone carbonyl groups and is affected by the secondary structure of the polypeptide chain. In the spectrum of monomeric emerin, amide I vibrations evolved at 1650 cm^{-1} (black in Figure 4B), which suggested the presence of random coil-like structure. After incubation during 1 hour at 338K, an extra shoulder at 1615 cm^{-1} (green in Figure 4B) was observed, which revealed formation of β-sheet structure related to emerin peptide self-assembly. Analysis using SDS-PAGE revealed that incubation during 4 hours at 338K led to the formation of high molecular weight oligomers unable to enter the gel (Figure 4C). The presence of these oligomers yielded to a substantial decrease in total protein staining. EM images of EmN incubated during 4 hours at 338K confirmed that these oligomers were either spherical particles or filaments with a diameter of about 10 nm (Figure 4D). Similar analyses carried out after incubation at 310K showed that after 4 hours the EmN sample can still be observed by SDS-PAGE (Figure 4C) and that well-structured filaments can only be observed after several days (Figure 4D).

Thus both formation of β-structure and assembly of filaments were facilitated when increasing the temperature from 310 to 338K. However, we also observed the EmN filaments after incubation at 310K.

EmN variants del95-99 and Q133H occurring in EDMD are impaired in their capacity to self-assemble – We produced five EmN variants, corresponding to emerin variants that occur in EDMD: S54F, del95-99 (the deleted sequence being YEESY), Q133H, P183T and P183H. These variants are similarly mainly monomeric after dialysis (Suppl. Figure 3). We further tested their capacity to self-assemble.

Following EmN self-assembly kinetics at pH 8.0 and 310K using ThT revealed that variants S54F and P183T behave as wild-type EmN (Figure 5A). In contrast, Q133H showed a reduced ability to form β-rich structures and del95-99 did not form any type of β-rich assembly within 1 day.

Further analysis of del95-99 after incubation during 4 hours at 338K showed that even in these conditions, the variant could be observed on SDS-PAGE gel, whereas the wild-type EmN was already assembled into high molecular weight oligomers (Figure 5B). Observation by negative staining EM of the different variants after incubation at a concentration higher than 300 μM and a temperature of 338K during 1 hour confirmed that, if del95-99 was unable to form filaments in these conditions (data not shown), Q133H and P183T self-assembled into filaments (Figure 6). Filaments obtained from variant P183T were indistinguishable from wild-type filaments, whereas filaments formed by Q133H at a similar time point were systematically shorter and less regular. In the case of P183H, the self-assembly kinetics was not significantly different from that measured for wild-type EmN. However, after 1 day, further self-assembly or aggregation systematically led to a decrease in detectable fluorescent signal. Such event was also observed in the case of wild-type EmN and variants S54F and P183T but only after several days. When wild-type EmN and the variants were dialyzed against a buffer at pH 6.8, closer to EmN isoelectric point (5.0), most of them could not be concentrated to 300 μM because of aggregation. Only Q133H and del95-99 could yield this concentration. Characterization of their capacity to form β-structure after 1 day using the ThT assay showed that, at pH 6.8, Q133H had an improved capacity to self-assemble as compared to pH 8.0, whereas del95-99 still did not form filament-like structures (Suppl. Figure 4). Thus, lowering the pH from 8.0 to 6.8 favored either aggregation or self-assembly, but did not change the relative capacities of the variants to self-assemble. Altogether, these experiments suggest that P183H has the strongest capacity to self-assemble or aggregate, whereas Q133H and then del95-99 have the lowest capacities to form filament-like structures.

Proximity ligation assays suggest the presence of emerin oligomers in cells and reveal the impact of EDMD-associated variants on emerin-emerin and emerin-lamin proximities at the nuclear periphery – We carried out experiments in HeLa cells in order to detect intermolecular emerin-emerin and emerin-lamin proximities. Cells were first transfected with plasmids coding for both wild-type green fluorescent protein (GFP)-full length emerin and FLAG-full length emerin (Figure 7A). GFP-FLAG proximities were then assessed using proximity ligation assays (PLA; Figure 7B). These experiments revealed emerin-emerin proximities mostly located in the nucleus, and particularly enriched at the nuclear periphery. Additional proximities were observed in the cytoplasm that could correspond to emerin located in the endoplasmic reticulum. The effects of three emerin variants with different filament assembly properties *in vitro*, del95-99, Q133H and P183T, were then tested on emerin architecture in cells. Cells were transfected with a couple of plasmids coding for GFP and FLAG fusions of each variant (Figure 7A).

Although HeLa cells express a pool of endogenous wild-type emerin, the design of our PLA assay allowed quantification of emerin-emerin proximity events occurring only between ectopic GFP and FLAG-tagged emerins. Quantification of PLA signals within the nuclear compartment demonstrated that variants del95-99 and Q133H created less emerin-emerin proximities than wild-type emerin at the nuclear periphery of transfected cells (Figure 7C). Variant P183T uniquely accumulated into foci in the endoplasmic reticulum.

Depending on the experiment and due to variability in the fraction of this variant to localize in the inner nuclear membrane, the amount of emerin-emerin proximities visualized by PLA was not reproducible.

Cells were also transfected with only one plasmid encoding GFP-emerin (Figure 8A). In these cells, proximities between GFP-emerin and endogenous lamin A/C were observed at the intranuclear periphery (Figure 8B). We further measured the proximities between the emerin del95-99, Q133H, P183T and lamin A/C by PLA. Variants del95-99 and Q133H were significantly less frequently close to lamin A/C in HeLa cells (Figure 8C). In the case of emerin P183T, which is not correctly targeted to the nuclear envelope, the amount of emerin-lamin A/C proximities visualized by PLA was not reproducible. However, intense PLA signals could be observed at the nuclear envelope suggesting that variant P183T is still capable of interacting with lamin A/C.

A specific role for proline 183 in regulating emerin oligomerization – Proline is a residue playing an important structural role in IDRs by hindering their folding into aggregation prone structures. Mutation of this residue is frequently described as favoring the formation of amyloid-like structures. In the case of EmN, amino acid substitutions P183T and P183H do not significantly modify EmN filament assembly properties *in vitro* at pH 8.0. The molecular impact of these mutations is limited to a lower solubility of variant P183H at pH 8.0. However, we show in HeLa cells and in the context of the whole emerin protein that GFP-emerin P183T shows a significant tendency to form cytoplasmic foci close to the nuclear membrane. Ellis et al. (1999) previously demonstrated that in patient cells, emerin variants P183T and P183H were no longer confined to the nuclear fraction, but were also distributed in other membranous fractions, with the majority being associated with “heavier” membrane fractions (lysosomes, plasma membrane, mitochondria), most particularly in the case of P183H [31]. Wilson and co-workers demonstrated that deletion of the region from amino acid 168 to 186 favored assembly of emerin oligomers stabilized by intermolecular interactions involving the region from amino acid 187 to 220 *in vitro* and in cells [29]. They showed that it caused a relocalization of emerin into perinuclear aggregates. These data suggest that the region from amino acid 168 to 186 and in particular Pro183 limits oligomerization through the region from amino acid 187 to 220 in full-length emerin. We propose that emerin can form two types of intermolecular interactions in the reducing environment of the nucleus, one resulting from EmN self-assembly and the other resulting from dimerization of the region from amino acid 187 to 220. In this frame, Pro183 would play a critical role in regulating emerin self-assembly, not by directly influencing EmN filament formation, but by limiting the formation of oligomers stabilized by self-assembly of the region from amino acid 187 to 220 in cells. Variants P183T and P183H would consistently show a stronger tendency to form dimers through region from amino acid 187 to 220, which would lead to aggregation and mislocalization in the cytoplasm.

Relationship between emerin oligomeric state and lamin binding capacity – Binding of emerin to lamin A, which has been described using several techniques including yeast 2-hybrid assays and co-immunoprecipitation experiments, involves the fragment from amino acid 70 to 164 of emerin [35]. Within this region, deletion del95-99 caused emerin to be present but significantly less associated to membranes and to the nuclear lamina components in patient cells [30], and hindered binding to lamin A *in vitro* as judged by blot overlay assays [10]. From our results, we now propose a model in which loss of function of emerin variant del95-99 (weaker lamin A/C binding by PLA) relies on its defective organization at the nuclear envelope (weaker emerin-emerin associations by PLA) and is related to its self-assembly defects (absence of oligomerization *in vitro*). For variant Q133H, no qualitative or quantitative differences in nuclear targeting were observed between mutant and normal emerin and quantitative BIAcore analysis showed no significant change in lamin A binding to emerin when the mutation was present [32]. However, our results showed that the amount of emerin–lamin A/C proximities is decreased, suggesting that emerin self-assembly defects impact lamin recognition in cells. It is also possible that emerin self-assembly defects affects its interaction with other binding proteins, such as lamina-associated polypeptide 1.

Variants P83T and P183H present a mutation outside of the lamin A/C binding region of emerin. In contrast to mutations del95-99 and Q133H, we observed that mutation P183T does not prevent binding to lamin A/C (as judged by PLA). Based on an *in vitro* blot overlay assay [10], Lee et al. proposed that variant P183H has an even increased lamin binding capacity as compared to wild-type emerin. We could not see such a clear effect for variant P183T in our cell assay. We propose that the observed mistargeting of P183T emerin within cytoplasmic aggregates might be favored upon specific stimuli in the muscular tissue and consequently impact (in an indirect fashion) the formation of lamin A/C-emerin complexes at the nuclear envelope. Further work is needed to determine to what extent amino acid substitutions P183T and P183H affect lamin A/C binding in various cellular contexts and to determine whether this can be related to the proposed increased dimerization capacity through region 170-220 of that variant (see above).

Conclusion – Our study provides a first structural description of the region from amino acid 1 to amino acid 187 of emerin and reveals the structural consequences of several mutations that occur in EDMD on the self-assembly capacity of this EmN fragment *in vitro* and the amount of emerin-emerin proximities in cells. Full-length emerin self-associates either through the EmN fragment, as shown in this study, or through the region between residues 187 and 220, as shown by others [29]. Oxydation of EmN hinders its self-assembly into high molecular weight oligomers. Mutations at positions 95-99 and 133 affect EmN self-assembly and impact emerin binding to lamin A/C in cells, whereas mutations of Pro183 affect emerin self-assembly probably mainly through the region between residues 187 and 220 and impact emerin localization in cells. Virtually all of the known EDMD-causing mutations lead to an absence of emerin in cells, suggesting that S54F, del95-99, Q133T, P183T and P183H are loss-of-function mutations.

Defects in oligomerization properties essential for emerin localization and function might explain the decreased functional capacities of the variants del95-99, Q133T, P183T and P183H. Variant S54F does not impact EmN self-assembly *in vitro*, and it is still correctly localized in cells [36]. However, Ser54 is either phosphorylated or *O*-GlcNAc-modified in wild-type emerin and mutation of Ser54 into phenylalanine impairs posttranslational modifications at several other positions in full-length emerin [36]. As the emerin regions proposed to mediate EmN self-assembly are hotspots for phosphorylation and *O*-GlcNAcylation in cells, we suggest that S54F might interfere with emerin-emerin interactions by perturbing the modification state of emerin *in vivo*. Overall our results suggest that emerin self-assembly is necessary for its proper function and that loss of either the protein itself or its ability to self-assemble causes muscular dystrophy.

Methods

Protein expression and purification - Human wild-type emerin fragments from amino acid 1 to amino acid 187 (EmN) and from amino acid 1 to amino acid 221 (EmN0) were expressed using a pETM13 vector as N-terminal octa-histidine fusions in *Escherichia coli* BL21 DE3 Star (Novagen). The emerin cDNAs were optimized for expression in *Escherichia coli* (GenScript). Mutations encoding emerin variants S54F, Q133H, Δ95-99, P183H and P183T were inserted in the EmN cDNA using a standard QuikChange Site-Directed Mutagenesis kit (Stratagene). The EmGCN4 protein was also expressed using pETM13. Its cDNA was obtained by synthetizing an optimized sequence coding for the N-terminal octa-histidine tagged EmN, a TEV cleavage site and a GCN4-derived peptide of 33 amino acids that efficiently dimerizes [37]. Bacteria were cultured in ^{15}N-labelled minimum medium, induced at an optical density of 1 with 0.1 mM isopropyl β-D-1-thiogalactopyranoside, grown overnight at 20 °C and lysed in 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 40 mM imidazole, 5% glycerol, 1% Triton X-100 and 1 mM phenylmethanesulfonylfluoride. After centrifugation at 20,000g for 20 min at 4 °C, the pellet was resuspended in buffer C8 (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 40 mM imidazole, 8 M urea). A second centrifugation step was performed at 20,000g for 20 min at 4 °C. The soluble extract was then filtered and loaded onto a Ni-NTA column (GE-Healthcare) equilibrated with buffer C8. Proteins were eluted directly with buffer E8 (50 mM Tris-HCl pH 8, 150 mM NaCl, 1 M imidazole, 8 M urea). Proteins were refolded by dialysis in the desired buffer. The final yield was typically 10 mg purified protein per litre of bacterial culture. We verified that purification either from the bacterial soluble fraction or from inclusion bodies yielded a protein with the same conformation by recording NMR ^{1}H - ^{15}N HSQC spectra on both preparations. These two NMR spectra were superimposable (data not shown), which validates the use of a protocol including denaturation and refolding steps to obtain a pure, soluble and concentrated preparation. Oligomerization studies were performed using proteins that were reduced after affinity purification in urea by adding 10 mM β-mercaptoethanol and refolded in a buffer containing 30 mM NaCl and 10 mM dithiothreitol (DTT) to observe oligomer formation from a monomeric preparation.

NMR spectroscopy - NMR experiments were performed on a 700 MHz Bruker Avance spectrometer equipped with a cryogenic probe. Two-dimensional $^{1}H - ^{15}N$ correlation spectra were acquired using a HSQC pulse sequence at 293K, on 3 mm diameter NMR sample tube containing 100 μM uniformly ^{15}N-labelled EmN in 50mM Tris (pH 6.8), 150 mM NaCl, 10 mM DTT and 90 %:10 %, H_2O:D_2O.

Analytical ultracentrifugation - 360 μL protein samples at a concentration of 20 to 30 μM were dialyzed against 20 mM Tris (pH 8.0) and 0 to 150 mM NaCl. They were subjected to sedimentation velocity runs at 293K and 45,000 rpm (147,280 g) on a XLA70 analytical ultracentrifuge (Beckman Coulter, Palo Alto, USA) using a An-60Ti rotor. Optical density scans recorded every 5 min at 280 nm were analyzed using the Sedfit software [38] to determine the sedimentation coefficients. The EmN sample was also subjected to a sedimentation equilibrium experiment in 50 mM Tris (pH 8.0) at two successive speeds (11,500 and 13,800 rpm / 9619 and 13851 g) and 293K. Data analysis was performed on the scans acquired on the equilibrated system using the Sedphat software [38] to determine the molecular masses of the species.

Analytical gel filtration – The proteins were first dialysed in the gel-filtration buffer (20 mM Tris-HCl pH 8.0, 30 mM NaCl) before being diluted to 1 mg/ml and loaded onto the column. Analytical gel-filtration experiments were carried out at 277K on a Superdex 75 10/30 column (GE Healthcare).

EM – To follow large oligomer assembly, EmN, its variants or EmGCN4 in buffer E8 and 10 mM β-mercaptoethanol were dialyzed against 20 mM Tris (pH 8.0), 30 mM NaCl, 10 mM DTT. They were then incubated at 20, 37 or 338K. Samples were taken at regular time points and observed by negative-staining transmission electron microscopy. For experiments at 20 and 338K, His-tagged wild-type EmN was used. For experiments at 310K, the His-Tag from the wild-type and mutated EmN constructs was cleaved using Tobacco Etch Virus (TEV) protease (see below), and proteins without Tag were incubated in 8 M urea and redialyzed against 20 mM Tris (pH 8.0), 30 mM NaCl, 10 mM DTT. For EM experiments, about 2-5 μl of the sample solution was deposited on a carbon-coated formvar copper grid. After 30 s, the sample droplet was blotted with filter paper. It was placed in contact with 10 μl of water and then uranyl acetate 0.5%. Samples were imaged using a JEOL MET 1,400 (120 keV) or a Tecnai G2 (200 keV) transmission electron microscope at 10,000x to 30,000x magnification.

Thioflavin kinetics – For these experiments, the His-tag of the wild-type and mutated EmN constructs was cleaved using Tobacco Etch Virus (TEV) protease during 2 h at room temperature. EmN mixed to the His-tagged TEV was then incubated with Ni-NTA beads during 1 h in buffer C0 (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 40 mM imidazole). It was collected by washing with C0, incubated in 8 M urea and redialyzed in 20 mM Tris (pH 8.0), 30 mM NaCl, 10 mM DTT. Filament assembly was followed with time by incubating EmN at the targeted concentration and temperature and by regularly taking protein aliquots to be analyzed by fluorescence. These aliquots were diluted in a thioflavin (ThT) containing buffer so as to obtain 20 μM protein and 2,5 μM ThT in 20 mM Tris (pH 8.0), 30 mM NaCl, 10mM β-mercaptoethanol.

The fluorescence measurements were carried out in a 60 µl cuvette at 293K using a fluorimeter JASCO ADP-303T. ThT fluorescence was monitored by excitation wavelength of 450 nm. Fluorescence emission was read at 480 nm.

Attenuated total reflection – Fourier-transform infrared (ATR-FTIR) spectrometry – ATR-FTIR spectra were recorded at 4 cm^{-1} resolution with a Bruker IFS 66 spectrophotometer equipped with a 45° N ZnSe attenuated total reflection attachment. For the FTIR experiments, EmN samples were concentrated to 800 µM. The concentrator flow-through was used as the reference buffer sample for measuring buffer signal. FTIR spectra of the EmN sample and its corresponding buffer solution were initially recorded just after concentration. Thereafter the EmN sample was incubated during 1 hour at 338K and additional spectra of both protein and buffer samples were recorded. Spectra of both protein and buffer samples were recorded with an average of 30 scans. The buffer spectra were subsequently subtracted from the protein spectra.

Transfection - HeLa cells were obtained from American Type Culture Collection and cultured in Minimum Essential Medium containing Glutamax (Gibco), 1% non-essential amino acids and 10% fetal bovine serum. HeLa cells were transfected using XtremeGene 9 (Roche). After 24 h, cells were processed for immunoblotting, immunofluorescence or Proximity ligation assay.

Immunoblotting - Whole cell protein extracts were suspended in Laemmli sample buffer, separated by SDS-PAGE and transferred to nitrocellulose membranes. Membranes were blocked for 1.5 h in TBS (10 mM Tris, pH 8.0, 150 mM NaCl, 0.05% Tween-20) containing 5% dry milk, incubated with mouse anti-emerin antibody (Leica; 1:300) and rabbit anti-lamin A/C (see [39]; 1:5000) for 1 h in TBS with 1% milk, washed 4 times and incubated with HRP-conjugated secondary antibodies. After 4 washes in TBS, proteins were detected by enhanced chemiluminescence.

Immunofluorescence microscopy - Cells were fixed with 3% paraformaldehyde for 12 min at room temperature, permeabilized with phosphate-buffered saline containing 0.5 % Triton for 5 min at room temperature and quenched with 2% bovine serum albumin diluted in phosphate-buffered saline containing 0.1% Triton X-1-00. Primary antibodies were rabbit anti-lamin A/C (1:500; [40]), rabbit anti-FLAG (Sigma, 1:150) and mouse anti-GFP (Roche, 1 :150). Fluorescent labeled secondary antibodies (donkey anti-mouse Cy2 1:60 and donkey anti-rabbit Cy3 1:200) were from Jackson ImmunoResearch. DNA was stained with Hoechst 33258 (1 µg/ml).

Proximity ligation assay (PLA) - PLA was used to detect GFP-emerin - lamin A/C interactions and GFP-emerin – FLAG-emerin interactions based on proximity (< 40 nm) of two secondary antibodies directed against these proteins. After cell fixation, cell permeabilisation and quenching (as above), pairs of primary antibodies, mouse anti-GFP 1:150 and rabbit anti-lamin A/C 1:500 or mouse anti-GFP 1:150 and rabbit anti-FLAG 1:150, were added to HeLa cells expressing GFP-emerin either alone or together with FLAG-emerin, respectively, for 30 min at room temperature. Next, Duolink PLA probe anti-rabbit plus, Duolink PLA probe anti-mouse minus and Duolink detection reagents orange (detected with a Cy3 filter) were used according to manufacturer's instructions (Olink, Bioscience). Confocal microscopy image acquisition was performed using a LSM 700 Laser scanning microscope (Zeiss). Quantitative analysis of PLA signals was done on images using Image J. Data were then analyzed by comparing median values for integrated densities of signals (Cy3) per nucleus, and statistical analysis were performed using Kruskall-Wallis tests.

Acknowledgments – The authors thank E. Kasotakis and M. Paternostre for constant help in carrying out ATR-FTIR and EM experiments, and F.X. Theillet and P. Selenko for fruitful discussions. They acknowledge the French national infrastructure FRISBI program for access to the NMR spectrometers at CEA Saclay (Gif-sur-Yvette, France), the Functional Biology and Adaptive (BFA) institute for access to the imaging facility at University Paris 7 and the Leibniz-Institut für Molekulare Pharmakologie (FMP Berlin, Germany) for access to the fluorescence and electron microscopy facilities of the institute.

Funding - This work was supported by CNRS and University Paris 7 and by funding from the French Association against Myopathies (AFM) (research grant n°17243 to S.Z.J. and PhD fellowship n°18159 to C.S.) and from the Foundation for Medical Research (FRM) (grant FDT20140931008 to I.H.). H.J.W. was supported by a grant (R01AR048997) from NIH/NIAMS.

References

[1] Bione, S., Maestrini, E., Rivella, S., Mancini, M., Regis, S., Romeo, G., and Toniolo, D. (1994) Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular dystrophy, *Nat Genet 8*, 323-327.

[2] Manilal, S., Nguyen, T. M., Sewry, C. A., and Morris, G. E. (1996) The Emery-Dreifuss muscular dystrophy protein, emerin, is a nuclear membrane protein, *Hum Mol Genet 5*, 801-808.

[3] Nagano, A., Koga, R., Ogawa, M., Kurano, Y., Kawada, J., Okada, R., Hayashi, Y. K., Tsukahara, T., and Arahata, K. (1996) Emerin deficiency at the nuclear membrane in patients with Emery-Dreifuss muscular dystrophy, *Nat Genet 12*, 254-259.

[4] Yorifuji, H., Tadano, Y., Tsuchiya, Y., Ogawa, M., Goto, K., Umetani, A., Asaka, Y., and Arahata, K. (1997) Emerin, deficiency of which causes Emery-Dreifuss muscular dystrophy, is localized at the inner nuclear membrane, *Neurogenetics 1*, 135-140.

[5] Ostlund, C., Ellenberg, J., Hallberg, E., Lippincott-Schwartz, J., and Worman, H. J. (1999) Intracellular trafficking of emerin, the Emery-Dreifuss muscular dystrophy protein, *J. Cell Sci. 112 (Pt 11)*, 1709-1719.

[6] Cartegni, L., di Barletta, M. R., Barresi, R., Squarzoni, S., Sabatelli, P., Maraldi, N., Mora, M., Di Blasi, C., Cornelio, F., Merlini, L., Villa, A., Cobianchi, F., and Toniolo, D. (1997) Heart-specific localization of emerin: new insights into Emery-Dreifuss muscular dystrophy, *Hum Mol Genet 6*, 2257-2264.

[7] Fairley, E. A., Kendrick-Jones, J., and Ellis, J. A. (1999) The Emery-Dreifuss muscular dystrophy phenotype arises from aberrant targeting and binding of emerin at the inner nuclear membrane, *J. Cell Sci. 112 (Pt 15)*, 2571-2582.

[8] Clements, L., Manilal, S., Love, D. R., and Morris, G. E. (2000) Direct interaction between emerin and lamin A, *Biochem. Biophys. Res. Commun. 267*, 709-714.

[9] Sakaki, M., Koike, H., Takahashi, N., Sasagawa, N., Tomioka, S., Arahata, K., and Ishiura, S. (2001) Interaction between emerin and nuclear lamins, *J. Biochem. 129*, 321-327.

[10] Lee, K. K., Haraguchi, T., Lee, R. S., Koujin, T., Hiraoka, Y., and Wilson, K. L. (2001) Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF, *J. Cell Sci. 114*, 4567-4573.

[11] Shin, J. Y., Mendez-Lopez, I., Wang, Y., Hays, A. P., Tanji, K., Lefkowitch, J. H., Schulze, P. C., Worman, H. J., and Dauer, W. T. (2013) Lamina-associated polypeptide-1 interacts with the muscular dystrophy protein emerin and is essential for skeletal muscle maintenance, *Dev Cell 26*, 591-603.

[12] Bonne, G., Di Barletta, M. R., Varnous, S., Becane, H. M., Hammouda, E. H., Merlini, L., Muntoni, F., Greenberg, C. R., Gary, F., Urtizberea, J. A., Duboc, D., Fardeau, M., Toniolo, D., and Schwartz, K. (1999) Mutations in the gene encoding lamin A/C cause autosomal dominant Emery-Dreifuss muscular dystrophy, *Nat Genet 21*, 285-288.

[13] Manilal, S., Nguyen, T. M., and Morris, G. E. (1998) Colocalization of emerin and lamins in interphase nuclei and changes during mitosis, *Biochem. Biophys. Res. Commun. 249*, 643-647.

[14] Vaughan, A., Alvarez-Reyes, M., Bridger, J. M., Broers, J. L., Ramaekers, F. C., Wehnert, M., Morris, G. E., Whitfield, W. G. F., and Hutchison, C. J. (2001) Both emerin and lamin C depend on lamin A for localization at the nuclear envelope, *J. Cell Sci. 114*, 2577-2590.

[15] Lattanzi, G., Cenni, V., Marmiroli, S., Capanni, C., Mattioli, E., Merlini, L., Squarzoni, S., and Maraldi, N. M. (2003) Association of emerin with nuclear and cytoplasmic actin is regulated in differentiating myoblasts, *Biochem. Biophys. Res. Commun. 303*, 764-770.

[16] Holaska, J. M., Kowalski, A. K., and Wilson, K. L. (2004) Emerin caps the pointed end of actin filaments: evidence for an actin cortical network at the nuclear inner membrane, *PLoS Biol 2*, E231.

[17] Ho, C. Y., Jaalouk, D. E., Vartiainen, M. K., and Lammerding, J. (2013) Lamin A/C and emerin regulate MKL1-SRF activity by modulating actin dynamics, *Nature 497*, 507-511.

[18] Chang, W., Folker, E. S., Worman, H. J., and Gundersen, G. G. (2013) Emerin organizes actin flow for nuclear movement and centrosome orientation in migrating fibroblasts, *Mol Biol Cell 24*, 3869-3880.

[19] Guilluy, C., Osborne, L. D., Van Landeghem, L., Sharek, L., Superfine, R., Garcia-Mata, R., and Burridge, K. (2014) Isolated nuclei adapt to force and reveal a mechanotransduction pathway in the nucleus, *Nat Cell Biol 16*, 376-381.

[20] Lammerding, J., Hsiao, J., Schulze, P. C., Kozlov, S., Stewart, C. L., and Lee, R. T. (2005) Abnormal nuclear shape and impaired mechanotransduction in emerin-deficient cells, *J. Cell Biol. 170*, 781-791.

[21] Segura-Totten, M., Kowalski, A. K., Craigie, R., and Wilson, K. L. (2002) Barrier-to-autointegration factor: major roles in chromatin decondensation and nuclear assembly, *J. Cell Biol. 158*, 475-485.

[22] Hirano, Y., Segawa, M., Ouchi, F. S., Yamakawa, Y., Furukawa, K., Takeyasu, K., and Horigome, T. (2005) Dissociation of emerin from barrier-to-autointegration factor is regulated through mitotic phosphorylation of emerin in a xenopus egg cell-free system, *J. Biol. Chem. 280*, 39925-39933.

[23] Bengtsson, L., and Wilson, K. L. (2006) Barrier-to-autointegration factor phosphorylation on Ser-4 regulates emerin binding to lamin A in vitro and emerin localization in vivo, *Mol Biol Cell 17*, 1154-1163.

[24] Yates, J. R., and Wehnert, M. (1999) The Emery-Dreifuss Muscular Dystrophy Mutation Database, *Neuromuscul Disord 9*, 199.

[25] Roberts, R. C., Sutherland-Smith, A. J., Wheeler, M. A., Jensen, O. N., Emerson, L. J., Spiliotis, II, Tate, C. G., Kendrick-Jones, J., and Ellis, J. A. (2006) The Emery-Dreifuss muscular dystrophy associated-protein emerin is phosphorylated on serine 49 by protein kinase A, *FEBS J 273*, 4562-4575.

[26] Wolff, N., Gilquin, B., Courchay, K., Callebaut, I., Worman, H. J., and Zinn-Justin, S. (2001) Structural analysis of emerin, an inner nuclear membrane protein mutated in X-linked Emery-Dreifuss muscular dystrophy, *FEBS Lett. 501*, 171-176.

[27] Cai, M., Huang, Y., Suh, J. Y., Louis, J. M., Ghirlando, R., Craigie, R., and Clore, G. M. (2007) Solution NMR structure of the barrier-to-autointegration factor-Emerin complex, *J. Biol. Chem. 282*, 14525-14535.

[28] Tifft, K. E., Bradbury, K. A., and Wilson, K. L. (2009) Tyrosine phosphorylation of nuclear-membrane protein emerin by Src, Abl and other kinases, *J. Cell Sci. 122*, 3780-3790.
[29] Berk, J. M., Simon, D. N., Jenkins-Houk, C. R., Westerbeck, J. W., Gronning-Wang, L. M., Carlson, C. R., and Wilson, K. L. (2014) The molecular basis of emerin-emerin and emerin-BAF interactions, *J. Cell Sci. 127*, 3956-3969.
[30] Ellis, J. A., Craxton, M., Yates, J. R., and Kendrick-Jones, J. (1998) Aberrant intracellular targeting and cell cycle-dependent phosphorylation of emerin contribute to the Emery-Dreifuss muscular dystrophy phenotype, *J. Cell Sci. 111 (Pt 6)*, 781-792.
[31] Ellis, J. A., Yates, J. R., Kendrick-Jones, J., and Brown, C. A. (1999) Changes at P183 of emerin weaken its protein-protein interactions resulting in X-linked Emery-Dreifuss muscular dystrophy, *Hum Genet 104*, 262-268.
[32] Holt, I., Clements, L., Manilal, S., and Morris, G. E. (2001) How does a g993t mutation in the emerin gene cause Emery-Dreifuss muscular dystrophy?, *Biochem. Biophys. Res. Commun. 287*, 1129-1133.
[33] Go, Y. M., and Jones, D. P. (2010) Redox control systems in the nucleus: mechanisms and functions, *Antioxid Redox Signal 13*, 489-509.
[34] Biancalana, M., Makabe, K., Koide, A., and Koide, S. (2009) Molecular mechanism of thioflavin-T binding to the surface of beta-rich peptide self-assemblies, *J. Mol. Biol. 385*, 1052-1063.
[35] Berk, J. M., Tifft, K. E., and Wilson, K. L. (2013) The nuclear envelope LEM-domain protein emerin, *Nucleus 4*, 298-314.
[36] Berk, J. M., Maitra, S., Dawdy, A. W., Shabanowitz, J., Hunt, D. F., and Wilson, K. L. (2013) O-Linked beta-N-acetylglucosamine (O-GlcNAc) regulates emerin binding to barrier to autointegration factor (BAF) in a chromatin- and lamin B-enriched "niche", *J. Biol. Chem. 288*, 30192-30209.
[37] Ciani, B., Bjelic, S., Honnappa, S., Jawhari, H., Jaussi, R., Payapilly, A., Jowitt, T., Steinmetz, M. O., and Kammerer, R. A. (2010) Molecular basis of coiled-coil oligomerization-state specificity, *Proc Natl Acad Sci U S A 107*, 19850-19855.
[38] Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling, *Biophys. J. 78*, 1606-1619.
[39] Duband-Goulet, I., Woerner, S., Gasparini, S., Attanda, W., Konde, E., Tellier-Lebegue, C., Craescu, C. T., Gombault, A., Roussel, P., Vadrot, N., Vicart, P., Ostlund, C., Worman, H. J., Zinn-Justin, S., and Buendia, B. (2011) Subcellular localization of SREBP1 depends on its interaction with the C-terminal region of wild-type and disease related A-type lamins, *Exp Cell Res 317*, 2800-2813.
[40] Vadrot, N., Duband-Goulet, I., Cabet, E., Attanda, W., Barateau, A., Vicart, P., Gerbal, F., Briand, N., Vigouroux, C., Oldenburg, A. R., Lund, E. G., Collas, P., and Buendia, B. (2015) The p.R482W substitution in A-type lamins deregulates SREBP1 activity in Dunnigan-type familial partial lipodystrophy, *Hum Mol Genet 24*, 2096-2109.
[41] Kozlowski, L. P., and Bujnicki, J. M. (2012) MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins, *BMC Bioinformatics 13*, 111.

Figure legends:

Figure 1. Emerin region from residue 1 to residue 187 (EmN) is predicted and observed to be mainly unstructured in solution. (A) Schematic representation of emerin. The two structured domains are represented in orange for the LEM domain (LEM 1-45) and in green for the transmembrane domain (TM 223-243). **(B)** Emerin fragments expressed for this study. **(C)** Metadisorder profile of emerin [41]: apart from the two structured domains (LEM and TM) that have a disorder probability lower than 0.5, most of the other residues are predicted as disordered. Nevertheless, three regions, between residues 74 to 106, residues 159 to 186 and residues 203 to 208, show intermediate scores between 0.5 and 0.7. The red dots mark positions mutated in EDMD causing missense or micro-deletion variants. **(D)** 2D NMR ^{1}H-^{15}N spectra recorded on a EmN sample at 100µM in 50mM Tris-HCl pH6.7, 150mM NaCl, 10 mM DTT at 293K and 700 MHz. Superposition of the ^{1}H-^{15}N HSQC spectrum in black (each peak corresponds to one residue; peaks corresponding to poorly structured residues have a ^{1}H chemical shift comprised between 7.8 and 8.4 ppm) and the saturated $^{1}H \rightarrow ^{15}N$ nOe spectrum in orange (positive peaks corresponding to structured residues) and blue (negative peaks corresponding to fully disordered residues). Residues with intermediate dynamical behavior have nOe values close to zero and thus are not visible on the saturated ^{1}H to ^{15}N nOe spectrum.

Figure 2. EmN can form dimers and higher molecular weight oligomers / aggregates in solution. (A) AUC sedimentation velocity results obtained on samples of EmN at 20µM in 20mM Tris-HCl pH8, in three different salt conditions (0mM, 30mM and 150mM NaCl). Three species are present in these conditions, which might correspond to an elongated monomer (sedimentation coefficient of 2.1S and frictional ratio of 1.52), an elongated dimer (sedimentation coefficient of 3.2S and frictional ratio of 1.58) and high molecular weight oligomers and/or aggregates. **(B)** Non-reducing SDS-PAGE of EmN in 20 mM Tris (pH 8.0), 30 mM NaCl and different concentrations of β-mercaptoethanol showing a disappearance of EmN dimers in presence of reducing agent. **(C)** Gel filtration profile obtained after dialysis of EmN: samples 1 (black full lines) and 2 (black dashed lines) were injected on the column at 1 and 3 mg/ml respectively. Clearly sample 2 contains oligomers and/or aggregates that could not pass the column prefilter. Both samples also contain monomers (elution volume corresponding to 33 kDa, the protein being a 25 kDa elongated protein), dimers (elution volume corresponding to 66 kDa) and oligomers (elution at the void volume).

Figure 3. EmN can form filaments in a concentration / temperature / reducing conditions dependent manner. (A-D) EM images acquired on a EmN sample concentrated at either 80 or 400 µM in a low salt concentration buffer (20mM Tris-HCl [pH8.0], 30mM NaCl, 10mM β-mercaptoethanol) and not heated or heated during 1 hour at 65ºC. **(E-F)** EM images acquired on a EmN sample concentrated at 400 µM with either **(E)** 10 mM H_2O_2 or **(F)** 10 mM β-mercaptoethanol and heated 1 hour at 65ºC in a low salt concentration buffer (20 mM Tris-HCl [pH8.0], 30 mM NaCl). All samples were observed by negative staining EM using uranyl acetate after 3 days at 293K. The black bars correspond to 200 nm.

Figure 4. EmN filament assembly at either 37 or 338K is associated with an increase of β-structure. EmN self-assembly kinetics followed by

(**A)** ThT fluorescence on a EmN sample at 300 μM in 20 mM Tris-HCl (pH8.0), 30 mM NaCl, 10 mM β-mercaptoethanol, heated at either 37ºC (orange) or 65ºC (green) during 4 hours, **(B)** ATR-FTIR on a EmN sample at 1.2 mM before (black) and after (green) heating during 1 hour at 65ºC, **(C)** SDS-PAGE on a EmN sample at 300 μM (S and P correspond to Supernatant and Pellet, respectively; the arrow corresponds to the emerin monomer) and **(D)** EM on a EmN sample at 700 μM (the black bar corresponds to 200 nm; negative staining was achieved using uranyl acetate).

Figure 5. EmN EDMD-causing variants del95-99 and Q133H have impaired oligomerization capacities. (A) Fluorescence signals measured at 480 nm after addition of ThT. The wild-type and mutated EmN samples were concentrated up to 300 μM in 20mM Tris-HCl pH8, 30mM NaCl, 10mM β-mercaptoethanol and heated during 1 day at 37ºC. **(B)** SDS-PAGE of wild-type and del95-99 EmN after heating the samples at 338K during either 1 or 4 hours (S and P correspond to Supernatant and Pellet, respectively; the arrows correspond to the protein monomers).

Figure 6. EmN variants Q133H and P183T form filaments after heating 1 hour at 338K. Filaments formed by Q133H were systematically shorter and less regular than filaments obtained from variant P183T and wild-type EmN at a similar time point. EM images were acquired on samples of wild-type EmN (left) and variants Q133H (center) and P183T (right), concentrated above 300 μM and heated during 1 hour at 338K. The black bars correspond to 100 nm. Negative staining was achieved using uranyl acetate.

Figure 7. Impact of emerin variants expressed in EDMD on nuclear emerin-emerin proximities as detected by PLA. (A) Whole cell protein extracts prepared from HeLa cells co-expressing either wild-type, del95-99, Q133H or P183T GFP and FLAG-emerins were analyzed by western blot using anti-emerin antibodies. The expression of endogenous A-type lamins revealed by anti-LA/C antibodies is also shown. **(B)** HeLa cells co-expressing either wild-type (WT), del95-99, Q133H or P183T GFP and FLAG-tagged emerins were fixed, labeled with anti-GFP and anti-FLAG antibodies, and processed for PLA before analysis at the confocal microscope. The PLA signals are shown alone (upper row) and merged with DNA staining and GFP fluorescence (lower row). **(C)** Quantification of PLA signals per nucleus for cells as shown in B). Are shown the median values for the total signal intensity per nucleus for 3 independent experiments (EXP1, EXP2, EXP 3; n = 150-330 nuclei per sample). Boxes show first and third quartiles (**$p < 0.001$, *$p<0.05$ with Kruskal and Wallis test).

Figure 8. Impact of emerin variants expressed in EDMD on nuclear emerin-lamin A/C proximities as detected by PLA. (A) Whole cell protein extracts prepared from HeLa cells expressing either WT, del95-99, Q133H or P183T GFP-emerin were analyzed by western blot using anti emerin antibodies. Shown is also the expression of endogenous A-type lamins revealed by anti-LA/C antibodies. **(B)** HeLa cells overexpressing either WT or mutant del95-99, Q133H and P183T GFP-emerin were fixed, labeled with anti-GFP antibodies, and processed for PLA before analysis at the confocal microscope. The PLA signals are shown alone (upper row) and merged with DNA staining and GFP fluorescence (lower row)**. (C)** Quantification of PLA signals per nucleus for cells as shown in (B). Are shown the median values for the total signal intensity per nucleus for 3 independent experiments (EXP1, 2, 3; n = 180-340 nuclei per sample). Boxes show first and third quartiles (**$p < 0.001$, *$p<0.05$ with Kruskal and Wallis test).

Figure 1

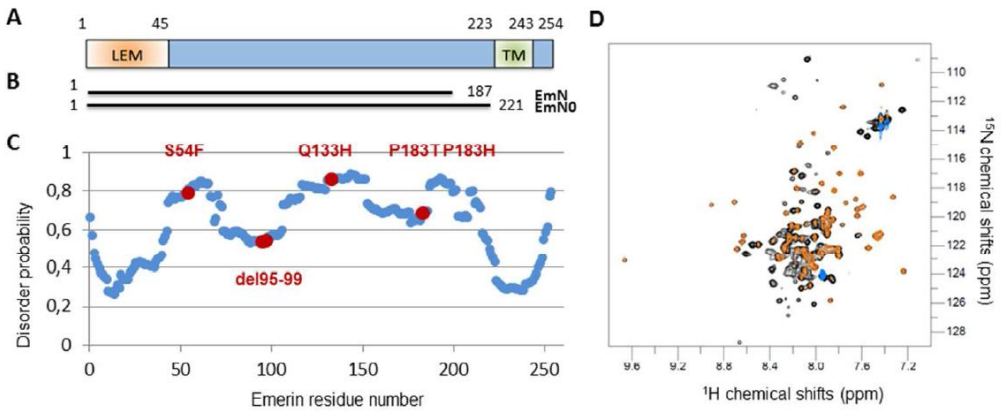
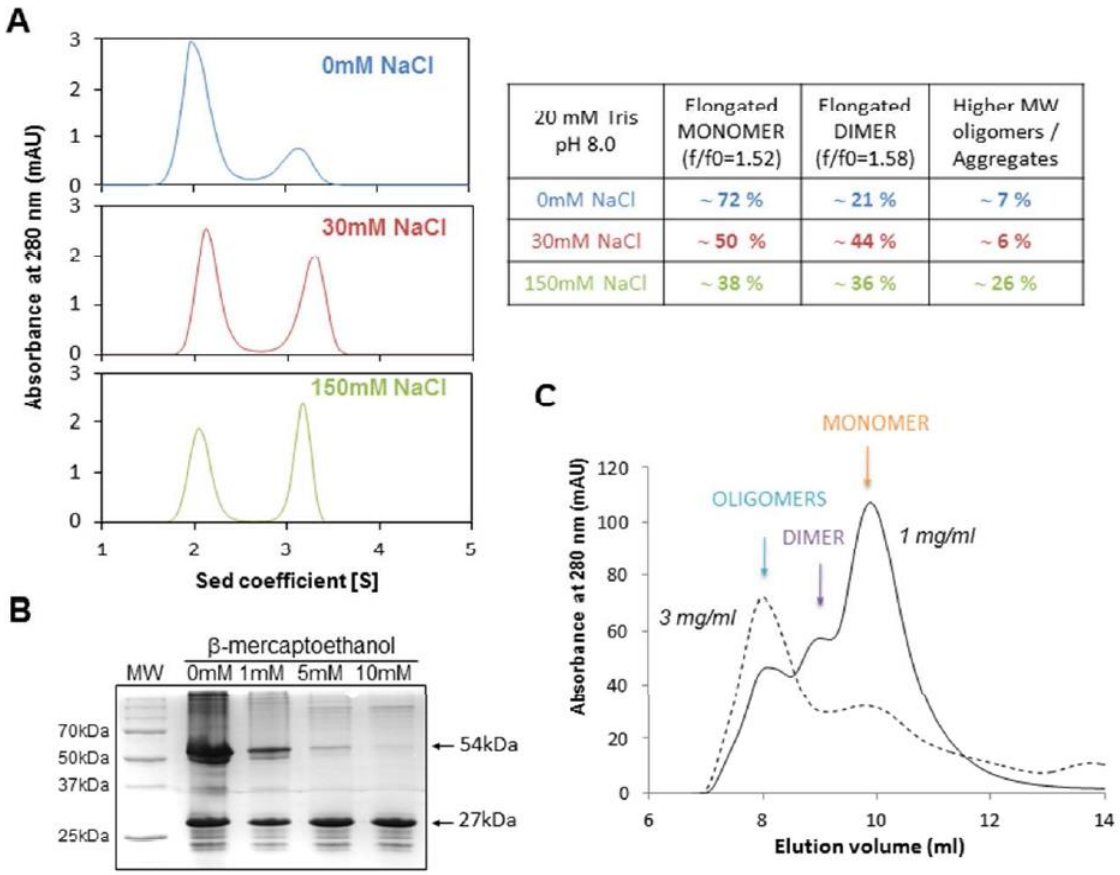


Figure 2



20 mM Iris pH 8.0	Elongated MONOMER (f/f0=1.52)	Elongated DIMER (f/f0=1.58)	Higher MW oligomers / Aggregates
0mM NaCl	~ 72 %	~ 21 %	~ 7 %
30mM NaCl	~ 50 %	~ 44 %	~ 6 %
150mM NaCl	~ 38 %	~ 36 %	~ 26 %

Figure 3

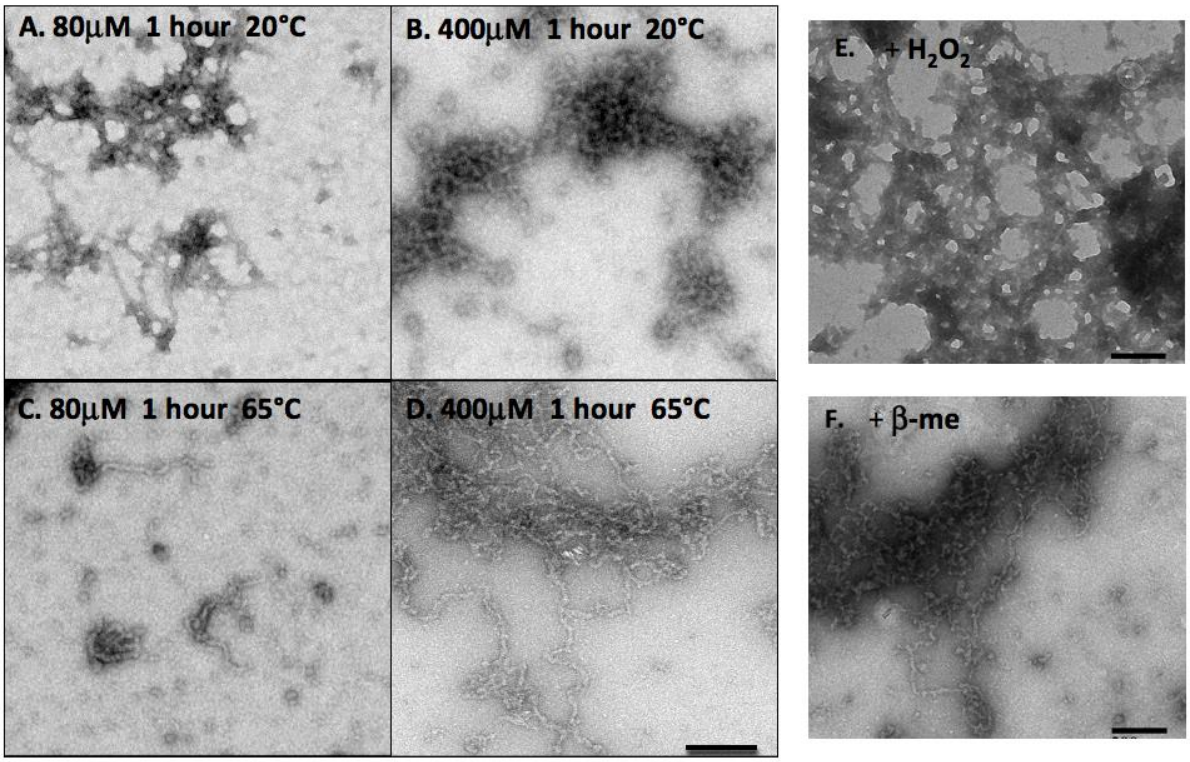


Figure 4

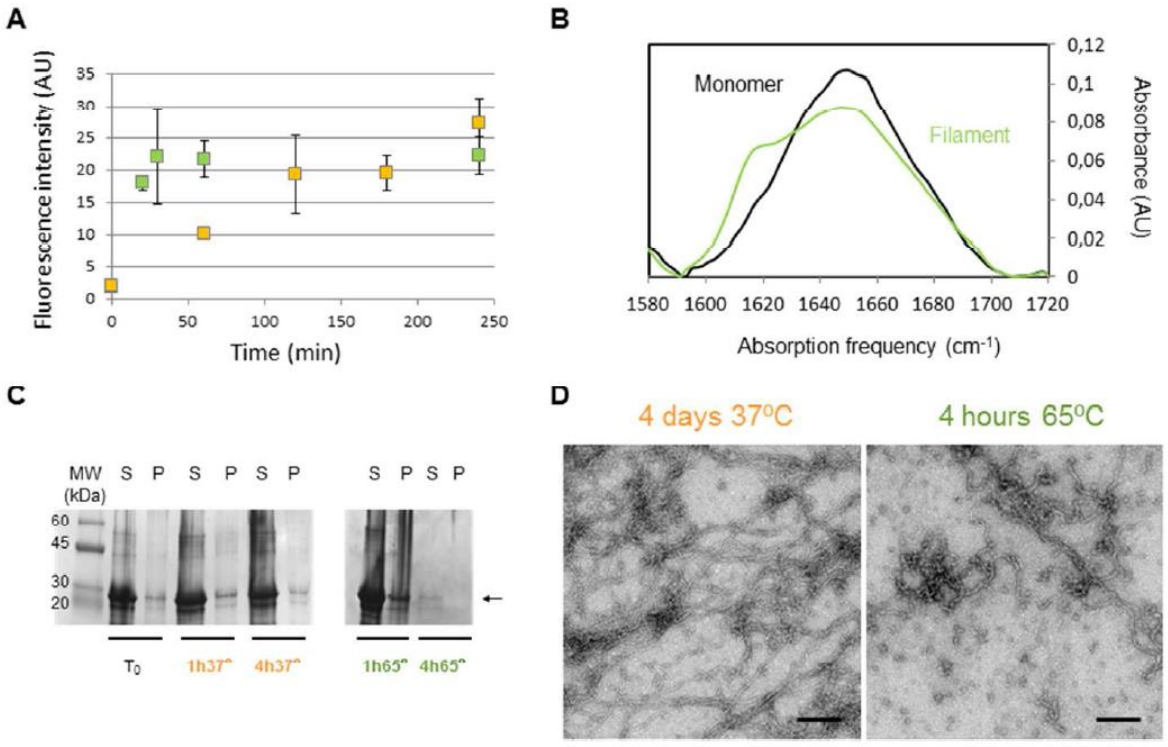


Figure 5

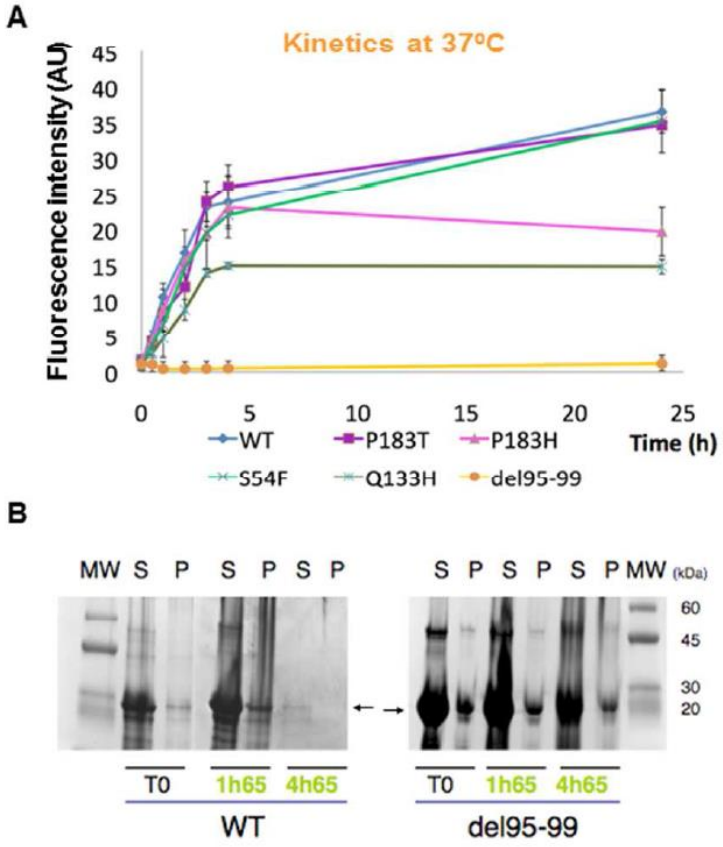


Figure 6

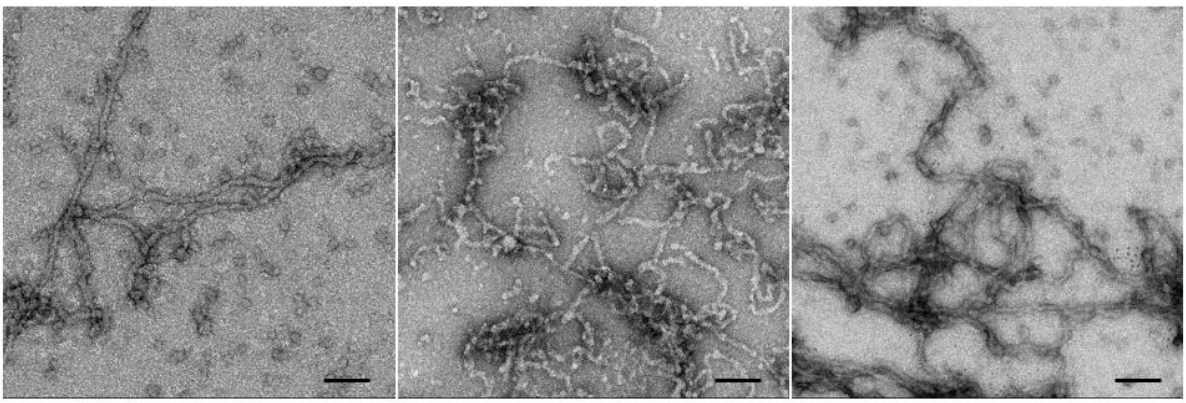


Figure 7

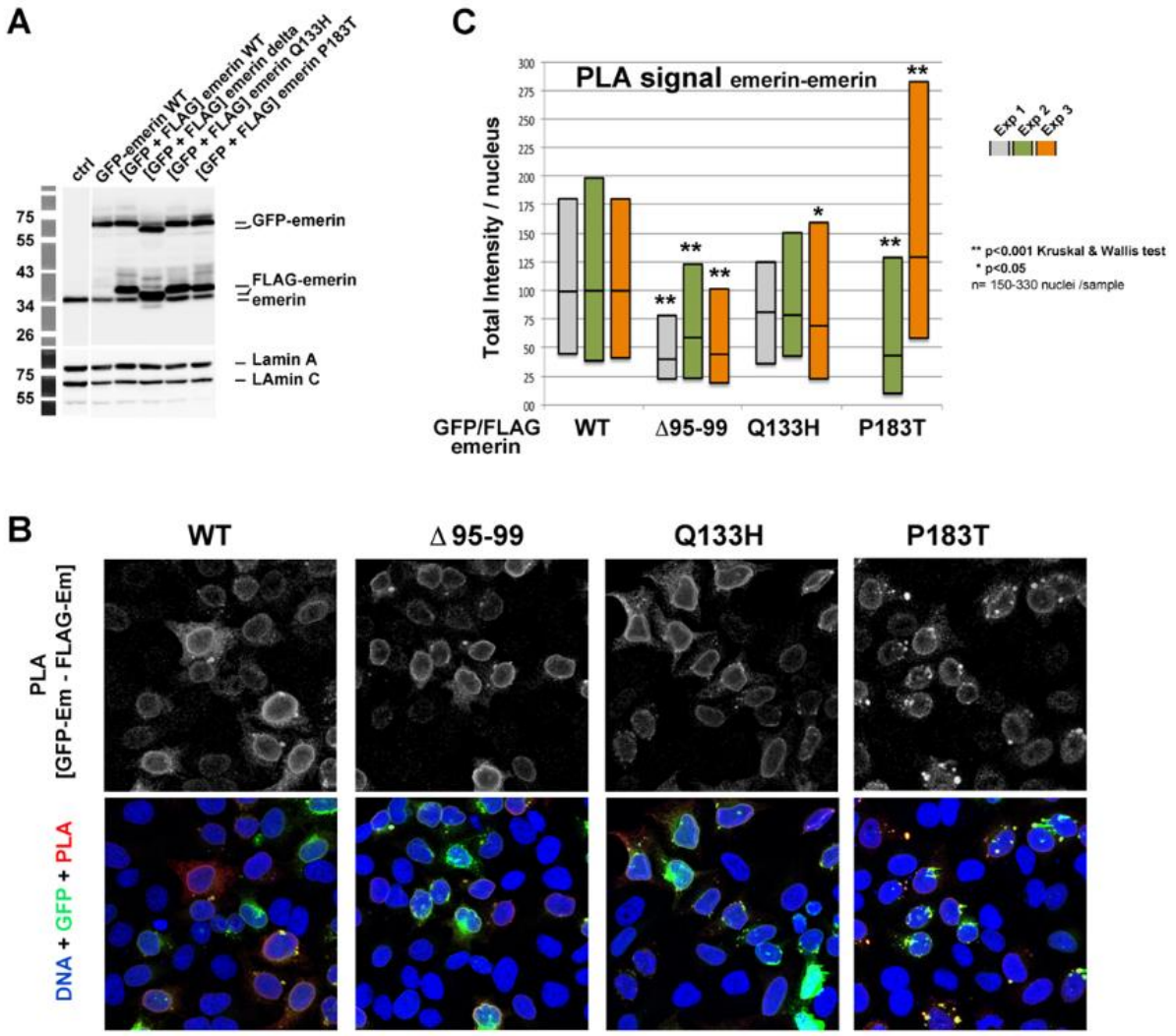
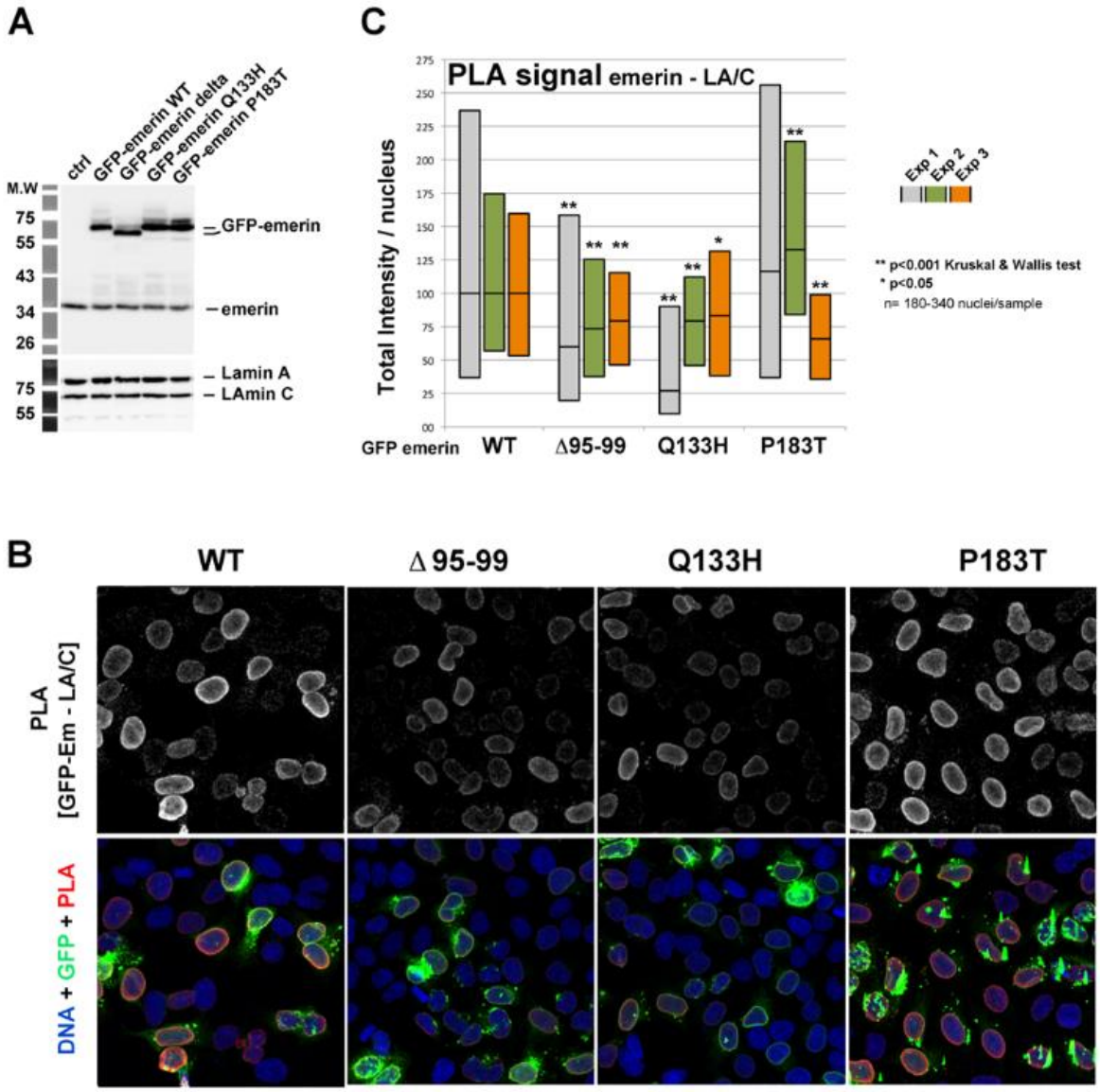

A
ctrl
GFP-emerin WT
[GFP + FLAG] emerin WT
[GFP + FLAG] emerin delta
[GFP + FLAG] emerin Q133H
[GFP + FLAG] emerin P183T
75
55
43
34
26
75
55
GFP-emerin
FLAG-emerin
emerin
Lamin A
LAmin C
C
PLA signal emerin-emerin
Total Intensity / nucleus
300
275
250
225
200
175
150
125
100
75
50
25
00
**
**
**
*
**
Exp 1
Exp 2
Exp 3
** p<0.001 Kruskal & Wallis test
* p<0.05
n= 150-330 nuclei /sample
GFP/FLAG emerin
WT
Δ95-99
Q133H
P183T
B
WT
Δ 95-99
Q133H
P183T
PLA [GFP-Em - FLAG-Em]
DNA + GFP + PLA

Figure 8


A
ctrl
GFP-emerin WT
GFP-emerin delta
GFP-emerin Q133H
GFP-emerin P183T
M.W
75
55
43
34
26
75
55
GFP-emerin
emerin
Lamin A
LAmin C
C
PLA signal emerin - LA/C
Total Intensity / nucleus
275
250
225
200
175
150
125
100
75
50
25
00
GFP emerin
WT
Δ95-99
Q133H
P183T
Exp 1
Exp 2
Exp 3
** p<0.001 Kruskal & Wallis test
* p<0.05
n= 180-340 nuclei/sample
B
WT
Δ 95-99
Q133H
P183T
PLA
[GFP-Em - LA/C]
DNA + GFP + PLA

Supporting Information: alignment of emerin sequences (Suppl. Fig. 1), Analytical Ultra Centrifugation plots (Suppl. Fig. 2), gel filtration profiles of EmN and its mutants (Suppl. Fig. 3) and fluorescence emission spectra recorded in the presence of ThT on EmN mutants at pH 6.8 (Suppl. Fig. 4).

Supplementary materials.

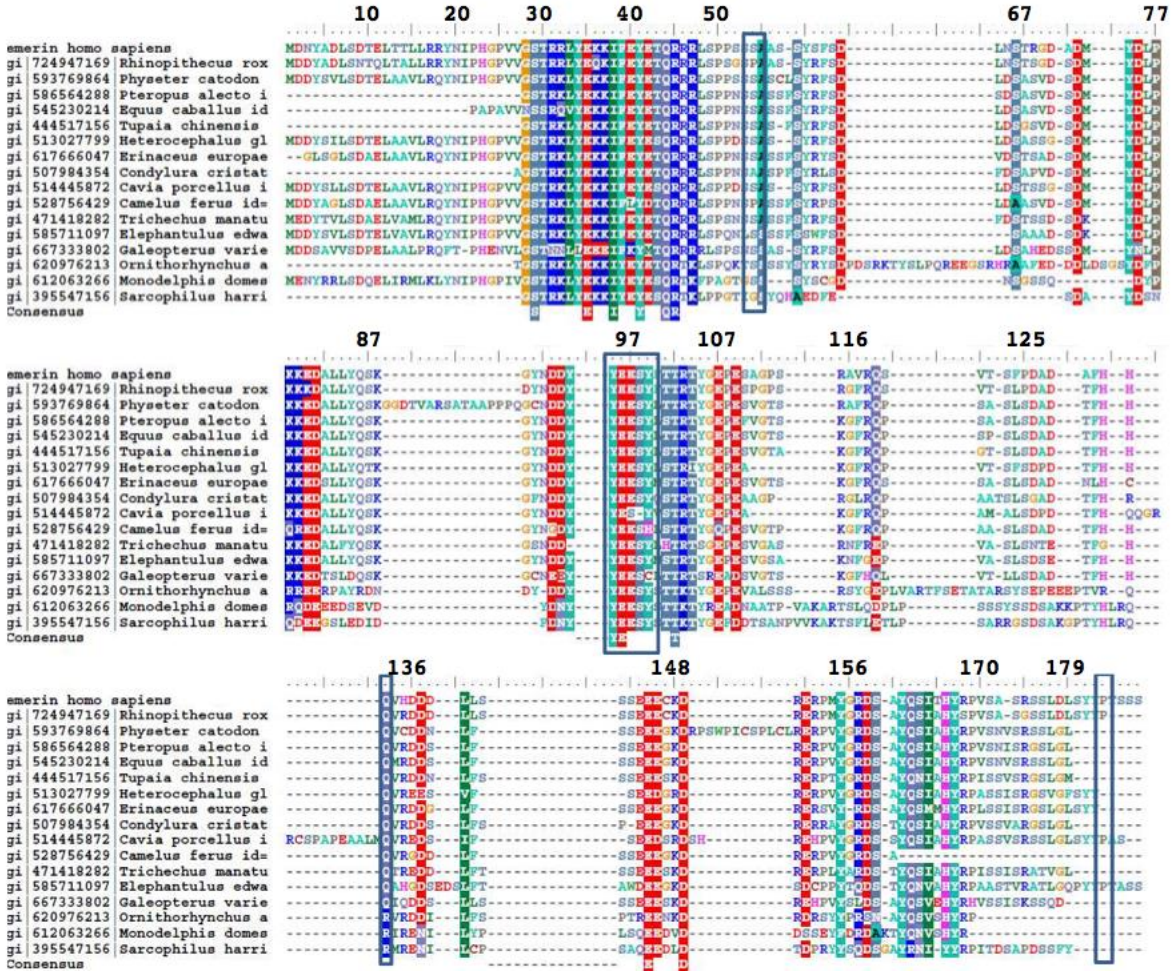
Supplementary Figure 1. Alignment of emerin sequences from 17 different mammalian species. Each sequence is identified by its Genbank identifier code. The numbering in black above the sequences corresponds to human emerin. Positions mutated in EDMD causing missense or micro-deletion variants are squared in blue.

Supplementary Figure 2. AUC sedimentation equilibrium analysis of EmN. This AUC plot was obtained on an EmN sample at 17.4μM centrifuged at either 9619 g (black) or 13851 g (orange).

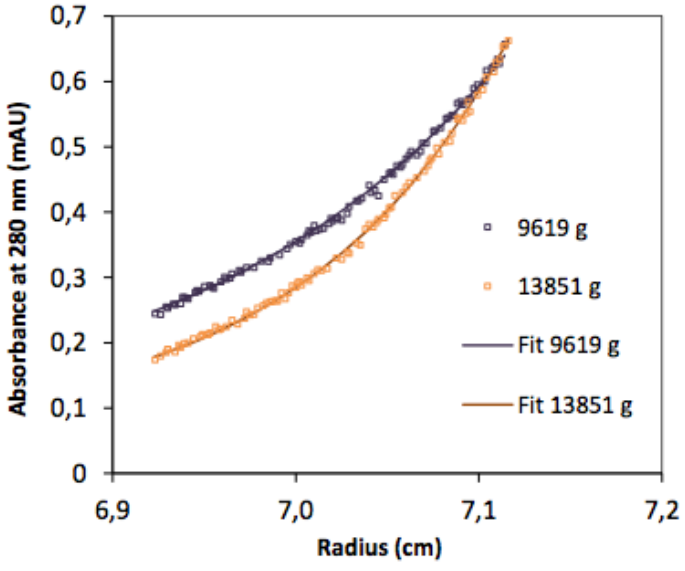
Supplementary Figure 3. Purification of an EmN sample in low salt and reducing conditions. Gel filtration of the EmN sample carried out immediately after dialysis shows that in these conditions EmN is mainly monomeric. (A) Schematic representation of EmN (Emerin 1-187) in blue and the dimeric construct EmGCN4 in red. This last construct possesses a C-terminal tag of 43 residues resulting from protein engineering and designed as to dimerize. (B) Gel filtration on EmN (blue), EmGCN4 (in red), and EmN EDMD-causing mutants (black lines) in a low salt concentration buffer (20mM Tris-HCl (pH8.0), 30mM NaCl). The EmGCN4 construct is used as a control corresponding to the dimeric form of EmN. (C) AUC Sedimentation velocity results obtained on EmN (blue) and EmGCN4 (red) in 20mM Tris-HCl (pH8.0), 30mM NaCl.

Supplementary Figure 4. Fluorescence emission spectra recorded after addition of ThT on samples of variants del95-99 (t=0 orange; t=24h brown) and Q133H (t=0 light green; t=24h dark green) incubated in 20mM phosphate buffer (pH6.8), 30mM NaCl at 310K. The corresponding reducing SDS-PAGE gel is shown at the upper right of the figure. The supernatant (S) and pellet (P) fractions of the samples were deposited before and after incubation at 310K during 24 hours. Formation of high molecular weight oligomers yields to a decrease in total protein staining.

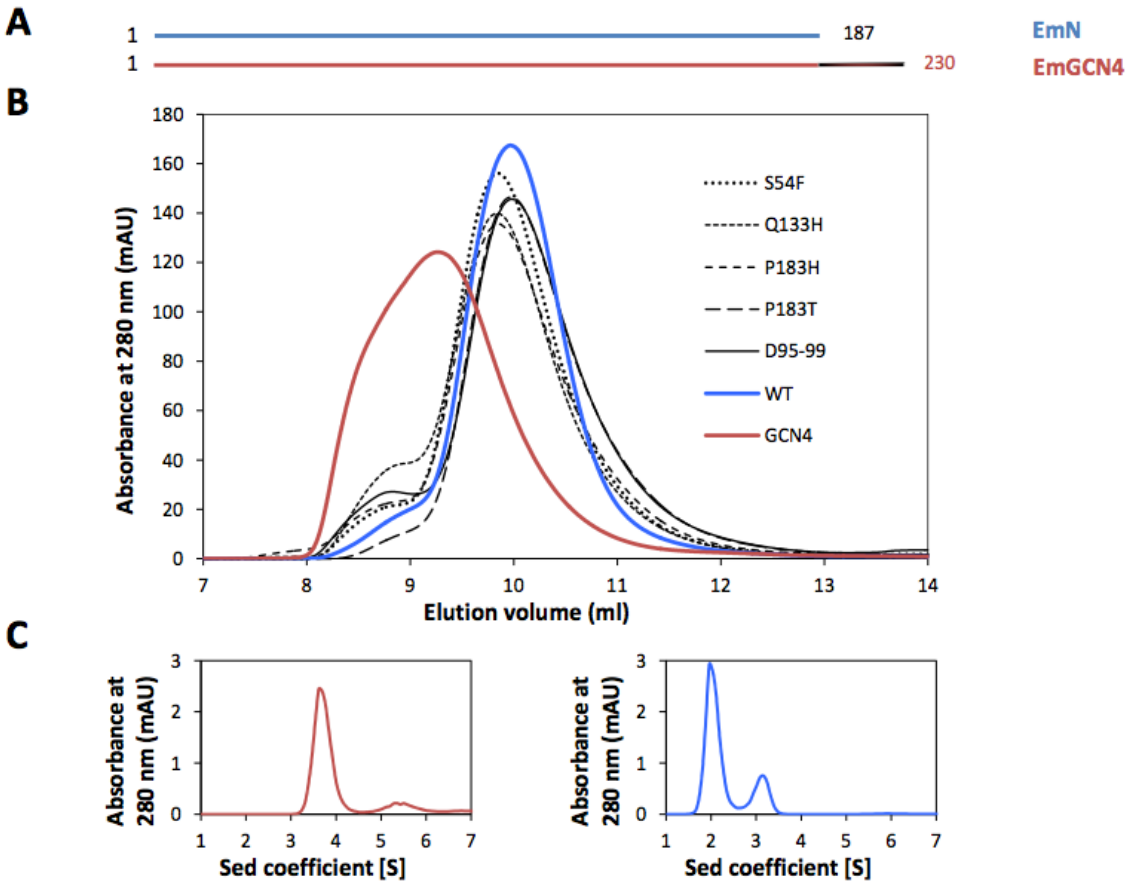
Supplementary Figure 1



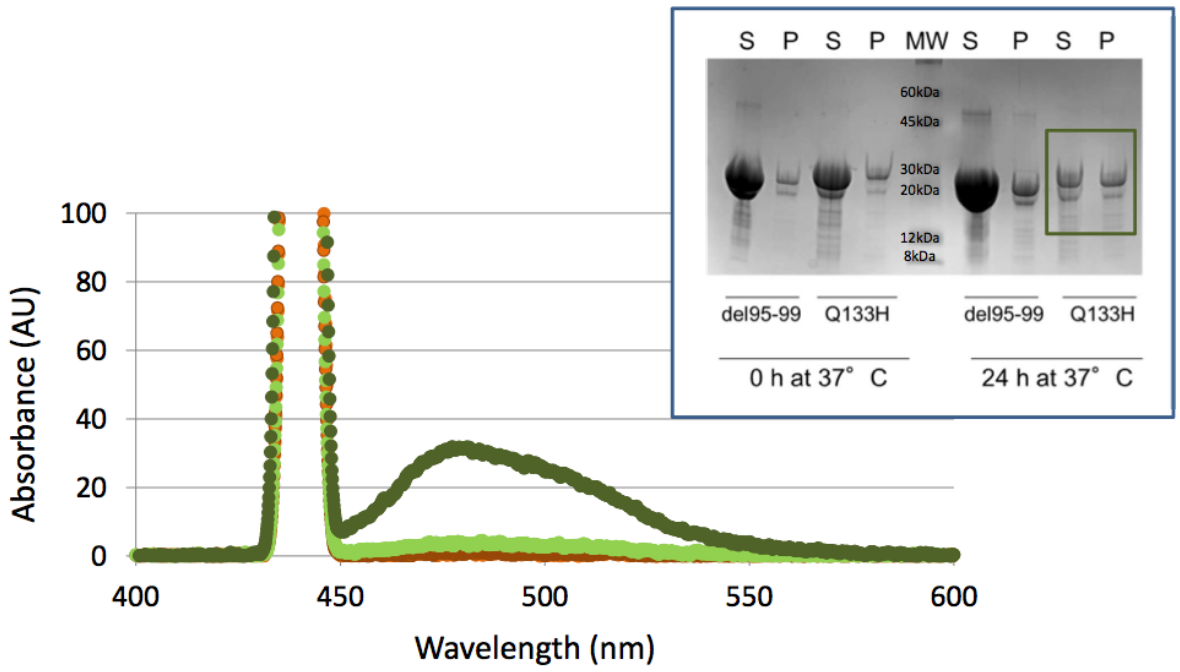
Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4



The second paper aims at describing the structural organization of EmN oligomers. I showed that emerin fragment from residue 1 to residue 132 is necessary and sufficient for self-assembly. The LEM domain and the unstructured region of this fragment both contribute to the structural core of the filaments, as shown by solid-state NMR (collab. with Prof Adam Lange, FMP Berlin) and mass spectrometry (collab. with Dr Jean Armengaud, CEA Marcoule). However the poor quality of the solid-state NMR spectra, probably due to conformational heterogeneity present in the EmN filaments, impaired a more detailed description of the 3D structure of emerin oligomers. Destabilization of the LEM domain favors filament formation. This is in particular observed in the case of emerin variant delK37, which causes a muscular dystrophy with cardiac defects. Our results predicted that, if emerin self-assembly regulates lamin binding, variant delK37 should still bind to lamin A/C. This was confirmed using NMR.

Finally another PhD student in the lab, Florian Celli, showed that EmC221 also self-assembles. This oligomerization process is so efficient that it was very difficult to obtain a sample of monomeric EmC221 for solution-state NMR. I will come back to the EmC221 oligomers later when discussing about the emerin – lamin interaction.

Emerin self-assembly mechanism: role of the LEM domain

Camille Samson[1], Florian Celli[1], Kitty Hendriks[2], Maximilian Zinke[2], Nada Essawy[3], Isaline Herrada[1], Ana-Andreea Arteni[4], François-Xavier Theillet[1], Béatrice Alpha-Bazin[5], Jean Armengaud[5], Catherine Coirault[3], Adam Lange[2,6], Sophie Zinn-Justin[1*]

[1] Laboratory of Structural Biology and Radiobiology, Institute for Integrative Biology of the Cell (CEA, CNRS, University Paris South), University Paris-Saclay, Gif-sur-Yvette, France

[2] Department of Molecular Biophysics, Leibniz-Institut für Molekulare Pharmakologie, Berlin, Germany

[3] Center for Research in Myology (INSERM, CNRS), Université Pierre et Marie Curie Paris 06, Sorbonne Universités, France

[4] Department of Structural Virology, Institute for Integrative Biology of the Cell (CEA, CNRS, University Paris South), University Paris-Saclay, Gif-sur-Yvette, France

[5] Laboratory "Innovative technologies for Detection and Diagnostics", Institute of Biology and Technology Saclay, CEA , Bagnols-sur-Cèze, France

[6] Institut für Biologie, Humboldt-Universität zu Berlin, Berlin, Germany

* Corresponding author: Dr Sophie Zinn-Justin. Address: Laboratory of Structural Biology and Radiobiology. CEA Saclay, Bât 144, 91191 Gif-sur-Yvette Cedex, France. Tel.: +33(0)169083026. Fax: +33(0)169084712. E-mail: Sophie.zinn@cea.fr.

Running title: Nuclear envelope emerin self-assembled structure.

Keywords: Nuclear envelope. Nucleoskeleton. Lamin. Oligomerization. Intrinsically disordered region. Folding.

Conflicts of interest: The authors declare no competing interest.

Abstract

At the nuclear envelope, the inner nuclear membrane protein emerin contributes to the interface between the nucleoskeleton and the chromatin. Emerin is an essential actor of the nuclear response to a mechanical signal. Genetic defects in emerin cause Emery-Dreifuss muscular dystrophy. It was proposed that emerin oligomerization regulates nucleoskeleton binding and impaired oligomerization contributes to the loss of function of emerin disease-causing mutants. We here report the first structural characterization of emerin oligomers. We identified an N-terminal emerin region from amino acid 1 to amino acid 132 that is necessary and sufficient for formation of long curvilinear filaments. In emerin monomer, this region contains a globular LEM domain and a fragment that is intrinsically disordered. Solid-state Nuclear Magnetic Resonance analysis identifies the LEM β-fragment as part of the oligomeric structural core. However the LEM domain alone does not self-assemble into filaments. Additional residues forming a β-structure are observed within the filaments that could correspond to the unstructured region in emerin monomer. We show that the delK37 mutation causing muscular dystrophy triggers LEM domain unfolding and increases emerin self-assembly rate. Similarly inserting a disulfide bridge that stabilizes the LEM folded state impairs emerin N-terminal region self-assembly whereas reducing this disulfide bridge triggers self-assembly. We conclude that the LEM domain, responsible for binding to the chromatin protein BAF, undergoes a conformational change during self-assembly of emerin N-terminal region. The consequences of these structural rearrangement and self-assembly events on emerin binding properties are discussed.

Introduction

The metazoan nuclear envelope is a highly regulated membrane barrier that separates the nucleus from the cytoplasm. It contains (i) a double lipid membrane, (ii) a large number of proteins anchored at either the outer nuclear membrane (as nesprins) or the inner nuclear membrane (as SUN1, SUN2 and emerin), and (iii) the intermediate filaments lamins forming the *nuclear lamina* that lines the inner nuclear membrane and interacts with inner nuclear membrane proteins. It critically contributes to the physical connection between the cytoskeleton and the nucleoskeleton. It determines the position of the nucleus, thus dividing the cytoplasm into forward and rear compartments whose internal pressure difference is a driving force for 3D cell migration [1]. It also senses and responds to mechanical forces either by modifying its protein composition or exhibiting post-translational modifications [2]. These events contribute to the tuning of cell cytoskeletal tension to tissue stiffness [3, 4] and to the increase in nuclear resistance in response to mechanical stress [5]. They initiate a signaling cascade and trigger the expression of mechanosensitive genes [5, 6]. The inner nuclear envelope also interacts with chromatin [7]. It directly binds to histones, histone modifiers and transcription regulators. Through these contacts it might regulate genome expression.

An increasing number of mutations in nuclear envelope protein genes are being discovered that causes a wide range of human diseases as for example muscular dystrophies, cardiomyopathies, lipodystrophies, neuropathies and premature aging syndromes (for review [8, 9]). There is growing evidence that integrity of the nucleus [10-12] and mechanotransduction signaling [11, 13] are impaired in several of these envelopathies. The structural and functional impacts of disease-causing missense mutations or short in-frame deletions in the nuclear envelope protein genes were analyzed in order to get insights into the molecular mechanisms of nuclear envelope functions. A very large number of missense mutations were found in the LMNA gene coding for the nuclear intermediate filaments lamins A and C (also called lamin A/C or A-type lamins). These mutations impact either lamin conformation, self-assembly or binding properties [14, 15]. Six missense mutations or short in-frame deletions causing the X-linked form of Emery-Dreifuss Muscular Dystrophy (EDMD) were identified in the *EMD* gene encoding the inner nuclear membrane emerin [16]. Five of them (i.e., S54F, del95-99, Q133H, P183T and P183H) correspond to a classical EDMD phenotype [17]. However the short in-frame deletion mutation delK37, which affects the LEM domain, is associated with a disease essentially affecting the cardiac muscle [18].

At the molecular level, the nuclear lamin A/C and the inner nuclear membrane protein emerin both interact with complexes formed by the outer nuclear membrane nesprins and the inner nuclear membrane SUN proteins, called the LINC complexes for LInker of Nucleoskeleton and Cytoskeleton [2, 19]. These interactions are key events of the nucleocytoskeletal coupling machinery and critically contribute to nuclear structure. However, very few studies have provided molecular details about complexes involving lamin, emerin and LINC. Recent X-ray crystallography studies have revealed that changes in SUN2 oligomeric states are associated with large structural rearrangements and regulate nesprin binding [20-23]. We and others showed that emerin self-associates *in vitro* and in cells and that defects in oligomerization were observed together with impaired lamin binding in the case of two mutants causing Emery-Dreifuss muscular dystrophy [24, 25]. However the molecular details of emerin conformational states and related binding properties are poorly understood. In this study, we focus on providing a molecular description of emerin oligomers, in order to propose mechanisms for emerin functional regulation through self-assembly.

Emerin is anchored at the inner nuclear membrane via its hydrophobic C-terminal domain (Figure 1A) [26, 27]. Its N-terminal nucleoplasmic region (residues 1 to 221) contains all six mutations described above. This region is responsible for emerin interaction with structural components of the nuclear envelope (SUN1, SUN2) and of the nucleoskeleton (lamins, actin), DNA binding proteins (barrier-to-autointegration factor (BAF), transcription factors) and signaling molecules (β-catenin). Fragment from residue 170 to residue 220 is sufficient to bind other emerin molecules homotypically [24]. Additional contacts between this fragment and either the globular LEM domain (residues 1 to 45; see Figure 1A) or the fragment between residues 153 and 169 were reported. We have found that fragment from residue 1 to residue 187 also self-associates *in vitro* and that residues 95 to 99 are important for self-association *in vitro* and in cells [25].

Here, we report a biophysical characterization of oligomers of fragment from residue 1 to residue 187 (called EmN; see Figure 1A). We show that the globular LEM domain, comprising residues 1 to 45 of emerin, plays a critical role in EmN oligomerization, and that the small in frame deletion delK37 significantly favors emerin self-assembly. On these bases, oligomerization mechanisms and their functional consequences are discussed.

Results

The N-terminal LEM domain is the only globular domain observed in the different monomeric emerin fragments – We have previously proposed that the monomeric emerin fragment from residue 1 to residue 187 (named EmN; see Figure 1A) is composed of a globular LEM domain, from residue 1 to residue 45 [28], and a large intrinsically disordered region, from residue 46 to residue 187 [25, 29]. EmN is also capable of oligomerizing into filaments rich in β-structure. In order to identify the minimal region of EmN that still self-associates into filaments, we produced three different recombinant fragments of EmN called EmN49, EmN132 and Em67C, corresponding to the LEM domain [28], the smallest N-terminal EmN fragment binding lamins [24] and the smallest C-terminal EmN fragment containing all residues whose mutation impairs lamin A binding [30], respectively (Figure 1A). Superimposition of the solution-state NMR ^{1}H-^{15}N HSQC spectra of fragments EmN49, EmN132 and EmN confirmed that the LEM domain fold of EmN49 [28, 31] is also present in monomeric EmN132 and EmN (Figure 1B). Significant differences are found between the chemical shifts of the 4 last residues of helix α3 in EmN132 or EmN as compared to EmN49. This is most probably due to edge effects related to the C-terminal extension in EmN132 and EmN, which is absent in EmN49. It may highlight a moderate difference in the stability of the helix α3 C-terminus. Analysis of the EmN132 NMR carbon signals consistently established that this fragment contains N-terminal helix α1 (3-6), helix α2 (8-19), strand β1 (22-28, with a bulge at Pro25) and helix α3 (29-48) (Figures 1C,D). It also contains an additional helix α4 (96-102) that is, as helix α1, only transiently observed: the corresponding region exists both as a α-helix and a disordered peptide. Our recent NMR analysis of the intrinsically disordered emerin fragment 67-170 also demonstrated α-helical propensities between residue 98 and residue 101 [29].

The LEM domain and the intrinsically disordered region are both needed to form filaments – We tested the self-assembly capacities of EmN49, EmN132 and Em67C. Therefore, we incubated the different fragments at a fixed concentration of 600 μM during one week at 293 K. We then observed the samples using negative staining electron microscopy. Figure 2A shows that long curvilinear filaments were detected for EmN, as reported earlier [25], but also for EmN132. No filaments could be identified for EmN49 and Em67C. Consistently, only EmN and EmN132 were capable of forming self-assembled particles interacting with thioflavin T, as detected by fluorescence on samples incubated at a concentration of 600 μM at 310 K (Figure 2B).

Thus both region 1 to 66, mostly restricted to the LEM domain, and region 50 to 132, which is intrinsically disordered, are necessary for filament assembly. This result is consistent with our previous study showing that residues from 95 to 99, in the intrinsically disordered region, are necessary for EmN self-assembly [25].

EDMD causing small in frame deletion delK37 impairs folding of the LEM domain – The emerin mutation delK37 was identified in a patient with an EDMD phenotype characterized by a strong cardiac dysfunction [18]. Whereas the previously reported missense mutations or small in frame deletions all corresponded to defects in emerin intrinsically disordered region, this mutation targets a LEM domain residue. We used this mutation in order to probe the role of the LEM domain in emerin self-assembly. Solution-state NMR 1D ^{1}H spectrum analysis of EmN mutant delK37 showed that this protein is completely unfolded (Figure 3A). Superimposition of the solution-state NMR 2D ^{1}H-^{15}N HSQC spectra of fragments EmN, wild-type and mutated, revealed that the peaks corresponding to folded LEM residues (in orange on Figure 1B) are absent in the case of the mutated EmN (Figure 3B). The NMR signals corresponding to protons of side chains buried in the LEM hydrophobic core, which are probes for domain folding, are absent in the 1D ^{1}H spectrum of the EmN mutant delK37 (Figures 3A,C). Thus, deletion of residue K37 in helix α3 impairs folding of the whole LEM domain.

LEM domain unfolding is associated with faster self-assembly in the case of the delK37 mutant – In order to probe for the role of the LEM domain in EmN self-assembly, we incubated in parallel EmN wild-type and delK37 at a concentration of 600 μM during one week at 293 K. Negative staining electron microscopy revealed that the mutant delK37 is also capable of forming filaments *in vitro* (Figure 4A). Comparison of the susceptibility to proteolysis by chymotrypsin of EmN wild-type and mutated as a function of their oligomeric states showed that both protein fragments are partially protected from degradation when they are self-assembled (Figure 4B). Following protein self-assembly at 310 K and 300 μM using thioflavin T as a probe, we demonstrated that the EmN mutant delK37 forms oligomers significantly faster than wild-type EmN, whereas the LEM domain delK37 alone does not form filaments (Figure 4C). To further probe the relationship between LEM domain folding and oligomerization kinetics, we introduced a disulfide bridge into the LEM domain of EmN132 (which contains no cysteine) by mutating Y4 and E35 into cysteines (Figure 4D). The presence of this disulfide bridge impairs EmN132 self-assembly, whereas reducing this disulfide bond triggers filament formation, as shown by fluorescence (Figure 4E). Solution state ^{1}H-^{15}N HSQC spectra of EmN132 wild-type (green on Figure 4F) and Y4C/E35C (mauve) were recorded in the absence of DTT to verify that the oxidized mutant adopts a well-defined LEM-like 3D structure (well-dispersed cross-peaks of homogeneous linewidths on Figure 4F). Analysis of the spectrum of the Y4C/E35C mutant in the presence of 5 mM DTT showed that in these conditions the mutant exhibits conformational exchange (heterogeneous linewidths on Figure 4F).

The LEM domain region contributes to the structural core of the EmN filaments – In order to identify residues forming the structural core of EmN132 and EmN filaments, we searched for residues protected from proteolysis by mass spectrometry. Therefore, we incubated filaments of these emerin fragments with either endoproteinase GluC or chymotrypsin during 4 hours at 310 K for a limited proteolysis. Degradation products were then resolved by SDS-PAGE and analyzed by means of high-resolution tandem mass spectrometry (Figure 5). First, two main bands with apparent molecular weights of 6 and 10 kDa are observed after proteolysis by endoproteinase GluC in the cases of both EmN132 and EmN. They all contain peptides with a LEM domain (boxed in yellow on Figure 5). The band at 6 kDa probably corresponds to the LEM region itself. The band at 10 kDa corresponds to: (i) the LEM region and part of the intrinsically disordered region from residue 48 to residue 103, in the case of EmN132; (ii) a mixed of peptides containing either the N-terminal LEM domain or the C-terminal peptide 158 to 187 in the case of EmN. Limited proteolysis by chymotrypsin led to a main degradation band with an apparent molecular weight of 9 kDa in the case of EmN132 and 10 kDa in the case of EmN. These bands again correspond to degradation products comprising the LEM region and an additional fragment: (i) fragment 47 to 74 in the case of EmN132; (ii) probably a slightly larger fragment in the case of EmN, even if only region 47 to 59 is detected by mass spectrometry. Thus an N-terminal region with an apparent molecular mass of about 10kDa is protected from proteolysis in self-assembled EmN132 and EmN with the LEM region being the most protected from cleavage by either endoproteinase GluC or chymotrypsin.

The EmN filaments, either wild-type or delK37, share a common structural core rich in β-structure comprising the LEM domain β1-strand - We finally characterized the filament structure of EmN, EmN132 and EmN delK37 using solid-state NMR. Superimposing the ^{13}C-^{13}C proton-driven spin-diffusion spectra of EmN wild-type and delK37 showed that the correlation peaks overlap nicely (Figure 6A). However, the spectra are dominated by random coil signals which are relatively broad. In a double quantum - single quantum experiment (Figure 6B), double quantum coherence is generated from and later reconverted to single quantum coherence by means of SPC-5 [32] excitation and reconversion periods. During these recoupling periods, an effective dipolar double-quantum Hamiltonian is active, that will be scaled down in the presence of molecular motions. Thus only very rigid parts of the protein are visible in the resulting spectrum. Approximately 20 amino acids contribute to the double quantum – single quantum ^{13}C-^{13}C correlation spectrum of EmN filaments (Figure 6B): these residues form the filament core structure. They comprise three threonines, two serines, one proline, one alanine, one valine, one leucine and one isoleucine. EmN132 shows a very similar double quantum – single quantum ^{13}C-^{13}C correlation spectrum, with a clear isoleucine signal, and possesses only two isoleucines, both in the LEM domain region, precisely at positions 21 and 38. This first observation confirms that the LEM domain region is part of the filament structural core. Valine residues are found at positions 26 and 27 and proline residues at positions 22 and 25 in this same LEM domain region.

This suggests that the structural core comprises at least fragment 21 to 26, which corresponds to the strand β1 of the LEM domain (Figure 1D). Analysis of the secondary chemical shifts of the isoleucine and valine $^{13}C_\beta$ measured in the double quantum – single quantum spectrum predicts that these residues also adopt a β-conformation in the filament (Figures 6B,C). Additionally secondary structure predictions for an alanine, two serines and two threonines indicate a β-conformation in the filament. Only one alanine is present in the LEM domain region, at position 5 (Figure 6D). Two serines are present at positions 8 and 29, and several threonines are present at positions 10, 13, 14, 30 and 43. These residues are all in α-helices in monomeric EmN. So either a fragment of the LEM domain region switches from an α- to a β-conformation (as fragment 5 to 14 for example; see Figure 6D) or a fragment of the intrinsically disordered region folds into a β-conformation during EmN self-assembly. Both processes could also take place simultaneously.

Discussion

Very few structural data are yet available on the architecture of the nuclear envelope. Inner nuclear membrane proteins exhibit large regions predicted to be intrinsically disordered that interact with the nuclear intermediate filaments called the lamins lining the inner nuclear membrane. However, the details of these intermolecular interactions and their mode of regulation through oligomerization and post-translational modifications are still unclear. At the nuclear envelope, LEM domain proteins share the ability to bind lamins and tether repressive chromatin at the nuclear periphery (for review, see [33]). Their LEM domain binds Barrier-to-Autointegration Factor (BAF), a metazoan DNA-binding protein [34, 35]. Here we focus on the inner nuclear envelope emerin, one of the most studied LEM-domain proteins. The group of Prof K. Wilson was the first to reveal, from biochemical analyses performed *in vitro* and in cells, that emerin could indeed be assembled as oligomers anchored at the inner nuclear membrane [24]. They showed that several intra and/or intermolecular interactions could be observed within emerin molecules, and suggested that changes in emerin oligomeric state were responsible for regulating emerin binding properties at the inner nuclear envelope. In particular, the LEM domain could be found in a “DOWN” state in which it interacts with BAF at the chromatin, but also in a “UP” state in which it participates to emerin self-assembly and potentially partner recognition at the nuclear envelope. In this study, we demonstrate that the LEM domain can indeed be found in two different quaternary structures, either as a globular domain in a monomeric EmN fragment, or as part of the structural core of long curvilinear filaments formed by EmN. The molecular details of these LEM structures and their potential functional role are discussed in the following.

The LEM domain plays a central role in EmN oligomerization – Our results reveal for the first time that emerin N-terminal region from residue 1 to residue 132 can self-assemble through interactions involving the LEM domain and part of the intrinsically disordered region 48 to 132. Region 21 to 26 might be involved in emerin self-assembly, as strongly suggested by solid-state NMR spectroscopy. This region corresponds to the LEM domain strand β1. Region 95-99 is also critical for emerin oligomerization: EDMD-causing small in frame deletion del95-99 in EmN impaired filament formation in our conditions [25]. It exhibits a serine residue (Ser98) and is close to a threonine-rich patch (Thr101, Thr102, Thr104), which correspond to amino acid types observed by solid-state NMR in the rigid part of the filaments.

LEM domain unfolding favors EmN oligomerization – Analysis of the solution structure of EmN mutant delK37 revealed that this single amino acid deletion in helix α3 destroys the LEM domain fold. However, delK37 self-assembles faster than wild-type EmN and is capable of forming oligomers with a similar structural organization, as shown by negative-staining electron microscopy, limited proteolysis and solid-state NMR. Figure 4B clearly shows that the first steps of self-assembly are significantly facilitated by deletion of K37, suggesting that structural rearrangement of the LEM domain is a rate-limiting step in filament assembly. Interactions between strand β1 and the rest of the LEM domain might have to be disrupted in order to enable intermolecular interactions with another emerin molecule. His23 is located in strand β1 and its side chain is directed towards helix α2. Protonation of this histidine under pH 6.5 might favor filament assembly. In our conditions, lowering the pH under 6.5 triggered EmN aggregation, thus preventing any EmN study under this pH. However, in the case of mutant Q133H, which self-assembles slower than wild-type EmN, going from pH 8.0 to pH 6.8 clearly favored self-assembly (Suppl. Figure 4 in [25]).

Functional role of the self-assembled form of emerin – The best described LEM domain function is to interact with BAF in order to create a physical link between the nuclear envelope and the chromatin. All studied LEM-domain proteins bind BAF. The 3D structure of a complex between emerin LEM domain and a dimer of BAF was determined from solution-state NMR data [31]. The two main α-helices (α2 and α3) as well as the β-strand (β1) all directly contact BAF dimer, suggesting that the LEM fold is critical for BAF recognition. Thus the emerin mutant delK37 and the filaments of EmN, either wild-type or delK37, might also have lost part of their BAF binding capacity. On the opposite, mutant del95-99, which is defective in EmN self-assembly [25], might be more often in a conformation efficient for BAF binding.

Emerin also binds lamin A/C, and this interaction is critical for emerin localization at the nuclear envelope. It was reported that region EmN132 is necessary and sufficient for binding to the tails of lamin A/C and lamin B1 [24]. We showed in this study that EmN132 is necessary and sufficient for filament formation. It was further reported that emerin mutant del95-99 does not bind lamin A/C as observed by blot overlay assays [30], and that in cells, proximities between this emerin mutant and lamin A/C are less frequently observed by Proximity Ligation assays [25]. We showed that EmN mutant del95-99 does not self-associate in our conditions [25]. Finally, several LEM domain mutants, and in particular a mutant with alanines at positions 24, 25, 26 and 27 (in strand β1) as well as a mutant with alanines at positions 34, 35, 36 and 37 (in helix α3), bind lamin A/C reproducibly better than wild-type emerin as observed by blot overlay assays [30]. Such large number of alanine substitutions in secondary structure elements probably destabilizes the LEM domain fold. As our results suggest that destabilization of the LEM fold favors EmN self-assembly, we propose that the alanine mutants have a less stable LEM domain, which favors emerin self-assembly and thus lamin A/C binding.

Conclusion – It was recently proposed that LEM domain proteins self-associate in order to form a 3D complex network at the nuclear envelope. We propose a first analysis of the emerin self-association mechanisms, which reveals that LEM domain residues interact with residues from the intrinsically disordered region from amino acid 50 to amino acid 132 in order to oligomerize. Our analysis strongly suggests that the LEM domain strand β1 undergoes a structural rearrangement in order to contact fragments of other emerin molecules, either at the N-terminus of the LEM fold, or in the intrinsically disordered region, thus triggering emerin self-association into filaments. Further studies will be needed in order to understand if this self-assembly pathway also occurs in cells, and how it is influenced by emerin membrane anchoring, post-translational modifications and partner binding. However, as emerin is highly phosphorylated in cells, and in particular at positions 4, 8, 10 and 28 at the N-termini of helices α1, α2 and α3, we can postulate that phosphorylation impacts the stability of the LEM domain and thus regulate emerin self-assembly. Moreover we propose that self-assembly regulates BAF and emerin binding. More generally, our results suggest that modification and oligomerization events modify the emerin interactome and regulate LEM domain protein network at the nuclear envelope. Deciphering these molecular mechanisms is now becoming an exciting question, directly relevant to human envelopathies.

Materials and methods

Protein Expression and Purification.

Human wild-type emerin fragments from amino acid 1 to amino acid 187 (EmN) and from amino acid 67 to amino acid 187 (EmN67C) were expressed in *Escherichia coli* BL21 DE3 Star (Novagen) using a pETM13 vector as N-terminal octa-histidine fusions. The emerin cDNAs were optimized in terms of codon usage for expression in *Escherichia coli* (Genscript). To encode emerin variants from amino acid 1 to amino acid 49 (EmN49) or amino acid 132 (EmN132), stop codons were inserted into the EmN cDNA using a standard QuickChange Site-Directed Mutagenesis kit (Stratagene). Mutations delK37 and Y4C/E38C were inserted into the EmN, EmN49 cDNAs and the EmN132 cDNA, respectively, using the same Mutagenesis kit. Bacteria were cultured in rich medium or ^{15}N-labeled minimum medium, induced at an optimal density of 1.0 with 0.5 mM isopropyl β-D-1-thiogalactopyranoside, grown overnight at 293 K, and lysed in 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 40 mM imidazole, 5% glycerol, 1% Triton X-100, and 1 mM phenylmethanesulfonylfluoride. After centrifugation at 20 000 g for 20 min at 277 K, the pellet was resuspended in buffer C8 (50 mM Tris-HCl pH8.0, 150 mM NaCl, 20 mM imidazole, 8 M urea). A second centrifugation step was performed at 20 000 g for 20 min at 293 K. The soluble extract was then loaded onto Ni-NTA beads (GE-Healthcare) equilibrated with buffer C8. Proteins were eluted directly with buffer E8 (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 M imidazole, 8 M urea). Proteins were refolded by dialysis in buffer D (20mM Tris-HCl pH8.0, 30mM NaCl, 2mM DTT). After refolding, they were incubated with the TEV protease during 3 hours at room temperature and finally dialyzed into the desired buffer. The final yield was typically 10 mg of purified protein per liter of bacterial culture [25].

Liquid-State NMR Spectroscopy.

After purification and His-Tag cleavage, proteins were dialyzed against 20 mM sodium phosphate pH 6.5, 30 mM NaCl, 5 mM DTT (20 mM Tris-HCl pH 7.5, 30-100mM NaCl, 5 mM DTT in Fig. 4F). NMR experiments were performed on 700 or 750 MHz Bruker Avance II spectrometers equipped with a cryogenic probe. 2D NMR ^{1}H ^{15}N correlation spectra were acquired using a HSQC pulse sequence at 303 K, on a 3-mm-diameter NMR sample tube containing 200 µM uniformly ^{15}N-labeled EmN, EmN67C, EmN49, EmN132, EmN delK37 or EmN132 Y4C/E35C. 3D ^{1}H ^{15}N ^{13}C correlation spectra were acquired using 3D HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO and HN(CO)(CA)NH pulse sequences at 303 K, on a 3-mm-diameter NMR sample tube containing 500-600 µM uniformly ^{15}N^{13}C-labeled EmN132. The data were processed using Topspin3.1 (Bruker) and analyzed with CCPNMR [36].

Solid-State NMR Spectroscopy.

After purification and His-Tag cleavage, proteins were dialyzed against 20 mM Tris-HCl pH 8, 30mM NaCl and 5mM DTT.

Solid-state NMR experiments for the EmN and EmN132 constructs were recorded on a Bruker Avance III HD 800 MHz spectrometer with 18.8 T external magnetic field or on a Bruker Avance III HD 700 MHz spectrometer with 16.4 T external magnetic field. All solid-state NMR experiments for the EmN delK37 mutant were recorded on a Bruker Avance III HD 700 MHz spectrometer with 16.4 T external magnetic field. The sample temperature was determined using the water-proton signal referenced to DSS and kept at 283 K. Double quantum – single quantum spectra were recorded with magic-angle spinning at 8 kHz and C-C correlation spectra with 11 kHz MAS rate. High-power ^{1}H decoupling was performed during evolution and detection periods using SPINAL-64 [37] at a radio frequency strength of 83 kHz. All spectra were processed using Bruker TopSpin 3.2 and analyzed with CcpNmr [36].

Double quantum – single quantum C-C correlation spectra were recorded using SPC5 [32] excitation and reconversion periods of 0.5 ms. Maximum acquisition times for the direct dimension were 10 ms and for the indirect dimension 1.2 ms. For the EmN and EmN132 samples the number of scans was set to 336 and it was set to 192 for the mutant delK37. C-C correlation spectra using proton-driven spin diffusion (PDSD) were recorded using 50 ms mixing time [38]. Maximum acquisition times were 15 ms for the direct and 10 ms for the indirect dimension. For the EmN and EmN132 samples the number of scans was set to 112 and it was set to 184 for the mutant delK37. Secondary structure classifications were based on the carbon chemical shifts as described by Wang et al. [39].

Electron microscopy.

To observe self-assembled proteins, EmN, EmN49, EmN132, EmN67C and EmN delK37 were dialyzed against 20 mM Tris-HCl (pH 8.0), 30 mM NaCl and 5 mM DTT. They were concentrated up to 600 μM and incubated at 293 K during 1 week. Specimens were prepared by negative staining with 2 % uranyl acetate on glow-discharged carbon-coated copper grids. Data collection was performed using a Tecnai Spirit transmission electron microscope (FEI) equipped with a LaB_6 filament, operating at 100 kV. Images were recorded on a K2 Base camera (Gatan, 4kx4k) at 15,000 or 4,400 magnification (pixel size at specimen level – 0.25 & 0.83 nm, respectively).

Thioflavin kinetics.

After purification and His-Tag cleavage, proteins were dialyzed against 20mM Tris-HCl pH 8.0, 30 mM NaCl and 5 mM DTT. Oligomer assembly was followed with time by incubating proteins at 300 μM and 310 K and by regularly taking protein aliquots to be analyzed by fluorescence. These aliquots were diluted in a thioflavin T (ThT) containing buffer so as to obtain 20 μM protein and 2.5 μM ThT in 20 mM Tris (pH 8.0), 30 mM NaCl, and 5 mM DTT.

The fluorescence measurements were carried out in a 60 μl cuvette at 293 K using a fluorimeter JASCO ADP-303T. ThT fluorescence was monitored using an excitation wavelength of 440 nm and fluorescence emission was read at 480 nm.

Limited Proteolysis.

After purification and His-Tag cleavage, EmN, EmN132 and mutant delK37 were dialyzed against 20mM Tris-HCl pH 8.0, 30 mM NaCl and 5 mM DTT. For experiments on monomers, proteins were concentrated up to 125 μM. Filament formation was performed by concentrating proteins until 600 μM and heating the samples during 1 hour at 338 K (338K) [25]. Before proteolysis, filaments were diluted at 125 μM and both monomers and filaments were incubated with Chymotrypsin or EndoproteinaseGluC (1:1000) during one hour at 310K. Protease reactions were stopped by the addition of 10 μl of Laemmli blue and sample heating during 5 min at 368 K (95°C). Degradation products were then resolved by SDS-PAGE and analyzed by means of high-resolution tandem mass spectrometry.

Tandem Mass Spectrometry.

For mapping fragments generated by limited proteolysis, excised polyacrylamide bands were first rinsed with MilliQ water, de-stained, and then subjected to in-gel protein reduction with DTT followed by protein alkylation with iodoacetamide, as described [40]. They were digested with either sequencing-grade Trypsin Gold and ProteaseMax surfactant (Promega) for 1h at 323 K or sequencing-grade chymotrypsin for 1h at 298 K as previously reported [41]. Trifluoroacetic acid was added to the samples at 0.5% final concentration to stop the enzymatic reaction. The resulting peptides were identified by tandem mass spectrometry with an LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific) coupled to an UltiMate 3000 nRSLC system (Dionex, ThermoFisher), under conditions similar to those previously described [42, 43]. Briefly, peptide mixtures (10 μL injection) were first desalted on-line in a reverse-phase Acclaim PepMap100 C18 μ-precolumn (5 μm, 100 Å, 300 μm inner diameter × 5 mm, LC Packings), and then were separated on a nanoscale Acclaim PepMap100 C18 capillary column (3 μm, 100 Å, 75 μm i.d. × 50 cm, LC Packings) using a 90-min gradient from 4 to 40% solvent B (CH_3CN, 0.01% formic acid) with solvent A being H_2O, 0.01% formic acid. The full-scan mass spectra of high mass accuracy were measured in the Orbitrap analyzer from m/z 300 to 1800 in data-dependent mode.

The MS/MS scans in the linear ion trap were acquired on the seven most abundant precursor ions. MS/MS data processing was done as described previously [44] against the Swissprot_Human database comprising 20,192 protein sequences using the following parameters: semi-tryptic peptides with a maximum of two miss-cleavages; mass error tolerances set at 5 ppm for parent ions and at 0.5 Da for fragments; fixed modification for carbamidomethylated cysteine and variable modification for methionine oxidation. Peptides were filtered with a p-value threshold below 0.05 and a rank set at 1.

Acknowledgments

The authors acknowledge Pascal Fricke for his help in setting up and analyzing the solid-state experiments and the FMP NMR facility (Leibniz-Institut für Molekulare Pharmakologie, Berlin, Germany) for access to the NMR spectrometers.

Funding

This work was supported by CEA, CNRS and University Paris South, by the French Infrastructure for Integrated Structural Biology (https://www.structuralbiology.eu/networks/frisbi, grant number ANR-10-INSB-05-01, Acronym FRISBI) and by the French Association against Myopathies (AFM; research grants no. 17243 and 20018 to S.Z.-J. and PhD fellowship no. 18159 to C.S.).

Author contributions

CS, FC, NE, IH, FXT performed the biochemistry, fluorescence and liquid-state NMR experiments. KH and MZ carried out the solid-state NMR experiments. CS and AAA were involved in the EM analyses. BAB and JA performed and analyzed the mass spectrometry experiments. CC, AL and SZJ supervised the project. CS, AL, SZJ wrote the paper.

Legends

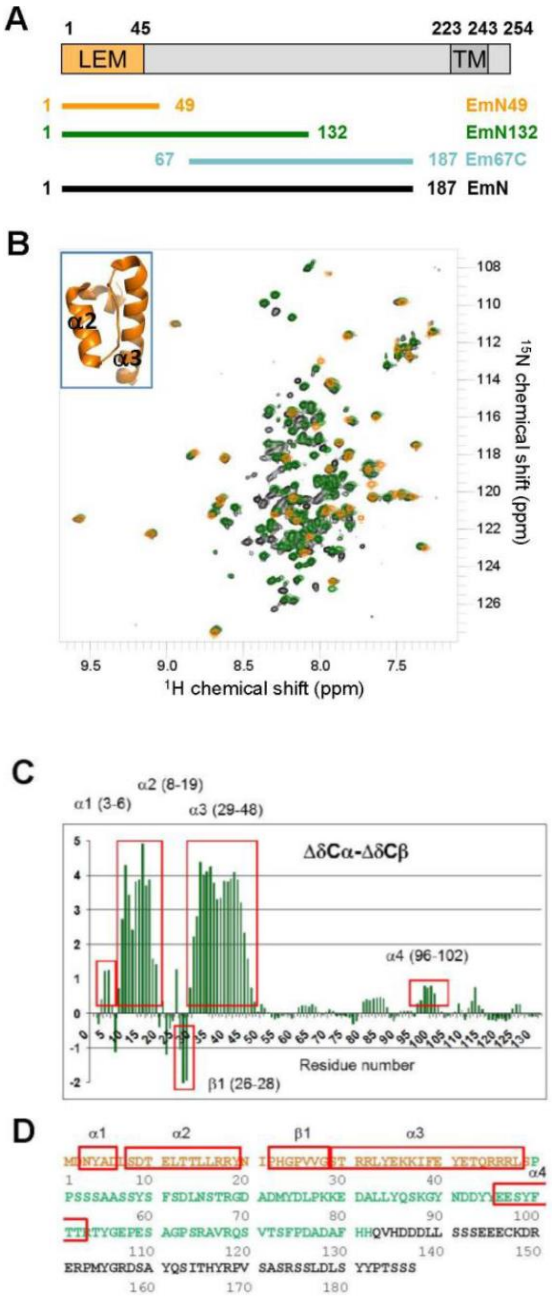


Figure 1. A LEM domain with a $\alpha\alpha\beta\alpha$ motif is present in the monomeric emerin fragments EmN49, EmN132 and EmN. **(A)** Emerin architecture and fragments studied by solution NMR. **(B)** Superimposition of the solution-state NMR ^{1}H-^{15}N HSQC spectra of EmN49 (the LEM domain, in orange; its 3D structure is boxed in the upper left corner of the figure (PDB references 1JEI and 2ODC)), EmN132 (in green) and EmN (in black), recorded at pH 6.5, 303 K and 750 MHz. Peak overlap reveals that the LEM domain adopts the same 3D structure in the 3 fragments. **(C)** Analysis of the solution-state NMR ^{13}C signals of EmN132 (pH 6.5, 303 K, 750 MHz), confirming that its N-terminal domain exhibits 3 helices and 1 extended strand, and further showing that region 50 to 132 is essentially unstructured.

Analysis of the NMR ^{13}C signals of emerin fragment 67-170 consistently showed that this fragment is intrinsically disordered (Samson et al., 2016). **(D)** Sequence of the emerin fragment EmN. The secondary structure elements deduced from (C) are boxed in red.

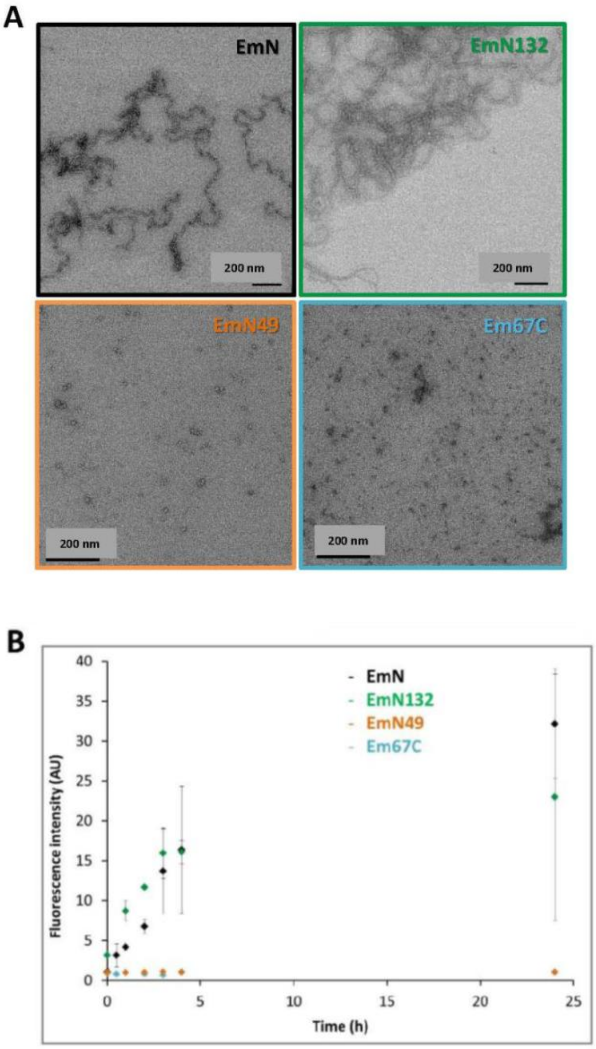


Figure 2. The LEM sequence and part of the unstructured region are necessary for emerin self-assembly. **(A)** Negative staining EM images of emerin fragments EmN, EmN132, EmN49 and Em67C concentrated at 600 µM and incubated at 293 K during one week. **(B)** ThT fluorescence as a function of the incubation time at 310 K, measured on emerin fragments concentrated at 300 µM. Only EmN132 and EmN form filaments. They self-assemble with similar kinetics. EmN49 and Em67C do not form filaments, as also observed by EM.

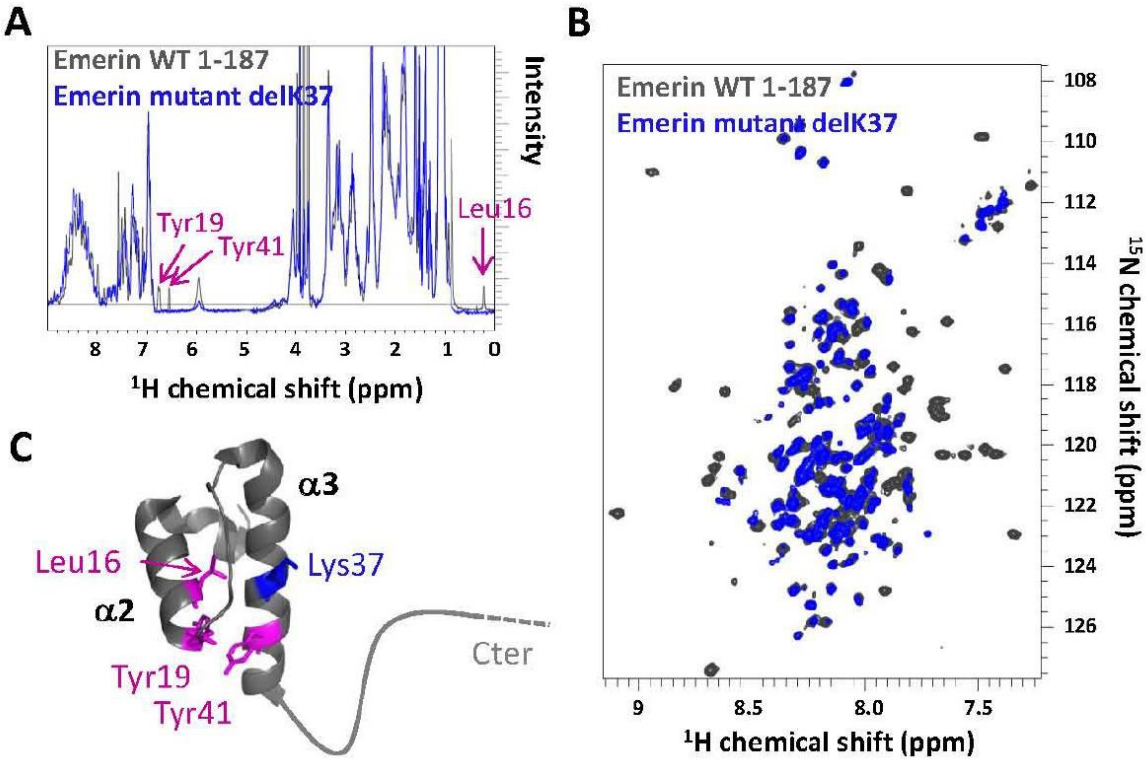


Figure 3. The LEM domain of disease-causing mutant delK37 is unstructured. **(A)** Superimposition of the solution-state NMR 1D spectra of EmN (in black) and delK37 (in blue), recorded at pH 6.5, 303 K, 750 MHz. Three signals corresponding to side-chain protons of buried residues Leu16, Tyr19 and Tyr41 are marked in magenta. These signals are strongly shifted in the spectrum of delK37. **(B)** Superimposition of the solution-state NMR 2D ^{1}H-^{15}N HSQC spectra of EmN (in black) and delK37 (in blue), recorded at pH 6.5, 303 K, 750 MHz. Signals corresponding to the LEM domain (in orange on Fig. 1B) are moved towards regions corresponding to random coil in the spectrum of delK37. **(C)** Representation of the 3D structure of EmN, with residue K37, lacking in the mutant, colored in blue, and residues Leu16, Tyr19 and Tyr41, exhibiting different chemical environments in the mutant, colored in magenta, as in (A).

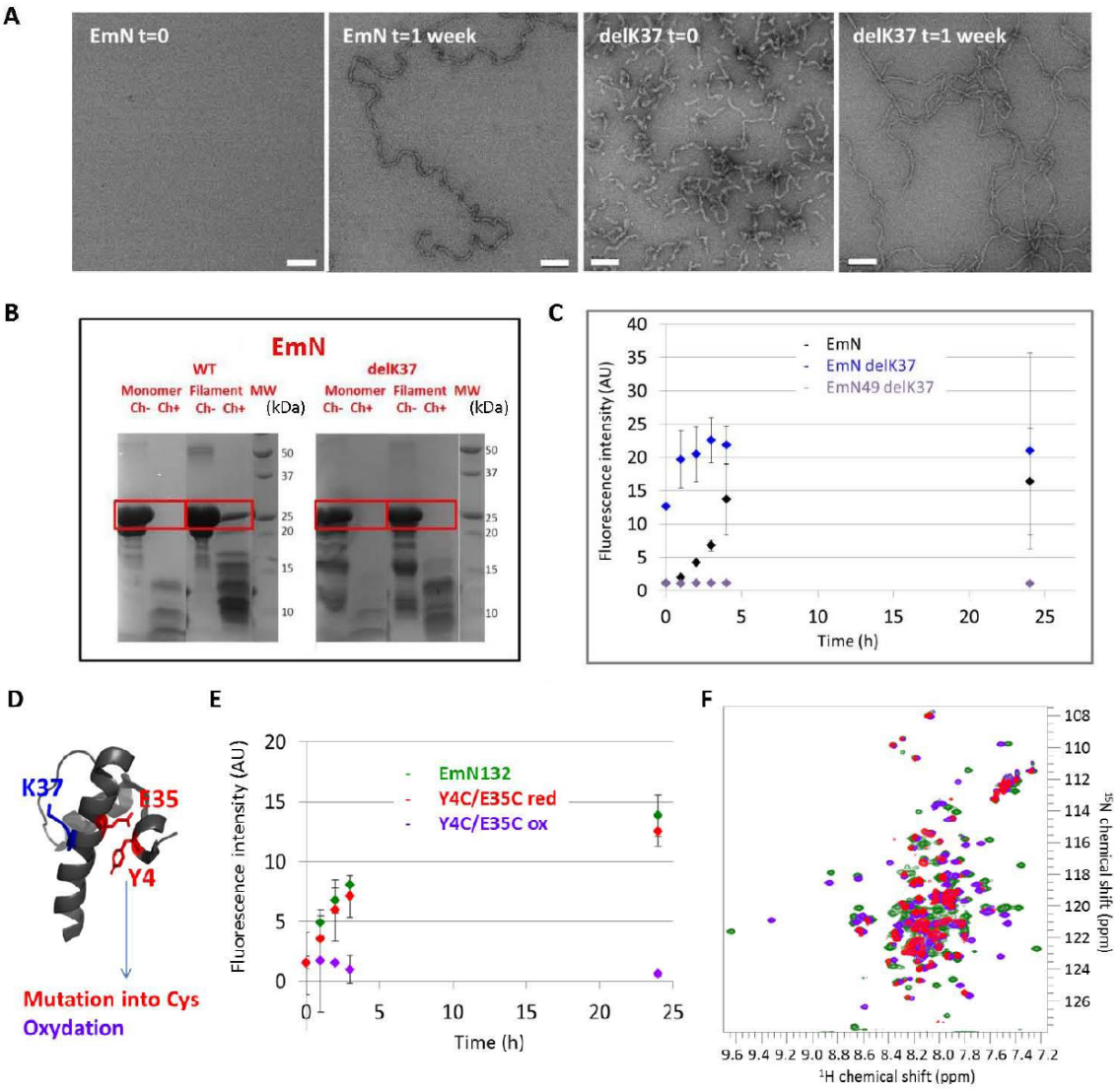
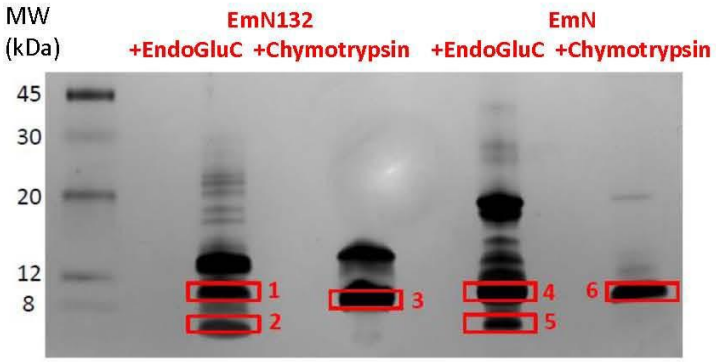


Figure 4. The EmN mutant delK37 self-assembles faster than wild-type EmN. **(A)** Negative staining EM images of emerin fragments EmN wild-type and delK37, before (t=0) and after (t=1 week) incubation during one week at 293 K and 600 µM. White bars represent 100 nm. **(B)** SDS-PAGE gels of EmN wild-type and delK37, corresponding to the samples observed in (A). The samples were deposited before and after incubation during one hour at 310 K with chymotrypsin 1:1000. Self-assembly protects both wild-type and mutated EmN from proteolysis. Degradation products migrate similarly in both cases, suggesting that the 3D structures of wild-type and mutated filaments are closely related. **(C)** ThT fluorescence as a function of the incubation time at 310 K, measured on EmN wild-type and delK37 concentrated at 300 µM. The mutant delK37 forms filaments faster than WT EmN. **(D)** 3D structure of the LEM domain with K37 in blue and the 2 residues Y4 and E35 mutated into cysteine in mutant EmN132 Y4C/E35C in red. **(E)** ThT fluorescence intensity as a function of time measured on EmN132 wild-type (green), Y4C/E35C in the presence of 5 mM DTT (red) and Y4C/E35C in the absence of DTT (mauve), measured at 310 K and 300 µM. The mutant forms filaments only in reducing conditions.

(F) Superimposition of the solution-state NMR 2D ^{1}H-^{15}N HSQC spectra of EmN132 wild-type (green, Tris pH 7.5 100 mM NaCl), Y4C/E35C in the presence of 5 mM DTT (red, Tris pH 7.5 30 mM NaCl) and Y4C/E35C in the absence of DTT (mauve, Tris pH 7.5 30 mM NaCl), recorded at 303 K, 700 MHz. Whereas in the absence of DTT the HSQC peaks are thin and well-dispersed, as for EmN132 wild-type, in the presence of DTT the HSQC peaks are only located in the central region of the spectrum and show heterogeneous linewidths.



			EmN132			EmN		
			band	band	band	band	band	band
Peptide sequence	start	stop	1	2	3	4	5	6
LTTLLR	12	17	+	+	-	+	-	-
RYNIPHGPVVGSTR	18	31	+	+	+	+	+	+
RYNIPHGPVVGSTRR	18	32	+	+	-	+	+	-
YNIPHGPVVGSTR	19	31	+	+	+	+	+	-
NIPHGPVVGSTR	20	31	-	-	+	-	-	+
KIFEYETQR	37	45	+	-	+	+	-	+
IFEYETQR	38	45	+	-	+	+	-	+
RLSPPSSSAASSY	47	59	-	-	+	-	-	+
LSPPSSSAASSY	48	59	-	-	+	-	-	+
LSPPSSSAASSYSFSDLNSTR	48	68	+	-	-	+	-	-
LSPPSSSAASSYSFSDLNSTRGDADMY	48	74	-	-	+	-	-	-
GDADMYDLPK	69	78	+	-	-	+	-	-
GDADMYDLPKKE	69	80	+	-	-	+	-	-
GYNDDYYEESYFTTR	89	103	+	-	-	-	-	-
DSAYQSITHYRPVSASR	158	174	-	-	-	+	-	-
SSLDLSYYPTSSS	175	187	-	-	-	+	-	-

Figure 5. The filament structural core contains the LEM domain. SDS PAGE gel of emerin fragments EmN132 and EmN after 4 hours of incubation at 310 K with either endoproteinase Glu-C or chymotrypsin (1:1000). The bands squared in red were characterized by mass spectrometry and their peptide composition is summarized in the table below. As judged by spectral count significant amounts of peptides are indicated by + whereas low and no signal are indicated by -. Clearly, all samples contain at least part of the LEM region. Sample 4 contains a mixed of peptides comprising either the N-terminus of the C-terminus of EmN, but further proteolysis (sample 5) shows that the LEM region is the most protected fragment of EmN. This experiment was performed in duplicate.

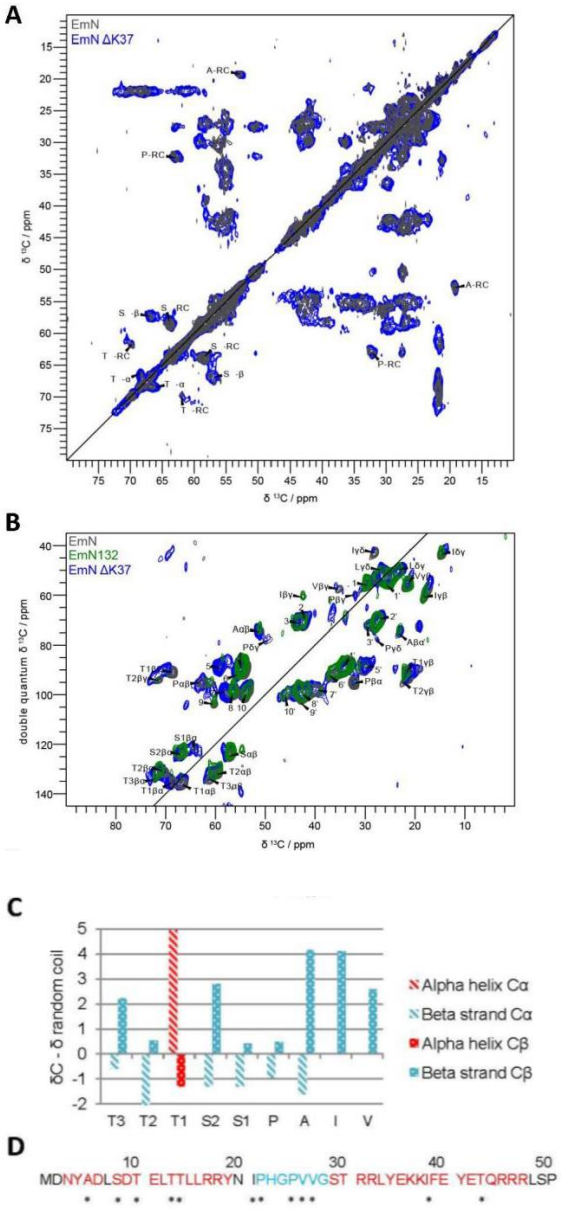


Figure 6. The filament structural core is similar in EmN wild-type and delK37 and contains β-structure. **(A)** Solid-state NMR 2D PDSD spectra recorded with a mixing time of 50 ms for EmN wild-type (black) and delK37 (blue); secondary structure classifications are indicated (α= alpha helical, β= beta sheet, RC = random coil). **(B)** Solid-state NMR double quantum - single quantum C-C correlation spectra for EmN wild-type (black) and delK37 (blue) and EmN132 (green), with the identified residue types indicated. **(C)** Secondary structure analysis for the assigned residues from the double quantum - single quantum C-C correlation spectrum of EmN as shown in (B). Depicted is the difference between ^{13}C observed chemical shift and random coil reference [39]. Negative differences for Cα and positive differences for Cβ indicate β-sheet secondary structure; corresponding bars are colored in cyan. Bars indicating α-helix secondary structure are colored in red. **(D)** Sequence of the LEM domain colored as a function of the secondary structure observed in solution (see Figures 1C,D), with α–helices displayed in red and β-strands in cyan. Asterisks indicate residues identified in (C) and discussed in the text.

References

1. Petrie, R. J., Koo, H. & Yamada, K. M. (2014) Generation of compartmentalized pressure by a nuclear piston governs cell motility in a 3D matrix, *Science.* **345**, 1062-5.
2. Osmanagic-Myers, S., Dechat, T. & Foisner, R. (2015) Lamins at the crossroads of mechanosignaling, *Genes & development.* **29**, 225-37.
3. Bertrand, A. T., Ziaei, S., Ehret, C., Duchemin, H., Mamchaoui, K., Bigot, A., Mayer, M., Quijano-Roy, S., Desguerre, I., Laine, J., Ben Yaou, R., Bonne, G. & Coirault, C. (2014) Cellular microenvironments reveal defective mechanosensing responses and elevated YAP signaling in LMNA-mutated muscle precursors, *Journal of cell science.* **127**, 2873-84.
4. Swift, J. & Discher, D. E. (2014) The nuclear lamina is mechano-responsive to ECM elasticity in mature tissue, *Journal of cell science.* **127**, 3005-15.
5. Guilluy, C., Osborne, L. D., Van Landeghem, L., Sharek, L., Superfine, R., Garcia-Mata, R. & Burridge, K. (2014) Isolated nuclei adapt to force and reveal a mechanotransduction pathway in the nucleus, *Nature cell biology.* **16**, 376-81.
6. Ho, C. Y., Jaalouk, D. E., Vartiainen, M. K. & Lammerding, J. (2013) Lamin A/C and emerin regulate MKL1-SRF activity by modulating actin dynamics, *Nature.* **497**, 507-11.
7. Harr, J. C., Gonzalez-Sandoval, A. & Gasser, S. M. (2016) Histones and histone modifications in perinuclear chromatin anchoring: from yeast to man, *EMBO reports.* **17**, 139-55.
8. Bertrand, A. T., Renou, L., Papadopoulos, A., Beuvin, M., Lacene, E., Massart, C., Ottolenghi, C., Decostre, V., Maron, S., Schlossarek, S., Cattin, M. E., Carrier, L., Malissen, M., Arimura, T. & Bonne, G. (2012) DelK32-lamin A/C has abnormal location and induces incomplete tissue maturation and severe metabolic defects leading to premature death, *Human molecular genetics.* **21**, 1037-48.
9. Azibani, F., Muchir, A., Vignier, N., Bonne, G. & Bertrand, A. T. (2014) Striated muscle laminopathies, *Seminars in cell & developmental biology.* **29**, 107-15.
10. Lammerding, J., Schulze, P. C., Takahashi, T., Kozlov, S., Sullivan, T., Kamm, R. D., Stewart, C. L. & Lee, R. T. (2004) Lamin A/C deficiency causes defective nuclear mechanics and mechanotransduction, *The Journal of clinical investigation.* **113**, 370-8.
11. Lammerding, J., Hsiao, J., Schulze, P. C., Kozlov, S., Stewart, C. L. & Lee, R. T. (2005) Abnormal nuclear shape and impaired mechanotransduction in emerin-deficient cells, *The Journal of cell biology.* **170**, 781-91.
12. Hale, C. M., Shrestha, A. L., Khatau, S. B., Stewart-Hutchinson, P. J., Hernandez, L., Stewart, C. L., Hodzic, D. & Wirtz, D. (2008) Dysfunctional connections between the nucleus and the actin and microtubule networks in laminopathic models, *Biophysical journal.* **95**, 5462-75.
13. Emerson, L. J., Holt, M. R., Wheeler, M. A., Wehnert, M., Parsons, M. & Ellis, J. A. (2009) Defects in cell spreading and ERK1/2 activation in fibroblasts with lamin A/C mutations, *Biochimica et biophysica acta.* **1792**, 810-21.

14. Worman, H. J. & Bonne, G. (2007) "Laminopathies": a wide spectrum of human diseases, *Experimental cell research.* **313**, 2121-33.
15. Krimm, I., Ostlund, C., Gilquin, B., Couprie, J., Hossenlopp, P., Mornon, J. P., Bonne, G., Courvalin, J. C., Worman, H. J. & Zinn-Justin, S. (2002) The Ig-like structure of the C-terminal domain of lamin A/C, mutated in muscular dystrophies, cardiomyopathy, and partial lipodystrophy, *Structure.* **10**, 811-23.
16. Bione, S., Maestrini, E., Rivella, S., Mancini, M., Regis, S., Romeo, G. & Toniolo, D. (1994) Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular dystrophy, *Nature genetics.* **8**, 323-7.
17. Ellis, J. A., Craxton, M., Yates, J. R. & Kendrick-Jones, J. (1998) Aberrant intracellular targeting and cell cycle-dependent phosphorylation of emerin contribute to the Emery-Dreifuss muscular dystrophy phenotype, *Journal of cell science.* **111 (Pt 6)**, 781-92.
18. Ben Yaou, R., Toutain, A., Arimura, T., Demay, L., Massart, C., Peccate, C., Muchir, A., Llense, S., Deburgrave, N., Leturcq, F., Litim, K. E., Rahmoun-Chiali, N., Richard, P., Babuty, D., Recan-Budiartha, D. & Bonne, G. (2007) Multitissular involvement in a family with LMNA and EMD mutations: Role of digenic mechanism?, *Neurology.* **68**, 1883-94.
19. Berk, J. M., Tifft, K. E. & Wilson, K. L. (2013) The nuclear envelope LEM-domain protein emerin, *Nucleus.* **4**, 298-314.
20. Sosa, B. A., Rothballer, A., Kutay, U. & Schwartz, T. U. (2012) LINC complexes form by binding of three KASH peptides to domain interfaces of trimeric SUN proteins, *Cell.* **149**, 1035-47.
21. Wang, W., Shi, Z., Jiao, S., Chen, C., Wang, H., Liu, G., Wang, Q., Zhao, Y., Greene, M. I. & Zhou, Z. (2012) Structural insights into SUN-KASH complexes across the nuclear envelope, *Cell research.* **22**, 1440-52.
22. Zhou, Z., Du, X., Cai, Z., Song, X., Zhang, H., Mizuno, T., Suzuki, E., Yee, M. R., Berezov, A., Murali, R., Wu, S. L., Karger, B. L., Greene, M. I. & Wang, Q. (2012) Structure of Sad1-UNC84 homology (SUN) domain defines features of molecular bridge in nuclear envelope, *The Journal of biological chemistry.* **287**, 5317-26.
23. Nie, S., Ke, H., Gao, F., Ren, J., Wang, M., Huo, L., Gong, W. & Feng, W. (2016) Coiled-Coil Domains of SUN Proteins as Intrinsic Dynamic Regulators, *Structure.* **24**, 80-91.
24. Berk, J. M., Simon, D. N., Jenkins-Houk, C. R., Westerbeck, J. W., Gronning-Wang, L. M., Carlson, C. R. & Wilson, K. L. (2014) The molecular basis of emerin-emerin and emerin-BAF interactions, *Journal of cell science.* **127**, 3956-69.
25. Herrada, I., Samson, C., Velours, C., Renault, L., Ostlund, C., Chervy, P., Puchkov, D., Worman, H. J., Buendia, B. & Zinn-Justin, S. (2015) Muscular Dystrophy Mutations Impair the Nuclear Envelope Emerin Self-assembly Properties, *ACS chemical biology.* **10**, 2733-42.
26. Manilal, S., Nguyen, T. M., Sewry, C. A. & Morris, G. E. (1996) The Emery-Dreifuss muscular dystrophy protein, emerin, is a nuclear membrane protein, *Human molecular genetics.* **5**, 801-8.
27. Yorifuji, H., Tadano, Y., Tsuchiya, Y., Ogawa, M., Goto, K., Umetani, A., Asaka, Y. & Arahata, K. (1997) Emerin, deficiency of which causes Emery-Dreifuss muscular dystrophy, is localized at the inner nuclear membrane, *Neurogenetics.* **1**,

135-40.
28. Wolff, N., Gilquin, B., Courchay, K., Callebaut, I., Worman, H. J. & Zinn-Justin, S. (2001) Structural analysis of emerin, an inner nuclear membrane protein mutated in X-linked Emery-Dreifuss muscular dystrophy, *FEBS letters.* **501**, 171-6.
29. Samson, C., Herrada, I., Celli, F., Theillet, F. X. & Zinn-Justin, S. (2016) 1H, 13C and 15N backbone resonance assignment of the intrinsically disordered region of the nuclear envelope protein emerin, *Biomolecular NMR assignments.* **10**, 179-82.
30. Lee, K. K., Haraguchi, T., Lee, R. S., Koujin, T., Hiraoka, Y. & Wilson, K. L. (2001) Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF, *Journal of cell science.* **114**, 4567-73.
31. Cai, M., Huang, Y., Suh, J. Y., Louis, J. M., Ghirlando, R., Craigie, R. & Clore, G. M. (2007) Solution NMR structure of the barrier-to-autointegration factor-Emerin complex, *The Journal of biological chemistry.* **282**, 14525-35.
32. Hohwy, M., Rienstra, C. M., Jaroniec, C. P. & Griffin, R. G. (1999) Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy, *J Chem Phys.* **110**, 7983-7992.
33. Barton, L. J., Soshnev, A. A. & Geyer, P. K. (2015) Networking in the nucleus: a spotlight on LEM-domain proteins, *Current opinion in cell biology.* **34**, 1-8.
34. Shumaker, D. K., Lee, K. K., Tanhehco, Y. C., Craigie, R. & Wilson, K. L. (2001) LAP2 binds to BAF.DNA complexes: requirement for the LEM domain and modulation by variable regions, *The EMBO journal.* **20**, 1754-64.
35. Cai, M., Huang, Y., Ghirlando, R., Wilson, K. L., Craigie, R. & Clore, G. M. (2001) Solution structure of the constant region of nuclear envelope protein LAP2 reveals two LEM-domain structures: one binds BAF and the other binds DNA, *The EMBO journal.* **20**, 4399-407.
36. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J. & Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline, *Proteins.* **59**, 687-96.
37. Fung, B. M., Khitrin, A. K. & Ermolaev, K. (2000) An improved broadband decoupling sequence for liquid crystals and solids, *J Magn Reson.* **142**, 97-101.
38. Bloembergen, N. (1949) On the Interaction of Nuclear Spins in a Crystalline Lattice, *Physica.* **15**, 386-426.
39. Wang, Y. & Jardetzky, O. (2002) Probability-based protein secondary structure identification using combined NMR chemical-shift data, *Protein science : a publication of the Protein Society.* **11**, 852-61.
40. Hartmann, E. M. & Armengaud, J. (2014) Shotgun proteomics suggests involvement of additional enzymes in dioxin degradation by Sphingomonas wittichii RW1, *Environ Microbiol.* **16**, 162-176.
41. Kish, A., Gaillard, J. C., Armengaud, J. & Elie, C. (2016) Post-translational methylations of the archaeal Mre11:Rad50 complex throughout the DNA damage response, *Mol Microbiol.* **100**, 362-378.
42. Clair, G., Armengaud, J. & Duport, C. (2012) Restricting Fermentative Potential by Proteome Remodeling AN ADAPTIVE STRATEGY EVIDENCED IN BACILLUS CEREUS, *Mol Cell Proteomics.* **11**.
43. Rubiano-Labrador, C., Bland, C., Miotello, G., Guerin, P., Pible, O., Baena, S. & Armengaud, J. (2014) Proteogenomic insights into salt tolerance by a halotolerant alpha-proteobacterium isolated from an Andean saline spring, *J Proteomics.* **97**, 36-47.
44. Christie-Oleza, J. A., Pina-Villalonga, J. M., Bosch, R., Nogales, B. & Armengaud, J. (2012) Comparative Proteogenomics of Twelve Roseobacter Exoproteomes Reveals Different Adaptive Strategies Among These Marine Bacteria, *Mol Cell Proteomics.* **11**.

2. IMPACT OF EMERIN LEM DOMAIN MUTATIONS ON EMERIN SELF-ASSEMBLY

In the first article, we studied the impact of five mutants (S54F, Δ95-99, Q133H and P183T/H) that show a defect in emerin unstructured region and cause Emery-Dreifuss Muscular Dystrophy, on EmN self-assembly, and found that mutant Δ95-99 cannot form filaments anymore. In the second paper I presented, we worked with another mutant showing a defect in emerin LEM domain and known to cause cardiac defects, mutant ΔK37. This mutant is particularly interesting because it exhibits a destabilized LEM domain and the EmN mutant ΔK37 is able to form filaments faster than EmN WT.

In parallel, we worked with two other mutants showing a defect in emerin LEM domain and associated to the same symptoms as ΔK37: mutants P22L and T43I (G. Bonne, personal communication). I here report my results on the 3 mutants of the LEM domain, which will be published in another article.

First, together with a PhD student in the lab, Nada Essawy, I studied the structure of EmN mutants (P22L and T43I) by NMR. After recording ^{15}N-^{1}H HSQC spectrum of each mutant (figure 35), we observed no chemical shift differences compared to the spectrum that we obtained for EmN WT, except for some peaks, corresponding to amino acids close to the mutated residues. We concluded that both mutations do not impact the structure of EmN monomers.

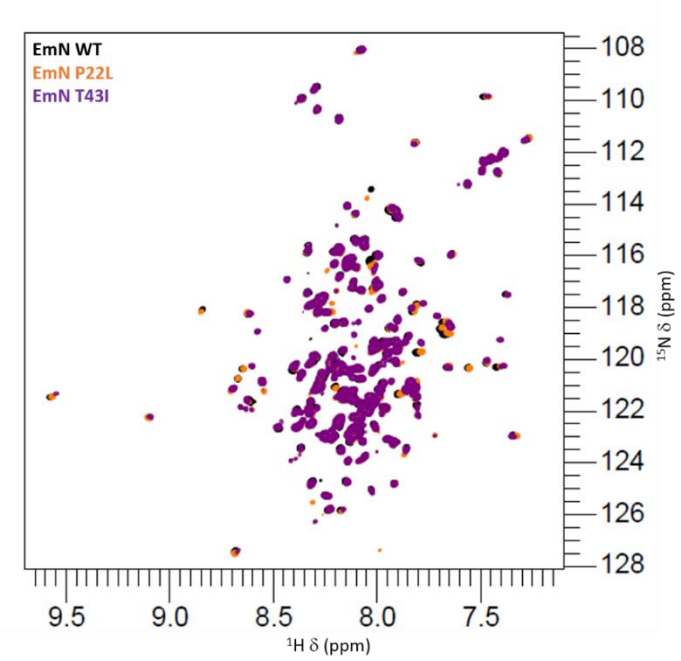


Figure 35 : Structure comparison between EmN monomer WT and mutants.

Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on an EmN WT sample at 100µM (in black), an EmN P22L sample at 100µM (in orange) and an EmN T43I sample at 100µM (in purple), in 20mM Phosphate pH6.5, 30mM NaCl, 5mM DTT, at 303K and 750MHz (FMP Berlin).

Second, I studied the self-assembly kinetics of the three mutated EmN fragments using thioflavin T fluorescence. As it is described in the second paper I presented, we used this technique because thioflavin T is a benzothiazole dye that increases in fluorescence upon binding to amyloid-like oligomers[246,247]. It gives a fluorescence signal at 480nm after excitation at 440nm, and this signal is significantly enhanced when thioflavin T is bound to these amyloid-like structures, especially the "cross-β" structures, characterized by an extended β-conformation with the strands running perpendicular to the long axis of the assembly of multiple β-sheets[248]. We first verified that thioflavin T fluorescence is not enhanced in the presence of EmN monomers, and we showed that it nicely increases during EmN self-assembly[4]. In the case of the three mutants, I concentrated them up to 300µM and incubated each protein sample at 310K, during one day. I took a sample for fluorescence measurement at 480nm every hour during 4h and then the next day. For each measurement, I diluted my protein sample to 40µM in the kinetic buffer (20mM Tris-HCl pH8, 30mM NaCl) and added 10µM of thioflavin T (figure 36).

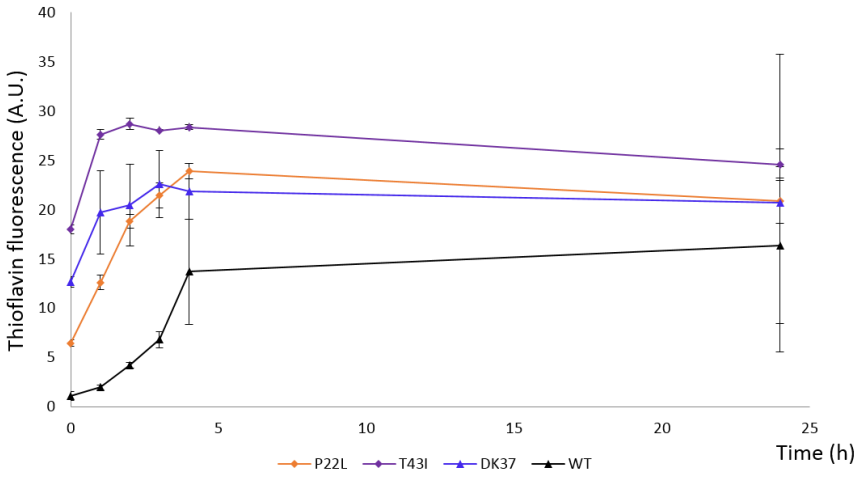


Figure 36 : Self-assembly kinetics of EmN T43I, P22L and ΔK37 mutants, compared to EmN WT, as followed by thioflavin T fluorescence.

The fluorescence intensities at 480nm of EmN WT, T43I, P22L and ΔK37 are represented in black, purple, orange and blue, respectively. Experiments were performed in 20mM Tris-HCl pH8, 30mM NaCl, at 310K.

My results show that the three EmN mutants self-assemble faster than EmN WT. This was never observed for any of the five mutants showing a defect in the disordered part of EmN[4]. To confirm that these three mutants could form filaments, we used negative staining electron microscopy. In more details, we concentrated the three proteins up to 600µM and observed the samples after one day at room temperature by EM (figure 37). We confirmed that these three EmN mutants are able to form filaments, which look like EmN WT filaments.

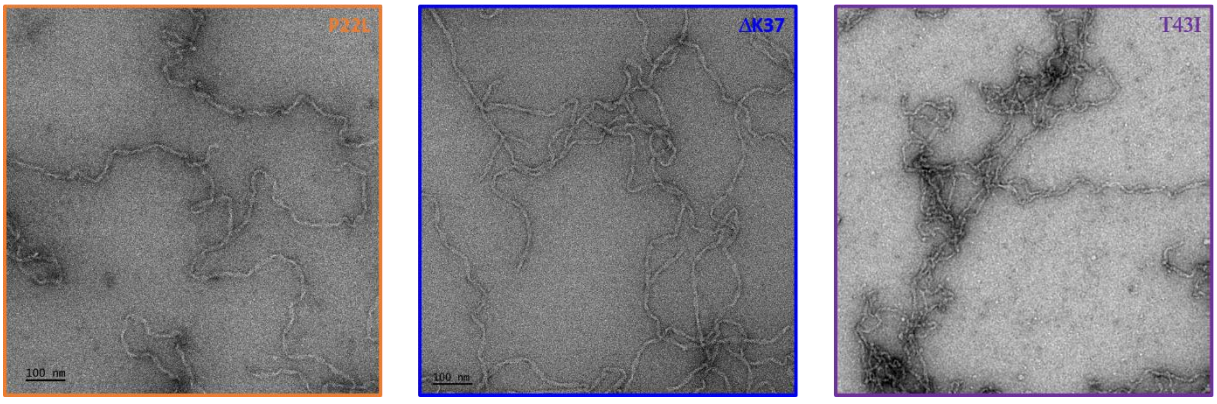


Figure 37 : Negative-staining EM pictures of the three EmN mutants.

Pictures were obtained in the presence of 2% uranyl acetate, using a Tecnai Spirit transmission electron microscope (FEI) equipped with a LaB6 filament (Dr Ana-Andreea Arteni, I2BC, Gif-sur-Yvette). Filaments were obtained after incubation at 600µM, during one day, at room temperature. P22L, ΔK37 and T43I images are framed in orange, blue and purple, respectively.

3. DESCRIPTION OF THE PHOSPHORYLATION OF EMN BY FUNCTIONALLY RELEVANT KINASES

Emerin is heavily phosphorylated in cells. Forty-two phosphorylation sites were reported on emerin *in vivo* (25 serine, 4 threonine and 11 tyrosine)[196]. However, responsible kinases, phosphorylation kinetics and functional consequences on emerin oligomerization and interaction with partners remain poorly understood. For example, emerin phosphorylation by the Src kinase during a mechanical stress was observed on isolated nuclei[107]. However, the consequences of this phosphorylation process on emerin oligomerization and lamin binding is not known. As, after a mechanical stress, lamin A/C was shown to be recruited to the LINC complex in order to reinforce nuclear stiffness[107], and because it was recently shown that application of a strain could increase emerin levels at the outer nuclear membrane without altering total emerin protein level[249], I did the hypothesis that emerin phosphorylation by Src could disrupt emerin binding to the nuclear lamina.

Here, first, I present the results I obtained by NMR and mass spectrometry, concerning emerin phosphorylation by Src *in vitro*. I tried to characterize which emerin residues are phosphorylated by Src kinase. Then, because Src phosphorylates emerin on residues 74 and 95, and because we found that this region is important for filament formation, I choose to study impact of Src phosphorylation on emerin oligomerization. Finally, I observed emerin phosphorylation by two other kinases predicted as able to phosphorylate emerin (CK1 and CK2) and in cell extracts.

a. Characterization of emerin phosphorylation in vitro

1) Emerin phosphorylation by Src kinase, *in vitro*

The Src kinase phosphorylates emerin tyrosines 74 and 95 during application of a mechanical force on the nucleus; mutation of these two residues impair nuclear response to a mechanical stress[107]. I first identified which emerin residues are phosphorylated *in vitro* by Src, through two different techniques which are NMR and mass spectrometry.

I also did some preliminary studies about the impact of Src phosphorylation on emerin oligomerization, following emerin self-assembly by fluorescence and electron microscopy. The final aim was to observe impact of Src phosphorylation on emerin binding to lamin A/C.

First, I studied emerin modifications in the presence of Src kinase by mass spectroscopy (figure 38). I worked with two emerin constructs: Emerin 1-187 (EmN) and 1-132 (EmN132). I mixed the emerin fragment with 10µl of Src kinase and 2mM ATP and incubated them during 12h at 303K. Then I observed the phosphorylated sample by SDS-PAGE and sent the band corresponding to the phosphorylated emerin fragment to the group of Dr Jean Armengaud (CEA Marcoule), who analyzed it by MS/MS mass spectroscopy with an LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific) coupled to an UltiMate 3000 nRSLC system (Dionex, ThermoFisher). This experiment revealed that only Tyr74 and Tyr95 are phosphorylated by Src at more than 20% in our conditions. Tyr161 was phosphorylated at about 15% and Tyr85, Tyr94, Tyr99 and Tyr105 were phosphorylated at about 5%. Tyrosine 74 and 95 are labelled in red in the emerin sequence from figure 38 and interestingly, it corresponds to both tyrosines found as phosphorylated during a mechanical stress. Other tyrosines are labelled in orange in the emerin sequence.

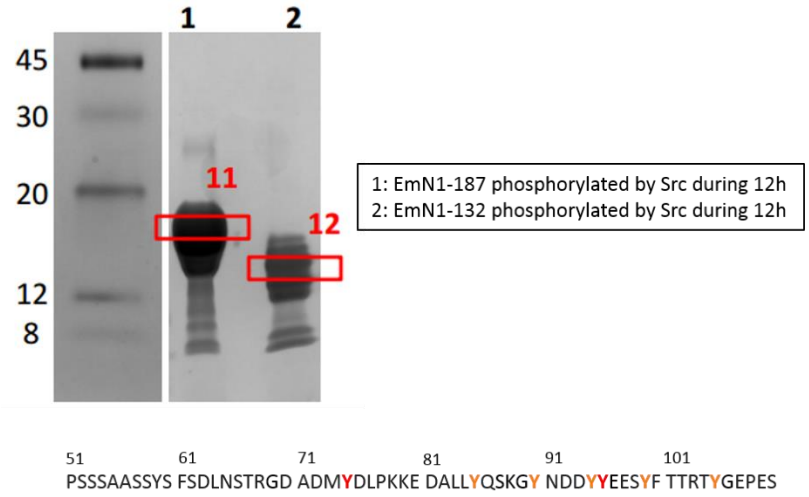


Figure 38 : SDS-PAGE gel of EmN and EmN132 phosphorylated by Src.

Samples labelled with number 11 and 12 were then analyzed by mass spectroscopy in order to identify which residues are modified by the kinase. Tyrosines found as phosphorylated are labelled on the emerin sequence (more than 20% phosphorylated: at red, less than 15% phosphorylated: orange).

To confirm the result obtained by mass spectrometry, I studied EmN phosphorylation kinetics by NMR. I recorded a ^{1}H-^{15}N HSQC spectrum on ^{15}N labelled EmN in the presence of 5mM $MgCl_2$ and 2mM ATP first, as a reference, and then I added 10µl of Src kinase in my NMR sample and recorded a spectrum every 2 hours (figure 39), using single tube. Several chemical shift modifications were observed during this kinetics. I had previously assigned the unphosphorylated EmN spectrum[250]. It was thus easy to follow the disappearance of the signal from the unphosphorylated tyrosine 95 (labelled on the figure 39), which is one of the tyrosines reported as modified by Src in the literature[107] (and shown as phosphorylated by Src in our mass spectrometry experiments). Contrariwise, it was difficult to assign all EmN modified residues because of too much spectral overlaps and changes during the kinetics. All these chemical shift differences were due to amino acid modifications and to changes in chemical environment of neighboring residues. The assignment problem is well illustrated though the observation of the peaks corresponding to the glycine residues of the protein (First Zoom on the figure 39). Indeed, many chemical shift differences were observed in this peak area, because of the presence of modified tyrosines (labelled in red and orange on the emerin sequence) next to several glycines, labelled in green on the emerin sequence (figure 39, B). The peak corresponding to residue E96 (Second Zoom on the figure 39) also totally shifts from its initial position after Src phosphorylation, probably due to phosphorylation of Y95.

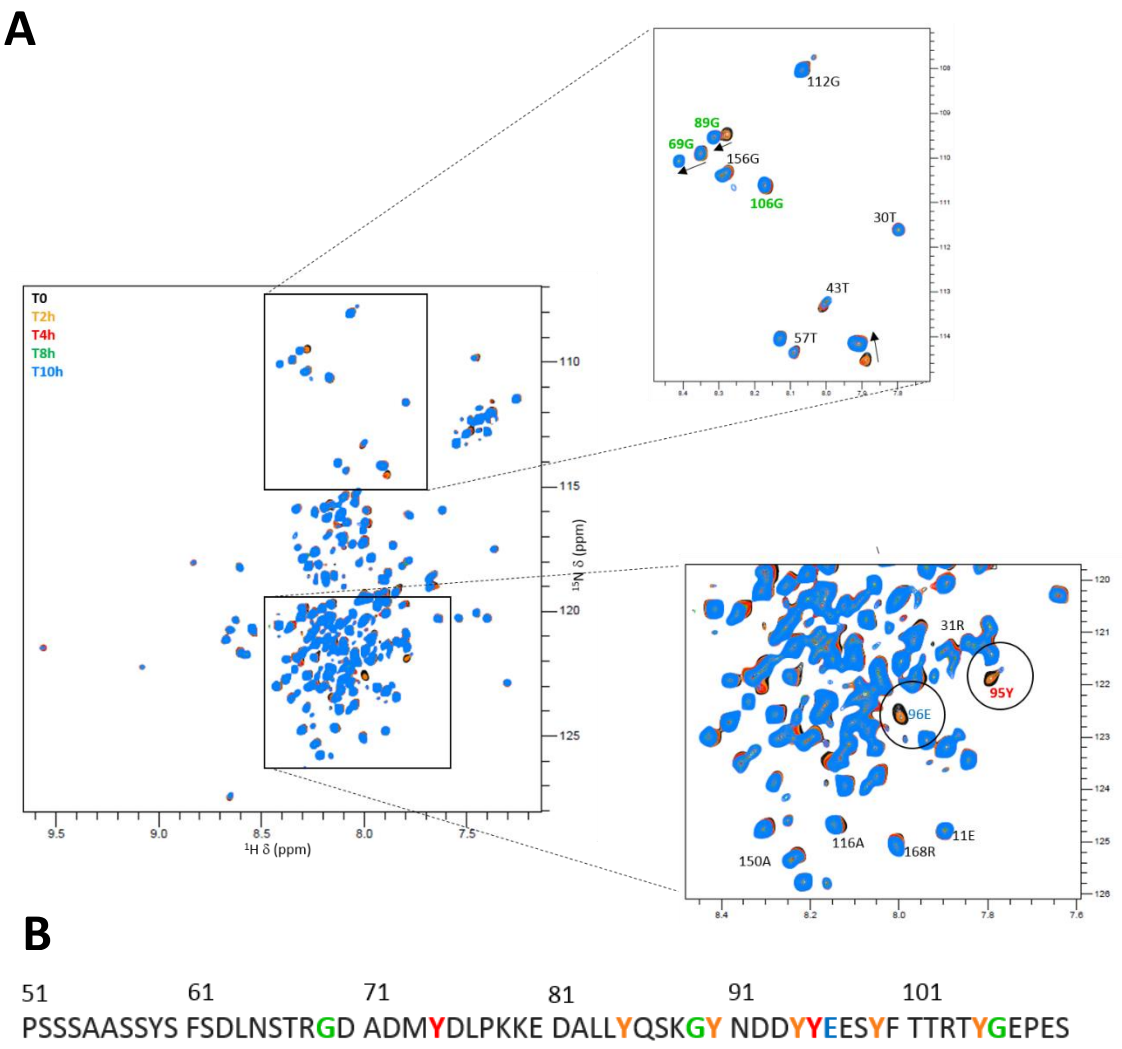


Figure 39 : EmN phosphorylation kinetics, in the presence of Src.

(A) Superimposition of 2D NMR ^{1}H–^{15}N spectra, recorded every two hours on a EmN sample at 100µM in 20mM Phosphate pH7, 30mM NaCl, 10mM β-mercaptoethanol, 2mM ATP, 5mM MgCl2 and 10µl of Src kinase, at 303K and 750MHz (FMP Berlin). (B) Emerin sequence with tyrosines identified as more than 20% phosphorylated by mass spectroscopy labelled in red, tyrosines less than 15% phosphorylated labelled in orange, glycines next to modified tyrosines labelled in green and the glutamate next to tyrosine 95 labelled in blue.

To override this problem, I studied EmN132 modifications as I did for EmN (figure 40), by NMR. As in the EmN study, several modifications appeared over time during the kinetics. Then, I produced a double labelled sample of EmN1-132 that I phosphorylated by Src kinase during 12h at 303K in order to record subsequently 3D NMR experiments to try to assign the phosphorylated sample. Because of too much signal superimpositions due to a lot of changes after phosphorylation and because the protein aggregated in phosphorylation conditions, it was not possible to assign all signals corresponding to amino acids affected by Src phosphorylation.

Only both tyrosines found as more than 20% phosphorylated by mass spectroscopy, Tyr74 and Tyr95 were assigned in their phosphorylated states. To continue on this phosphorylation study by NMR and to succeed in protein signal NMR assignment, we need to produce EmN mutants, in which each tyrosine, one per one, is mutated.

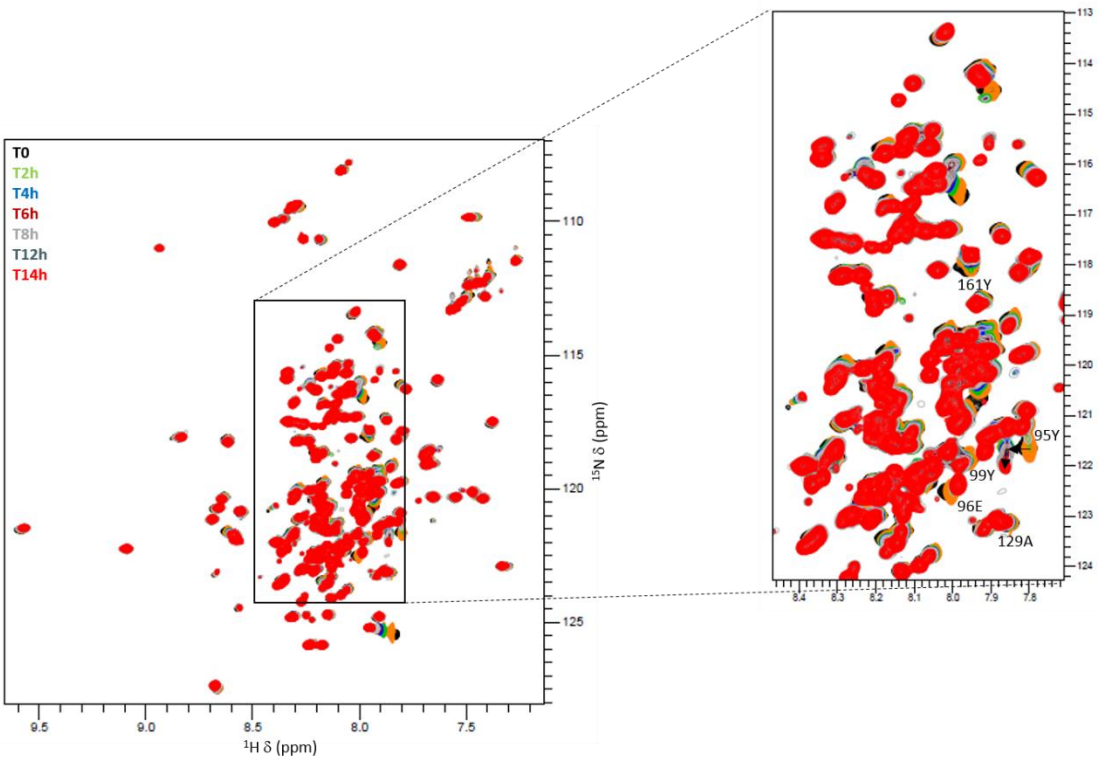


Figure 40 : Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded, every two hours, on an EmN132 sample at 100µM in 20mM Phosphate pH7, 30mM NaCl, 10mM β-mercaptoethanol, 2mM ATP, 5mM MgCl2 and 10µl of Src kinase, at 303K and 750MHz (FMP Berlin).

After trying to identify which amino acids are modified by Src, we did preliminary studies on the impact of Src phosphorylation on EmN oligomerization. We studied this by SDS-PAGE, EM and thioflavin T fluorescence. Results obtained using thioflavin T fluorescence experiments are not presented here because they were complicate to reproduce and interpret. One question here in particular is: can thioflavin T bind to phosphorylated filaments, given that it is reported as preferentially interacting with tyrosine rich cross–β surfaces (see for example Biancalana and Koide[248])? I studied EmN oligomerization in the presence of Src using SDS-PAGE, because I observed that EmN filaments could not enter the SDS-PAGE gels in the absence of β-mercaptoethanol and heating.

I first dialyzed the EmN sample in 20mM Tris-HCl pH8, 30mM NaCl and 10mM β-mercaptoethanol, and after concentrating EmN to 600µM, I prepared three different samples: one with EmN, 5mM $MgCl_2$ and 2mM ATP, one with EmN, $MgCl_2$, ATP and 10µl of Src kinase and one containing just the concentrated EmN. Afterwards, I heated each sample during one hour at 338K and I prepared SDS-PAGE samples of the supernatant and the pellet, after 10 minutes of centrifugation at 12000 g, for each condition, before and after heating at 338K (figure 41, A).

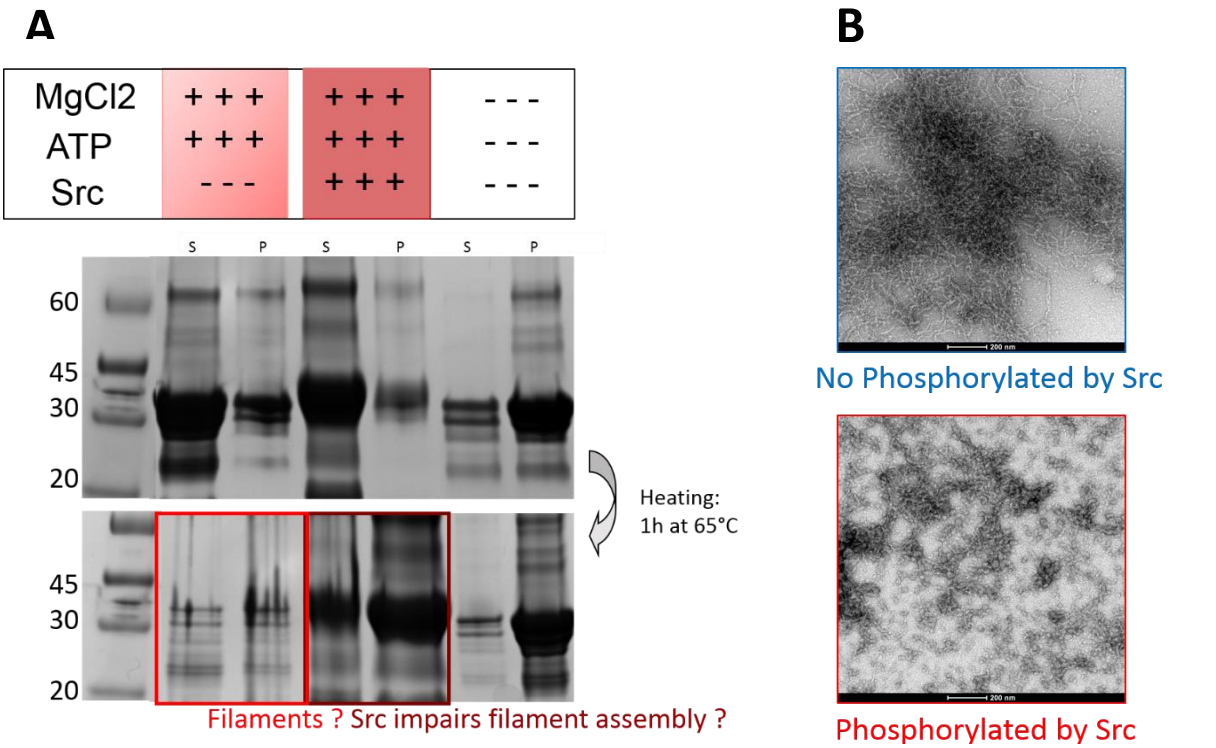


Figure 41 : Src phosphorylation impact on EmN filament assembly, as observed by SDS-PAGE and EM.

(A) Picture of SDS-PAGE gels obtained with EmN samples, phosphorylated or not by Src, in the presence or absence of $MgCl_2$ and ATP, before and after heating samples during one hour at 338K. (B) Picture obtained by negative staining EM using 0.5% of uranyl acetate on an EmN sample heated during one hour at 338K, phosphorylated or not during 15h by Src, using a JEOL MET 1400 (120KeV) (Dmytro Puchkov, FMP Berlin).

I observed the disappearance of EmN on SDS-PAGE gels only in one condition, i.e. EmN in the presence of MgCl2 and ATP after one hour of heating. In the presence of Src, it was only the solubility of the protein sample that was different before and after heating. This experiment suggested that phosphorylation by Src impairs EmN filament assembly.

To confirm this, I observed the impact of Src phosphorylation using electron microscopy (figure 41, B). First, I dialyzed EmN in 20mM phosphate pH7 and 30mM NaCl and I concentrated it until 250µM.

Then, I splitted the EmN sample in two and added 5mM MgCl2 and 2mM ATP to both of them. I added the Src kinase in one of the two samples and incubated both samples at 303K during 15h. Afterwards, samples were dialyzed against buffer with 20mM Tris-HCl pH8 and 30mM NaCl, concentrated until 600µM and put at 338K during one hour before being placed at room temperature during three days. Finally, both samples were observed by electron microscopy (figure 41, B). We observed the presence of filaments in our non-phosphorylated sample whereas aggregates were observed in the phosphorylated sample. This experiment also suggested that Src phosphorylation inhibits EmN filament formation or decreases the filament self-assembly rate. To confirm this, I have to repeat all these experiments. Secondly, I would like to try another protocol to phosphorylate EmN. Indeed, because we have to be in specific conditions to phosphorylate a protein, depending on the kinase, I would like to phosphorylate EmN before cleavage of its tag, in order to purify again EmN after phosphorylation and to see impact of this phosphorylation on filament assembly without the presence of kinase additives.

2) Emerin phosphorylation by CK1 kinase, *in vitro*

Analysis of emerin sequence using KinasePhos2[251] revealed that different serines are predicted as phosphorylated by the CK1 kinase, but two with a better confidence index: Ser120 (SVM score = 0.886429) and Ser123 (SVM score = 0.871596). These two serine are phosphorylated *in vivo*[196]. Therefore I studied by NMR phosphorylation of the EmN132 construct by CK1 (Appendix 2, figure 100). After 12h of phosphorylation, no clear chemical shift differences were observed between the unphosphorylated emerin 1-132 spectrum (in black) and the spectrum obtained after 12h of phosphorylation (in red). I concluded that CK1 could not phosphorylate EmN132 in our conditions.

3) Emerin phosphorylation by CK2 kinase, *in vitro*

With the help of the KinasPhos2 software, I found that 6 serines are predicted as phosphorylated by CK2: Ser8, 123, 141, 142, 143 and 175. All these serines are phosphorylated *in vivo*[196]. I studied by NMR phosphorylation of EmN132 by CK2 (Appendix 2, figure 101).

After superimposition of the non-phosphorylated emerin spectrum (in black) and the spectrum of emerin that was phosphorylated during 12h by CK2 (in green), only some minimal chemical shift differences were observed, so I concluded that CK2 does not seem to phosphorylate EmN132 in our conditions.

4) Characterization of EmN132 modifications in cell extracts

I did one preliminary study of EmN132 modifications in cell extracts. In more details, I produced our ^{15}N labelled EmN132 construct with a Histidine Tag at the N-terminus. Then, I concentrated the protein up to 200µM and incubated it in 293T cell extracts (10mg/ml) in the presence of 10mM ATP, 2mM $MgCl_2$, 2mM DTT and protease and phosphatase inhibitors. Finally, the protein was purified again by affinity chromatography and concentrated in 20mM Phosphate buffer pH6.5 and 30mM NaCl. NMR $^{1}H–^{15}N$ HSQC spectra were recorded on the EmN132 sample before (figure 42, in green) and after (figure 42, in black) addition of cell extracts.

Different changes were observed between the two conditions. For example, it was easy to assign peaks corresponding to Tyr4 and Tyr41 (labelled in blue) as well as surrounding residues (labelled in green) and to observe that peaks corresponding to these residues were shifted, suggesting that these LEM domain tyrosines (or their neighboring residues) were phosphorylated in cell extracts.

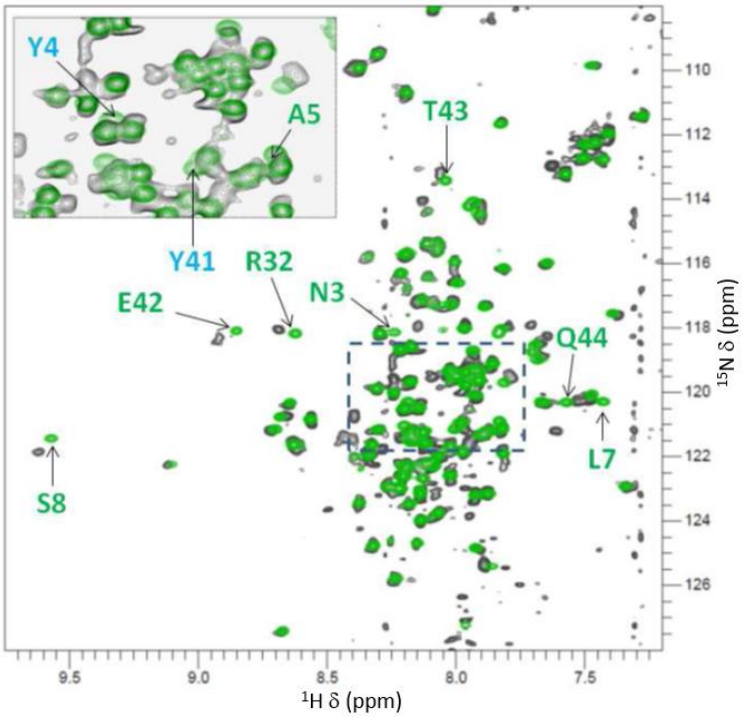


Figure 42 : Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a non-phosphorylated EmN132 sample (in black) or an EmN132 sample phosphorylated by 293T cell extracts (in green) at 100µM in 20mM Phosphate pH6.5, 30mM NaCl, 10mM β-mercaptoethanol, 2mM ATP and 5mM $MgCl_2$, at 303K and 700MHz (CEA Saclay).

II. Emerin-lamin interactions

Lamins are known to be essential for nuclear functions. On the one hand, as cytoplasmic intermediate filaments, nuclear lamins play structural functions by forming a nucleoskeletal scaffold to provide physical stability to the nucleus, to define nuclear shape and to position nuclear pore complexes[15,16]. On the other hand, lamins are found as interacting with chromatin-associated proteins, which means that lamins are implicated in higher order chromatin organization[252]. Lamins are proposed to be involved in epigenetic control mechanisms and may also take part in DNA replication.

Since several years, a lot of studies have been initiated in order to understand how lamins interact with their partners at the molecular level. However, when I arrived at the lab, nothing was known at this level. In particular, an interaction between A-type lamins and emerin was confirmed by several techniques but the molecular details of this functionally critical interaction were not known. Co-immunoprecipitation assays demonstrated that an interaction, direct or indirect, exists between lamin A and emerin[166]. Then, surface plasmon experiments showed that this interaction is direct[170]. In addition, yeast two hybrid[166] and blot overlays[245] assays confirm these results. Finally, only binding regions on both sides were characterized: the C-terminal tail of lamins, and more particular region 384-664 which contains the globular Igfold domain, was reported to bind to emerin and mutations in emerin region between residues 70 to 164 were reported to impair the interaction between A-type lamins and emerin.

Our collaborators, the team of Dr Brigitte Buendia from Diderot University in Paris, observed, using the technique of Proximity Ligation Assay (PLA) that GFP-emerin and endogenous A-type lamins were close, i.e. separated by less than 30-40 nm, in HeLa cells (figure 43).

In my laboratory, the study of a direct interaction between EmN and the lamin A tail from amino acid 389 to 646 had been performed *in vitro* by the previous PhD student, Isaline Herrada. She did not observe a direct interaction between these two protein fragments by NMR and ITC.

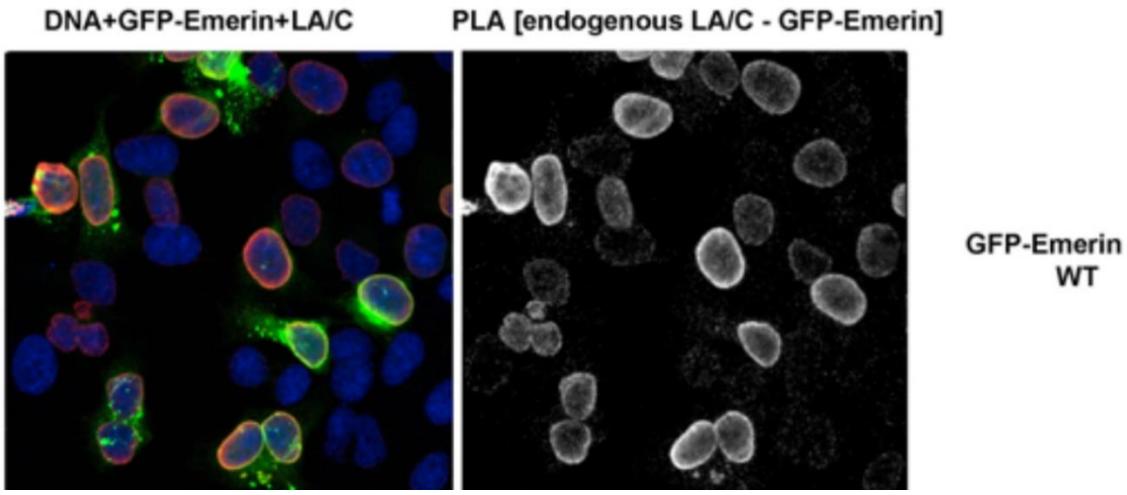


Figure 43 : Localization of emerin-lamin complexes in cells.

HeLa cells overexpressing WT GFP-emerin were fixed, labeled with anti-GFP and anti-LA/C antibodies and processed for PLA before analysis using a confocal microscope. The PLA signals are displayed either merged with DNA staining and GFP fluorescence (on the left) or alone (on the right).

At the beginning of my thesis, I tried again to observe a direct interaction between EmN and the tail of the lamin A/C. To test this interaction, different constructs of the lamin A/C tail were available in our laboratory. I worked with a construct obtained thanks to a collaboration with the laboratory of Prof. Howard J. Worman from Columbia University. It was a GST-tagged construct, with a thrombin cleavage site and coding only for the globular domain of the lamin A/C, from amino acid 411 to 566. We called this construct Igfold. Production and purification of this protein fragment were efficient (the final yield was typically 10 mg of purified protein per liter of bacterial culture) and it was important to add some reducing agents in the sample because of the presence of a solvent-exposed cysteine (Cys522).

First, I recorded the ^{1}H-^{15}N HSQC spectrum of the Igfold alone, which was ^{15}N labeled during bacterial expression, and then the ^{1}H-^{15}N HSQC spectrum of the ^{15}N labeled Igfold in presence of one equivalent of non-labeled EmN monomer (figure 44). After superimposition of both spectra, no difference was detected. We concluded that the Igfold does not directly interact with EmN, in our conditions.

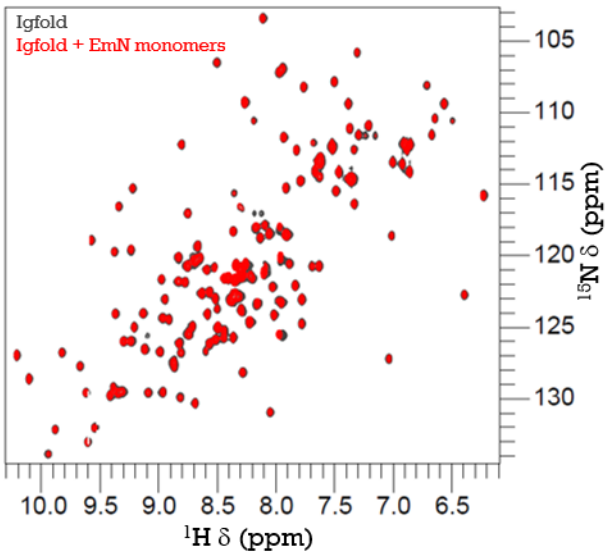


Figure 44 : Study of the EmN monomer/Igfold interaction, by NMR.

Superimposition of 2D NMR ^{1}H–^{15}N spectra recorded on an Igfold sample at 100µM alone (in grey) or in presence of 100µM of EmN monomer (in red), in 20mM Tris-HCl pH8, 30mM NaCl, at 303K and 700MHz (CEA Saclay).

1. A FIRST MODEL OF THE INTERACTION BETWEEN THE NUCLEOSKELETON AND THE CHROMATIN

Because we did not observe a direct interaction between emerin and the lamin A/C Igfold, we thought that this interaction could be mediated by a third partner. As both proteins were shown as interacting with the BAF protein independently, we decided to look at the formation of a ternary complex between emerin, BAF and the globular domain of the A-type lamin (Igfold). The formation of this complex was already suggested by several studies[55,253,254].

a. Interaction between emerin and BAF

Interaction between emerin and BAF was already well characterized and the 3D structure between the LEM domain of emerin and BAF was solved (Cf. Introduction). In more details, it was shown that the LEM domain of emerin was able to bind a dimer of BAF.

First, I tried to reproduce this interaction *in vitro*, by NMR, ITC and size exclusion chromatography. I worked with two emerin fragments: EmN and the LEM domain alone (fragment EmN49 comprising residues 1 to 49). Contrary to EmN, EmN49 was soluble in bacteria and during purification.

Concerning the BAF protein, because of a lot of oligomerization problems during purification and because BAF protein possesses 4 cysteines, which are not completely conserved between species, we chose to work at the end with a new BAF construct, with all cysteines being replaced by alanines. We called this construct BAF_{CtoA}. As for EmN, BAF_{CtoA} production was efficient but the protein was produced in inclusion bodies, requiring a purification in urea. After a first purification step, using a nickel column, the protein was renatured by dialysis. Then, BAF_{CtoA} was injected onto a gel filtration column in order to observe its oligomerization states (figure 45, A). BAF is a protein of 10kDa and just after purification, BAF_{CtoA} formed a dimer (one peak was observed at 11.8ml, which corresponds to a protein of 25kDa, on a Gel Filtration Superdex 75 10/300 GL), but we observed that the number of BAF_{CtoA} oligomers increased with time. The dimeric protein was observed on a SDS-PAGE gel and corresponded to a protein of 10kDa (figure 45, B).

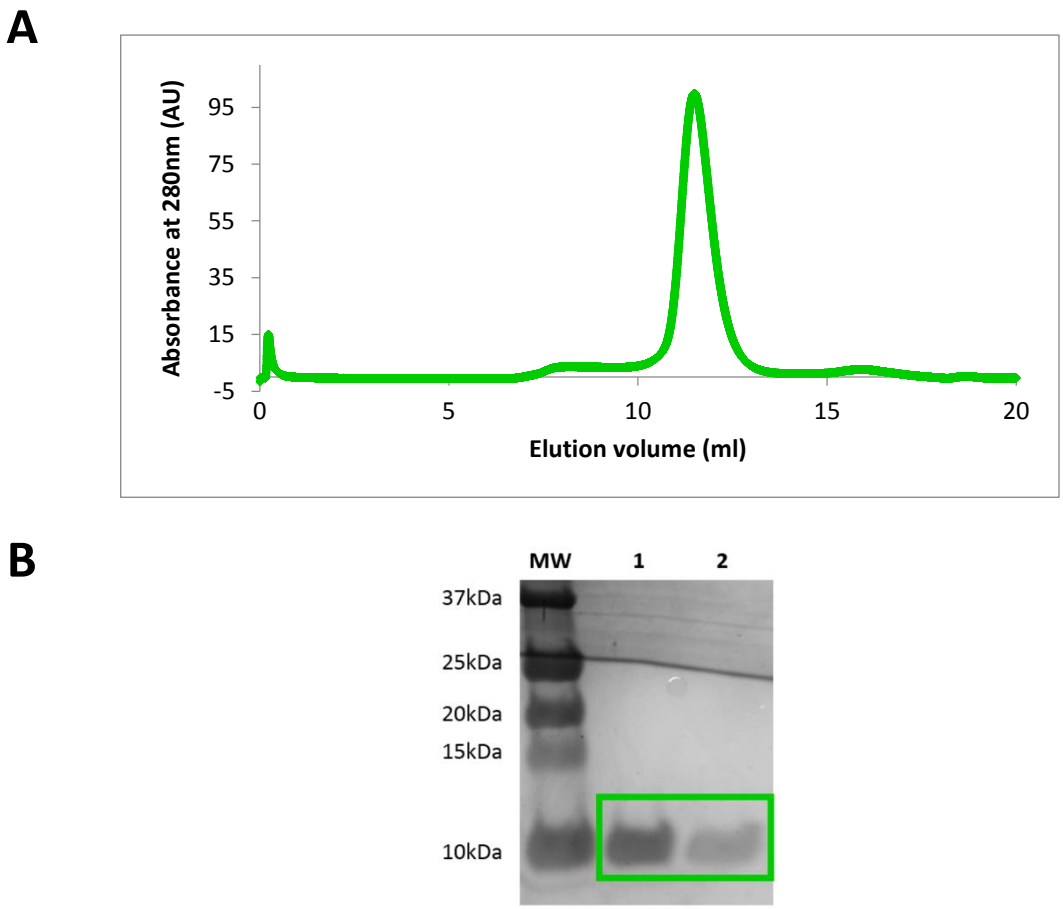


Figure 45 : BAF_{CtoA} purification.

(A) Size exclusion chromatography using a GF Superdex 75 10/300 GL, equilibrated with 20mM Tris-HCl pH8 and 100mM NaCl. (B) SDS-PAGE gel obtained after size exclusion chromatography.

After protein production, I first tried to observe an interaction between the BAF_{CtoA} dimer and EmN by NMR. To do so, I labeled each protein (BAF_{CtoA} and EmN) with ^{15}N separately and recorded a ^{1}H-^{15}N HSQC spectrum of each protein alone and with its non-labeled partner (figure 46). I observed a decrease of the signal intensity due to addition of the partner in both cases, which was characteristic of an interaction between both proteins. I thus confirmed that BAF_{CtoA} and EmN interact *in vitro* and therefore, I obtained a first proof that mutation of BAF cysteines into alanines and refolding by dialysis for both proteins (BAF_{CtoA} and EmN) did not impair the interaction. So, I estimated that using these purification protocols, I produced functional proteins.

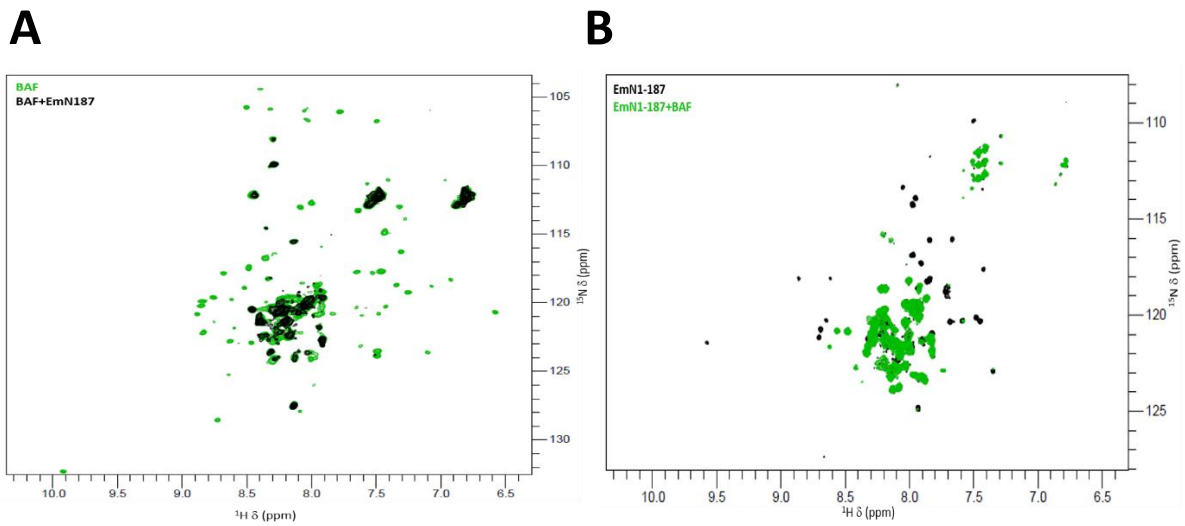


Figure 46 : Study of BAF_{CtoA} /EmN monomer interaction, by NMR.

(A) Superimposition of 2D NMR ^{1}H–^{15}N spectra recorded on a BAF_{CtoA} sample at 100µM alone (in green) or in presence of 100µM of EmN (in black), in 20mM Phosphate pH6.5, 150mM NaCl, at 293K and 700MHz (CEA Saclay). (B) Superimposition of 2D NMR ^{1}H–^{15}N spectra recorded on a EmN sample at 100µM alone (in black) or in presence of 100µM of BAF_{CtoA} (in green), in 20mM Tris-HCl pH7.5, 100mM NaCl, at 303K and 600MHz (CEA Saclay).

To confirm the affinity and stoichiometry of the interaction reported by the group of Dr Marius Clore, I studied this interaction by ITC (figure 47). Like this other group, I found a dissociation constant of around 0.5µM, a favorable enthalpic contribution to the binding free energy of about 4 kcal/mol and a defavorable entropic energy of about 10 cal/mol.K[52]. This result confirmed that mutation of BAF cysteines into alanines did not change its affinity for emerin.

However, surprisingly, my stoichiometry was different from 2 as found by Clore and co-workers. Indeed, I found that one emerin could interact with 3 to 4 BAF, depending on the experiment. But this can be explained by the fact that BAF forms oligomers, so part of my BAF protein was not able to interact, which distorted the stoichiometry.

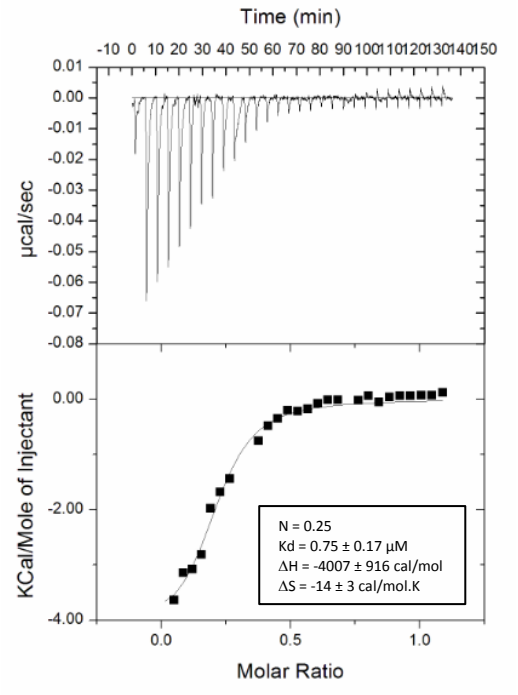


Figure 47 : Characterization of the EmN/BAF interaction by ITC.

Successive injections of 10µl of EmN concentrated at 200µM in a BAF sample concentrated at 40µM. Both proteins were dialyzed against the same buffer (20mM Tris-HCl pH8, 100mM NaCl, 10mM β-mercaptoethanol, protease inhibitors Roche). The experiment was done three times, at 288K, on a VP-ITC and we found in average a Kd of 0.75µM. All ITC values are listed in appendix 2 (Table 5).

b. Interaction between the globular domain of lamin A/C and BAF

Interaction between BAF and the globular domain of lamin A/C (Igfold) was suggested by different studies but no direct interaction was previously described. I tried to observe this interaction by NMR, ITC and size exclusion chromatography. I worked with our lamin fragment called Igfold (from residues 411 to 566) and the same BAF_{CtoA} already produced to analyze its interaction with EmN.

First, I studied this interaction by NMR. I produced the ^{15}N labeled Igfold and added non-labeled BAF_{CtoA} and I produced the ^{15}N labeled BAF_{CtoA} and added non-labeled Igfold domain (figure 48). I observed a decrease of intensity after addition of the partner in both cases.

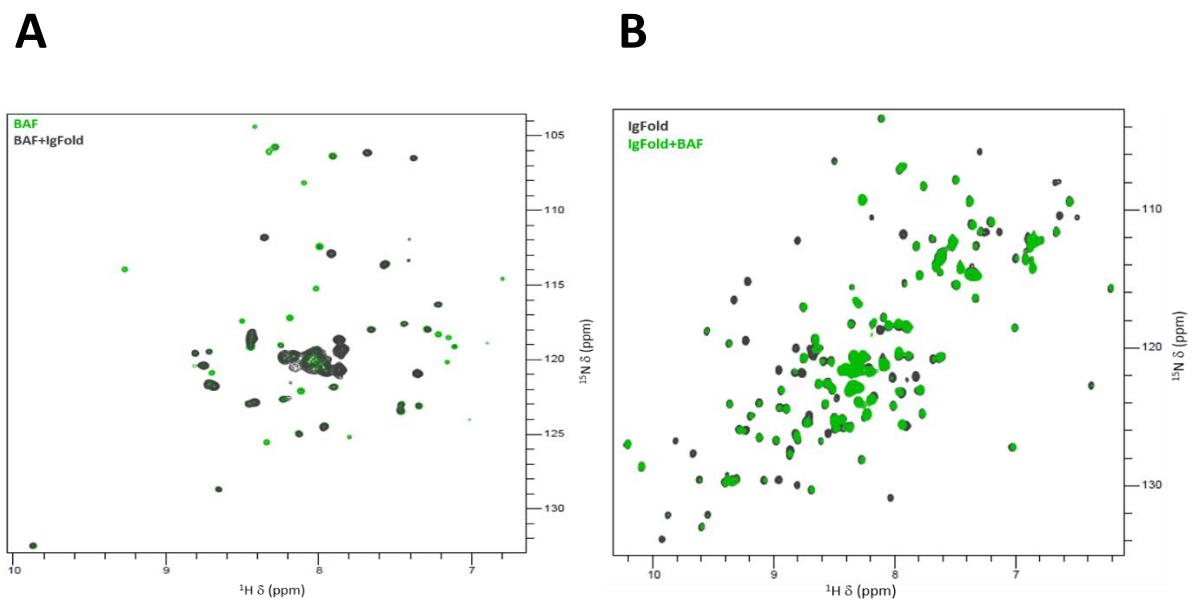


Figure 48 : Study of BAF_{CtoA} /Igfold interaction by NMR.

(A) Superimposition of 2D NMR $^{1}H-^{15}N$ spectra recorded on a BAF_{CtoA} sample at 115µM alone (in green) or in presence of 115µM of Igfold (in grey), in 20mM Phosphate pH6.5, 150mM NaCl, at 293K and 600MHz (CEA Saclay). (B) Superimposition of 2D NMR $^{1}H-^{15}N$ spectra recorded on a Igfold sample at 150µM alone (in grey) or in presence of 150µM of BAF_{CtoA} (in green), in 20mM Phosphate pH7, 100mM NaCl, at 303K and 600MHz (CEA Saclay).

To identify which part of the Igfold was implicated in this interaction, I produced a $^{15}N/^{13}C$ labeled Igfold sample and assign its NMR signals (on the basis of the published Igfold assignment)[42]. Therefore, I recorded and analyzed 3D HNCO, HNCACO, CBCACONH and HNCACB spectra. Afterwards, I quantified intensity differences for each ^{1}H-^{15}N HSQC peak corresponding to the Igfold 411-566 amino acids, before and after addition of one BAF equivalent. In more details, I did the following calculation for each peak:

$$\text{Intensity Ratio} = \frac{\text{Peak Intensity on the Igfold spectrum after addition of one BAF equivalent}}{\text{Peak Intensity on the Igfold alone spectrum}}$$

In addition, I calculated an average of the error to substrate it for each peak. To do this, I measured the average intensity of 10 dots belonging to the background noise (BNA). I did the following calculations for each peak:

$$\text{Maximum ratio value} = \frac{\text{Peak intensity on the Igfold spectrum after addition of one BAF equivalent} - \text{BNA}}{\text{Peak intensity on the Igfold alone spectrum} + \text{BNA}}$$

$$\text{Minimum ratio value} = \frac{\text{Peak intensity on the Igfold spectrum after addition of one BAF equivalent} + \text{BNA}}{\text{Peak intensity on the Igfold alone spectrum} - \text{BNA}}$$

With these calculations, I obtained the ratio of intensity between peaks corresponding to the same amino acid in the free and bound protein states and I identified peaks showing more than 50% and 70% of intensity decrease (in orange and red, respectively). Finally, I localized the residues corresponding to these peaks on the known Igfold 3D structure obtained by X-ray crystallography (PDB: 1IFR) using the same color labelling (figure 49, B).

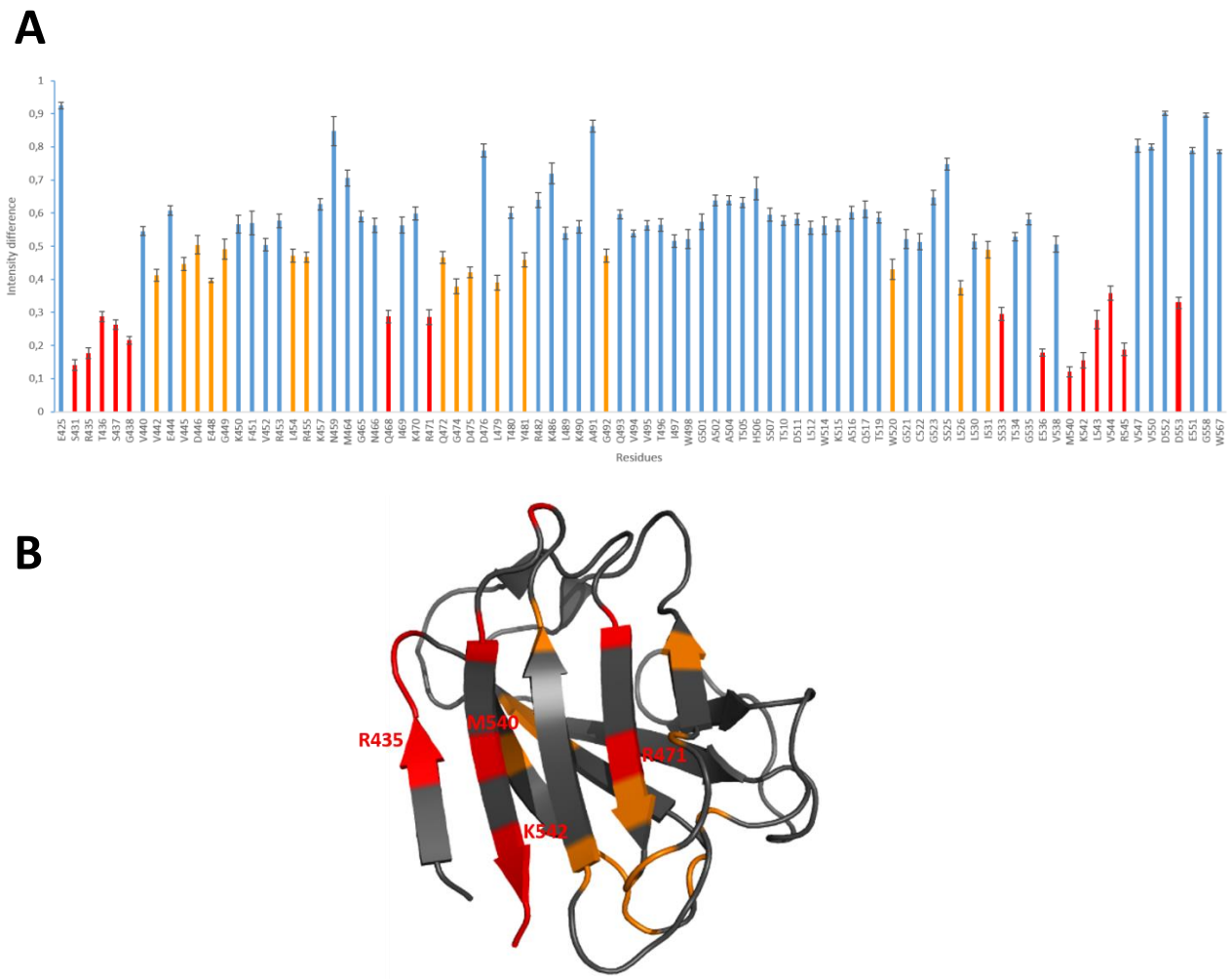


Figure 49 : Intensity ratio measurement after addition of one BAF_{CtoA} equivalent onto the ^{15}N labeled Igfold.

(A) Ratio as a function of the sequence. Bars corresponding to peaks losing more than 50% and 70% of intensity after BAF_{CtoA} addition are labeled in orange and red, respectively. (B) Three-dimensional-structure of the Igfold domain (PDB: 1IFR) with residues colored as a function of (A). Residue labelled correspond to residues that are mutated in progeroid syndromes.

Interestingly, ^{1}H-^{15}N HSQC peaks which were the most impacted by BAF_{CtoA} addition (in red) corresponded to several residues that are mutated in progeroid syndromes: R435, M540, K542 and R471. Thus, our NMR data demonstrate that BAF can directly interact with the globular domain of lamin A/C *in vitro* and suggest that mutations of the Igfold that cause progeroid syndrome impair this interaction.

I used ITC to measure the affinity and stoichiometry of the Igfold / BAF_{CtoA} interaction (figure 50). I found that the Igfold of lamin A/C can interact with BAF_{CtoA} with a micro molar affinity, (Kd = 3.2µM +/- 1.2µM at 288K, as calculated from 5 independent experiments). Concerning the stoichiometry, I found a number of sites between 0.21 and 0.43 (0.32 in average) as a function of the experiment. This again suggests that the Igfold binds to a BAF_{CtoA} dimer and that all the purified BAF_{CtoA} protein is not active.

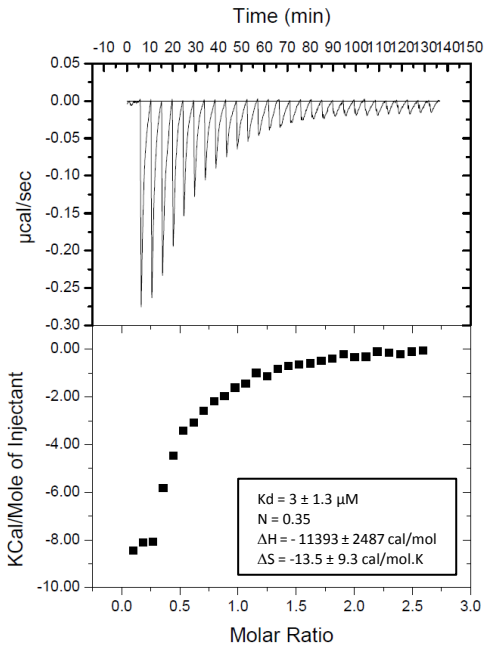


Figure 50 : Characterization of the Igfold/ BAF_{CtoA} interaction by ITC.

Successive injections of 10µl of Igfold concentrated at 270µM in a BAF_{CtoA} sample concentrated at 22.5µM. Both proteins were dialyzed against the same buffer (20mM Tris-HCl pH8, 150mM NaCl, 10mM β-mercaptoethanol, protease inhibitors Roche). The experiment was done five times, at 288K on a VP-ITC and we found in average a Kd of 3µM. All ITC values are listed in appendix 2 (Table 5).

Because B-type lamins also possess an Igfold domain, I asked if the C-terminal domain of lamin B1 could also interact with BAF. Thus, I produced the lamin B1 fragment from amino acid 409 to 586 of lamin B1 (LB1) and I tested its interaction with BAF_{CtoA} by NMR. LB1 was ^{15}N labeled (figure 51, black) and I added non-labeled BAF_{CtoA} (figure 51, green).

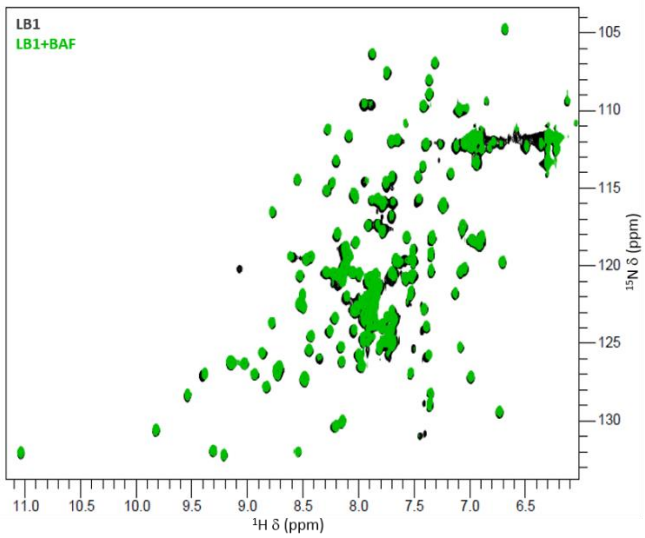


Figure 51 : Study of BAF_{CtoA} /LB1 interaction by NMR.

Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a LB1 sample at 200µM alone (in grey) or in presence of 200µM of BAF_{CtoA} (in green), in 20mM Phosphate pH7, 150mM NaCl, at 293K and 600MHz (CEA Saclay).

No intensity change or no chemical shift difference was observed after BAF_{CtoA} addition, so we concluded that LB1 cannot interact with BAF_{CtoA}, *in vitro*, in our conditions (which are the same as the conditions in which BAF_{CtoA} can interact with lamin A/C Igfold). This result is consistent with a study that showed that in cells, in presence of BAF RNAi, A-type lamins and emerin signals at the nuclear envelope were significantly weaker whereas the nuclear envelope localization of lamin B1 was unchanged[253].

c. *Existence of a ternary complex bridging nuclear membrane, nucleoskeleton and chromatin*

The existence of a ternary complex involving lamin A/C, BAF and emerin was postulated in cells[55]. I tried to observe the existence of this complex *in vitro*. To do this, I used size exclusion chromatography (figure 52) and I tested if the three protein fragments that I produced, Igfold, EmN and BAF_{CtoA}, could co-elute. In parallel, I performed this experiment with LB1.

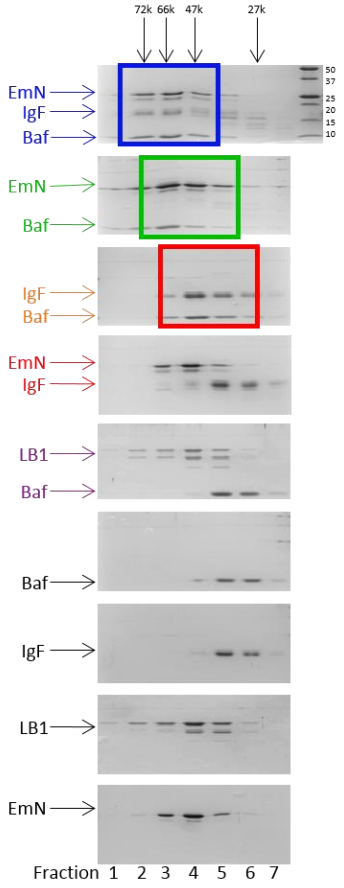


Figure 52 : Identification of a ternary complex between emerin, BAF and lamin A/C by size exclusion chromatography.

Each gel corresponds to the result of one size exclusion chromatography. Proteins were injected in a volume of 500µl and the GF column Superdex 75 10/300 GL was equilibrated in a buffer containing 20mM Tris-HCl pH8 and 30mM NaCl. Bands revealing an interaction are boxed with different colors.

Different conclusions were drawn from this study. First, we confirmed that interaction between lamin Igfold and BAF_{CtoA} is specific to lamin A/C, because an interaction was observed between lamin A/C Igfold and BAF_{CtoA} (boxed in red), whereas no interaction was observed between LB1 and BAF_{CtoA}. Furthermore, no direct interaction was observed between the Igfold of lamin A/C and emerin monomers. This result confirmed what we had already observed by NMR. Then, interaction between BAF_{CtoA} and EmN was confirmed once again (boxed in green) and the most important was the observation of a ternary complex between EmN/BAF_{CtoA}/Igfold (boxed in blue). Indeed, after injection of the three proteins pooled together at a concentration of 150µM, a predominant pic was observed that corresponded to a complex of approximatively 60kDa, which is in accordance with the weight of a complex between a dimer of BAF_{CtoA} (20kDa) in interaction with one EmN (25kDa) and one Igfold (15kDa).

After observation of a ternary complex formation between BAF_{CtoA}, EmN and Igfold by gel filtration, we wanted to confirm that this ternary complex exists in solution and that this complex was composed of two BAF, one Igfold and one EmN, as it was supposed by ITC results. For this reason, we analyzed our complexes by analytical ultracentrifugation. In addition, because BAF was shown to interact with EmN49 and not with another part of emerin, we decided to look at the presence of EmN49 only and not EmN, in the ternary complex. Analysis was done by Dr Christophe Velours (CNRS, LEBS, Gif-sur-Yvette), using an analytical ultracentrifuge XL470 (Beckman Coulter, Palo Alto, USA) and an An-50Ti rotor. I produced several samples: BAF_{CtoA} samples at 32, 57 and 86µM, Igfold samples at 17, 26 and 40µM, a BAF_{CtoA}/Igfold sample, with a 2/1 ratio and an EmN49/BAF_{CtoA}/Igfold sample, with a 1/2/1 ratio. All proteins were first dialyzed in the same buffer: 50mM Tris-HCl pH8, 100mM NaCl. Then, sedimentation velocity experiments were performed on these different samples, with an optical path of 1.2cm, at 42000 rpm (128 297g), at 293K. Both complex samples were observed at an OD of 0.5, 0.8 and 1.2. Absorbance and interference were measured at 280nm, every 8 minutes. 400µl of sample were used for each experiment and 410µl of our dialysis buffer was used as a reference. At the end, results were analyzed using the Sedfit software[255].

Because the same results were obtained for the different protein concentrations, I will only present results obtained with 32µM of BAF_{CtoA}, 17µM of Igfold and both complexes (Igfold/ BAF_{CtoA} and LEM/ BAF_{CtoA}/Igfold) at an OD of 1.2 (figure 53, Table A).

After analyzing the figures 53 and table A, we observed that the BAF_{CtoA}/Igfold sample contains a majority of trimers (2.8S) in solution, which could correspond to 2 BAF_{CtoA} and one Igfold. In more details, resulting trimers possessed an elongated shape with an f/f0 value of 1.4.

Unfortunately, analysis of the EmN49/BAF_{CtoA}/Igfold complex did not give more information than analysis of the BAF_{CtoA}/Igfold complex, because no real modifications were observed after addition of EmN49. Indeed, same species that the one observed for the BAF_{CtoA}/Igfold sample were observed, with majority of trimers.

Regarding the results obtained for the BAF_{CtoA}/Igfold sample by analytical ultracentrifugation, we concluded that only one ternary complex seems to exist in solution and is composed, at least, of two BAF_{CtoA} and one Igfold.

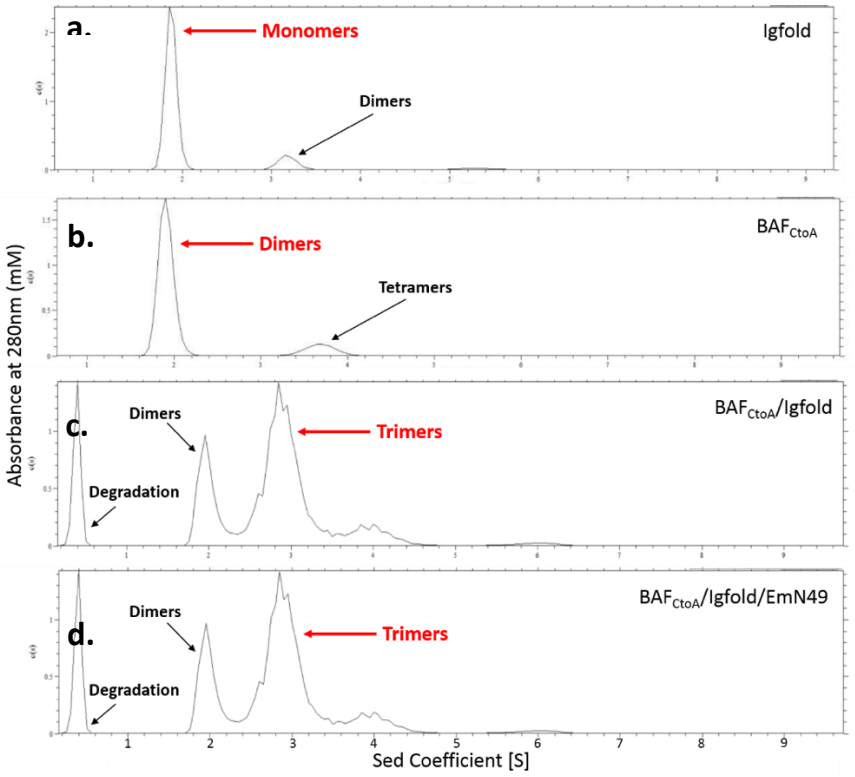


Figure 53 : Analytical ultracentrifugation results obtained using absorbance at 280nm, with a XL470 ultracentrifuge (Beckman Coulter, Palo Alto, USA) and an An-50Ti rotor.

(a) Sedimentation velocity profile of a BAF_{CtoA} sample, in 50mM Tris-HCl pH8, 100mM NaCl, injected at 32µM, at 293K. (b) Sedimentation velocity profile of an Igfold sample, in 50mM Tris-HCl pH8, 100mM NaCl, injected at 17µM, at 293K. (c) Sedimentation velocity profile of a BAF_{CtoA}/Igfold sample, in 50mM Tris-HCl pH8, 100mM NaCl, injected at an OD of 1.2, at 293K. (d) Sedimentation velocity profile of an EmN49/BAF_{CtoA}/Igfold sample, in 50mM Tris-HCl pH8, 100mM NaCl, injected at an OD of 1.2, at 293K. Each corresponding species are indicated by arrows; in red are represented the majority species.

	Degradation	Monomers	Dimers	Trimers	Tetramers	Higher MW oligomers/ Aggregates
Igfold alone	-	**85%**	12%	-	3%	2%
BAF_{CtoA} alone	-	-	**87%**	-	11%	2%
BAF_{CtoA}/Igfold	8%	-	25%	**54%**	11%	2%
EmN49/BAF_{CtoA}/ Igfold	14%	-	20%	**55%**	10%	1%

Table 1 : Analytical ultracentrifugation results obtained using absorbance at 280nm, for four different protein samples (BAF_{CtoA}, Igfold, BAF_{CtoA}/Igfold and EmN49/BAF_{CtoA}/Igfold).

Percentage values correspond to the total of absorbance. In red is represented the percentage of the majority species for each sample; (-) no species detected.

In parallel, we tried to obtain the structure of the ternary complex, using crystallography. To achieve this goal, after size exclusion chromatography, I pooled fractions corresponding to the complex peak (fractions 2 to 4 on the figure 52) and I concentrated this sample until 3mg/ml. Then, I conserved the concentrated sample during one week at 277K. I sent one part of my complex sample to the European crystallogenesis facility in Grenoble (HTX laboratory), in order to screen a lot of conditions to get first crystals. Six different primary crystallogenesis sets, used for initial screening, were tested on my complex sample, at 277K (Classics Suite and PEGs-I from Qiagen/Nextal, JCSG+ and PACT *premier* from Molecular Dimensions, Salt Grid from Hampton Research and Wizard I&II from Rigaku Reagents). After one week, I obtained microcrystals in two conditions (0.2M NH4Cl and 16% of PEG3.350 pH8; 0.1M MgCL2 and 16% of PEG3.350 pH8; 0.2M of NH4SO4, 20% of PEG3.350) and crystals in one condition (100mM of Bis-Tris phosphate pH5.5).

I tried to obtain similar crystals of this sample, in our laboratory, using the hanging-drop during vapor diffusion step. I prepared one crystallization screen with buffers related to conditions which gave microcrystals (listed on the table B) and one related to condition in which I got crystals (listed on the table C). After doing a crystallization screen at 277K, I got crystals in three different conditions.

The first condition contained 0.2M NH4Cl and 14% of PEG3.350 (condition 1), the second condition contained 0.1M MgCL2 and 20% of PEG3.350 (condition 2) and the last condition, which was relatively different, contained 0.2M of NH4SO4, 23% of PEG3.350 and 100mM of Bis-Tris phosphate pH5,5 (condition 3). A picture of crystals obtained with the last condition is inserted below (figure 54).

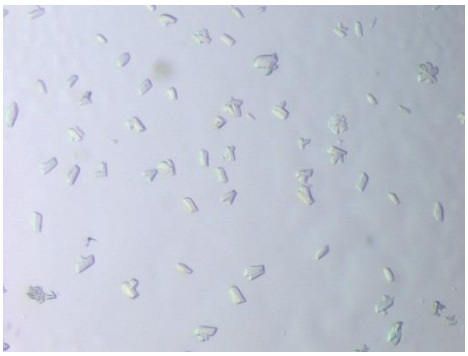


Figure 54 : Picture of crystals obtained from the purified EmN/BAF/Igfold sample, at 277K, using condition 3.

	1	2	3	4	5	6
A	PEG 3.350 12% TrispH8 100mM NH4CL 0.2M	PEG 3.350 14% TrispH8 100mM NH4CL 0.2M	PEG 3.350 16% TrispH8 100mM NH4CL 0.2M	PEG 3.350 18% TrispH8 100mM NH4CL 0.2M	PEG 3.350 20% TrispH8 100mM NH4CL 0.2M	PEG 3.350 22% TrispH8 100mM NH4CL 0.2M
B	PEG 3.350 18% TrispH8 100mM NH4CL 0.1M	PEG 3.350 20% TrispH8 100mM NH4CL 0.1M	PEG 3.350 22% TrispH8 100mM NH4CL 0.1M	PEG 3.350 18% TrispH8 100mM NH4CL 0.3M	PEG 3.350 20% TrispH8 100mM NH4CL 0.3M	PEG 3.350 22% TrispH8 100mM NH4CL 0.3M
C	PEG 3.350 12% TrispH8 100mM MgCl2 0.2M	PEG 3.350 14% TrispH8 100mM MgCl2 0.2M	PEG 3.350 16% TrispH8 100mM MgCl2 0.2M	PEG 3.350 18% TrispH8 100mM MgCl2 0.2M	PEG 3.350 20% TrispH8 100mM MgCl2 0.2M	PEG 3.350 22% TrispH8 100mM MgCl2 0.2M
D	PEG 3.350 18% TrispH8 100mM MgCl2 0.1M	PEG 3.350 20% TrispH8 100mM MgCl2 0.1M	PEG 3.350 22% TrispH8 100mM MgCl2 0.1M	PEG 3.350 18% TrispH8 100mM MgCl2 0.3M	PEG 3.350 20% TrispH8 100mM MgCl2 0.3M	PEG 3.350 22% TrispH8 100mM MgCl2 0.3M

Table 2 : Crystallization screen designed from conditions which led to microcrystals formation in HTX laboratory.

Conditions for which we got microcrystals are surrounded by a red rectangle.

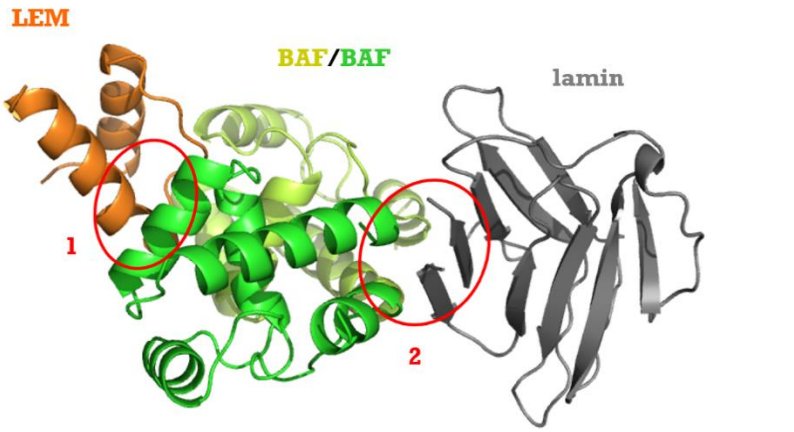
	1	2	3	4	5	6
A	PEG 3.350 23% Bis-Tris 5.5 100mM NH4SO4 0.2M	PEG 3.350 25% Bis-Tris 5.5 100mM NH4SO4 0.2M	PEG 3.350 27% Bis-Tris 5.5 100mM NH4SO4 0.2M	PEG 3.350 23% Bis-Tris 5.5 100mM NH4SO4 0.1M	PEG 3.350 25% Bis-Tris 5.5 100mM NH4SO4 0.1M	PEG 3.350 27% Bis-Tris 5.5 100mM NH4SO4 0.1M
B	PEG 3.350 23% Bis-Tris 5.5 100mM Li2SO4 0.2M	PEG 3.350 25% Bis-Tris 5.5 100mM Li2SO4 0.2M	PEG 3.350 27% Bis-Tris 5.5 100mM Li2SO4 0.2M	PEG 3.350 23% Bis-Tris 5.5 100mM Li2SO4 0.1M	PEG 3.350 25% Bis-Tris 5.5 100mM Li2SO4 0.1M	PEG 3.350 27% Bis-Tris 5.5 100mM Li2SO4 0.1M

Table 3 : Crystallization screen designed from conditions which led to crystals formation in HTX laboratory. The condition for which we got crystals is surrounded by a red rectangle.

After shooting with X-ray on these crystals, on the PROXIMA1 beamline of synchrotron Soleil (with the help of Dr Virginie Ropars), we obtained a good diffraction pattern for crystals obtained from condition 3, with a resolution of 2.1 Angstrom (Å). By using two servers, Molrep and Phenix, and starting from structures already available for each component of the complex, I solved the 3D structure of the ternary complex. As starting structures, I used the X-ray structure of the Igfold of the lamin A/C (PDB: 1IFR), the NMR structure of the LEM domain (PDB: 2ODC) and the NMR structure of the BAF dimer in interaction with the LEM of emerin (PDB: 2ODG). Interestingly, in our crystals, only the LEM domain of emerin was present, which means that after one week at 277K, all the unfolded nucleoplasmic part of emerin was degraded in my size exclusion chromatography sample and only the LEM domain, which is the important region for binding to BAF, was conserved.

One ternary complex was present per asymmetric unit of the crystal (figure 55). I analyzed interfaces in my model, using the software PISA[256], and found an interface of 594 $Å^2$ between the LEM domain and BAF and an interface of 465 $Å^2$ between BAF and the Igfold domain.

A



B

Data collection	
Space group	P 1
Cell dimensions	
a, b, c (Å)	57.77 62.75 64.34
a, b, γ (°)	66.44 66.55 88.05
Resolution (Å)	49.152 – 2.1
Anisotropy direction	
Resolution where $CC_{1/2} > 0.3$	2.1
R_{meas} (%)	15.9
I/s(I)	5.9
$CC_{1/2}$	99.1
Completeness (%)	95.67
Redundancy	3.52
B Wilson (Å²)	34.74
Refinement	
No. reflections	83937
R_{work} / R_{free}	0.1879/0.2444
No. atoms	5666
Protein	675
DNA	0
SO_4^{2-}	8
Water	279
B factors	43.38
Protein	42.97
Ligand/ion	80.03
Water	45.37
R.m.s. deviations	1.953
Bond lengths (Å)	0.008
Bond angles (°)	1.16
One crystal per structure.	

Figure 55 : X-ray crystallography structure of the ternary complex formed between the LEM domain of the emerin (EmN49), the BAF_{CtoA} dimer and the Igfold domain.

(A) 3D structure of the ternary complex with the LEM (EmN49) in orange, BAF_{CtoA} dimer in green and the Igfold in grey. (B) Data collection and refinement statistics.

Concerning the interface between BAF_{CtoA} and the Igfold, my X-ray structure shows, consistently with my NMR results, that the β-sheet of the Igfold implicated in the interaction is the same as the one exhibiting the majority of mutations leading to progeroid syndromes.

Concerning the interface between BAF_{CtoA} and EmN49, we superimposed our X-ray structure with the structure obtained by NMR by the group of Dr Marius Clore, in Bethesda (figure 56) and first, we measured the rmsd values between each corresponding chain (named on the figure 56) of the complex, taken alone. We found a rmsd value of 1.3 Å between chain C and G (which correspond to the LEM domain), a rmsd value of 1.42 Å between chain A and D (which correspond to the first monomer of BAF) and a rmsd value of 0.9 Å between chain B and E (which correspond to the second monomer of BAF). Then, the rmsd value was measured between the two BAF chains of the complex with the LEM domain fixed and we found a value of 1.3 Å. On the contrary, the rmsd value was measured between the two LEM domains, with the two BAF chains fixed and we found a value of 2.75 Å. This result showed that the interaction is significantly tighter in the crystal structure. We concluded that the presence of the Igfold in the complex did not change the structure of the LEM domain in interaction with BAF (and that the 3D structure of BAF_{CtoA} was highly similar to that of wild-type BAF).

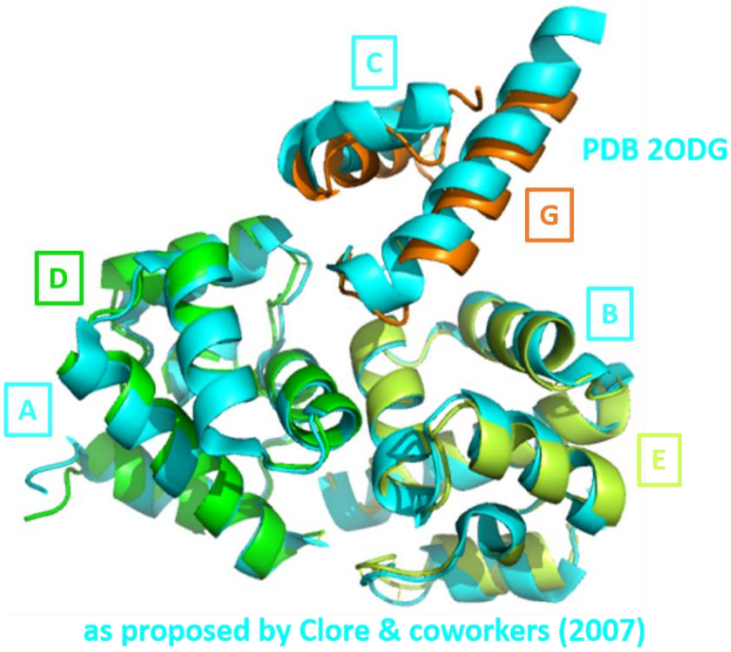


Figure 56 : Superimposition of our 3D structure of EmN49 in interaction with the dimer of BAF_{CtoA}, obtained by X-ray crystallography, with the 3D structure of the same complex obtained by NMR, by the group of Dr Marius Clore in 2007 (PDB: 20DG)[52].

Finally, we superimposed the 3D structure of our ternary complex with the 3D structure that was already solved by X-ray crystallography between BAF and DNA (Cf. Introduction) (figure 57). Our model suggests that BAF is able to simultaneously bind to lamin A/C, emerin and DNA and provides a first model of the interaction between the nucleoskeleton and DNA.

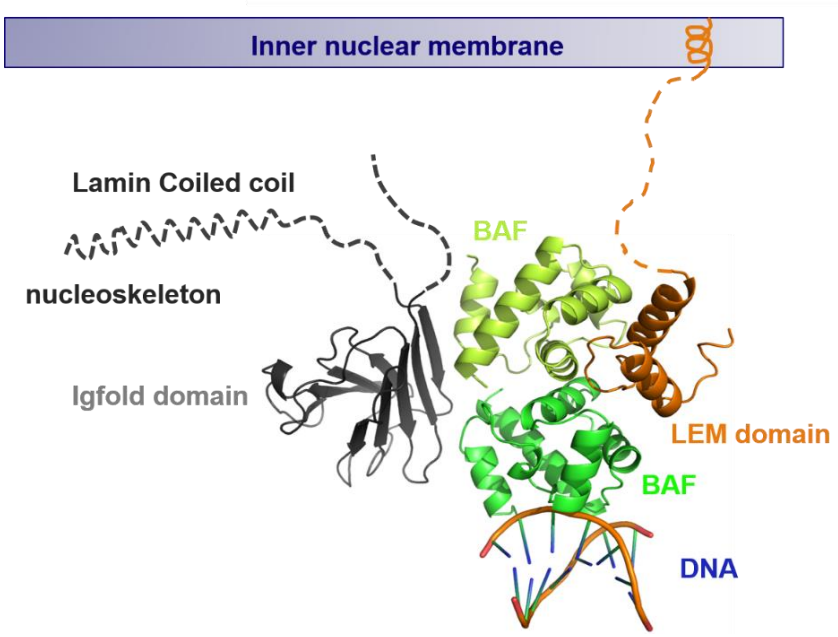


Figure 57 : A first model of the interaction between the nucleoskeleton and the chromatin.

This model results from the superimposition of our X-ray structure with that of BAF in interaction with DNA (PDB: 2BZF). The BAF dimer binds to two DNA molecules but I represented here only one DNA for simplicity.

To conclude this part, as it was shown in the literature, I confirmed the presence of an interaction between BAF and the LEM domain of emerin and between BAF and the Igfold domain of lamin A/C. Moreover, I demonstrated that this second interaction is specific to lamin A/C, because no interaction was observed, through two different techniques, between BAF_{CtoA} and LB1. Finally, I revealed the existence of a ternary complex between EmN49, Igfold and BAF_{CtoA} dimer and I solved the 3D structure of this complex.

d. Several mutants associated with progeroid syndromes disrupt the interaction between lamin A/C and BAF

We produced a set of Igfold mutants associated with progeroid syndromes. These mutants are listed in table D: homozygous mutant R435C causes a restrictive dermopathy[230], homozygous mutant K542N[238] causes HGPS, homozygous mutants R471C[203] and R527H[233] are associated with HGPS and/or a severe form of mandibuloacral dysplasia and homozygous mutants A529T[236] and A529V[257] cause mandibuloacral dysplasia. Localization of the mutated residues is represented on figure 58. Mutants were cloned with a mutagenesis kit (Stratagene QuikChange) based on our WT Igfold 411-566 expression vector by Ambre Petitalot in the lab.

Production and purification of each mutant was performed using the same protocol as for the WT Igfold 411-566.

Igfold Mutants	Diseases	Publications
R435C	Restrictive Dermopathy	Madej-Pilarczyk A et al., 2009
K542N	HGPS	M Plasilova et al., 2004
R471C	HGPS and/or a severe form of Mandibuloacral Dysplasia	Birgit Zirn et al., 2008
R527H	HGPS and/or a severe form of Mandibuloacral Dysplasia	Novelli G et al., 2002
A529T	Mandibuloacral Dysplasia	Kosho T et al., 2007
A529V	Mandibuloacral Dysplasia	Garg A et al., 2005

Table 4 : Igfold mutants causing different, more or less severe, progeroid syndromes and their associated disease and publication.

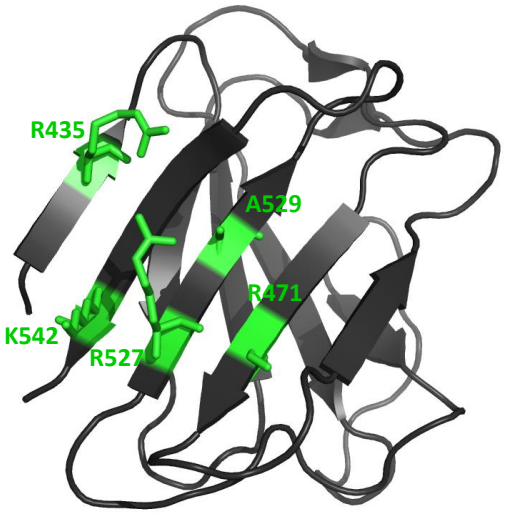


Figure 58 : Localization of the different Igfold residues mutated in progeroid syndromes (in green) (PDB: 1IFR)[43].

First, we checked by NMR if the five mutants R435C, K542N, R471C, R527H and A529V are well-folded. After recording ^{1}H-^{15}N HSQC spectrum of these mutants, we confirmed that they are well-folded (figure 59), because we obtained a well-dispersed spectrum for each of them as for the Igfold WT.

Then, we compared the stability of the Igfold WT with that of the five mutated Igfold (R435C, K542N, R471C, A529V and A529T) based on a Fluorescence-based Thermal Shift Assay performed by the group of Dr Eric Jacquet (ICSN, Gif-sur-Yvette).

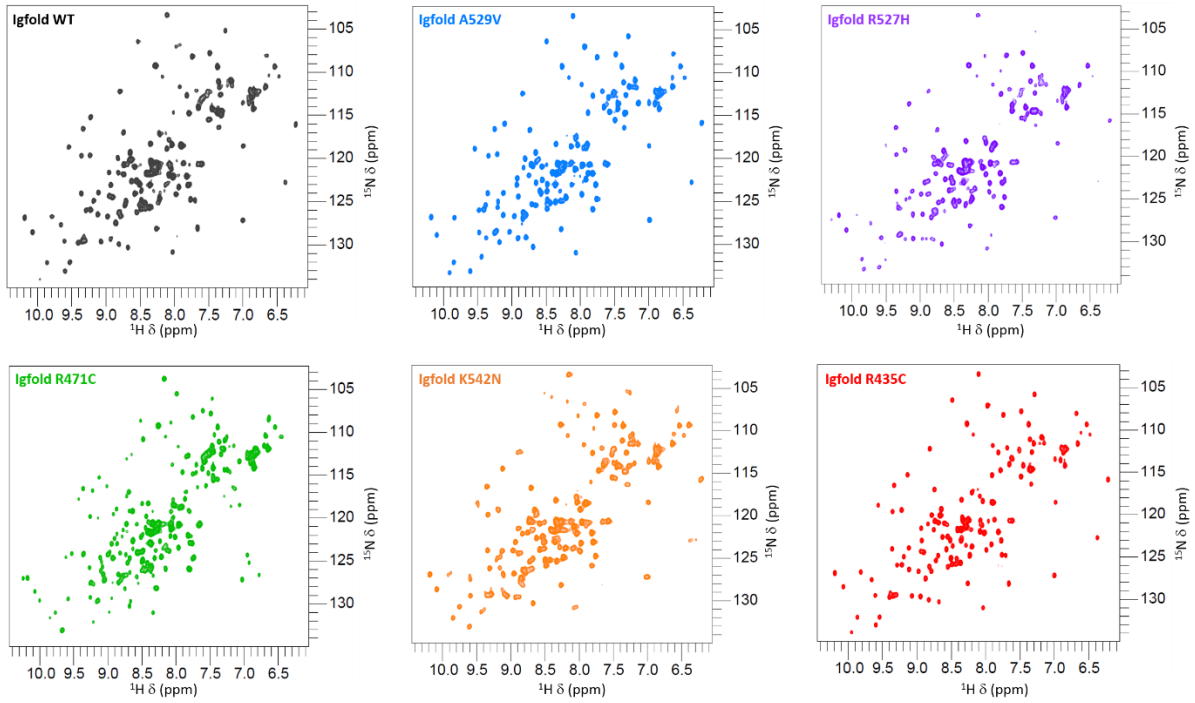


Figure 59 : Folding study of different Igfold mutant causing progeroid syndrome, by NMR.

^{1}H-^{15}N HSQC spectra recorded on Igfold WT sample at 200µM (in dark grey), on Igfold A529V sample at 200µM (in blue), on Igfold R527H sample at 200µM (in purple), on Igfold R471C sample at 200µM (in green), on Igfold K542N sample at 200µM (in orange) and on Igfold R435C sample at 200µM (in red). All proteins were dialyzed in 20mM Tris-HCl pH8, 30mM NaCl, 2mM DTT and spectra were recorded at 303K and 700MHz spectrometer (CEA Saclay).

First, the group of Dr Eric Jacquet compared the stability of the Igfold WT in several buffers, in order to observe the stability difference between pH6 and 8.5 (figure 60). Interestingly, they found that Igfold WT shows a first unfolding event around 323K at a pH between 7 and 8.5.

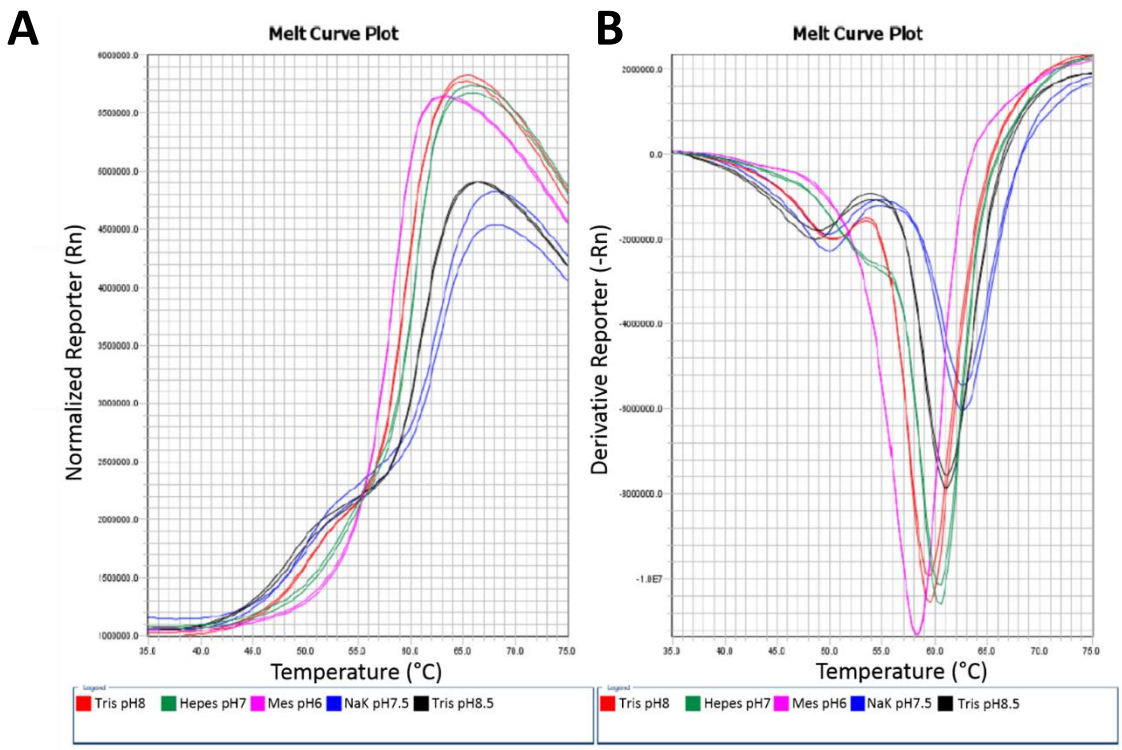


Figure 60 : Stability test of Igfold WT in different buffers.

(A) Profiles of denaturation curves. (B) Derivatives of denaturation curves. Denaturation tests were done, according to the temperature, in different buffers: Tris-HCl pH8 and NaCl 100mM (in red), Hepes 50mM pH7 and NaCl 100mM (in green), MES 50mM pH6 and NaCl 100mM (in pink), NaKP 50mM pH7.5 and NaCl 100mM (in blue) and Tris-HCl pH8.5 and NaCl 100mM (in black).

Then, the stability of the Igfold WT was compared to the stability of the 5 mutated Igfold causing progeroid syndromes (R435C, K542N, R471C, A529T and A529V) (figure 61). Stability was studied in buffer A, composed of 50mM Mes pH6 and 100mM NaCl and 10µl of each protein were used for the different measurements and diluted two times with buffer A, according to the experiment (figure 61). The tested mutants related to progeroid diseases do not show a large modification of their thermal stability. Indeed, only R471C shows a 4 degree decrease of its thermal stability relatively to WT Igfold.

Knowing that my mutants are well folded and stable, I could go further and observe the impact of the different mutations on the formation of a complex between the dimer of BAF_{CtoA} and Igfold. First, I used the technique of ITC and worked with the same 5 mutants (R435C, K542N, R471C, R527H and A529V) (figure 62). I measured the affinity and stoichiometry of the Igfold mutants / BAF_{CtoA} interactions.

A

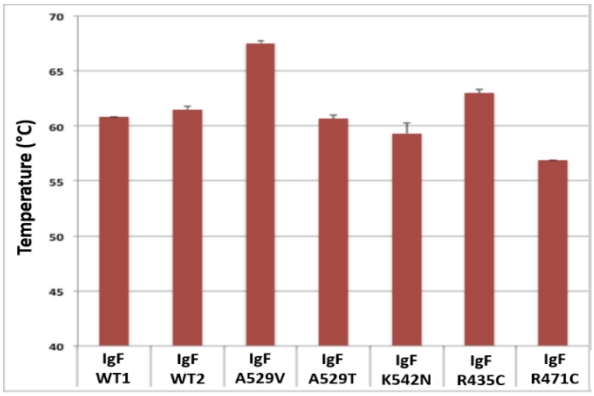


Figure 61 : Stability study of Igfold WT and mutants, using a fluorescence-based thermal shift assay.

(A) Graphical representation of Igfold WT and mutant thermal stability in buffer A (50mM Mes pH6, 100mM NaCl). (B) Table of Tm and standard error values obtained for the different Igfold WT and mutants in buffer A.

B

Protein	Tm (°C)	Standard error (AU)
Igfold WT 1	60,8	0,0
Igfold WT 2	61,5	0,3
Igfold A529V	67,5	0,2
Igfold A529T	60,7	0,3
Igfold K542N	59,3	0 ,9
Igfold R435C	63,0	0,3
Igfold R471C	56,9	0,1

In conditions in which one Igfold WT can interact with a dimer of BAF_{CtoA} with a micro molar affinity (Kd = 2µM, see Table 6 in Appendix 2), we observed that two mutants (R471C and A529V, in blue on the figure 62) were still able to bind a dimer of BAF_{CtoA}, but with a lower affinity (and a decreased heat signal). Then, no affinity could be measured for two mutants (K542N and R527H, in purple on the figure 62) and no binding could be detected for one mutant (R435C, in red on the figure 62). Interestingly, the most pathogenic mutations are those which have the strongest impact on the interaction with BAF_{CtoA}.

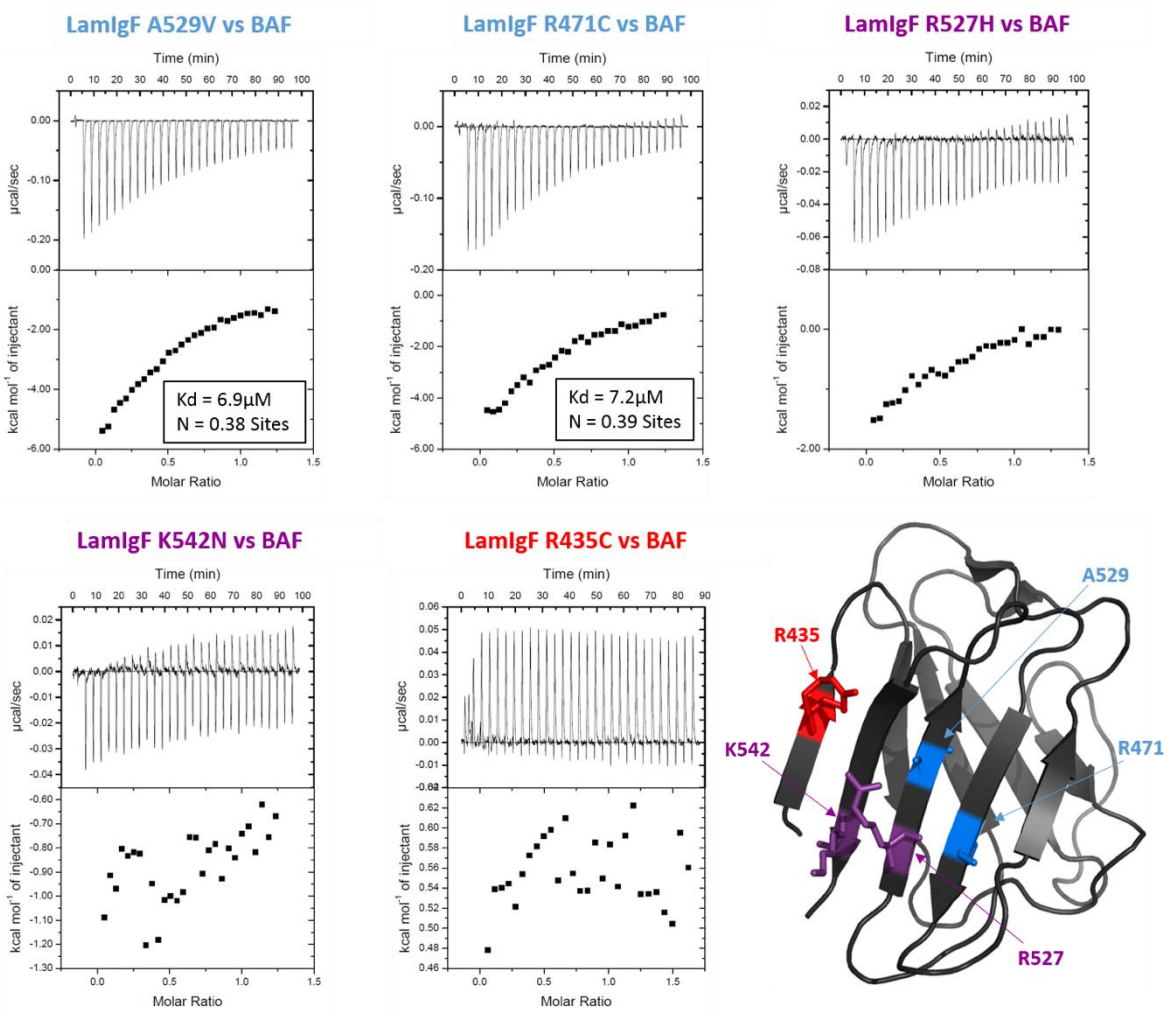


Figure 62 : Characterization of the impact of 5 Igfold mutants on the interaction with BAF_{CtoA}, by ITC.

(A) Results of ITC experiments. For each experiments, successive injections of 10µl of Igfold concentrated at 120µM in a BAF_{CtoA} sample concentrated at 20µM. Both proteins were dialyzed against the same buffer (20mM Tris-HCl pH8, 150mM NaCl, 10mM β-mercaptoethanol, protease inhibitors Roche). Each experiment was done two times, at 283K, on a VP-ITC. (B) Igfold structure (PDB: 1IVT)[42] with mutants which decreased affinity of the interaction represented in blue and mutants which strongly impact the interaction represented in red. All ITC values are listed in appendix 2 (Table 6).

To confirm these results, we used size-exclusion chromatography and observed which of these five mutants could still bind to BAF_{CtoA} (figure 63). A decrease of the affinity was observed with four mutants (K542N, R471C, R527H and A529V, in blue on the figure 63) and no binding was observed with the last mutant (R435C, in red on the figure 63).

We concluded that, in vitro, the more severe is the syndrome induced by the mutant, the more decreased is the affinity of the interaction between the mutated Igfold and BAF_{CtoA}.

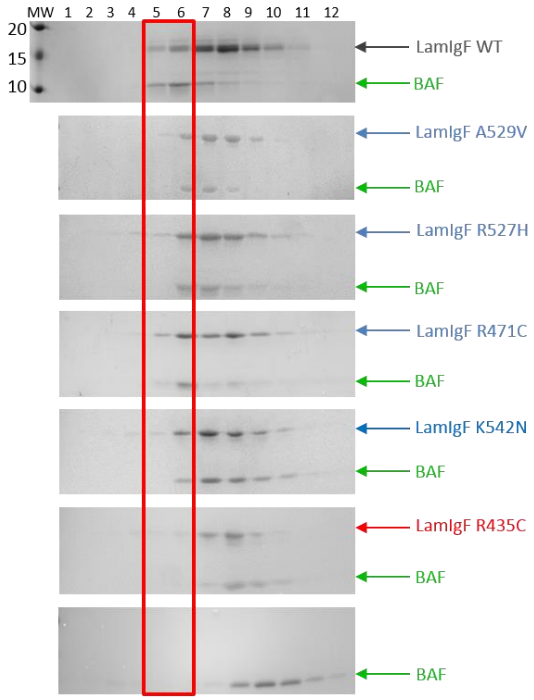


Figure 63 : Observation of the impact of 5 Igfold mutants on the interaction with BAF_{CtoA}, by size exclusion chromatography.

Each gel corresponds to the result of one size exclusion chromatography. Proteins were injected in a volume of 500µl and GF column Superdex 75 10/300 GL was equilibrated in a buffer containing 20mM Tris-HCl pH8 and 150mM NaCl. Igfold WT is represented in dark grey and BAF in green. Mutants which decreased affinity of the interaction are represented in blue and mutant which strongly impact the interaction is represented in red.

In parallel, we expressed and purified a lamin A/C fragment corresponding to the tail of the mutant progerin, which causes a well-studied form of dominant HGPS. This mutant progerin, as it was explained in the introduction, is due to a mutation that triggers an alternative splicing event leading to a deletion of 50 residues including the proteolytic cleavage site important for lamin A/C maturation. We expressed a large fragment from amino acid 389 to amino acid 614, corresponding to the whole progerin tail.

To simplify, I will call this fragment IgfoldProgerin. My idea was to determine if the presence of region 567 to 614 (absent in the Igfold fragment) was impairing BAF binding.

Purification of IgfoldProgerin was performed as described for lamin A/C Igfold. Then I studied the interaction between IgfoldProgerin and the dimer of BAF_{CtoA}. I produced a ^{15}N labeled BAF_{CtoA} protein and added the same equivalent of non-labeled IgfoldProgerin and on the other hand, I produced a ^{15}N IgfoldProgerin and added the same equivalent of non-labeled BAF_{CtoA} dimer (figure 64).

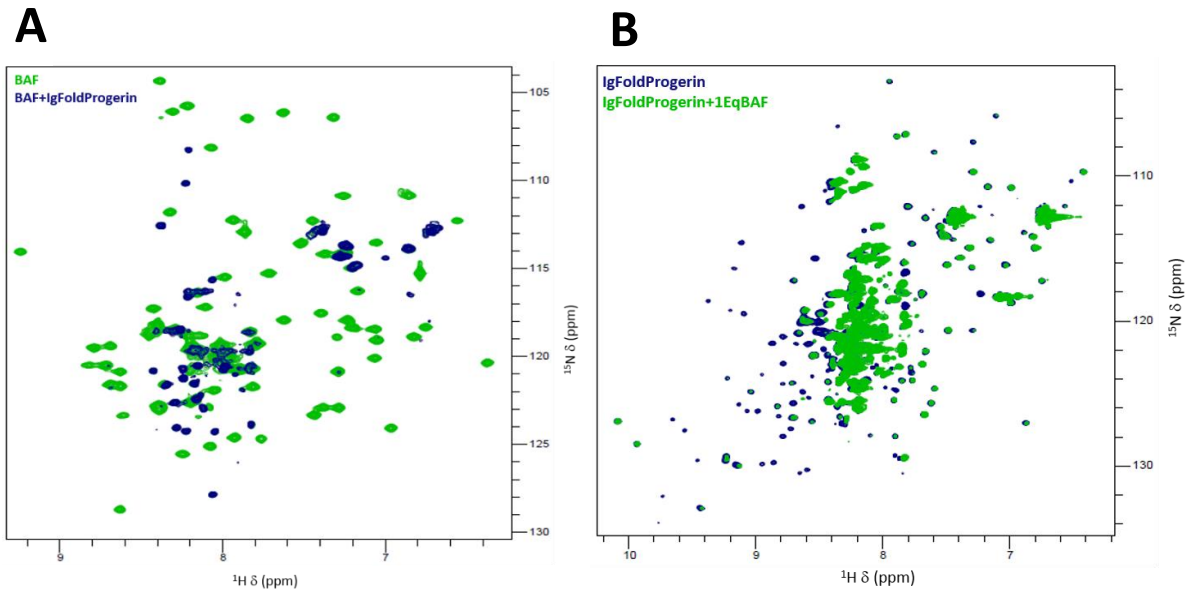


Figure 64 : Study of BAF_{CtoA} /IgfoldProgerin interaction, by NMR.

(A) Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a BAF_{CtoA} sample at 100µM alone (in green) or in presence of 100µM of IgfoldProgerin (in navy), in 20mM Phosphate pH6.5, 150mM NaCl, 1mM DTT, at 283K and 600MHz (CEA Saclay). (B) Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a IgfoldProgerin sample at 75µM alone (in navy) or in presence of 150µM of BAF_{CtoA} (in green), in 20mM Phosphate pH6.5, 150mM NaCl, 1mM DTT, at 283K and 600MHz (CEA Saclay).

In both cases, I observed an decrease of intensity for several peaks. Based on my Igfold assignment, I did the same analysis as for the interaction between the ^{15}N labeled Igfold 411-566 and the BAF_{CtoA} dimer and I observed which NMR signals were the most impacted by the interaction with BAF_{CtoA}.

I compared the amino acids found at the interaction surface in both cases (figure 65). I found the same interaction surface than the one obtained in presence of Igfold alone.

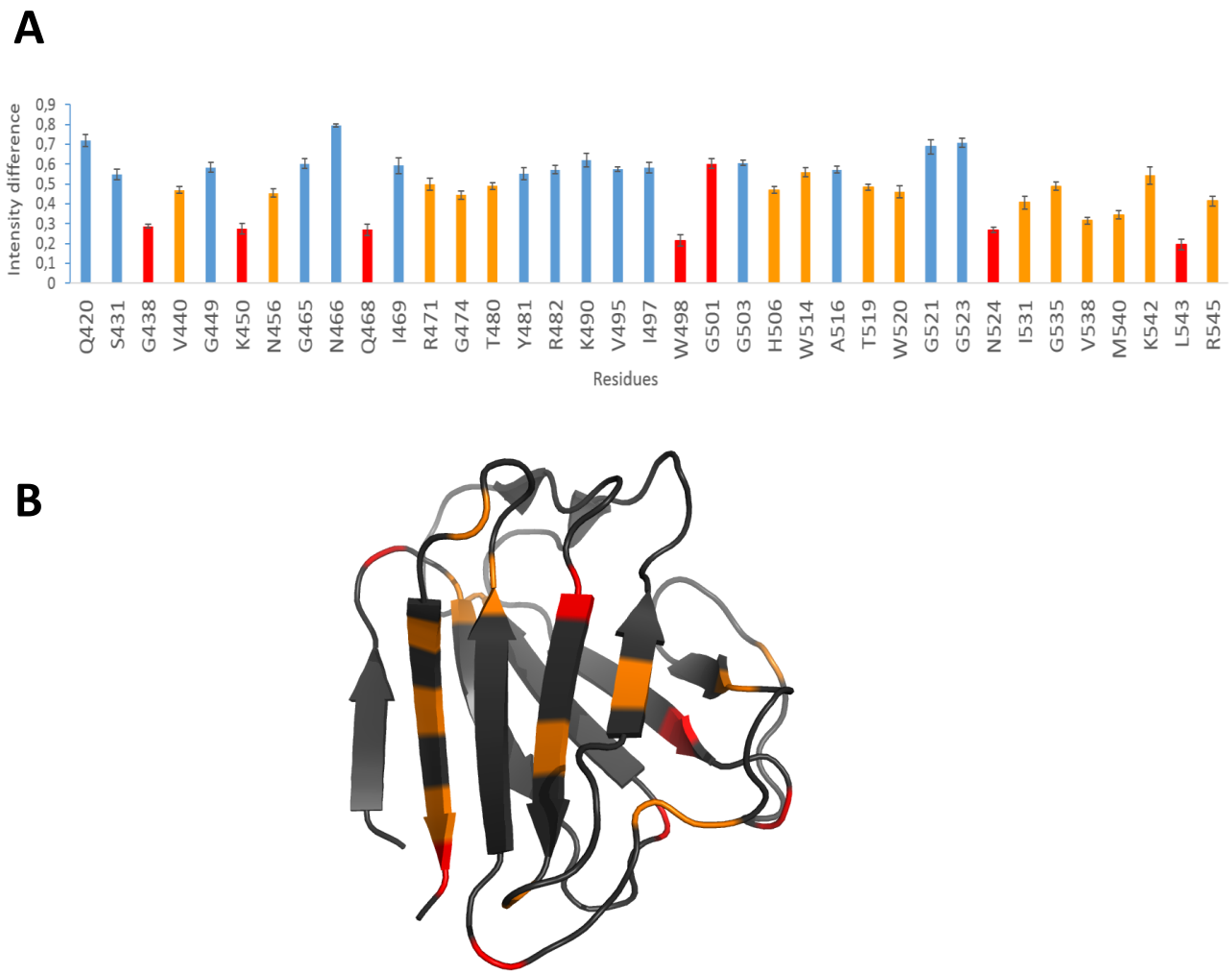


Figure 65 : Intensity ratio measurement after addition of one BAF_{CtoA} equivalent onto the ^{15}N labeled IgfoldProgerin.

(A) Ratio as a function of the sequence. Bars corresponding to peaks losing more than 50% and 70% of intensity after BAF_{CtoA} addition are labeled in orange and red, respectively. (B) Three-dimensional-structure of the Igfold domain (PDB: 1IFR) with residues colored as a function of (A).

To measure the affinity between IgfoldProgerin and BAF_{CtoA}, I studied this interaction by ITC (figure 66). Concerning the stoichiometry, like for other ITC studies with BAF_{CtoA}, I did not found a number of sites of 0.5, but of 0.22, certainly because of some inactive BAF_{CtoA} in our protein sample. Concerning the affinity, I found the same range of values as found for the interaction between BAF_{CtoA} and Igfold. On the other side, a compensation between enthalpic and entropic distribution was observed after comparison of both experiments (Interaction between BAF and Igfold or BAF and IgfoldProgerin).

Indeed, an enthalpic contribution to the binding free energy of about 11000 cal/mol and an entropic energy of about 13 cal/mol.K was found in presence of the Igfold whereas an enthalpic contribution of about 16000 cal/mol and an entropic energy of about 31 cal/mol.K was found in presence of the IgfoldProgerin, which suggests that additional contacts exist in the case of IgfoldProgerin.

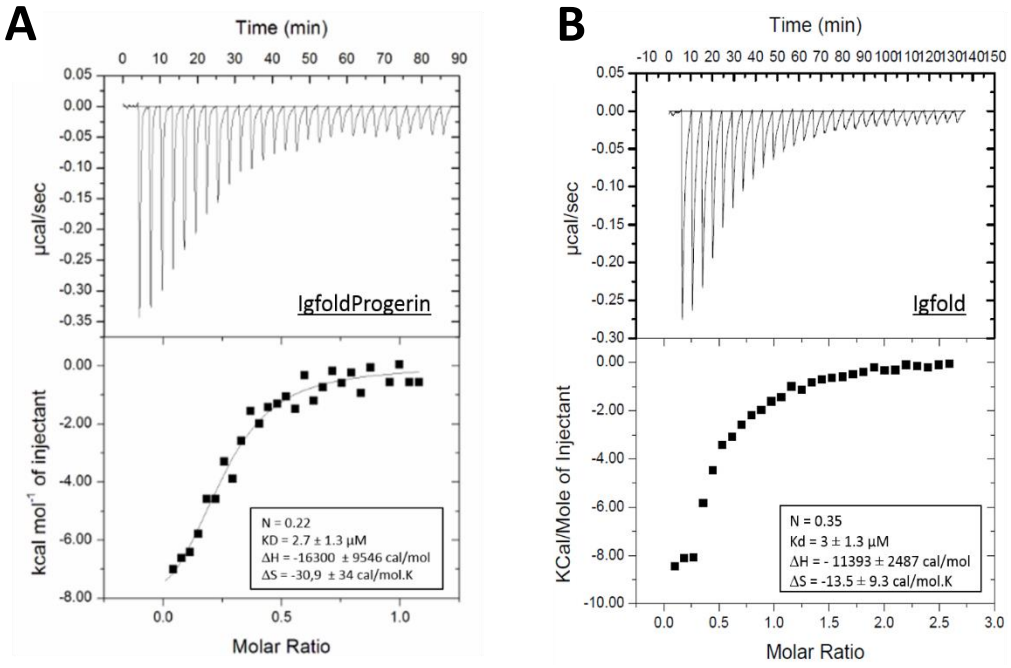


Figure 66 : Characterization of IgfoldProgerin/ BAF_{CtoA} interaction by ITC.

(A) Successive injections of 10µl of IgfoldProgerin concentrated at 120µM in a BAF_{CtoA} sample concentrated at 24µM. A Kd of 2.7 µM was found in average. (B) Successive injections of 10µl of Igfold concentrated at 270µMin a BAF_{CtoA} sample concentrated at 22.5µM. A Kd of 3.0 µM was found in average. All proteins were dialyzed against the same buffer (20mM Tris-HCl pH8, 150mM NaCl, 10mM β-mercaptoethanol, protease inhibitors Roche). Experiments were done at 288K, on a VP-ITC. All ITC values are listed in appendix 2 (Table 5).

e. Mutants of the emerin LEM domain associated to cardiac disease still binds to BAF

Because the emerin LEM domain is necessary for emerin interaction with BAF, we hypothesized that mutations in the LEM domain, and especially ΔK37 that impacts the LEM 3D structure, impair binding of EmN to BAF.

To study this, I teamed up with Nada Essawy, a PhD student in the lab, and we first used NMR and size exclusion chromatography to determine the impact of the mutations on the EmN affinity for BAF.

For the NMR experiments, we produced ^{15}N labeled BAF_{CtoA} protein and non-labeled emerin mutants. Then, I recorded an ^{1}H-^{15}N HSQC spectrum of BAF_{CtoA} alone (in green, on figure 67) and spectra of BAF_{CtoA} in contact with each LEM domain mutant, added in a 1:1 ratio (figure 67, A for P22L, B for ΔK37 and C for T43I). For the three mutants, I observed that the majority of the peaks, which correspond to the amino acids of BAF_{CtoA}, lost a significant fraction of their intensity. We concluded that these three mutants seem to interact with BAF_{CtoA}. I was surprised that mutant ΔK37 could still interact, because as it was explained in the first part of my thesis results, this mutant possesses a destabilized LEM domain. Thus, we did two new hypotheses: first, this mutant could still interact but with a lower affinity than EmN WT, or this mutant could still interact by forming a new complex structure compared to EmN WT in interaction with BAF_{CtoA}.

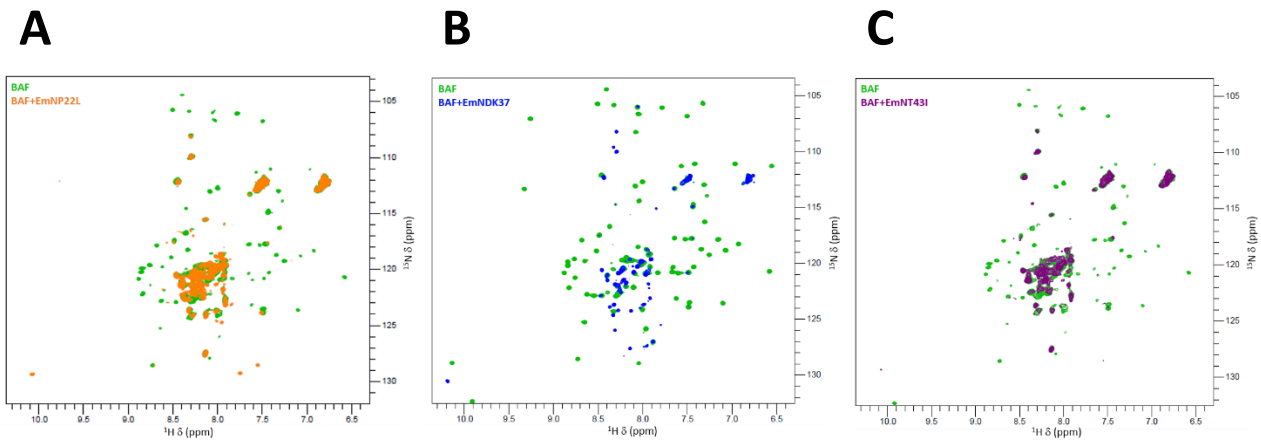


Figure 67 : Study of BAF_{CtoA}/EmN LEM domain mutant interaction, by NMR.

(A) Superimposition of 2D NMR ^{1}H–^{15}N spectra recorded on a BAF_{CtoA} sample at 150µM alone (in green) or in presence of 150µM of P22L (in orange), in 20mM Phosphate pH6.5, 30mM NaCl, at 303K and 700MHz (CEA Saclay). (B) Superimposition of 2D NMR ^{1}H–^{15}N spectra recorded on a BAF_{CtoA} sample at 150µM alone (in green) or in presence of 150µM of ΔK37 (in blue), in 20mM Phosphate pH6.5, 30mM NaCl, at 303K and 700MHz (CEA Saclay). (C) Superimposition of 2D NMR ^{1}H–^{15}N spectra recorded on a BAF_{CtoA} sample at 150µM alone (in green) or in presence of 150µM of T43I (in purple), in 20mM Phosphate pH6.5, 30mM NaCl, at 303K and 700MHz (CEA Saclay).

We confirmed our NMR results using size exclusion chromatography. Figure 68 shows that the three EmN mutants still coelute with BAF_{CtoA}, demonstrating that indeed these mutants interact with BAF_{CtoA}.

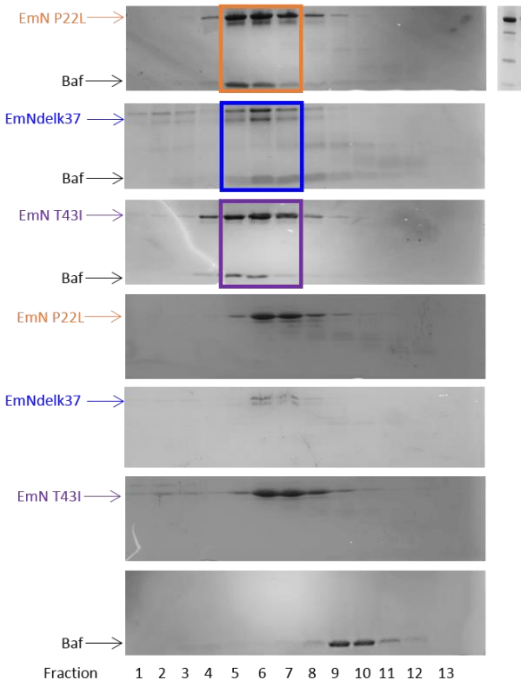


Figure 68 : Observation of the impact of 3 emerin LEM domain mutations on the interaction with BAF, by size exclusion chromatography.

Each gel corresponds to the result of one size exclusion chromatography. Proteins were injected in a volume of 500µl and the GF column Superdex 75 10/300 GL was equilibrated in a buffer containing 20mM Tris-HCl pH8 and 30mM NaCl.

To see if these different mutations have an impact on the 3D structure of the EmN/ BAF_{CtoA}/Igfold complex, we did the same size exclusion chromatography experiment as presented in figure 68, but adding the Igfold to the EmN/BAF_{CtoA} complex, and we kept the elution fractions corresponding to the ternary complex formed by the EmN mutant, BAF_{CtoA} and the Igfold, in order to concentrate them and try to obtain crystals.

We tried to crystallize the different complexes using conditions in which we obtained crystals with the WT complex. Thus, we obtained crystals of the T43I/BAF/Igfold complex and the P22L/BAF_{CtoA}/Igfold complex. After shooting with X-ray on these crystals, on the PROXIMA1 beamline of synchrotron Soleil (with the help of Virginie Ropars), we obtained a good diffraction pattern for the complex containing T43I, corresponding to a resolution of 2.3Å, whereas we did not obtain any diffraction pattern for the complex containing P22L, because of twinned crystals.

Concerning the complex containing the mutation T43I, I solved its 3D structure, using two servers, Molrep and Phenix, and starting from the structure of the WT ternary complex that I obtained before. After superimposition of this structure with the one obtained with EmN WT, I found that mutation T43I did not affect the structure of the ternary complex. Indeed, as it is shown on figure 69, the 3D structure of the complex containing emerin WT is superimposable to the one containing emerin T43I and after superimposition of both complexes (EmN49 WT/BAF_{CtoA}/Igfold and EmN49 T43I/ BAF_{CtoA}/Igfold), the rmsd value measured between the Calpha carbons of the 3 proteins was 0.55 Å.

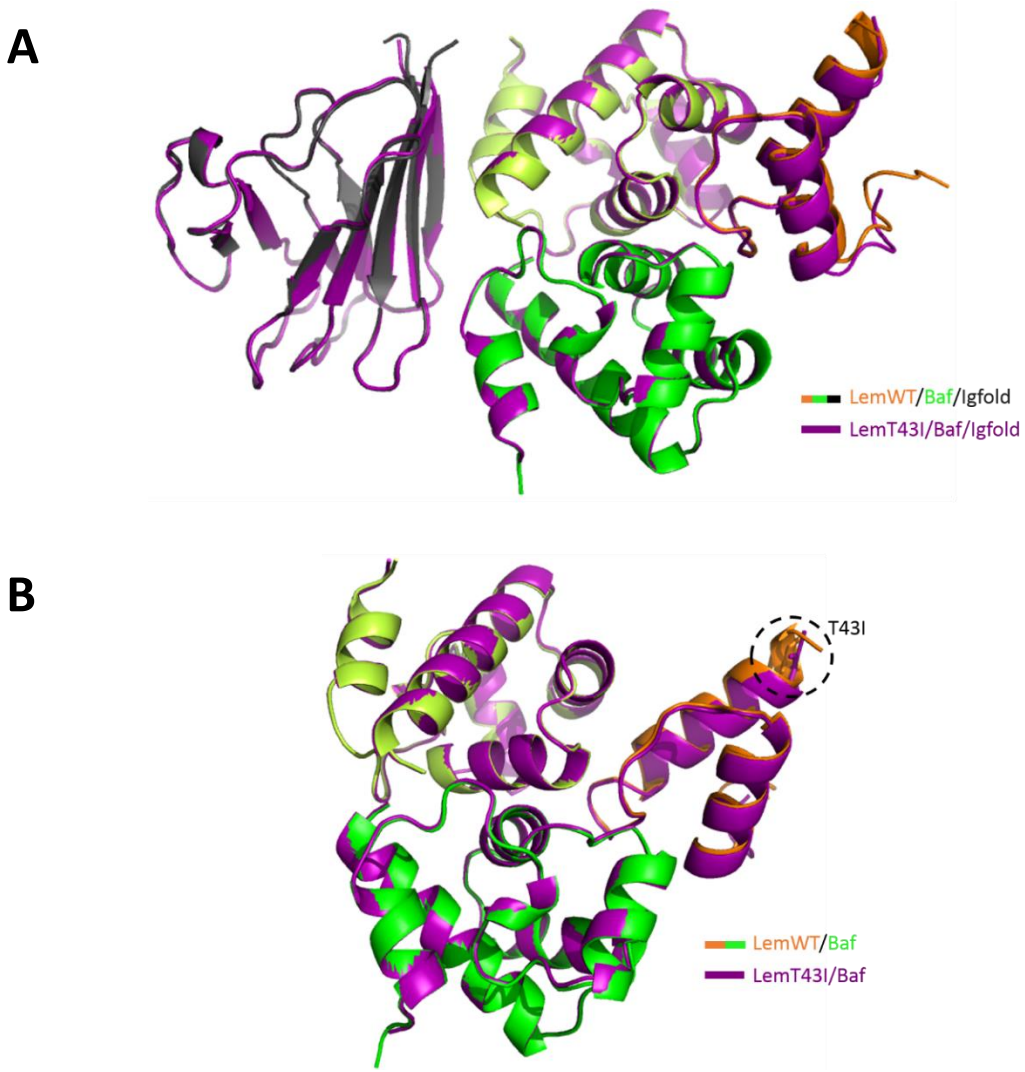


Figure 69 : X-ray crystallography structure of the ternary complex formed between the mutated T43I LEM domain of the emerin, the BAF_{CtoA} dimer and the Igfold domain.

(A) Superimposition of the X-ray crystallography structures of the ternary complexes formed between the LEM domain of emerin WT, BAF and lamin A/C Igfold and between the LEM domain of emerin T43I, BAF and lamin A/C Igfold. The resolution of the two complexes are 2.1Å and 2.3Å, respectively, and the rmsd measured between the Calpha carbons of the two complexes yields was 0.547Å. (B) Zoom on the superimposition of the BAF_{CtoA}/LEM WT and BAF_{CtoA}/LEM T43I complexes.

Concerning the ternary complexes containing mutants ΔK37 and P22L, we choose to send a sample of these complexes to the European crystallogenesis facility in Grenoble (HTX laboratory), in order to screen a lot of conditions to get crystals.

To conclude with this part of my thesis, I now propose a first interaction mechanism involving lamin A/C and emerin. Indeed, I found that this interaction could be mediated by a third partner, the BAF protein, which is a DNA-binding protein. From this conclusion and because I obtained the X-ray structure of this ternary complex, I am able to present a first structural model of the interaction between the inner nuclear membrane, the nucleoskeleton and DNA (figure 57). Moreover, I found that some mutants of the Igfold domain that cause progeroid syndromes have a decreased affinity for BAF_{CtoA} *in vitro*. Ambre Petitalot in the lab also introduced the same mutations in a vector for expression in mammalian cells, and our collaborator Dr Brigitte Buendia at Univ. Paris Diderot will now observe *in situ* the proximities between mature lamin A and BAF as a function of the mutation.

We also reasoned that in the classical dominant HGPS syndrome, aging is caused by an incorrect processing of lamin A. If the mechanism of this dominant disease is related to the mechanism of the recessive diseases caused by homozygous mutations in the lamin A/C Igfold domain, then the incorrect processing of lamin A might also cause a loss of Igfold-BAF interaction. We observed that the progerin tail is able to bind BAF with an affinity similar to that of the lamin A/C Igfold alone. In cells, the progerin C-terminus is farnesylated. Thus it would now be interesting to test what is the impact of progerin farnesylation on these interactions.

Finally, we observed that mutants of the LEM domain of emerin, including one that exhibits a destabilized LEM domain, are capable of forming a ternary complex with BAF and the lamin A/C Igfold, as WT emerin (see for example figure 69 below).

2. THE LAMIN A/C IGFOLD DOMAIN DIRECTLY INTERACTS WITH SELF-ASSEMBLED EMN

We found that A-type lamins interact indirectly with EmN, *in vitro*, whereas several studies confirmed the existence of a direct interaction between both proteins[166,170]. Moreover, the team of Dr K. Wilson demonstrated that mutations in emerin region from amino acid 70 to amino acid 170 inhibit lamin binding[245]. We did not observe a direct interaction between the Igfold and EmN monomers and moreover, our indirect interaction does not involve region 70-170 of emerin. We showed that EmN can form oligomers and filaments *in vitro*. For this reason, we chose to test the impact of emerin oligomerization on the interaction with the Igfold domain of lamin A/C.

a. EmN oligomers bind to the globular domain of lamin A/C

First, I studied this interaction by NMR. I produced a ^{15}N labeled Igfold and recorded its ^{15}N-^{1}H HSQC spectrum alone, at 303K. Then, I added different oligomeric forms of non-labeled EmN. To do this, I concentrated my emerin sample at 600µM, which was a sufficient concentration to obtain oligomers and filaments, and I stored the sample at 293K, with 5mM DTT. Finally, every day during one week, I recorded a spectrum of my ^{15}N labeled Igfold sample in the presence of the same amount of EmN protein, taken from the EmN sample stored at 293K (figure 70). In parallel, I observed each NMR sample by negative staining EM to see which oligomeric forms of emerin were present in the sample.

I reproduced this experiment three times, and every time, I observed no change in the spectrum of the Igfold in the presence of emerin before day 6. By EM, I also observed that before day 6, only spheres or aggregates are observed in the NMR sample, which could correspond to some emerin inactive oligomers (figure 70, A and B). Contrariwise, after day 6, I observed that the majority of the peaks, which correspond to the amino acids of the Igfold, lost a significant fraction of their intensity (figure 70, C). Interestingly, I observed the NMR sample by EM, and found the presence of emerin filaments.

At this point, I checked if the interaction observed by NMR was not due to an aggregation of the Igfold in the presence of emerin oligomers. Therefore, I centrifuged my NMR sample, obtained after interaction, during 10 minutes at 12000 g and observed if the Igfold was still soluble or not on a SDS-PAGE gel (figure 71).

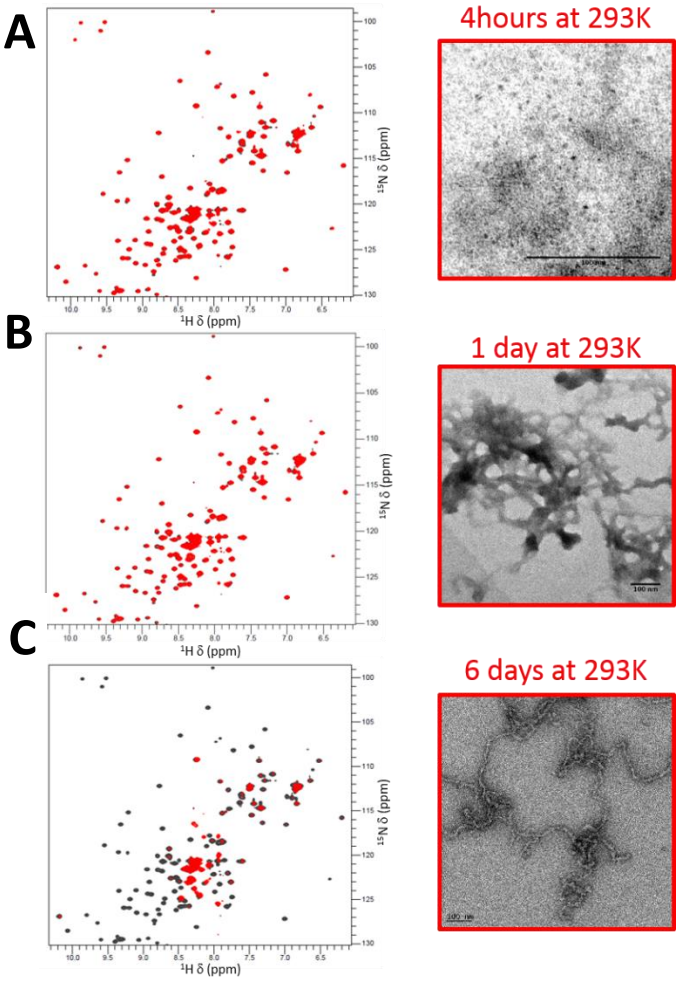


Figure 70 : Study of Igfold/EmN oligomers interaction.

Proteins were dialyzed in 20mM Tris-HCl pH8, 30mM NaCl, 2mM DTT and spectra were recorded at a temperature of 303K on a 700MHz spectrometer. (A) Superimposition of the ^{1}H-^{15}N HSQC spectra of ^{15}N labeled Igfold at 250µM alone in grey, or in the presence of 250µM of emerin incubated at 600µM during 4h at 293K, in red. (B) Superimposition of the ^{1}H-^{15}N HSQC spectra of ^{15}N labeled Igfold at 250µM alone in grey, or in the presence of 250µM of emerin incubated at 600µM during 24h at 293K, in red. (C) Superimposition of the ^{1}H-^{15}N HSQC spectra of ^{15}N labeled Igfold at 250µM alone in grey, or in the presence of 250µM of emerin incubated at 600µM during 6 days at 293K, in red. EM pictures correspond to the NMR sample of each experiment.

It was clearly visible on this SDS-PAGE gel that half of both proteins were still soluble whereas the other part was in the pellet. Even if we lost some Igfold because of some precipitation of the sample, the amount of Igfold that was insoluble was not sufficient to explain all the intensity decrease observed on the spectrum. We thus confirmed that the signal intensity diminution observed by NMR is due to an interaction with emerin oligomers and not to precipitation of the Igfold sample.

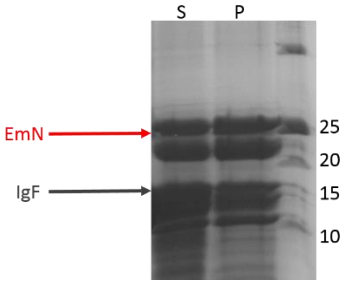


Figure 71 : SDS-PAGE picture of the supernatant and the pellet obtained after centrifugation of the NMR sample corresponding to figure 69, C.

I concluded that a direct interaction between the lamin A/C Igfold and EmN can be observed, but only after EmN self-assembly. I did not know at this moment if the Igfold interacted with globular emerin oligomers, or only with emerin filaments, and if this interaction favored emerin filament formation.

To observe the impact of the Igfold on emerin self-assembly, we used thioflavin T fluorescence, like we did to follow filament formation (figure 72). Thus, I concentrated both EmN and Igfold proteins until 600µM and I incubated half of the EmN sample with the Igfold to obtain a final protein concentration of 300µM and I incubated the other part of my EmN sample with the kinetics buffer (20mM Tris-HCl pH8, 30mM NaCl) to be also at a final concentration of 300µM. Afterwards, I incubated both samples (EmN alone and in the presence of Igfold) at 310K and I took a fluorescence measurement at 480nm, after excitation at 440nm, every hour during 4h and then the day after. For each measurement, I diluted my protein sample to 40µM with the kinetics buffer and added 10µM of thioflavin T. I did this experiment once and I still need to reproduce it.

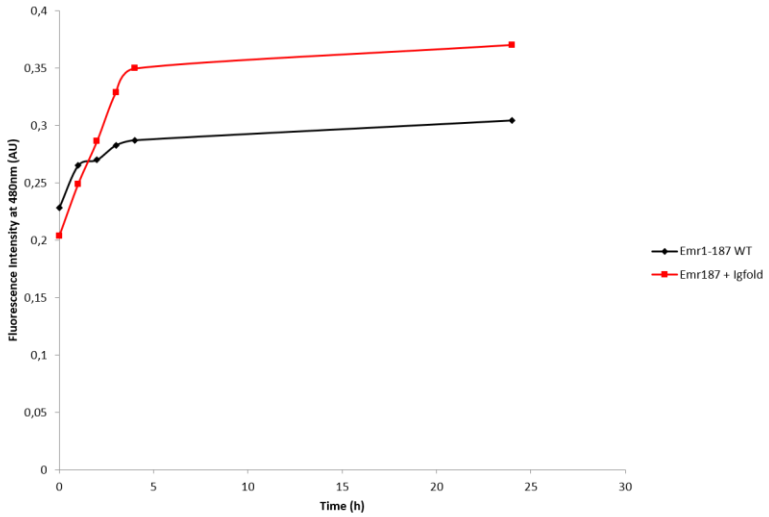


Figure 72 : Kinetics of EmN self-assembly in the absence (black) or presence (red) of the Igfold, as followed through thioflavin T fluorescence measurements.

I observed a small increase in the self-assembly rate of EmN in the presence of the Igfold, which should now be confirmed. It is also not clear how the Igfold favors EmN self-assembly.

To confirm that the interaction between EmN oligomers and the Igfold is specific, we decided to search for some Igfold mutants that cannot bind to the EmN oligomers. Because we did not know at all the surface of the Igfold binding to the EmN oligomers, Jinchao Yu, a PhD student in the team of Dr Raphaël Guérois, in our lab, used a server called Consurf[258], to analyze a lamin multiple sequence alignment and identify a hot spot of best conserved residues, which could be implicated in lamin interactions with partners. He found a conserved patch on the Igfold surface, which contains an apolar core, and from this observation he proposed several mutants that could disrupt the Igfold/EmN interaction. He proposed mutants T496E, W498E, H506E, P508A, P509A and V513A. The mutated residues are represented on the X-ray structure of the Igfold (PDB: 1IFR) in figure 73.

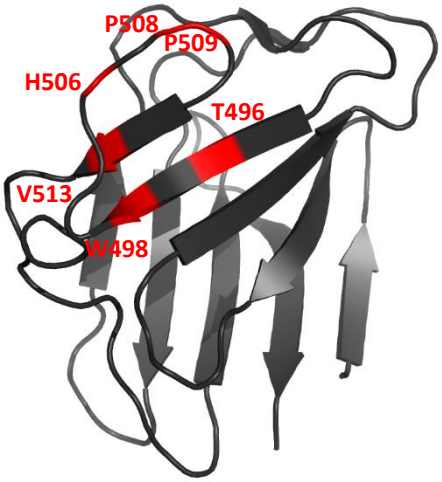


Figure 73 : Localization of the different Igfold residues predicted as being important to interact with EmN oligomers (in red) (PDB: 1IFR).

For the moment, four mutants were obtained using the QuikChange kit (Stratagene) by Ambre Petitalot in the lab. Together with Ambre, I managed to produce and purify these four mutants: T496E, H506E, P509A and V513A.

First, we observed by NMR, recording ^{1}H-^{15}N HSQC spectrum of each mutant, that these four mutants are well-folded (figure 74) whereas using fluorescence-based Thermal Shift Assay, we observed that the thermal stability of the Igfold is largely decreased by some of these mutations (figure 75). Indeed, only H506E shows a WT-like stability and large decreases in stability (up to 18 degrees) are observed for the others.

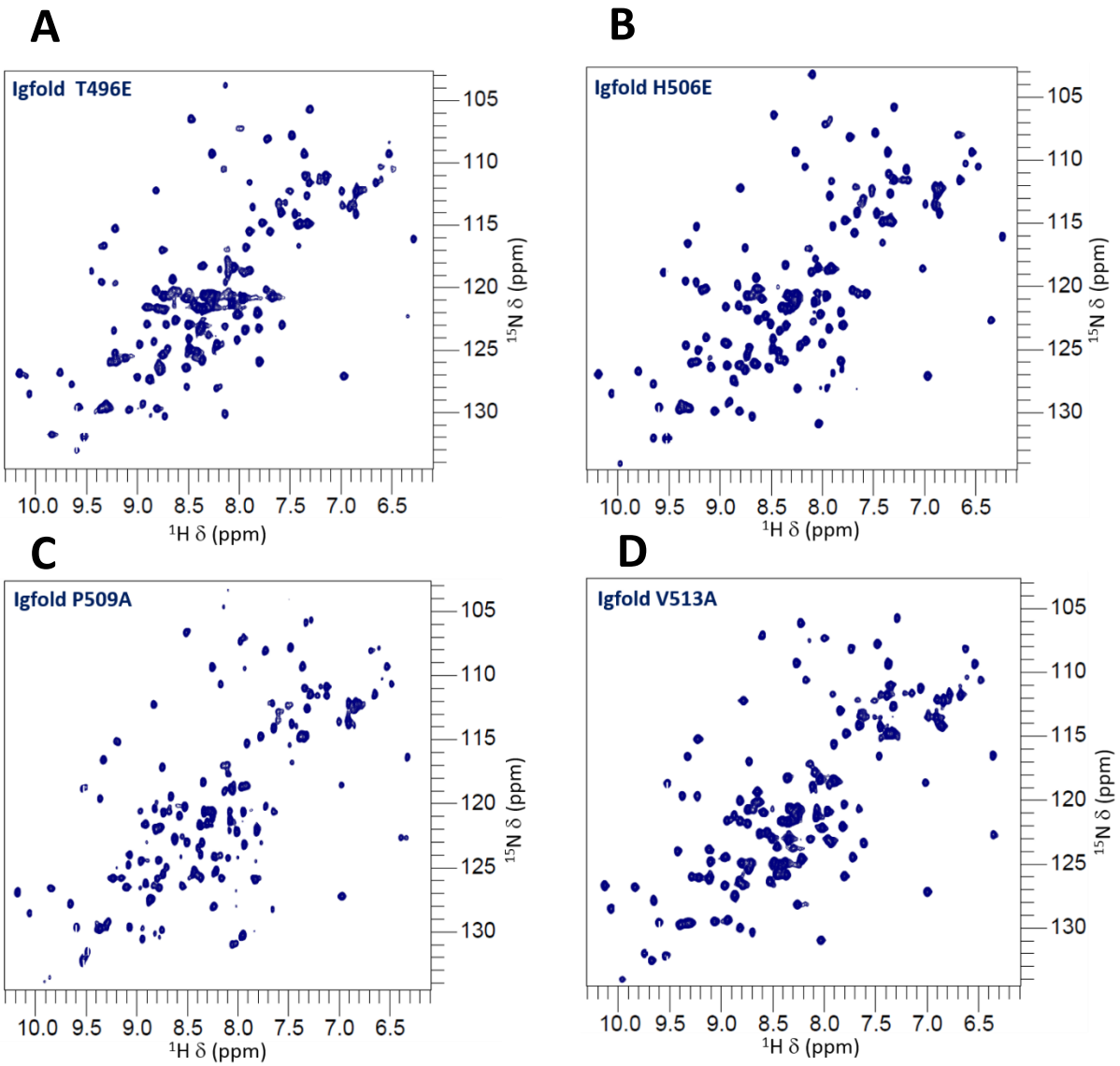
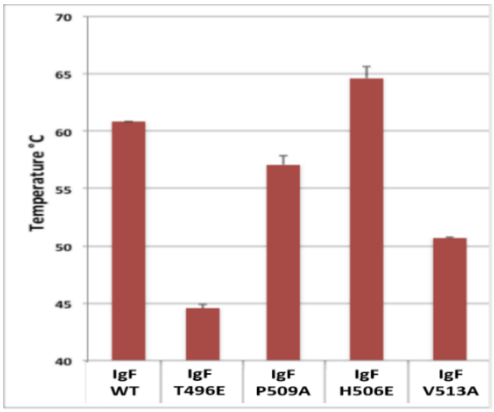


Figure 74 : NMR analysis of the different Igfold mutants.

^{1}H-^{15}N HSQC spectra recorded on an Igfold T496E sample at 200µM (A), on an Igfold H506E sample at 200µM (B), on an Igfold P509A sample at 200µM (C) and on an Igfold V513A sample at 200µM (D). All proteins were dialyzed in 20mM Tris-HCl pH8, 30mM NaCl, 2mM DTT and spectra were recorded at 303K and 700MHz spectrometer (CEA Saclay).

A



B

Protein	Tm (°C)	Standard error (AU)
Igfold WT 1	60,8	0,0
Igfold WT 2	61,5	0,3
Igfold T496E	44,6	0,4
Igfold P509A	57,1	0,8
Igfold H506E	64,6	1,0
Igfold V513A	50,7	0,1

Figure 75 : Thermal stability of Igfold WT and mutants, measured using a fluorescence-based thermal shift assay.

(A) Graphical representation of Igfold WT and mutant thermal stabilities, in Buffer A (50mM Mes pH6, 100mM NaCl). (B) Table of Tm and standard error values obtained for different Igfold WT and mutants, in Buffer A.

Then, I observed interaction of these ^{15}N labeled mutants with EmN by NMR. For these experiments, the Igfold mutants were dialyzed in 20mM Tris-HCl pH8, 30mM NaCl, 2mM DTT and concentrated until 400µM. In parallel, we produced non-labeled EmN, dialyzed it in the same buffer and after concentration of the protein until 600µM, the EmN sample was incubated at 293K during one week in order to obtain EmN filaments. Then, we recorded HSQC spectra of each Igfold mutant at 200µM alone or in the presence of the same ratio of EmN oligomers (figure 76). After each NMR experiment, samples containing the Igfold and EmN oligomers were observed by EM, in order to confirm the presence of filaments. EM pictures were obtained thanks to Dr Ana Arteni (I2BC, Gif-sur-Yvette).

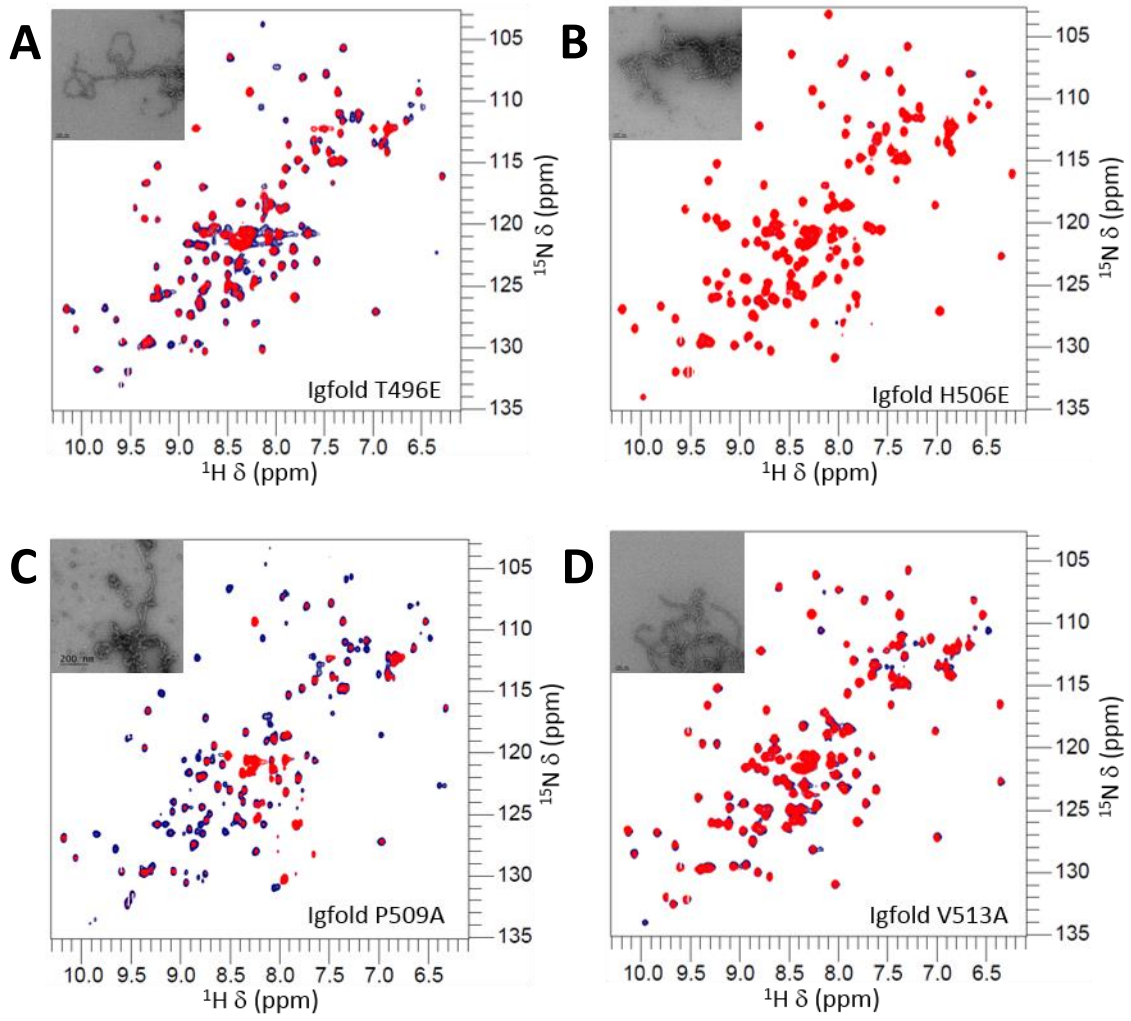


Figure 76 : NMR and EM interaction study between different Igfold mutants and EmN oligomers.

Proteins were dialyzed in 20mM Tris-HCl pH8, 30mM NaCl, 2mM DTT and spectra were recorded at 303K and 700MHz spectrometer (CEA Saclay). (A) Superimposition of the ^{1}H-^{15}N HSQC spectra of ^{15}N labeled T496E Igfold at 200µM alone in navy, or in the presence of 200µM of EmN oligomers, in red. (B) Superimposition of the ^{1}H-^{15}N HSQC spectra of ^{15}N labeled H506E Igfold at 200µM alone in navy, or in the presence of 200µM of EmN oligomers, in red. (C) Superimposition of the ^{1}H-^{15}N HSQC spectra of ^{15}N labeled P509A Igfold at 200µM alone in navy, or in the presence of 200µM of EmN oligomers, in red. (D) Superimposition of the ^{1}H-^{15}N HSQC spectra of ^{15}N labeled V513A Igfold at 200µM alone in navy, or in the presence of 200µM of EmN oligomers, in red. EM pictures correspond to the NMR sample of each experiment.

We reproduced these NMR experiments two times at least and then, we measured intensity differences for every peak corresponding to the bound or unbound Igfold mutant and did an average for each experiment; then values corresponding to experiments involving the same Igfold mutant were again averaged (figure 77). We observed that T496E, P509A and V513A significantly impaired binding to EmN oligomers, whereas H506E completely abolished it.

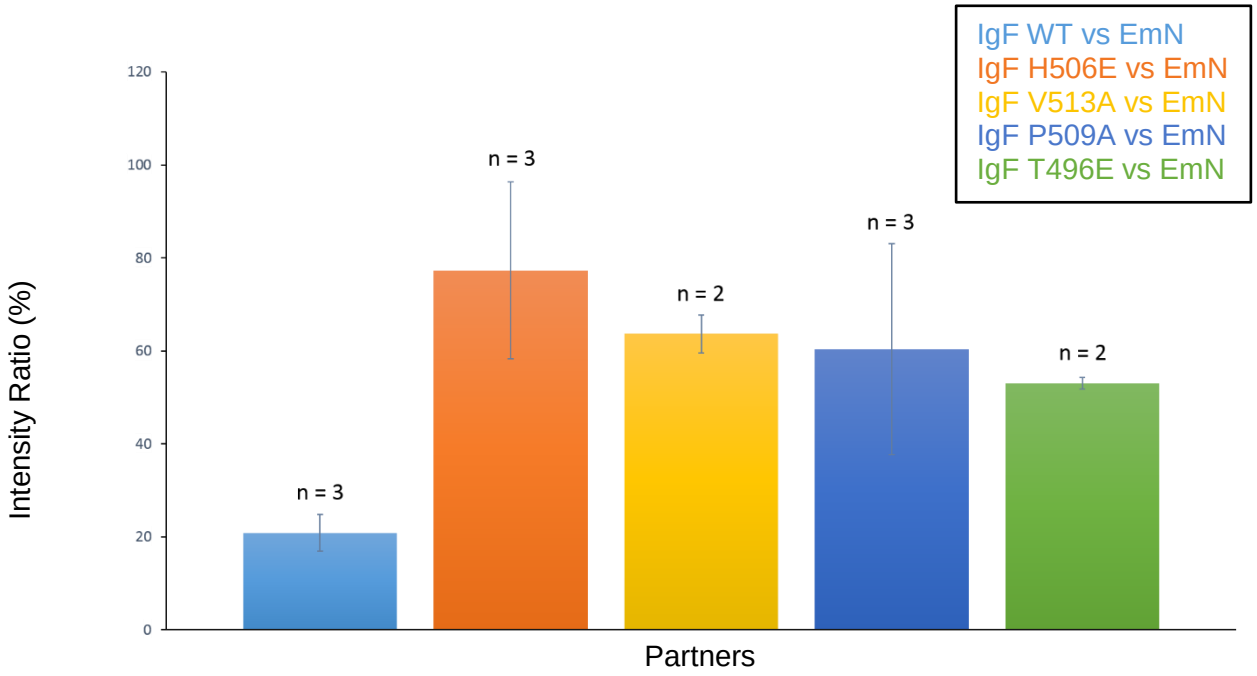


Figure 77 : Intensity ratio between the NMR HSQC signals measured in the absence or presence of EmN oligomers on the different Igfold mutants.

In light blue, the Igfold WT, in orange, Igfold H506E, in yellow, Igfold V513A, in dark blue, Igfold P509A and in green, Igfold T496E. "n" corresponds to the number of times the experiment was done.

As the fragment from Val494 to Trp514 (containing our four mutated positions) is mostly apolar, this lamin-emerin interaction could be driven by the hydrophobic effect. In agreement with this observation, we did preliminary isothermal titration calorimetry experiments in order to obtain stoichiometry and affinity of this lamin-emerin interaction (figure 78) and we observed that at 288K, the binding energy is provided by a favorable entropic contribution (7.4 ± 0.1 kcal mol $^{-1}$) partially compensated by an unfavorable enthalpic contribution (0.3 ± 0.1 kcal mol $^{-1}$). However, because we obtained a low heat signal, we did not conclude about the affinity and stoichiometry of this interaction.

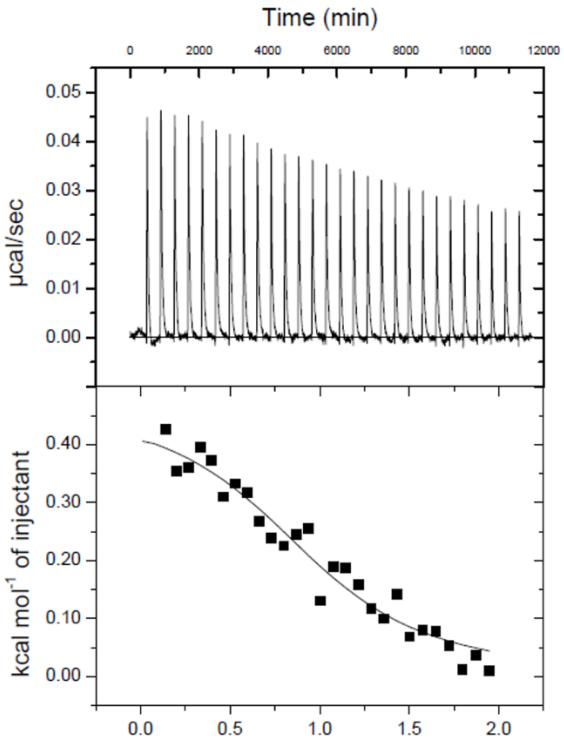


Figure 78 : Characterization of EmN oligomers/Igfold interaction by ITC.

Successive injections of 10µl of Igfold concentrated at 315µM in a EmN oligomer sample concentrated at 31.5µM. Both proteins were dialyzed against the same buffer (20mM Tris-HCl pH8, 30mM NaCl, 10mM β-mercaptoethanol, protease inhibitors Roche). The experiment was done at 288K, on a VP-ITC.

In parallel, we observed interactions of two ^{15}N mutants causing progeroid syndromes (R435C, which abolishes the interaction with BAF and R471C, which decreases the affinity for BAF) with EmN by NMR, in order to observe whether the mutants localized on another part of the Igfold than the one we proposed to be the interaction surface, could destabilize the interaction with EmN too. We did exactly the same experiments as the one described for the four other mutants (figure 79).

After intensity differences measurement, we observed that both mutants impaired binding to EmN oligomers, but because we only did the experiment once, we have to reproduce it to conclude.

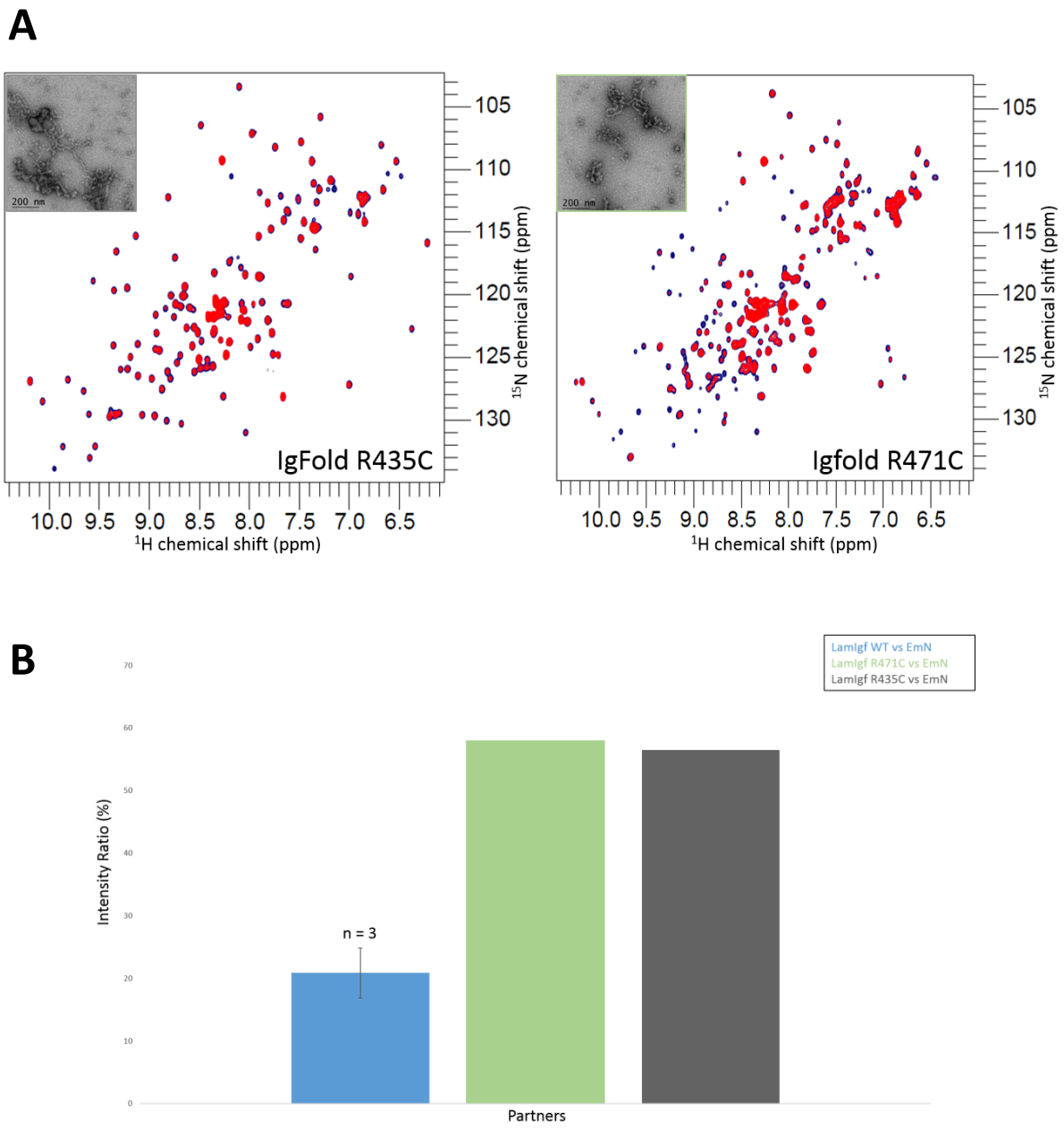


Figure 79: NMR and EM interaction study between Igfold mutants causing progeroid syndromes and EmN oligomers.

(A) On the left, superimposition of the ^{1}H-^{15}N HSQC spectra of ^{15}N labeled R435C Igfold at 200µM alone in navy, or in the presence of 200µM of EmN oligomers, in red and on the right, superimposition of the ^{1}H-^{15}N HSQC spectra of ^{15}N labeled R471C Igfold at 200µM alone in navy, or in the presence of 200µM of EmN oligomers, in red. Proteins were dialyzed in 20mM Tris-HCl pH8, 30mM NaCl, 2mM DTT and spectra were recorded at 303K and 700MHz spectrometer (CEA Saclay). EM pictures correspond to the NMR sample of each experiment. (B) Intensity ratio between the NMR HSQC signals measured in the absence or presence of EmN oligomers on, Igfold WT (in light blue), R435C Igfold mutant (in grey) or on R471C Igfold mutant (in light green).“n” corresponds to the number of times the experiment was done.

b. *The EmN / Igfold interaction is specific to the emerin 1-187 region and A-type lamins*

In the first part of my results, I presented two mechanisms of emerin filament formation. Indeed, I showed that EmN, which comprises a LEM domain and a region that is intrinsically disordered, is able to form monomers and oligomers and these oligomers bind to the Igfold of lamin A/C. But another PhD student of the lab, Florian Celli, found that another part of emerin, emerin 67-221 (EmC221), which is entirely unstructured, is able to oligomerize immediately after purification. This oligomerization necessitates region 188-221, because fragment 67-187 does not oligomerize[259].

We decided to observe if the lamin A/C Igfold could also interact with the filaments formed by fragment 67-221 (EmC221). By NMR, we recorded ^{1}H-^{15}N HSQC spectra of the ^{15}N labeled Igfold alone or in the presence of EmC221 oligomers.

Whereas addition of EmN oligomers provoked the disappearance of most NMR signals, addition of EmC221 oligomers did not modify the Igfold ^{1}H-^{15}N HSQC spectrum (figure 80, A and C). We concluded that interaction between emerin oligomers and lamin A/C Igfold is specific to the EmN oligomers.

In addition, because the surface identified by Jinchao Yu is also conserved in B-type lamins, we hypothesized that LB1 was able to bind to EmN oligomers, but no interaction was detected (figure 80, B and C) by NMR. We concluded that the detected emerin-lamin interaction is specific to A-type lamins. Other Igfold residues that are not conserved in B-type lamins probably also contribute to EmN oligomer binding.

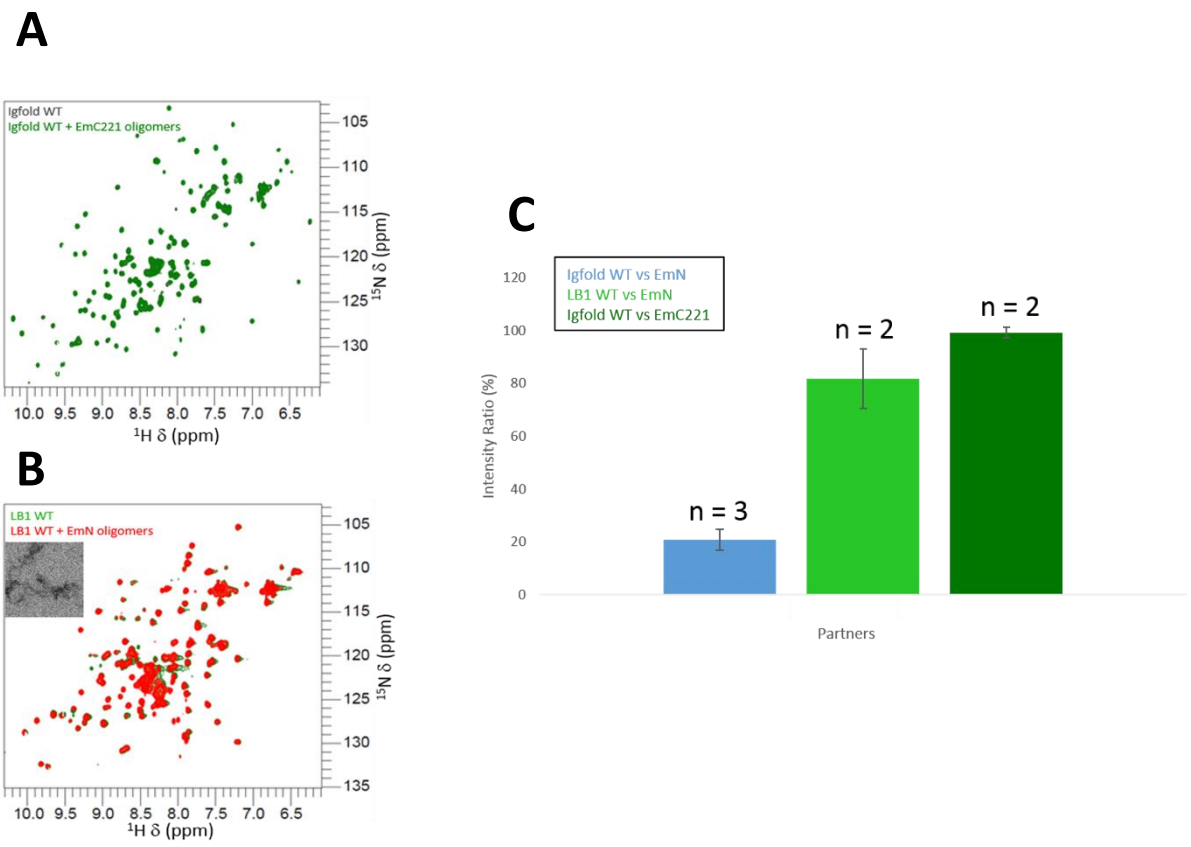


Figure 80 : No binding was observed between EmC221 oligomers or LB1 and Igfold, by NMR.

Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a Igfold sample at 200µM alone (in grey) or in presence of 200µM of EmC221 (in dark green), in 20mM Tris-HCl pH8, 30mM NaCl, at 303K and 700MHz (CEA Saclay). (B) Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a LB1 sample at 200µM alone (in middle green) or in presence of 200µM of EmN (in red), in 20mM Tris-HCl pH8, 30mM NaCl, at 303K and 700MHz (CEA Saclay). (C) Intensity ratio between the NMR HSQC signals measured in the absence or presence of EmN oligomers (in blue) or EmC221 (in dark green) on Igfold or of EmN oligomers on LB1 (in middle green). "n" corresponds to the number of times the experiment was done.

Our results are consistent with the previously published results of the group of K. Wilson, who proposed that mutations around positions 70, 76, 95, 112, as well as 141 and 164, decrease binding of emerin to lamin A/C[165]. In addition, we did not observe any significant binding of EmN to the lamin B1 tail, and they reported only a weak binding to lamin B1 tail[165]. Altogether, these results demonstrate that a hydrophobic surface present on the lamin A/C Igfold domain binds to self-assembled emerin fragment EmN.

c. *Two in-frame deletion mutants of EmN have two different impacts on the direct lamin/emerin interaction*

Because we found that two emerin mutations (ΔK37 and Δ95-99) which cause two different types of diseases have a different impact on EmN filament assembly, we chose to study the impact of these mutations on the interaction of EmN with the Igfold of lamin A/C. As it is explained in the first part of my results, mutant ΔK37 forms filaments faster than EmN WT, whereas mutant Δ95-99 cannot form filaments *in vitro*.

I tested these interactions by NMR. To do so, I produced the ^{15}N labeled Igfold and recorded ^{1}H-^{15}N HSQC spectra of this protein alone or in presence of ΔK37 or Δ95-99 in conditions favoring oligomerization (one spectrum per day to observe evolution of mutant oligomerization) (figure 81). Concerning the mutant Δ95-99, after 6 days at 600µM and at 293K, which were the same conditions in which WT EmN could interact, this mutant cannot bind to the Igfold, because after superimposition of the spectra of the Igfold alone or in presence of this concentrated mutant, no intensity or chemical shift difference was observed. Contrariwise, the mutant ΔK37, which can form filaments faster than the emerin 1-187 WT, could interact with the Igfold.

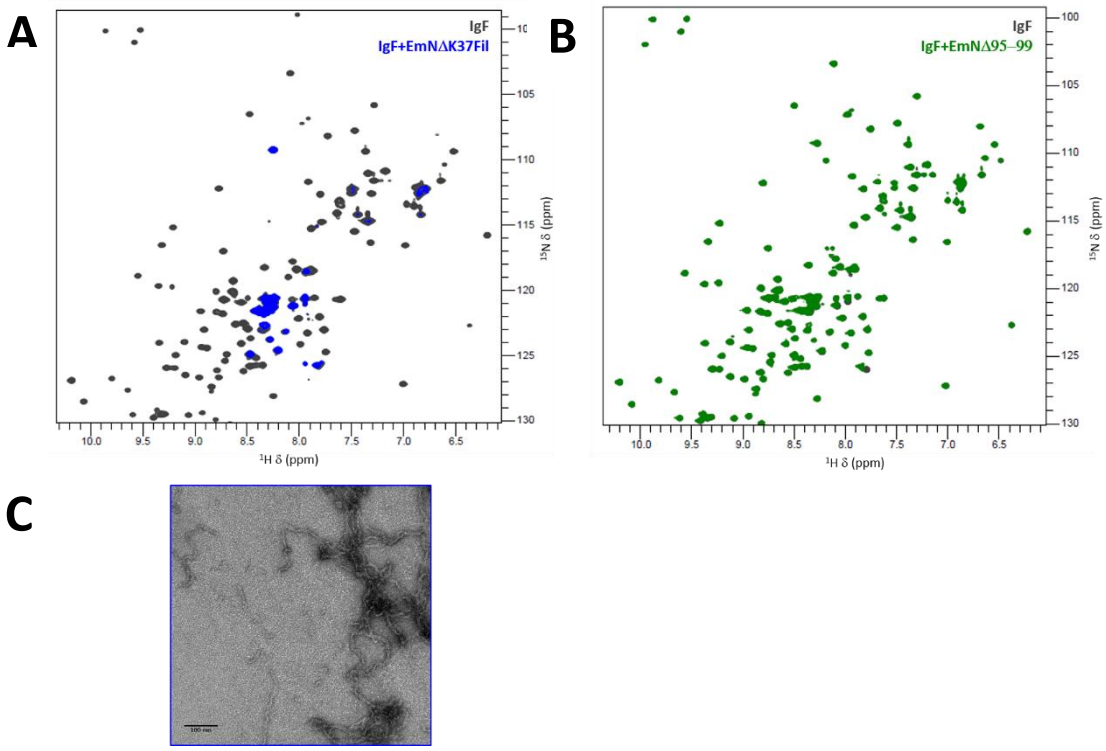


Figure 81 : Study of Igfold/EmN ΔK37 oligomers and Igfold/EmN Δ95-99 oligomers interaction.

Proteins were dialyzed in a 20mM Tris-HCl pH8, 30mM NaCl, 2mM DTT and spectra were recorded at a temperature of 303K on a 700MHz spectrometer. (A) Superimposition of the ^{1}H-^{15}N HSQC spectra of ^{15}N labeled Igfold at 250µM alone in grey, or in the presence of 250µM of EmN ΔK37 incubated at 600µM during 6 days at 293K, in blue. (B) Superimposition of the ^{1}H-^{15}N HSQC spectra of ^{15}N labeled Igfold at 250µM alone in grey, or in the presence of 250µM of EmN Δ95-99 incubated at 600µM during 6 days at 293K, in green. (C) EM picture of the NMR sample from experiment (A).

Like I did for EmN WT, I checked if the interaction observed by NMR was not due to aggregation of the Igfold. Again, I centrifuged my NMR sample, obtained after interaction, during 10 minutes at 12000 g and observed if the Igfold was still soluble or not on a SDS-PAGE gel (figure 82). Interestingly, all proteins were still soluble after the experiment, so we confirmed that observed NMR peak intensity decrease was not due to aggregation of the sample.

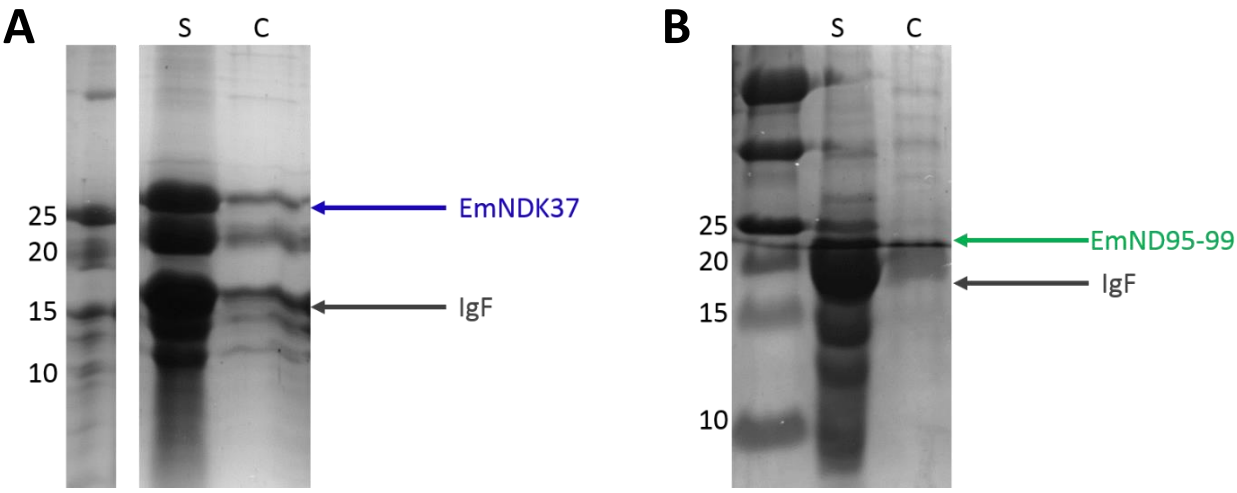


Figure 82 : SDS-PAGE picture of the supernatant and the pellet obtained after centrifugation of the NMR samples containing the Igfold and (A) EmN ΔK37 oligomers and (B) EmN Δ95-99 monomers.

Then, we studied these interactions in HeLa cells, thanks to our collaborators from Paris Diderot University (Dr Brigitte Buendia's team), who did PLA assays to quantify the amount of closed proximities between GFP-emerin Δ95-99 mutant or FLAG-emerin ΔK37 mutant and endogenous lamin A/C, in HeLa cells (figure 83). They confirmed that proximities between emerin Δ95-99 and lamin A/C were very much lower compared to the WT situation[4], whereas for the mutant ΔK37, proximities between both proteins were still visible in cells.

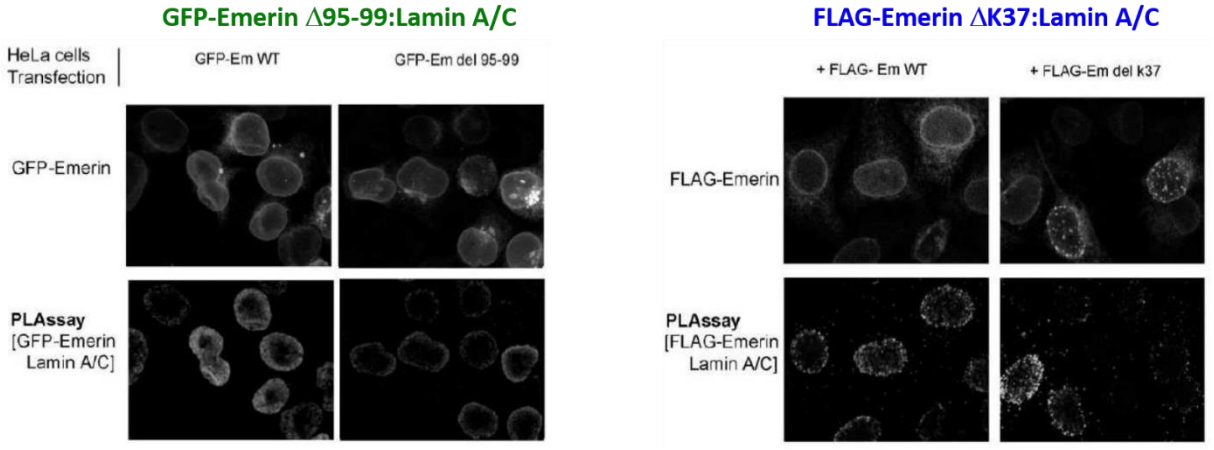


Figure 83 : Analysis of the impact of the Δ95-99 and ΔK37 mutations on Emerin/lamin A/C binding in cell.

Localization and endogenous lamin proximity observed for GFP-emerin WT, GFP-Δ95-99, Flag-emerin WT and Flag-emerin ΔK37 (collaboration with Brigitte Buendia, Univ. Paris Diderot). Whereas Δ95-99 localization is similar to that of the WT, ΔK37 forms nuclear foci in subpopulation of cells. On the other hand, less emerin-lamin proximities are observed in the case of Δ95-99 when compared to WT emerin[4]; in the case of ΔK37, quantification of the observed emerin-lamin proximities is currently being performed.

To conclude, I found a second interaction mechanism involving the lamin A/C Igfold and emerin. I demonstrated that EmN oligomerization is important for direct binding to lamin A/C and to confirm this hypothesis, I observed that one emerin mutant, which cause EDMD and which cannot form EmN oligomers, was unable to bind to the Igfold *in vitro* and in cells, whereas a mutant of the emerin LEM domain, which can form filament faster than WT EmN and causes a cardiac disease, is able to bind directly to the Igfold. In addition, we found that EmN oligomers interact specifically with the lamin A/C Igfold and not with the lamin B1 tail. In the same vein, we showed that this interaction is specific to EmN oligomers and that EmC221 oligomers could not bind to the Igfold. Finally, we observed that the surface of the Igfold domain, which directly contacts emerin oligomers, contains the parallel β-sheet formed by strands β6 and β7, which was described as characteristic of the lamin Igfold structure. It is centered on lamin fragment from Val494 to Trp514 containing residues Thr496, P509, His506 and Val513. Because this fragment is mostly apolar, we did the hypothesis that the lamin-emerin interaction is driven by the hydrophobic effect.

III. Other interaction studies involving emerin

1. INTERACTION WITH ANOTHER LEM DOMAIN PROTEIN: MAN1

Emerin and MAN1 are two LEM domain proteins reported to play overlapping functional roles[260]. A direct interaction between the two LEM Domain proteins, MAN1 and Emerin, was observed by blot overlay and affinity purification from HeLa cell lysates[261]. More in details, they demonstrated that the N-terminal nucleoplasmic region of MAN1 interacts with emerin 1-222, BAF, lamin A and lamin B1. They also demonstrated that the C-terminal nucleoplasmic region of MAN1 binds to BAF. This last interaction was shown to be mediated by DNA by others[57].

In 2012, Benjamin Bourgeois, a PhD student of the team, showed that, using a HisTag-1U2 construct (from amino acid 755 to 911) corresponding to the C-terminal domain of MAN1, it was possible to purify endogenous emerin from 293T cells. Isaline Herrada, another PhD in the team, later tried to test *in vitro* this interaction but the 2 proteins systematically precipitated when they were mixed.

As I had worked with different emerin constructs and had some experience in purifying these constructs in different oligomeric states, I decided to look at this interaction as a side project. Purification of 1U2 was already established by other students from our laboratory. This construct was cloned in a pET-M13 vector, with a histidine tag, so I purified it by nickel affinity chromatography in 50mM Tris-HCl pH7.5 and 150mM NaCl. After this first purification step, I used a heparin column in order to remove DNA and then I used a gel filtration step, in the buffer of my choice, to obtain a pure sample.

First, I studied the interaction between monomeric EmN and 1U2 by NMR. I produced ^{15}N labeled samples and after recording the NMR HSQC spectra of the free proteins, on one hand, I added non-labeled EmN onto ^{15}N labeled 1U2 and on the other hand, I added non-labeled 1U2 onto ^{15}N labeled EmN (figure 84). After spectra superimposition, I observed chemical shift differences and peak intensity decreases due to the addition of the partner. I concluded that EmN interacts with 1U2.

After the NMR experiments, I also did a SDS-PAGE gel of our NMR samples centrifuged during 5 minutes at 12000 g, in order to observe if the proteins were in the supernatant or in the pellet.

As already observed by Isaline Herrada, half of the proteins were in the pellet. Also, in one of the samples, unlabeled 1U2 seemed degraded, and in the other, EmN seemed degraded. So I need to do these experiments again.

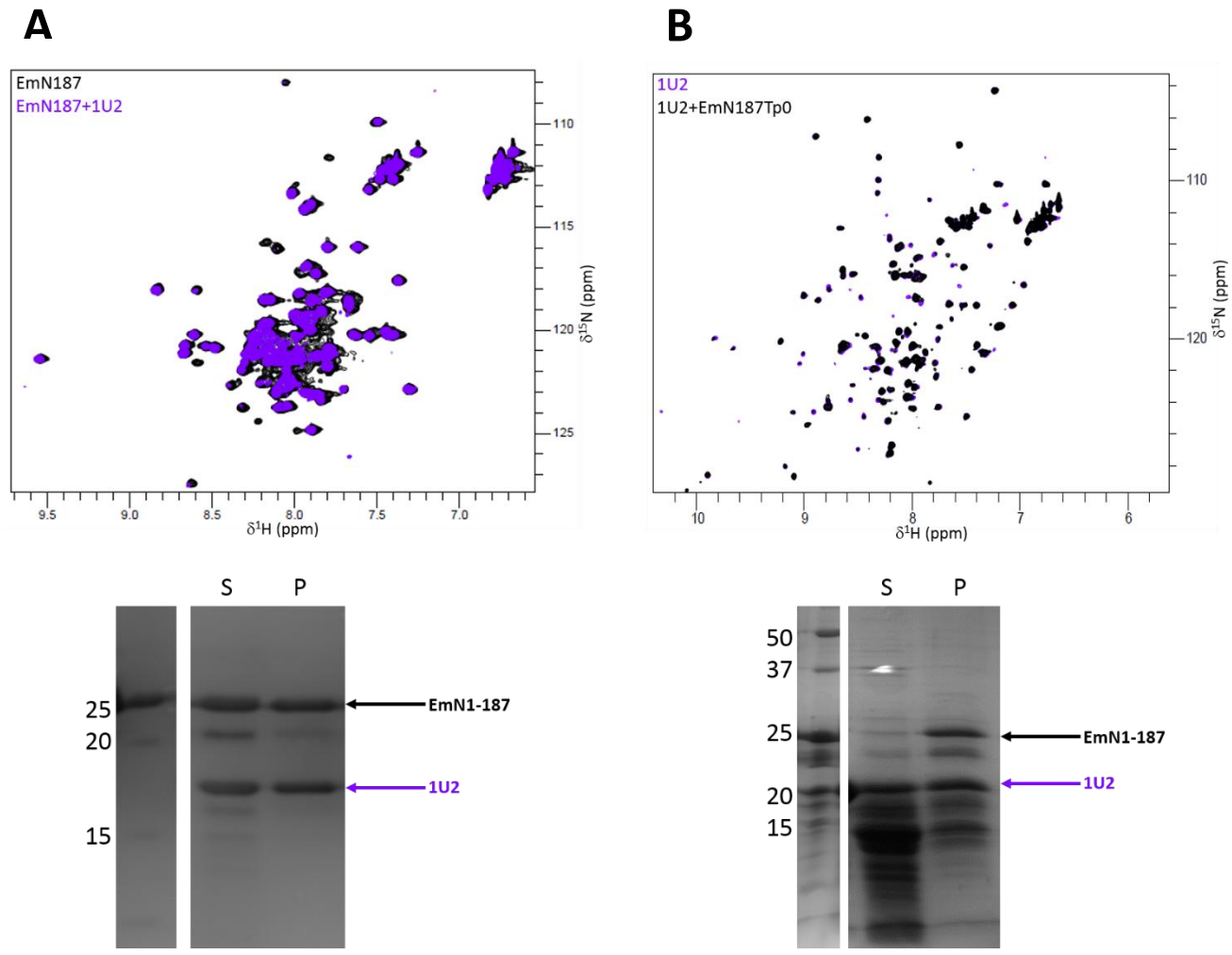


Figure 84 : Study of EmN/1U2 interaction, by NMR.

(A) Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a EmN sample at 100µM alone (in dark) or in presence of 100µM of 1U2 (in purple), in 20mM Phosphate pH6.8, 30mM NaCl, 10mM EDTA, Roche inhibitors, at 303K and 700MHz (CEA Saclay). (B) Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a 1U2 sample at 100µM alone (in purple) or in presence of 100µM of EmN (in dark), in 20mM Phosphate pH6.8, 30mM NaCl, 10mM EDTA, Roche inhibitors, at 303K and 700MHz (CEA Saclay). SDS-PAGE gels corresponding to NMR samples (pellet and supernatant) are displayed under the NMR spectra.

To try to stabilize this complex, I tested the interaction between 1U2 and a smaller emerin construct, EmN132. By NMR, I looked at the interaction between ^{15}N labeled 1U2 and monomeric EmN132 (figure 85). Like for EmN, chemical shift differences were observed after spectra superimposition, but half of the NMR sample was aggregated after experiment.

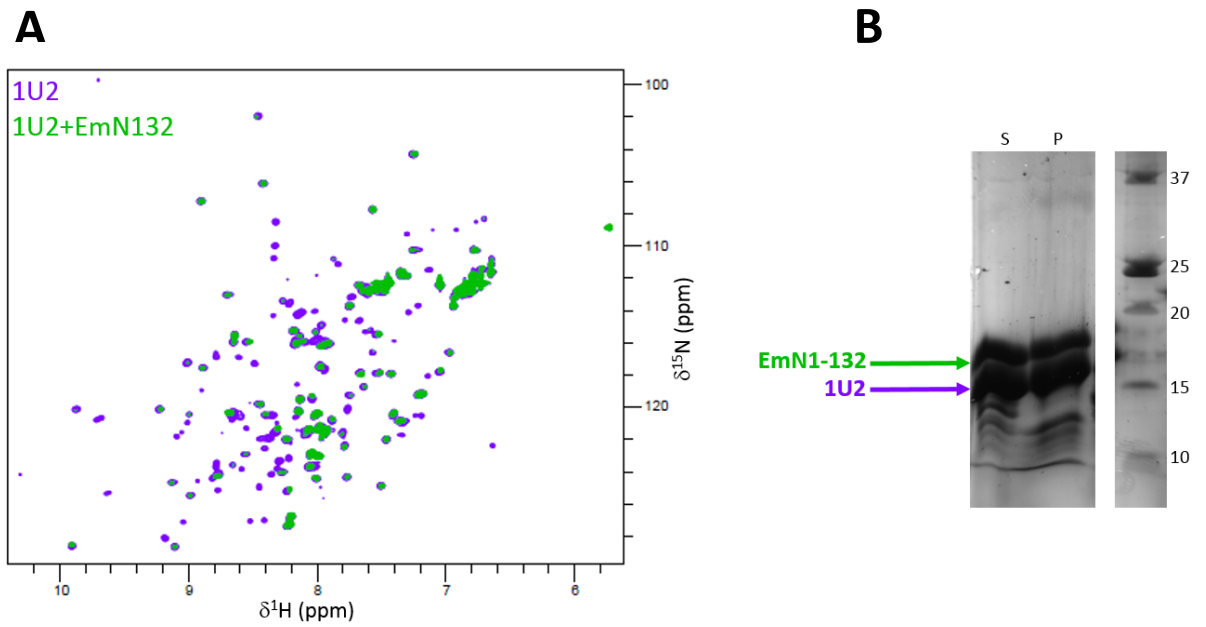


Figure 85 : Study of EmN132/1U2 interaction, by NMR.

(A) Superimposition of 2D NMR ^{1}H–^{15}N spectra recorded on a EmN132 sample at 100µM alone (in green) or in presence of 100µM of 1U2 (in purple), in 20mM Phosphate pH6.8, 30mM NaCl, 10mM EDTA, Roche inhibitors, at 303K and 700MHz (CEA Saclay). (B) SDS-PAGE gel picture of NMR sample after centrifugation (S = Supernatant, P = Pellet).

To go further on this study, I calculated peak intensity ratio like I did for other interaction analyses and thanks to the 1U2 assignment already done in the laboratory, I found which amino acids were the most implicated in this interaction (figure 86, A). Moreover, thanks to a 1U2 structure model obtained in my team[58], I localized residues impacted by the interaction (figure 86, B). Finally, no clear interaction surface was observed on the 1U2 domain. Indeed, peaks found as affected by the interaction with EmN132 correspond to amino acids that are dispersed on the 1U2 structure.

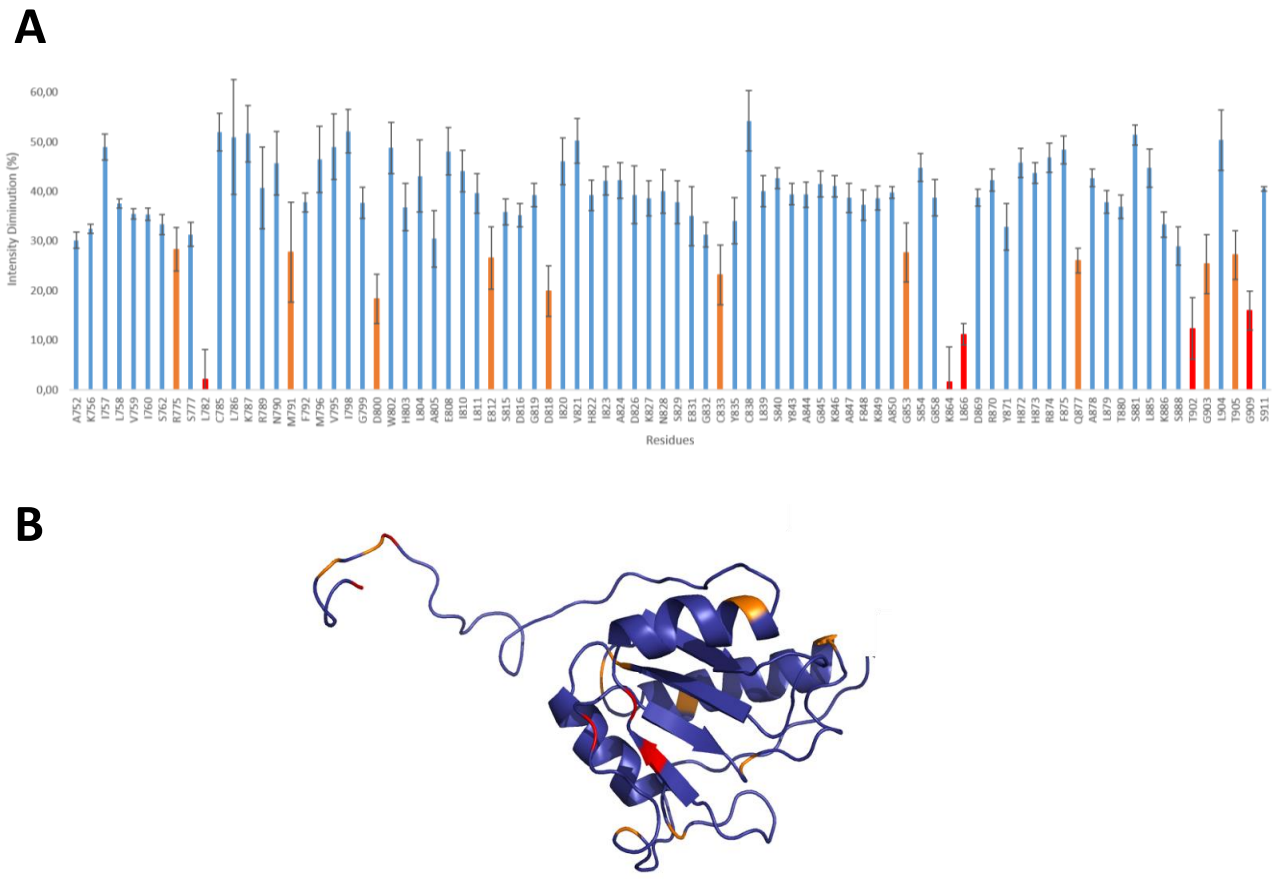


Figure 86 : Intensity ratio measurement after addition of EmN132 equivalent onto the ^{15}N labeled 1U2.

(A) Ratio as a function of the sequence. Bars corresponding to peaks losing more than 70% and 85% of intensity after EmN132 addition are labeled in orange and red, respectively. (B) Three-dimensional-structure of the 1U2 domain with residues colored as a function of (A).

We did not conclude about this interaction. My results suggest that binding between EmN132 and the 1U2 domain of MAN1 exists *in vitro*, but we have to find a way to stabilize this complex. An idea would be to use a fragment of our emerin protein which does not form oligomers or filaments as the mutant EmN Δ95-99 or the LEM domain alone.

2. INTERACTION WITH MYOSIN 1B (IN COLLABORATION WITH THE TEAM OF DR A. HOUDUSSE, INSTITUT CURIE)

First evidence of the presence of myosin in the nucleus was reported in 1980s by a group who demonstrated that myosin, like actin, was involved in nucleoplasmic transport[262]. Then, ultrastructural studies were performed to confirm localization of myosin in the nucleus. In particular, 4 different myosins are in the nucleus: nuclear myosin I[263], myosin VI[264], myosin 16b[265] and myosin Va[266].

Proteomic studies revealed that emerin binds to nuclear myosin I[267] and we decided to work on this interaction through a collaboration with the group of Dr Anne Houdusse, from Curie Institute, who was working on the structure of myosin 1B and 1C tails.

We searched for a direct interaction between different fragments of our emerin (1-187, 1-132, 1-49 and 67-187) and 3 fragments of their myosin tails (M1b01, M1b04 and M1c01), associated with calmodulin, *in vitro*. The different myosin fragments are represented on figure 87 and for each experiment, the group of Dr Anne Houdusse provided them to us.

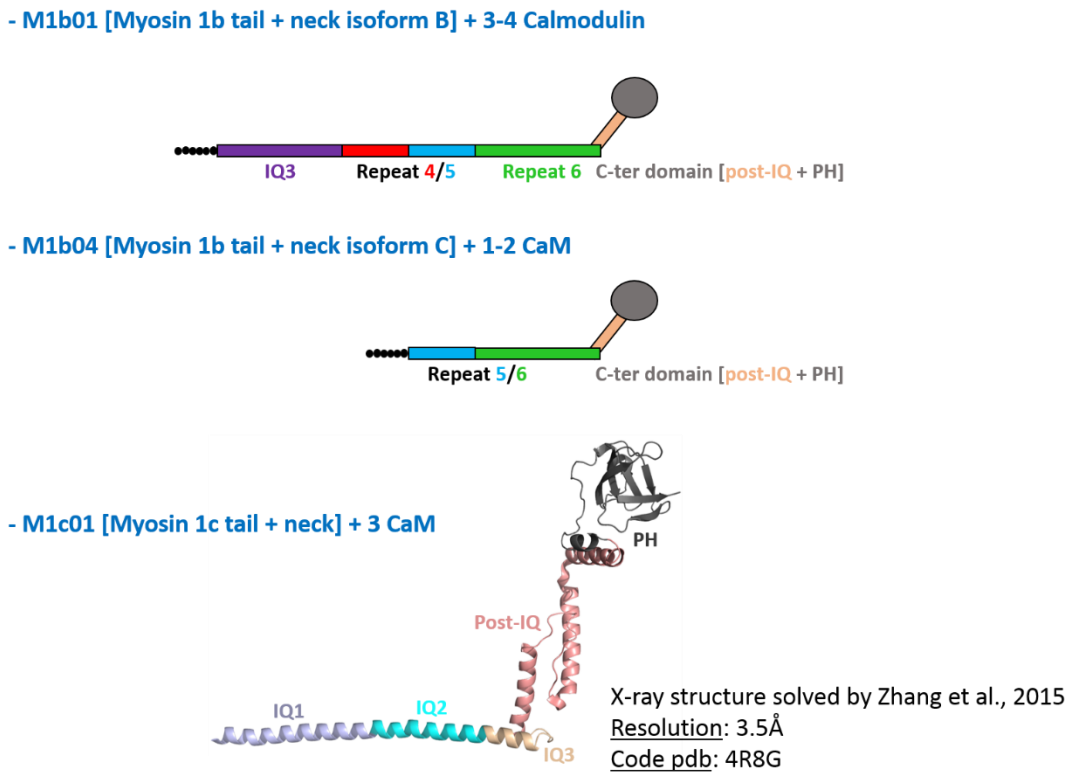


Figure 87 : Scheme of the different myosin fragments provided by the team of Dr Anne Houdusse.

To observe if a direct interaction exists between emerin and one of the myosin tails, we used NMR. I first produced ^{15}N labeled EmN in order to add the three non labelled myosin constructs one by one (figure 88). After spectra superimposition, I observed an intensity signal decrease for several EmN peaks in the presence of M1b01 and M1b04, whereas no change was observed with M1c01. I concluded that M1b tail constructs, but not the M1c tail construct, directly bind to EmN.

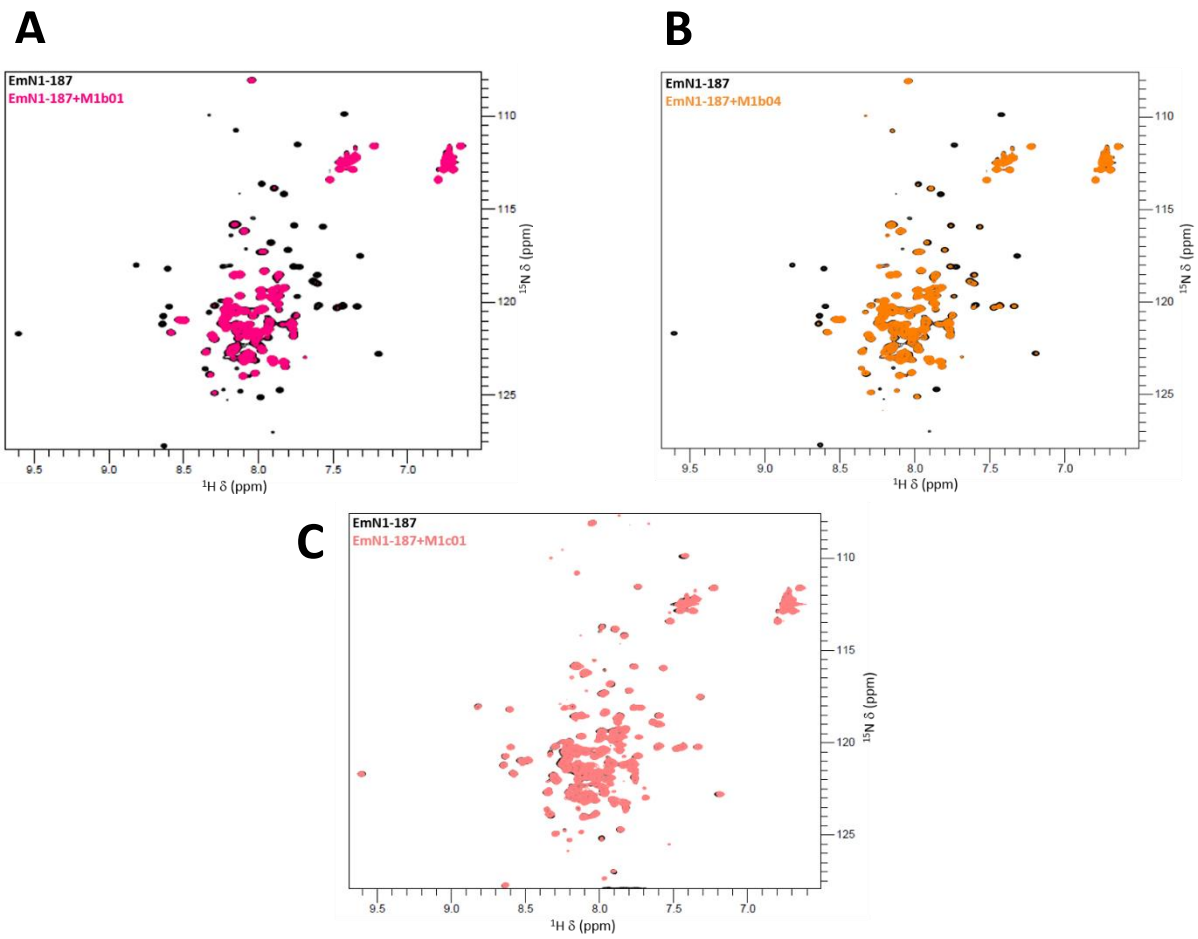


Figure 88 : Study of the interaction between EmN and 3 myosin 1 fragments.

(A) Superimposition of 2D NMR ^{1}H–^{15}N spectra recorded on a EmN sample at 100µM alone (in dark) or in presence of 100µM of M1b01 (in pink), in 50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT, at 293K and 700MHz (CEA Saclay). (B) Superimposition of 2D NMR ^{1}H–^{15}N spectra recorded on a EmN sample at 100µM alone (in dark) or in presence of 100µM of M1b04 (in orange), in 50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT, at 293K and 700MHz (CEA Saclay). (C) Superimposition of 2D NMR ^{1}H–^{15}N spectra recorded on a EmN sample at 100µM alone (in dark) or in presence of 100µM of M1c01 (in coral), in 50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT, at 293K and 700MHz (CEA Saclay).

To find which region of EmN was important for interaction with M1b tails, I assigned several peaks on the emerin spectrum and calculated, like I did for the interaction between emerin and Igfold, the intensity ratio after addition of one equivalent of M1b01 (figure 89). I did not assign all the emerin spectrum because our published emerin assignment was not done at the same pH and temperature, so it was difficult to find all the corresponding peaks and I could not reassign the spectrum without a new ^{15}N, ^{13}C labeled sample, notably for the unfolded region. After this study, I observed that the region that could be the most important for this interaction was the LEM domain of emerin and I localized residues corresponding to peaks that had lost more than 50% of their initial intensity on the known LEM domain structure solved by NMR (PDB: 2ODC).

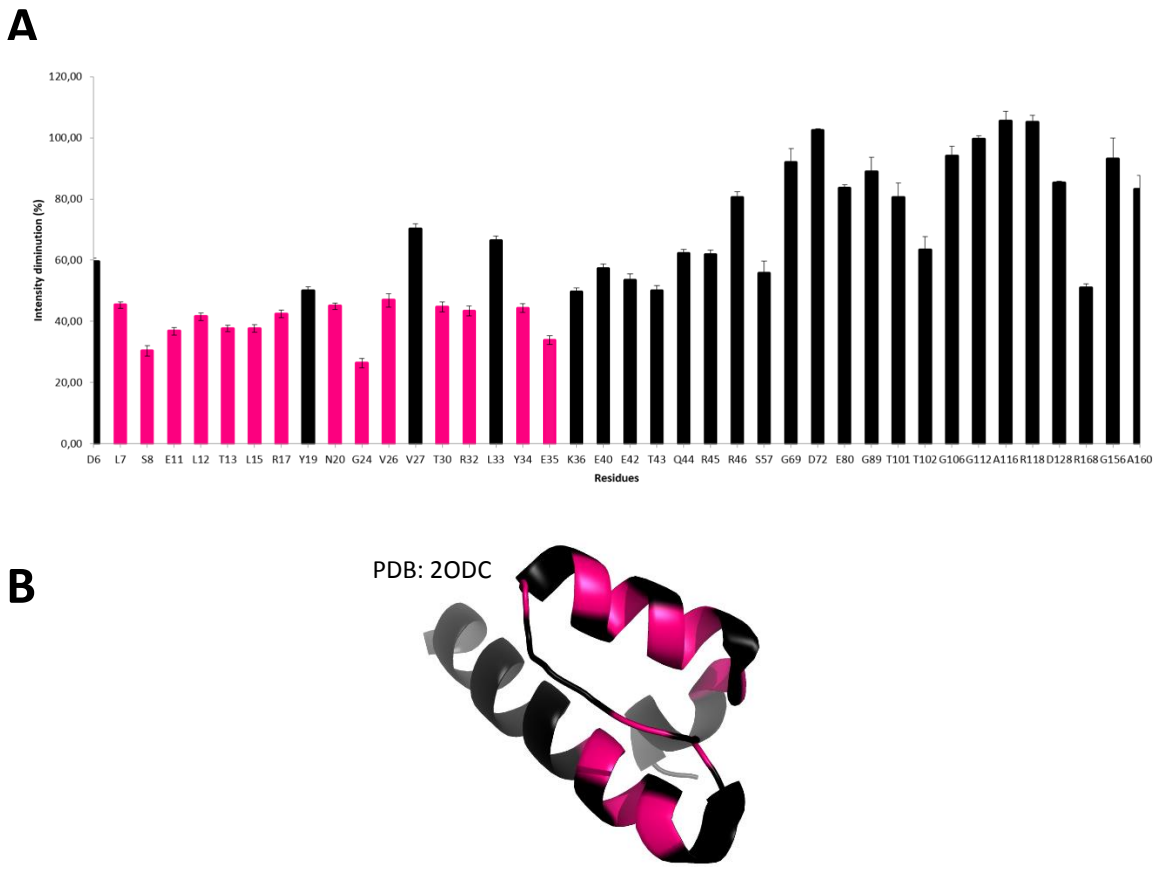


Figure 89 : Intensity ratio measurement after addition of one M1b01 equivalent onto ^{15}N EmN.

(A) Ratio as a function of the sequence. Bars corresponding to peaks losing more than 50% of intensity after M1b01 addition are labelled in pink. (B) Three-dimensional structure of the emerin LEM domain (PDB: 2ODC) with residues colored as in (A).

During the interaction analysis by NMR, another important information was observed. Indeed, I discovered that after 12h of contact between EmN and M1b01 or M1b04, the peak intensities on the interaction spectra continued to decrease (figure 90, A and B). This effect was really visible for the interaction with M1b01: indeed, after 12h, most peaks had totally disappeared. We have different hypotheses from this observation. A first hypothesis was that there was a precipitation of our NMR sample. A second hypothesis was that this intensity diminution after several hours could be due to an oligomerization of emerin induced by the interaction with M1b tails. Our last hypothesis was that this diminution could be due to an oligomerization of myosin1b which eventually carried the emerin.

To exclude our first hypothesis, I did a SDS-PAGE gel with the supernatant and the pellet of my NMR sample after a centrifugation of 10 minutes at 12000 g (figure 90, C). For M1b01, all the proteins were still in the supernatant after 12h of interaction, so protein precipitation was not the explanation for the signal decrease. Contrariwise, for M1b04, half of the sample was in the pellet and the diminution of peak intensity was not so important so we cannot exclude this hypothesis.

After these different results, I choose to study more in details this interaction through two different approaches. First, I wanted to confirm that the LEM domain was the binding site for the interaction with M1b tails and to study this, I used different emerin constructs containing or not the LEM domain and observed the impact on the interaction. In a second part, I studied emerin oligomerization in the presence of M1b01, using emerin constructs forming or not oligomers and using several techniques.

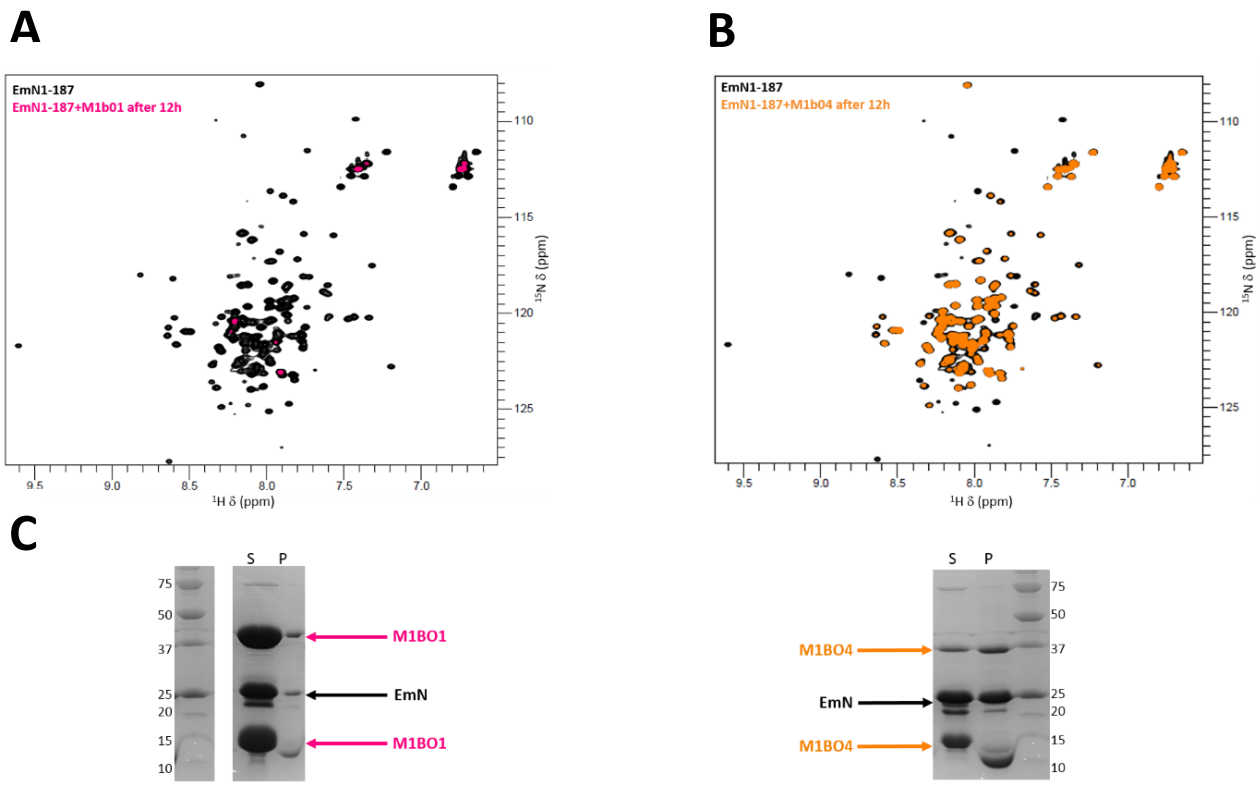


Figure 90 : Study EmN/M1b tails interaction, by NMR.

(A) Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a EmN sample at 100µM alone (in dark) or in presence of 100µM of M1b01 during 12h (in pink), in 50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT, at 293K and 700MHz (CEA Saclay). (B) Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a EmN sample at 100µM alone (in dark) or in presence of 100µM of M1b04 during 12h (in orange), in 50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT, at 293K and 700MHz. (C) SDS-PAGE gels of both NMR samples after 10 minutes of centrifugation.

a. Characterization of the emerin region important for M1b tail binding

First, to confirm that the LEM domain was necessary for this interaction, we produced the ^{15}N emerin 67-187 fragment (EmC187), which did not contain the globular domain, and added non labelled M1b01 or M1b04 (figure 91). After spectra superimposition of emerin protein alone or in presence of myosin, no difference was observed. We concluded that the unfolded region of EmN cannot bind M1b01 or M1b04 and that the LEM domain of emerin is necessary to mediate the interaction with M1b tails.

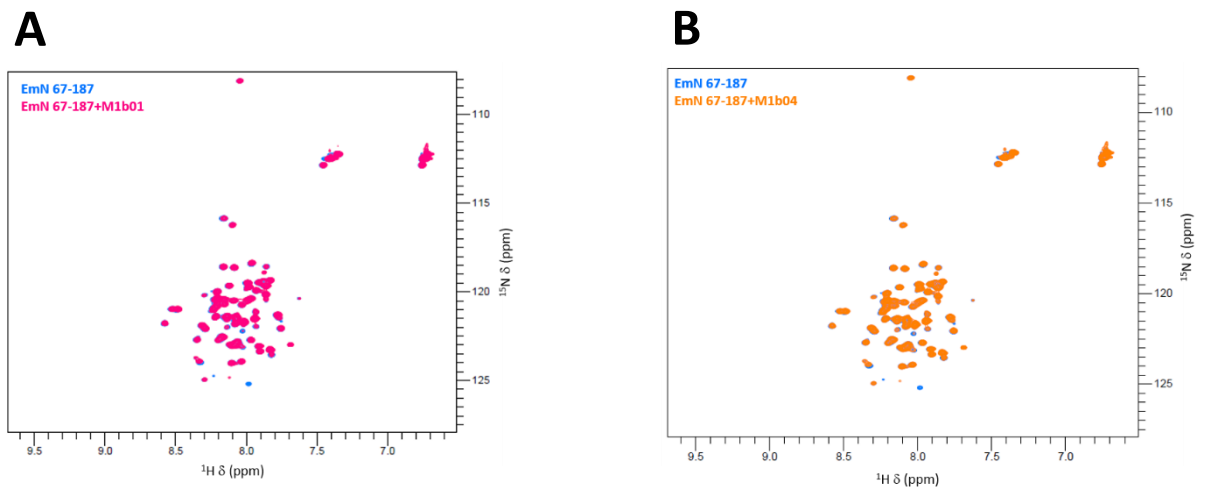


Figure 91 : Study EmC187/M1b tails, by NMR.

(A) Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a EmC187 sample at 100µM alone (in blue) or in presence of 100µM of M1b01 (in pink), in 50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT, at 293K and 700MHz (CEA Saclay). (B) Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a EmC187 sample at 100µM alone (in blue) or in presence of 100µM of M1b04 (in orange), in 50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT, at 293K and 700MHz (CEA Saclay).

To confirm my result, I produced the LEM domain of emerin (EmN49) and tested its interaction with M1b01 by NMR (figure 92, A). After having observed an intensity decrease following addition of M1b01, I quantified this decrease (figure 92, B): it yielded more than 50% for the majority of the peaks corresponding to residues of the LEM domain. I concluded that it was not possible to identify the binding site on the LEM domain; however it was clear that the LEM domain was necessary and sufficient for binding to M1b01.

Our collaborators tried to characterize this interaction between the LEM domain and M1b01 in fusion with YFP by Microscale Thermophoresis (MST) (figure 93). They confirmed that this interaction exists but obtained just an estimation of the affinity. Indeed, the LEM interacted through a weak affinity (Kd > 10µM) with M1b01.

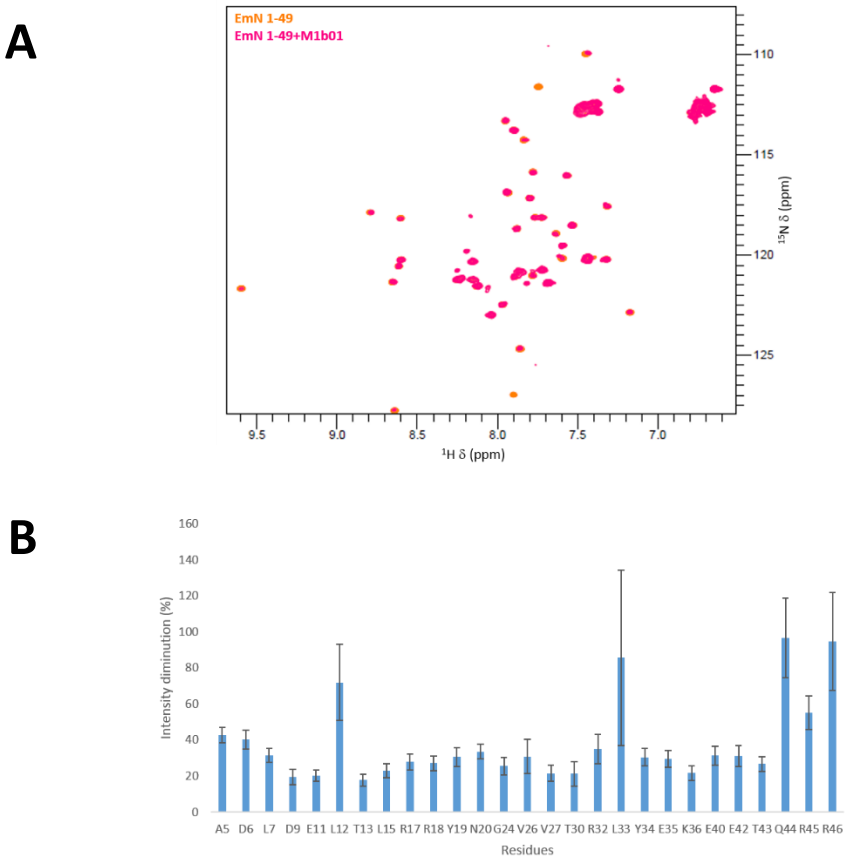


Figure 92 : Study of EmN49/M1b01 interaction, by NMR.

(A) Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a EmN49 sample at 100µM alone (in orange) or in presence of 100µM of M1b01 (in pink), in 50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT, at 293K and 700MHz (CEA Saclay). (B) Intensity ratio measurement after addition of one M1b01 equivalent onto ^{15}N EmN49.

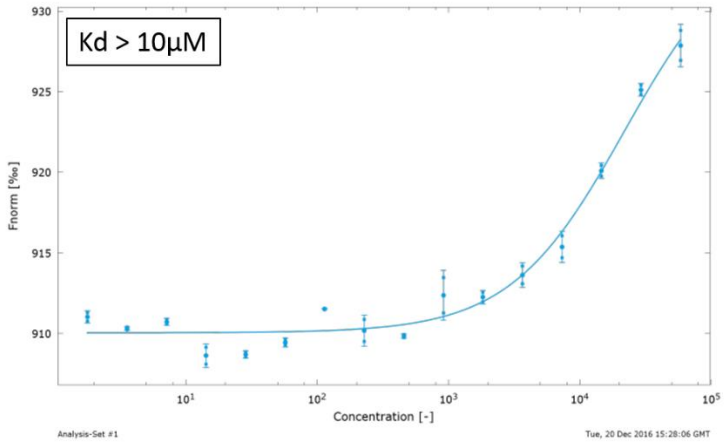


Figure 93 : Characterization of the interaction between EmN49 and M1b01, by MST.

Proteins were dialyzed in 50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT. MST experiments were carried with 0.05% Tween. Proteins were both concentrated at 100nM.

b. *Study of protein oligomerization during interaction between EmN and M1b tails*

To study this particularity of the interaction, first, we used size exclusion chromatography to observe the complex between EmN and M1b01 after 1 or 12 hours of incubation. To do this, we prepared four different samples: one containing EmN alone at 100µM, one with M1b01 alone at 100µM, one with both proteins at 100µM, in contact during 1h and the last one, with both proteins at 100µM, in contact during 12h (figure 94). Interestingly, we observed a displacement of the emerin elution peak from its initial position, to the same position as M1b01 after 1h of contact between both proteins, and the amount of displaced emerin seemed to increase over time.

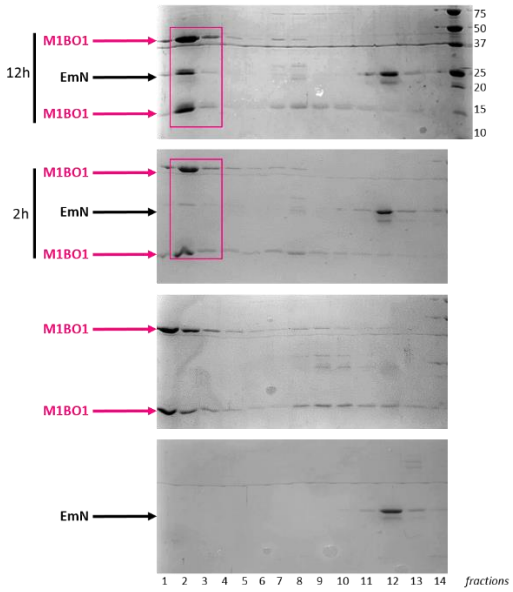


Figure 94 : Observation of the interaction between EmN and M1b01 after 1 or 12h of contact, by size exclusion chromatography.

Each gel corresponds to the result of one size exclusion chromatography. Proteins were injected in a volume of 500µl on a column GF Superdex 200 10/300GL, equilibrated in a buffer containing 50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT.

To understand this, first we observed interaction of M1b tails with an emerin construct that can form oligomers and filaments, EmN132, and with an EmN mutant that cannot form filaments, mutant Δ95-99. Both emerin constructs were able to interact (figure 95). This result was in accordance with our other results, which showed that the presence of the LEM domain is sufficient for binding.

Concerning the oligomerization question, it was interesting to observe that after 1h, signal intensity decrease was the same for both emerin constructs, whereas after 12h of interaction, the signal decreased dramatically in presence of EmN132 but did not change in presence of EmN Δ95-99. We hypothesized that the presence of M1b01 could induce emerin oligomerization.

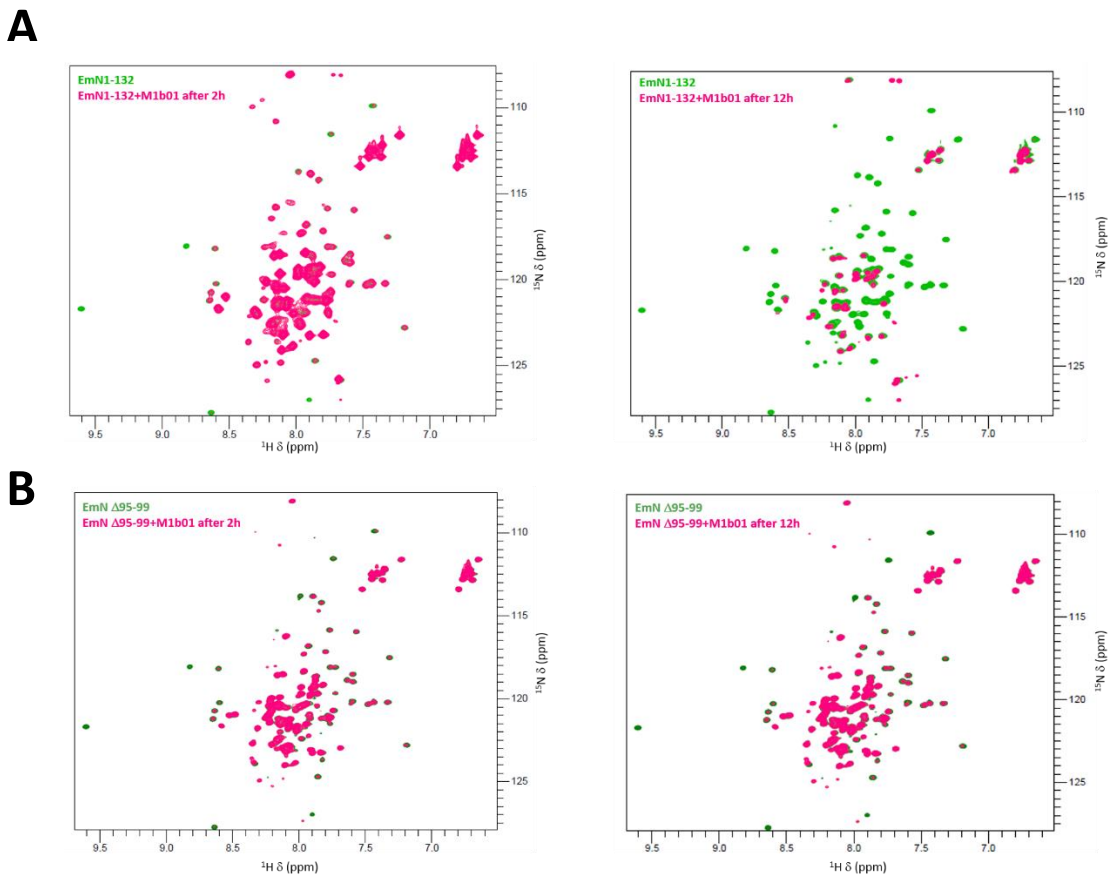


Figure 95 : Study of EmN132 and EmN Δ95-99 interaction with M1b01, by NMR.

On the left, spectrums were recorded after 1h of interaction and on the right, after 12h. (A) Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a EmN132 sample at 100µM alone (in light green) or in presence of 100µM of M1b01 (in pink), in 50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT, at 293K and 700MHz (CEA Saclay). (B) Superimposition of 2D NMR $^{1}H–^{15}N$ spectra recorded on a Δ95-99 sample at 100µM alone (in dark green) or in presence of 100µM of M1b01 (in pink), in 50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT, at 293K and 700MHz (CEA Saclay).

To study the impact of M1b01 on emerin oligomer formation, I followed the self-assembly kinetics of EmN WT alone or in the presence of M1b01, through thioflavin T fluorescence measurement (figure 96). For this experiment, both proteins were dialyzed in the kinetics buffer (50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT), and then EmN was concentrated up to 600µM and M1b01 was concentrated up to 200µM.

I prepared two samples: one with emerin alone at 300µM and one with emerin at the same concentration, in the presence of 100µM of M1b01. Samples were heated during one day at 310K and I took a fluorescence measurement at 480nm, after excitation at 440nm, every hour during 4h and the next day. For each measurement, we diluted the protein sample to 40µM in the kinetics buffer and added 10µM of thioflavin T.

Interestingly, we observed that in presence of M1b01, EmN seemed to form oligomers faster than in absence of this partner.

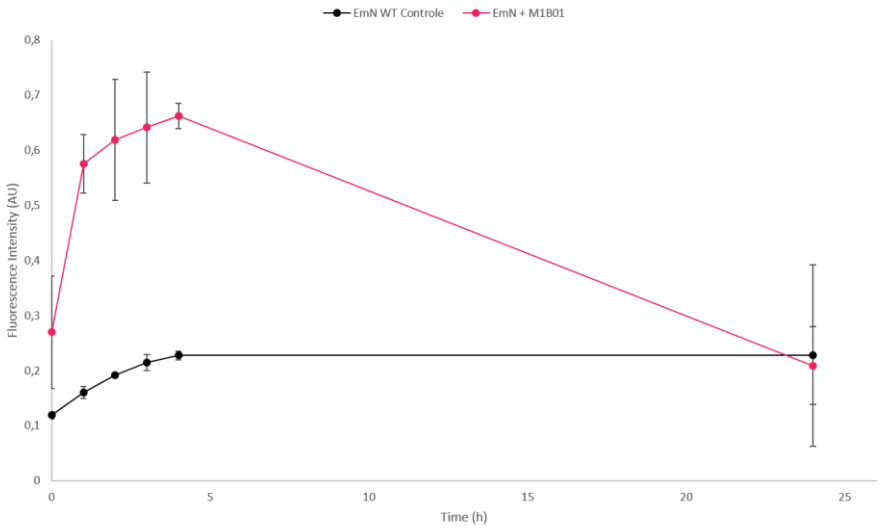


Figure 96 : Kinetics of emerin self-assembly in the absence (black) or presence (pink) of M1b01, as followed by thioflavin T fluorescence.

To confirm our hypothesis, we observed, thanks to a collaboration with the group of Dr Jean Lepault, the impact of M1b01 on emerin oligomerization by negative staining EM. We analyzed a sample of M1b01 alone concentrated at 100µM, of emerin 1-187 alone concentrated at 100µM and of both proteins, concentrated at 100µM (figure 97).

We observed the presence of filaments in the EmN alone sample, and we found that M1b01 alone also seemed to form oligomers. But when emerin was in the presence of M1b01, oligomers of this myosin seemed to be present but no emerin filaments were present anymore. We concluded that the most probable hypothesis was that M1b01 could form oligomers and these kinds of oligomers could bind to EmN overtime.

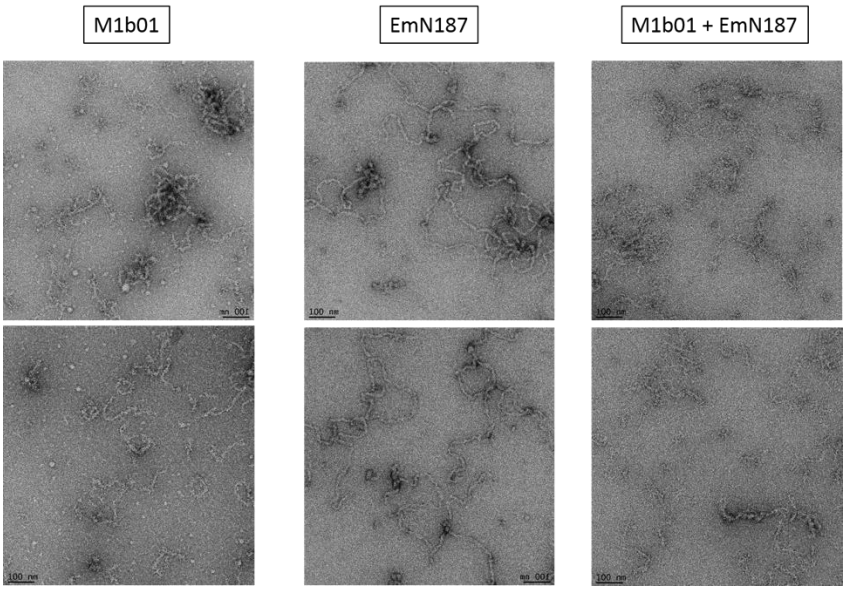


Figure 97 : Study of the interaction between emerin 1-187 and M1b01, observed by electron microscopy.

We used ammonium acetate to labelled our sample, using the technique of negative staining. All proteins were dialyzed in the same buffer (50mM Tris-HCl pH7.5, 100mM NaCl, 5mM EGTA, 0.2mM DTT) and concentrated at 100µM. 0.6mg/ml of proteins were put on the grids.

To conclude, we found an interaction between EmN and myosin, and more specifically, we found that the LEM domain of emerin could interact with the tested M1b tail constructs (M1b01 and M1b04). Then, we found that M1b01 alone seemed to form oligomers that could carry EmN overtime, but we need to confirm this result. Finally, we can hypothesize that an interaction between emerin and myosin is possible at the nuclear envelope and contributes to the formation of a complex with actin and the LINC complex during the nucleus response to a mechanical stress.

DISCUSSION AND CONCLUSION

Proteins of the inner nuclear envelope have several essential functions in different cellular processes. Some of these proteins contribute to the nuclear shape and positioning; they play a role in the regulation of nuclear mechanical properties. Some contribute to genome organization and gene expression regulation. Finally, they may participate to the DNA damage response.

During my thesis, I mainly focused on the role of three proteins, emerin, lamin A and BAF, first because those three were shown as implicated in the regulation of nuclear shape and mechanical resistance. Indeed, emerin is phosphorylated during a mechanical stress, in order to increase nucleus stiffness[107]. In response to a mechanical stress, lamin A is post-translationally modified and exhibits a partial unfolding of its Igfold domain[94]. Lamin A/C deficiency is associated with both defective nuclear mechanics and impaired mechanically activated gene transcription[95,97]. Finally, BAF interacts with chromatin in order to regulate nuclear assembly and organization[116,268]. In addition, the three proteins interact together at the inner nuclear envelope[55,172].

Despite of this, only few molecular details are known concerning these proteins and their functions. I chose to focus first on emerin structure, oligomerization and post-translational modifications in order to understand how this protein could play a role in the nucleus response to mechanical stress and to understand how its function is regulated in this situation. Then, I studied emerin interaction with lamin and BAF in order to understand how these interactions contribute to emerin function.

I. Emerin nucleoplasmic region forms oligomers *in vitro*

As a large number of other inner nuclear membrane proteins, emerin exhibits large regions predicted to be intrinsically disordered[4]. In addition, this protein possesses an N-terminal globular domain, the LEM domain, which connects emerin to a group of proteins possessing this LEM domain and sharing the ability to bind lamins and tether repressive chromatin at the nuclear periphery[269]. Finally, emerin LEM domain is known to interact with the DNA binding Barrier-to-Autointegration protein (BAF)[52].

When I began my thesis, a proteome study focused on liver nuclear envelope had shown, using a semi-quantitative measurement of protein abundance by mass spectrometry that the emerin/A-type lamin ratio was about 0.03[270]. From this ratio, it can be estimated that there is about 50,000 emerin molecules per nucleus (which corresponds to an approximate concentration of 10-100 μM). These molecules are concentrated at the inner nuclear membrane. They were reported to self-associate, *in vitro* and *in cell*, through different mechanisms[165]. Several intra and/or intermolecular interactions were observed within emerin molecules, suggesting that changes in emerin oligomeric states could be responsible for the regulation of emerin binding properties at the inner nuclear envelope. Two different configurations were proposed depending on emerin LEM domain interactions. A down configuration in which the LEM domain interacts with BAF and an up configuration, in which the LEM domain is implicated in emerin self-assembly.

At the beginning of my thesis, I tried to characterize emerin structure in solution, working together with another PhD student, Dr Isaline Herrada. For this, we used different emerin constructs (1-49, 1-132, 67-170, 67-187, 67-221 and 1-187). After assignment of emerin 1-170 NMR chemical shifts, we confirmed that emerin monomers are composed of only one globular part, the LEM domain, which is followed by a large disordered fragment[250]. Then, we showed that more than monomers, emerin 1-187 (EmN) and emerin 67-221 (EmC221) can form higher molecular weight oligomers *in vitro*.

First, I only focused on EmN oligomers. By electron microscopy, Isaline and I observed that EmN forms some spherical particles and curvilinear filaments of 10nm of diameter. Then, using Fourier Transform Infra-Red spectroscopy and thioflavin T fluorescence, we observed that these filaments are composed of β-structure [4]. To find which part of emerin was important for filament formation and to understand the role of the LEM domain in emerin self-assembly, we followed the oligomerization kinetics of different emerin constructs (see above) and found that EmN132 (1-132) was sufficient for oligomerization whereas EmC187 (67-187) cannot form filaments anymore. In addition, using solid-state NMR as well as limited proteolysis and mass spectrometry, we demonstrated that the LEM domain, as it was suggested by the group of Dr K. Wilson, can be found either as a globular domain in a monomeric EmN fragment or as part of the structural core of long curvilinear filaments formed by EmN[259].

It is interesting to observe that despite the amount of intrinsic disorder present outside of the LEM domain, a 10 kDa fragment is proteolytically protected. This could be due to an interaction between the LEM domain and a part of the unfolded region that forms a β-sheet with the LEM region, thus creating the EmN filament core.

Finally, we wondered if EmN oligomers are reversible. Indeed, we observed that dilution did not change our EmN oligomer structure *in vitro*, so our hypothesis is that maybe, *in cell*, their stability could be modified by the presence of the membrane or by post-translational modifications such as phosphorylations.

Another important experiment could have been done. Because thioflavine fluorescence was a good method to study our oligomers *in vitro*, we thought that it could be interesting to observe if thioflavine could bind EmN oligomers *in vivo*. Our collaborator, Brigitte Buendia, from Diderot University, tried once to observe thioflavine fluorescence *in vivo* in presence of EmN oligomers, without success. But first, it could be due to a protocol problem because no positive controls were done or it could be due to transitory interactions between thioflavine and EmN oligomers.

NMR analysis of the EmN mutant ΔK37 - this mutation causing cardiac defects - revealed first that the LEM domain of this mutant is destabilized and, interestingly, that this mutant forms oligomers faster than EmN WT.

After confirmation that the resulting oligomers have a structural organization similar to that of EmN WT oligomers (using solid-state NMR, electron microscopy and limited proteolysis), we suggested that because the first step of emerin self-assembly is largely facilitated by deletion of K37, a rearrangement of the LEM domain structure is necessary for EmN oligomer assembly.

In parallel, together with another PhD student from the Institut de Myologie (Paris), Nada Essawy, I worked with two other mutants showing a defect in emerin LEM domain and associated with the same symptoms as ΔK37: mutants P22L and T43I (G. Bonne, personal communication). We observed, using thioflavin T fluorescence and negative staining EM, that these two mutants form, as ΔK37, filaments faster than EmN WT, and that their filaments are structurally similar to that of EmN WT. Contrariwise to the mutant ΔK37, these mutants exhibit a well folded LEM domain. We concluded that the kinetics rate increase observed for the three LEM domain mutants could be due to the fact that in these mutants the interaction between the LEM domain and the unfolded part of EmN is favored.

In addition, we studied self-assembly of 5 EmN mutants causing Emery-Dreifuss Muscular Dystrophy (S54F, Q133H, P183H/T and Δ95-99), using thioflavin T fluorescence and negative staining EM. In these cases, mutants self-assembled similarly or less than EmN WT. Even if the mutant Q133H seemed to form oligomers slower than EmN WT, only EmN Δ95-99 could not form oligomers anymore, *in vitro*. Interestingly, deleted amino acids of this mutant are localized in a region predicted as partially folded by Disopred2.0 software and by NMR chemical shift analysis.

Self-assembly of emerin WT and mutants was then studied in cells, using *in situ* Proximity Ligation Assay (PLA). Consistently with our *in vitro* results, emerin WT and two mutants (P183T and ΔK37) were found to self-associate, whereas emerin Δ95-99 and Q133H created less emerin-emerin proximities than emerin WT at the nuclear periphery of transfected cells.

Finally, together with another PhD student Florian Celli, I observed that another construct of emerin, which did not contain the LEM domain, EmC221 (67-221), could form oligomers *in vitro*. Based on thioflavin T fluorescence and EM studies, we found that EmC221 could form similar curvilinear filaments of 10nm of diameter as EmN, but with a faster self-assembly kinetics. Thus, two types of self-assembly mechanisms can be described for emerin. We could never purify the whole nucleoplasmic region of emerin, from aa 1 to aa 221, comprising both EmN and EmC221 fragments.

This region systematically aggregated during the dialysis against a buffer without urea. Such aggregation process might be due to a bad timing of the different self-assembly interaction steps in the absence of membrane anchoring.

1. EMERIN OLIGOMERS DIRECTLY BIND LAMIN A/C

Co-immunoprecipitation assays[166], surface plasmon experiments[170] and yeast two hybrid assays[166] showed that a direct interaction exists between lamin A and emerin.

Then, blot overlay assays revealing emerin binding to lamin A showed that mutations in emerin region from aa 70 to aa 170 inhibit lamin A binding[245]. On the opposite, mutations of LEM domain residues 24 to 27 into alanine seem to favor lamin A binding. A later report from the same group revealed that two emerin regions can independently bind to mature lamin A tail: regions 1-132 and 159-220, with fragments 119-132 and 159-169 being necessary for these interactions, respectively[165].

During my thesis, I studied this emerin-lamin interaction and found first, that emerin monomers could not bind directly to the lamin A/C tail. It was a very contradictory result compared to results obtained by other groups. Our first hypothesis was that the putative presence of emerin aggregates in the samples with a low concentration of urea at the end of the purifications reported in the litterature[165,190] could favor a direct interaction with the nuclear lamina. A second hypothesis was that because one of these experiments used a GST-emerin construct purified in urea[165], the fusion protein was not well folded. Also, we tried to produce this kind of construct and it was not soluble in our conditions. Finally, our last hypothesis was that emerin forms oligomers that could bind lamin A/C.

Finally, we confirmed that EmN oligomers can directly bind to the Igfold domain of lamin A/C. In more details, only EmN oligomers, and not EmC221 oligomers, were found to interact with the Igfold. The LEM domain is necessary for EmN self-assembly, which is essential for EmN binding to lamin A. Destabilization of the LEM domain through deletion of K37 favors EmN self-assembly. This is consistent with the observation of the group of K. Wilson that mutation of the hydrophobic core of the LEM domain (G24-P25-V26-V27 into AAAA) favors lamin A binding[245]. Altogether these data strongly suggest that destabilization of the LEM domain is necessary for EmN self-assembly and subsequent lamin A binding. We did not observe such emerin – lamin interaction in the case of laminB1 tail; K. Wilson and co-workers observed only a weak binding of emerin regions 1-132 and 159-220 to the lamin B1 tail[165]. I conclude that this interaction between emerin and lamins is specific to EmN oligomers and the Igfold of lamin A/C.

As EmN mutant ΔK37 self-assembles faster than EmN WT, I hypothesized that it still binds to the Igfold of lamin A/C. This was confirmed *in vitro* by NMR. Concerning the two other LEM domain mutants P22L and T43I, Nada Essawy will now try to characterize their interaction with the Igfold of lamin A/C.

Then, because EmN Δ95-99 does not self-assemble in our conditions, we thought that this mutant will not interact with lamin A/C *in vitro* and *in cell*. Concerning the *in vitro* part, we showed by NMR that this mutant cannot interact directly with the Igfold of laminA/C. Then, *in cells*, by PLA, while emerin WT and P183T were shown to be in closed proximities with lamin A/C, for variants Q133H and Δ95-99, the amount of emerin-lamin A/C proximities is decreased, suggesting that emerin self-assembly defects impact lamin recognition in cells[4].

In order to identify which part of the Igfold is involved in EmN oligomer binding, we decided to produce several ^{15}N labelled mutated Igfolds and to observe the impact of addition of EmN oligomers on their ^{1}H-^{15}N HSQC spectrum. This idea was to find mutants that could not bind to EmN oligomers. We tested mutants that cause diseases (R435C and R471C) and mutants of the parallel β-sheet region that is characteristic of the lamin fold (T496E, H506E, P509A and V513A)[42]. We observed by NMR, that R435C, R471C, T496E, P509A and V513A significantly impaired binding to EmN oligomers, whereas H506E completely abolished the binding.

Altogether, my results demonstrate that a surface present on the lamin A/C Igfold domain binds to self-assembled emerin fragment EmN and not EmC221 and this lamin-emerin interaction seemed to be specific to the Igfold domain of lamin A/C.

2. IMPACT OF PTM ON EMN SELF-ASSEMBLY

Emerin was proposed to be phosphorylated by the Src kinase in response to a mechanical stress[107]. In more details, Src kinase phosphorylates emerin on tyrosines 74 and 95 and mutations of these two residues impair nuclear response during application of a mechanical force on the nucleus. In addition, lamin A/C is recruited to the LINC complex in order to reinforce nuclear stiffness. More recently, it was also proposed that application of a force on the nucleus increases emerin level at the outer nuclear membrane without altering emerin protein level[249].

During the first year of my thesis, I confirmed *in vitro*, using NMR and mass spectrometry, that EmN is phosphorylated by Src on tyrosines 74 and 95. Interestingly, the emerin region phosphorylated by Src is important for EmN self-assembly. Indeed, EmN Δ95-99 cannot form oligomers in our conditions. Moreover, this region is important for direct binding between emerin and lamin A/C.

Indeed, mutation of D70-D72-Y74 into alanines as well as deletion Δ95-99 impair lamin A binding[245]. On the bases of these different observations, I expressed two opposite hypotheses:

(1) Because the region containing tyrosine 95 is important for EmN self-assembly, because only EmN oligomers bind lamin A/C and because recently, a study suggested that during a mechanical stress, emerin level at the outer nuclear membrane increases, I hypothesized that phosphorylation of tyrosines 74 and 95 by Src could impair emerin oligomer assembly and thus disrupt emerin binding to lamin A/C. In these conditions, monomeric emerin could be displaced to the outer nuclear membrane in order to let lamin A/C free to interact with the LINC complex, and thus to reinforce nuclear stiffness.

(2) I have a second hypothesis that does not take into account the fact that emerin level at the outer nuclear membrane increases during a mechanical stress (indeed this observation was reported only in one recent paper). Because EmN Δ95-99 cannot self-assemble or interact with lamin A/C, and because emerin phosphorylation by Src was shown to be important for nuclear response during a mechanical stress, Src phosphorylation could favor emerin self-assembly in order to increase emerin-lamin binding at the inner nuclear envelope and thus allow interaction between this complex and the LINC complex in order to increase nuclear stiffness.

To understand the impact of Src phosphorylation, I tried to observe the influence of these phosphorylations on EmN oligomer assembly, with the aim of subsequently studying their effect on lamin binding.

First, I tried to follow EmN oligomerization kinetics using thioflavin T fluorescence on both a non-phosphorylated EmN sample and an EmN sample phosphorylated by Src. I did this experiment several times, but it was difficult to reproduce and interpret it.

This can be explained by two phenomena; first, because Src phosphorylations impact tyrosines and because it was reported that thioflavin T interacts preferentially with tyrosine-rich cross–β surfaces, we thought that maybe, phosphorylation on tyrosines could impact thioflavin T fluorescence in the presence of phosphorylated EmN. In this case, this technique is not adapted to follow impact of Src on filament assembly.

But a second idea could be that our experimental conditions are not adapted. Indeed, to do this experiment, I used every time EmN samples that were in presence of $MgCl_2$, ATP and then, with or without Src, during 12h at 303K and before following oligomer kinetics, I did not change the buffer nor re-purify protein samples. This means that during kinetics, other factors, such as the presence of $MgCl_2$, could have an impact on oligomer assembly, or Src kinase present in the sample could sometimes still be active. For this reason, Nada Essawy, who will now work on this project, will try to reproduce this experiment, but will first phosphorylate EmN before cleavage of its tag, in order to re-purify each EmN sample, after the phosphorylation kinetics and before following self-assembly by thioflavin T fluorescence.

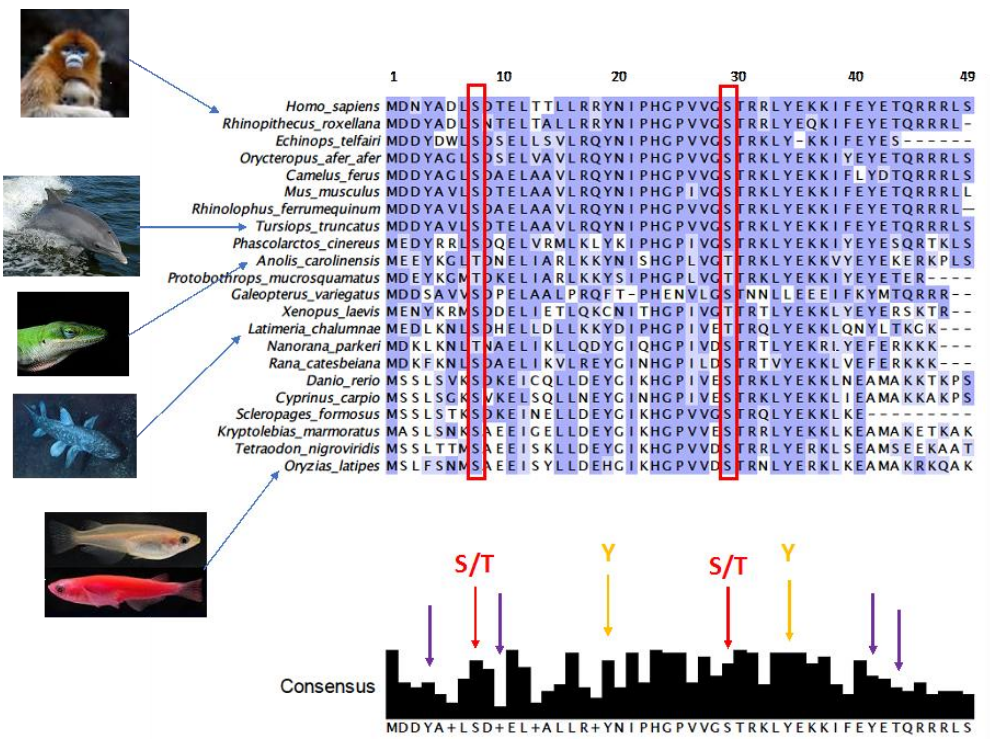
I also tried to observe the impact of Src phosphorylation on EmN self-assembly using SDS-PAGE and negative-staining EM. My first results suggested that Src phosphorylation impacts EmN oligomerization, because I observed the presence of filaments in my non-phosphorylated sample whereas aggregates were observed in the phosphorylated sample. This experiment has to be reproduced by Nada Essawy and again, re-purification of EmN after phosphorylation will be introduced in the protocol to avoid modifying the oligomerization kinetics because of the presence of additives related to the kinase in the buffer.

My preliminary results are in favor of my first hypothesis; Src phosphorylation seems to impair filament assembly, which could disrupt emerin-lamin binding during a mechanical stress and favor emerin displacement from the inner nuclear envelope to the outer nuclear envelope, in order to allow lamin A/C interaction with the LINC complex and thus reinforce nuclear stiffness.

To finish on emerin phosphorylation, I thought that in addition to their implication during a mechanical stress, phosphorylations could regulate emerin oligomerization and interactions also in a cell cycle context, like assembly or disassembly of the nuclear membrane during mitosis.

The LEM domain of emerin plays a critical role in EmN self-assembly. It is highly phosphorylated in cells (see https://www.phosphosite.org/proteinAction.action?id=2624&showAllSites=false) and protein sequence analysis of several emerin LEM domain from distinct animals (figure 98, A) shows that several of its serines, threonines and tyrosines are well conserved.

A



B

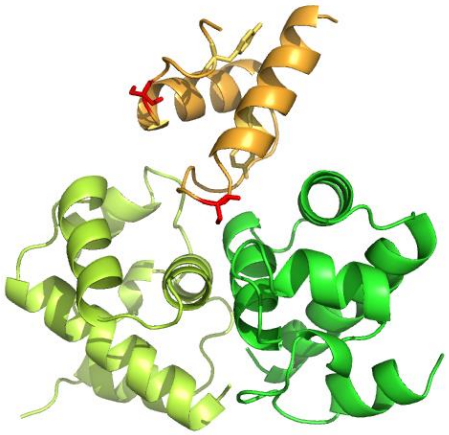


Figure 98 : Emerin LEM domain phosphorylation sites.

(A) The Phosphosite database reveals that Tyr4, Ser8, Thr10, Tyr19, Ser29, Tyr34, Tyr41 and Thr43 are phosphorylated in cells. (B) Emerin LEM domain 3D structure (PDB code: 2ODC) in orange, with conserved serines observed as phosphorylated represented in red.

In the case of emerin LEM domain, we were interested by Ser8 and Ser29, because they are conserved and phosphorylated in cells. Moreover, Ser8 is predicted to be phosphorylated by CK2; indeed, CK2 phosphorylates sites characterized by the motif S/T-X-X-E. Both serines are located at the N-terminus of a large α-helix (figure 98).

Phosphoserine is the most stabilizing amino acid for α-helices when located at the Ncap, N1, N2, and N3 positions, whereas it is destabilizing in the interior or at the C-terminus of the helix[271]. As LEM domain folding and EmN self-assembly are related, I hypothesized that phosphorylation of these serines might disfavor EmN self-assembly.

To go further, we hypothesized that EmN phosphorylation by CK2 could lead to the stabilization of the LEM domain structure, which will be in disfavor of filament assembly. Then, if CK2 phosphorylation disfavors EmN filament assembly, this phosphorylation could impair lamin-emerin binding.

During my thesis, I did preliminary studies of EmN phosphorylation by CK1 and CK2 as well as in cell extracts in vitro using NMR. No phosphorylation was observed *in vitro* with CK1 and CK2, in our conditions. We have to improve our phosphorylation conditions to go further into the study of the impact of phosphorylation by CK2 on emerin self-assembly and lamin binding.

Concerning emerin phosphorylation in cell extracts, I confirmed by NMR that EmN132 is modified after addition of cell extracts (293T cell extracts) and more, that the LEM domain seemed highly modified after addition of cell extracts. To go further into this study, we can imagine to test different protocols of phosphorylation in cell extracts. Because emerin mutations cause myopathies, it would be interesting to observe emerin modifications in different cell types, notably in myoblasts. Then, we can imagine to observe emerin modifications during specific cell cycle phases or after a mechanical stress.

II. Nuclear envelope interactions that could be important for structure and mechanical properties of the nucleus

1. A FIRST MODEL OF THE INTERACTION BETWEEN NUCLEAR ENVELOPE AND DNA

An important part of my thesis was focused on the emerin-lamin interaction. Indeed, it was well established that both proteins were interacting directly *in vitro* and in cells and but no molecular details were known. As I already explained, we confirmed that a direct interaction exists between emerin oligomers and the Igfold of lamin A/C. In addition, it was reported that the two proteins could interact through a third one, BAF, in order to form a ternary complex[175,253,254]. Concerning the emerin-BAF interaction, the 3D structure of this complex was already solved by NMR[52]; it was shown that the LEM domain of emerin can bind a dimer of BAF.

We decided to work on these interactions using several techniques such as NMR, ITC or gel filtration and we confirmed first that the LEM domain of emerin is able to bind BAF with a ratio 1:2. Then, we found that the Igfold of lamin A/C can bind a dimer of BAF too and more, we observed that the three proteins form a ternary complex. In addition, no interaction between the lamin B1 tail and BAF was observed, which demonstrated that the interaction between lamin and BAF is specific to the Igfold of lamin A/C. Finally, we obtained the X-ray structure of the ternary complex by crystallography and confirmed that the lamin A/C Igfold domain, in addition to directly interact with EmN oligomers, could also interact with the emerin LEM domain through a dimer of BAF. Our result are consistent with the observation by the group of K. Wilson that binding of ^{35}S-emerin to prelamin A tail-containing strips was enhanced fourfold by wild-type BAF[172].

I superimposed our X-ray structure with the one already available for BAF in interaction with DNA (PDB code: 2BZF)[111] in order to propose a first model of the interaction between the nuclear envelope and chromatin (figure 98). This model suggests that the BAF dimer is able to simultaneously bind to DNA, emerin and lamin A/C. In addition, it led me to another observation. Indeed, whereas BAF WT increases binding of ^{35}S-emerin to prelamin A tail, BAF S4E has no effect on this binding[172].

In addition, binding of BAF to the prelamin A tail is also significantly reduced in the case of BAF mutant S4E. And finally, in cells that overexpress BAF (that is phosphorylated by endogenous kinases) or BAF-S4E, but not BAF-S4A, emerin fails to localize at the nuclear envelope. Because of these different results. I observed the position of Ser4 on BAF to define its localization compared to the emerin and lamin binding surfaces (figure 99). Surprisingly, this serine did not localize on BAF binding surfaces to the emerin LEM domain or the Igfold of lamin A/C but on the BAF binding surface to DNA. Consistently, it was reported that phosphorylation of BAF Ser4 impairs BAF binding to DNA[172]. I now wonder why a phosphorylation that affects BAF binding to DNA can disrupt formation of our ternary complex. In addition, BAF interacts directly with histones H3 and H4 and more, three methyl marks (H3-K4-Me2, H3-K9-Me3, H3-K79-Me2) are increased in BAF-overexpressing cells[272,273]. To go further into our structural analysis of the interaction between the nuclear envelope and DNA, it would be important to understand how our ternary complex lamin/BAF/emerin interacts with DNA and histones.

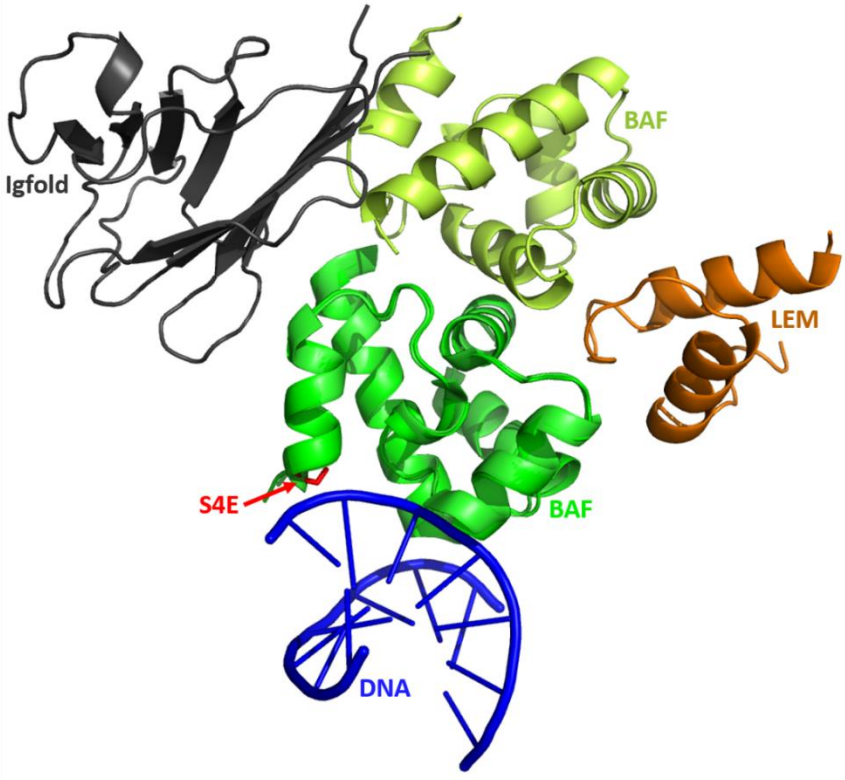


Figure 99 : A first model of the interaction between nucleoskeleton and DNA at the inner nuclear envelope.

To obtain this model, I superimposed my X-ray structure to the 3D structure of BAF in interaction with DNA (PDB: 2BZF). Here, only one DNA fragment is represented, whereas in the crystal, both BAF monomers interact with DNA. In red is represented BAF Ser4 that is phosphorylated during the cell cycle.

My structural study also revealed that the Igfold of lamin A/C interacts with BAF through a surface containing residues mutated in progeroid syndromes. During the last part of my thesis, working with Ambre Petitalot, we tried to characterize the impact of these Igfold mutants, which cause different progeroid syndromes (R435C, K542N, R471C, R527C and A529V) on BAF binding *in vitro*. In parallel, Dr Brigitte Buendia (Uni Paris Diderot) was working on the impact of these mutations in cells, using *in situ* Proximity Ligation Assays. We observed *in vitro* that four mutations (R471C, R527H and A529V and K542N) decreased the affinity of the BAF/Igfold interaction whereas one mutation (R435C) seemed to completely disrupt the interaction. *In vivo*, less proximities between BAF and each mature lamin A mutants (R435C, K542N, R527C and A529V) were observed compared to lamin A WT.

To go further, in the well-studied progeria syndrome called Hutchinson-Gilford Progeria Syndrome (HGPS), because of a deletion of 50 amino acids in the C-terminal region of prelamin A, containing a critical site for maturation[37,38] or because of a mutation in the ZMPSTE24 enzyme, involved in lamin A maturation[36,274], lamin A exhibits a maturation defect. The presence of a permanently farnesylated prelamin A mutant called progerin is then observed in cells. The lamin mutation causing progerin is heterozygous, which means that only part of the lamin A proteins possesses a maturation defect in HGPS, but this is sufficient to cause the accelerated aging disease. After my structural study, I thought that maybe a common mechanism exists between autosomal recessive progeroid syndromes due to a homozygous mutation in the BAF binding region and the autosomal dominant HPGS disease. First, I hypothesized that (1) the farnesylated prelamin A (progerin) could bind to BAF, which could induce a defect in mature lamin A binding to BAF. Consistent with this hypothesis, different prelamin A forms were shown to interact with BAF *in vivo*[218]. In addition, it was demonstrated that progerin could bind BAF in HEK293 cells[219]. Finally, it was shown that a mutation in ZMPSTE24 (which normally cleaves the farnesylated lamin C-terminus) favors BAF nuclear localization[275]. It was also shown that the presence of progerin induces a loss of heterochromatin and H3K9 trimethylation[151,152,213,214], whereas BAF was shown to bind H3 in order to increase methyl marks[272,273]. We deduced that presence of progerin could impair, in addition to BAF-lamin A/C binding, BAF binding to histone H3.

But a second hypothesis is possible. (2) Indeed, a defect in lamin-BAF interaction could induce a partial defect in prelamin A maturation that would provoke the accelerating aging syndrome.

Consistently with this idea, it was observed that prelamin A is accumulated in cells from a patient with mutation R527H causing mandibuloacral dysplasia[275].

During the end of my thesis, we did preliminary studies about the impact of lamin maturation defects on the BAF-lamin interaction. I produced two different prelamin A constructs (IgfoldProgerin [389-614] and ProgerinCter [567-614]) in order to test their ability to bind BAF *in vitro*, using ITC and NMR. First, I observed that the fragment IgfoldProgerin, corresponding to the lamin A tail in HGPS, could bind BAF with the same affinity and stoichiometry than the Igfold alone and more, the same binding surfaces were identified. Then, I tried to characterize the interaction between BAF and the C-terminus of the progerin alone (ProgerinCter), by NMR and ITC, but no clear and strong interaction was identified. These experiments have to be reproduced. In addition, both progerin peptides that I used were not farnesylated, so Florian Celli will try to co-express the two constructs of progerin with a farnesyltransferase, in order to understand the impact of farnesylation on progerin-BAF binding and then, to observe if farnesylation induces a defect in BAF-histone binding.

To conclude about emerin-lamin binding, I showed that the lamin A/C Igfold domain interacts with the inner nuclear membrane protein emerin through two mechanisms: it either binds directly to self-assemble EmN or it interacts with monomeric EmN through the chromatin-associated protein BAF (figure 100). We hypothesized that both mechanisms could be regulated by post-translational modifications.

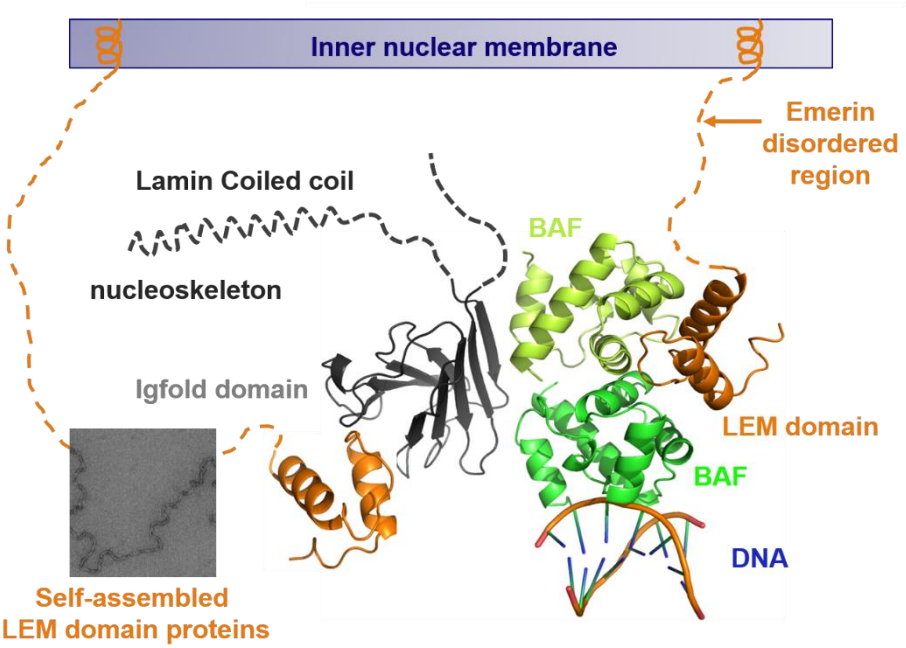


Figure 100 : Lamin A/C Igfold domain interacts with the inner nuclear membrane protein emerin through two mechanisms.

2. EMERIN INTERACTS WITH MYOSIN1B

During my thesis, I also did a collaboration with the group of Dr Anne Houdusse (Institut Curie, Paris), in order to observe if emerin could bind myosin 1 tail *in vitro*. Several results have prompted us to be interested in this possible interaction. First, nuclear myosin I was shown to be localized in the nucleus, by ultrastructural studies[263]. Then, proteomic studies revealed that emerin binds to nuclear myosin I[267].

We searched for a direct interaction between different constructs of our emerin (1-187, 1-132, 1-49 and 67-187) and 3 constructs of their myosin tails (M1b01, M1b04 and M1c01), associated with calmodulin, *in vitro*. First, we found a direct interaction between emerin and myosin 1 tail, and more, we observed that this interaction was specific to myosin 1b; indeed, no interaction was observed between M1c01 and emerin. In addition, we found that the LEM domain was necessary and sufficient for binding to myosin 1b tail. Furthermore, we observed protein oligomerization during emerin binding to myosin1b. Two hypotheses were possible. (1) First, we thought that myosin 1b tail binding to emerin could favor emerin oligomerization. (2) Then, we supposed that instead of emerin, maybe myosin 1b tail could form oligomers that could carry emerin overtime.

After several *in vitro* studies, using thioflavin T fluorescence, NMR and electron microscopy, it seemed that the second hypothesis is the real one, but we have to confirm our results.

Our hypothesis concerning the existence of an interaction between emerin and myosin at the nuclear envelope is that this complex could contribute to the formation of a higher molecular weight complex with actin in order to maintain the structural integrity of the nucleus. Consistent with our hypothesis, a very recent study demonstrated that emerin binds to non-muscle myosin IIA at the outer nuclear membrane in order to increase F-actin polymerization at this site[249].

MATERIAL AND METHODS

I. PLASMID CONSTRUCTS

1. EMERIN NUCLEOPLASMIC FRAGMENTS

All emerin fragments (1-187, 67-221, 67-187 and 1-49) were expressed as N-terminal octa-histidine fusions with a TEV cleavage site between the tag and the emerin fragment. Expression vectors coding for these fragments were purchased from Genscript, who cloned an optimized emerin gene into a pETM-13 vector provided by the lab. All emerin mutant expression vectors (ΔK37, P22L, T43I, S54F, Q133H, P183H, P183T and Δ95-99) were obtained by mutagenesis using the Quikchange (Agilent) kit from the WT emerin 1-187 expression vector. Only mutant Y4C-E35C was expressed by mutating the vector coding for WT emerin 1-132.

2. LAMIN A/C IGFOLD FRAGMENT (TAIL REGION COMMON TO LAMINS A AND C)

The pGEX-4T vector coding for the Igfold of WT lamin A/C (411-566) was a gift from the lab of Prof Howard J. Worman (Department of Medicine, College of Physicians & Surgeons, Columbia University, New York). The resulting protein contained a GST tag, a thrombin cleavage site and a lamin A/C fragment from amino acid 411 to amino acid 566. All mutant expression vectors (A529V, A529T, K542N, M540T, T528M, R435C, R471C, T496E, W498E, P508A, P509A, H506E, V513A and R482W) were obtained by mutagenesis using the Quikchange (Agilent) kit from the WT expression vector.

3. IGFOLDPROGERIN (PROGERIN TAIL)

The tail region from progerin was also expressed using a pGEX-2T expression vector provided by the laboratory of Dr Brigitte Buendia from Université Paris Diderot. The resulting protein contained a GST tag, a thrombin cleavage site and the C-terminal progerin fragment from amino acid 389 to amino acid 614.

4. LAMINB1 TAIL

The tail region of lamin B1 was also expressed using a pGEX-4T expression vector provided by the laboratory of Prof Howard J. Worman (Department of Medicine, College of Physicians & Surgeons, Columbia University, New York). The resulting protein contained a GST tag, a thrombin cleavage site and the C-terminal progerin fragment from amino acid 409 to amino acid 586.

5. BAF

BAF_{CtoA} was expressed using a pET-M13 vector. The BAF_{CtoA} gene was optimized for *E. coli* production (Genscript). The resulting protein contained an N-terminal octa-histidine tag, a TEV cleavage site and BAF_{CtoA}. All BAF mutants (S4E, A12T and G47E) were expressed after mutagenesis with the Quikchange (Agilent) kit of the BAF_{CtoA} expression vector.

6. 1U2

The 1U2 domain of Man1 was expressed using a pTEM10 vector, optimized for *E. coli* production (Genscript). The resulting protein contained an N-terminal octa-histidine tag, a TEV cleavage site and 1U2.

7. MYOSIN

The different myosin fragments were produced by the team of Dr Anne Houdusse (Institut Curie, Paris).

II. BACTERIAL CULTURES

1. IN RICH MEDIUM

For all proteins, expression was performed in *E. coli* BL21 (DE3) Star. The starter culture was done at 310K overnight in LB medium containing antibiotics.

For Igfold, IgfoldProgerin, LaminB1 and 1U2 fragments, bacterial cultures were grown at 310K in LB medium with antibiotics, and after addition of the starter culture and induction at an OD of 0.8 with 0.5mM IPTG, cultures were grown at 293K overnight. For the ProgerinCter construct, bacterial cultures were induced at an OD of 0.8 with 0.5mM IPTG, and were grown at 310K during 4 hours. For the other fragments, bacterial cultures were induced at an OD of 1, with 0.5mM IPTG, and grown at 293K overnight.

Afterwards, cultures were centrifuged at 3500 rpm during 30 minutes and each pellet was resuspended in 20ml of Lysis buffer (50mM Tris-HCl pH 8, 300mM NaCl, 5% glycerol, 1% Triton TX-100 and 10mM PMSF) before being frozen at -80°C.

2. IN MINIMAL MEDIUM

This production protocol is used in two cases: to increase protein solubility and to label the protein with ^{15}N or ^{13}C for NMR experiments.

The production protocol was the same as in rich medium, except that the LB medium was replaced by a minimal medium which is composed of:

- 100ml M9 10X (60g Na_2HPO4, 30g KH_2PO4, 5g NaCl)
- 2ml Trace Element 500X (EDTA 5 g /L 13.4mM, $FeCl_3$-$6H_2O$ 0.83 g/L 3.1mM, $ZnCl_2$ 84 mg/L 0.62mM, $CuCl_2$-$2H_2O$ 13 mg/L 76µM, $CoCl_2$-$2H_2O$ 10mg/L 42µM, H_3BO_3 10mg/L 162µM, $MnCl_2$-4H2O 1.6 mg/L 8.1 µM)
- 1ml $MgSO_4$ 1M
- 0.3ml $CaCl_2$ 1M
- 1mM of Biotine
- 1mM of Thiamine
- 2g Glucose
- 0.5g $^{15}NH_4CL$

III. PROTEIN PURIFICATIONS

1. EMERIN AND BAF PURIFICATION

Both proteins were produced in inclusion bodies. Therefore, purification protocols for the different emerin and BAF fragments are almost the same.

The first step of purification consisted of a lysis step.

1) Lysis step

- Pellet thawing.
- Sonication in 3 cycles of 30 seconds, separated by 1 second, with a power of 70% and without any temperature constraint.
- Centrifugation at 20 000 rpm during 20 minutes.
- Pellet solubilization in buffer C8 (50mM Tris-HCl pH8, 150mM NaCl, 20mM Imidazole and 8M urea) during one hour at room temperature on a wheel.
- Centrifugation at 20 000 rpm during 20 minutes.
- Recovery of the supernatant (which contains urea).

2) Purification

After the lysis step, the soluble extract is loaded onto a Ni-NTA column (GE-Healthcare) equilibrated with buffer C8. After two hours of incubation at room temperature, the column is washed with buffer C8 and then, proteins are eluted with buffer E8 (50mM Tris-HCl pH8, 150mM NaCl, 1M Imidazole and 8M urea). Proteins are then diluted at an OD of 1 for emerin fragments and an OD of 0.5 for BAF fragments, before being refolded by dialysis in D1 buffer (50mM Tris-HCl pH8, 30mM NaCl) for emerin fragments and buffer D2 for BAF fragments (50mM Tris-HCl pH8, 150mM NaCl), overnight.

The next day, proteins are dialyzed twice during two hours in the same buffer as before. Then they are concentrated to obtain 20ml. After addition of 2mM DTT and one TEV aliquot at 1mg/ml, proteins are incubated during 2h at room temperature in a falcon tube and deposited onto a Ni-NTA column. After column incubation during 30 minutes at room temperature, proteins are recovered in the flow-through and dialyzed in the selected buffer depending on the experiments.

2. IGFOLD PURIFICATION

Every Igfold construct (WT, mutants and IgfoldProgerin) was soluble in bacteria. Like for all proteins, the first purification step consists of a lysis step.

1) Lysis step

- Pellet thawing.
- Sonication in 3 cycles of 30 seconds, separated by 1 second, with a power of 70% and at 277K.
- Incubation with 2µl of benzonase (Sigma) and 100mM $MgCl_2$, during 20 minutes at room temperature.
- Centrifugation at 20 000 rpm during 20 minutes.
- Recovery of the supernatant.

2) Purification

Because of their GST-Tag, the Igfold fragments were purified on a GSTrap FF column (GE Healthcare). The column is first washed with buffer A (50mM Tris-HCl pH8, 150mM NaCl, 1mM DTT). Then, proteins are added onto on the column and incubated during 2h at 277K on a wheel. After incubation, the column is washed with buffer A, followed by buffer B (50mM Tris-HCl pH8, 1M NaCl, 1mM DTT) to remove in particular DNA and again with buffer A. Finally, proteins are incubated in 10ml of buffer A and the thrombin protease is added overnight at 277K.

The next day, proteins are eluted in the flow-through, dialyzed in a selected buffer corresponding to further experiments and kept at 277K.

3. PROGERINCTER PURIFICATION

In the case of this construct, the lysis step is performed using the protocol described for the Igfold construct.

After the lysis step, the supernatant is loaded onto a GSTrap FF column (GE-Healthcare) equilibrated with buffer A. Then, proteins are incubated during 2h at 277K, on a wheel. After incubation, column is washed with buffer A, followed by buffer B (50mM Tris-HCl pH8, 1M NaCl, 1mM DTT) to remove DNA and again with buffer A. Afterwards, proteins are incubated in 10ml of buffer A and the TEV protease is added during 2h at 277K. Finally, proteins are recovered in the flow-through, dialyzed in a selected buffer corresponding to further experiments and kept at 277K.

4. LAMINB1 AND 1U2 PURIFICATION

In the case of these two constructs again, the lysis step is performed using the protocol described for the Igfold construct.

After the lysis step, the supernatant is loaded onto a Ni-NTA column (GE-Healthcare) equilibrated with buffer A. Then, proteins are incubated during 2h at 277K, on a wheel. After incubation, column is washed with buffer A, followed by buffer B (50mM Tris-HCl pH8, 1M NaCl, 1mM DTT) to remove DNA and again with buffer A. Afterwards, proteins are eluted with buffer E (50mM Tris-HCl pH8, 150mM NaCl, 1M Imidazole and 1mM DTT) and dialyzed against buffer A 3 times during 2 hours. The next day, the TEV protease is added during 2h at room temperature in a falcon tube, and the protein sample is loaded onto a Ni-NTA column during 30 minutes at room temperature. After column incubation, proteins are eluted in the flow-through and dialyzed in the selected buffer depending on experiments.

IV. LIQUID NMR

NMR experiments were performed on 600, 700 and 750 MHz Bruker Avance spectrometers, all equipped with a triple resonance cryogenic probe. Two-dimensional ^{1}H-^{15}N correlation spectra were acquired using a HSQC pulse sequence at a temperature of 283K to 303K, on a 3-mm-diameter NMR sample tube, containing 80:20% H_2O/D_2O. All NMR data were processed using Topspin3.1 (Bruker) and analyzed with CCPNMR[276].

1. ASSIGNMENTS OF THE NMR SIGNALS

a. Emerin 1-170

We assigned the NMR signals of the emerin fragment from residues 1 to 170 using two approaches. We acquired three-dimensional ^{1}H-^{15}N-^{13}C correlation spectra, using 3D HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO and HN(CO)(CA)NH experiments (table 5) at 303 K, on a 5-mm-diameter NMR sample tubes containing, on one hand, 500 µM uniformly $^{15}N/^{13}C$-labeled Emerin67-170, in 20mM Tris-HCl pH8, 30mM NaCl, 5mM DTT and 80:20% H_2O/D_2O and on the other hand, 500 µM uniformly $^{15}N/^{13}C$-labeled EmN132, in 20mM Tris-HCl pH 8, 30mM NaCl, 5mM DTT and 80:20% H_2O/D_2O.

Spectrum	Dimensions	Observed atoms
HSQC	2D	N-H
HNCACB	3D	N-H+Cα+Cβ+(Cα)-1+ (Cβ)-1
CBCA(CO)NH	3D	N-H+(Cα)-1+(Cβ)-1
HNCO	3D	N-H+C from liaison (C=O)-1
HN(CA)CO	3D	N-H + C from liaison C=O and (C=O)-1
HN(CO)(CA)NH	3D	N-H+(N-H)+1

Table 5 : List of spectra used for emerin 67-170 and 1-132 backbone assignment.

Using our 3D experiment combination and using the emerin 67-170 sample, we success to assign 96% of ^{1}H-^{15}N pairs (95 out of 99), 99% of $^{13}C\alpha$ (104 out of 105), 99% of $^{13}C\beta$ (98 out of 99) and 98% of ^{13}CO (103 out of 105) resonances of emerin 67-170 sequence.

The chemical shift data have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu) under the accession number 26654 for the emerin 67-170 fragment. And the assignment of this unstructured peptide is further described in Samson et al., *Biomol NMR Assign.* 2016 (Appendix 1).

b. *Igfold*

For the assignment of the NMR signals of the lamin A/C Igfold fragment, three-dimensional ^{1}H-^{15}N-^{13}C correlation spectra were acquired using 3D HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO and HN(CO)(CA)NH experiments at 303 K, on a 3-mm-diameter NMR sample tube containing 500 µM uniformly $^{15}N/^{13}C$-labeled IgFold411-566, in 20mM sodium phosphate pH 7, 30mM NaCl and 80:20% H_2O/D_2O. The assignment procedure took advantage of the already published assignment of lamin A/C fragment 428-547 at pH 6.3 and 303K (BMRB 5224)[42].

2. PHOSPHORYLATION KINETICS

a. *In vitro*

To study emerin phosphorylation kinetics, the protocol was the same, regardless the kinase. The only differences were the pH value of the NMR buffer and the temperature of the NMR spectra acquisition, which were chosen according to the kinase. Here are reported the conditions used for phosphorylation by the Src kinase: the first step was to produce ^{15}N labelled emerin (1-187 or 1-132) and to dialyze it, at the end of purification, in the NMR buffer (20mM Phosphate pH7, 30mM NaCl, 10mM β-mercaptoethanol, 2mM ATP and 5mM $MgCl_2$). Then, emerin protein was concentrated until 200µM and the sample was stored at 277K.

The first NMR step was to record a ^{1}H-^{15}N HSQC spectrum of the emerin protein alone, at 200µM, in presence of all kinase additives (5mM MgCl2, 2mM ATP, 2mM DTT), at 303K. This spectrum served as a reference for the rest of the study.

Afterwards, a certain amount of kinase (depending of its activity) was added to the reference NMR tube and the pH was adjusted again to 7.0. Finally, the phosphorylation kinetics was followed by acquisition of ^{1}H-^{15}N HSQC spectra every hour, at 303K.

b. In cell extracts

To study emerin phosphorylation kinetics in cell extracts, the first important condition was to purify the protein in two steps. First, emerin was purified in 8M urea and refolded by dialysis. Then, instead of cutting the His-Tag with the TEV protease, the protein was directly modified in cell extracts and repurified by affinity chromatography on an Ni-NTA column.

To prepare cell extracts, the protocol was:

- Extracts were obtained using 200µl of lysis buffer (50mM Tris-HCl pH7.5, 150mM NaCl, 1% Triton X-100, 1mM EDTA, 1mM DTT, 10mM $MgCl_2$), and protease inhibitors (we used SigmaFAST pastille) as well as phosphatase inhibitors (we used Phosphostop pastille from Roche) were added to these extracts.
- Extracts were incubated during 15 minutes on ice.
- They were centrifuged during one hour at 13000g.
- The protein amount present in the cell extracts was quantified using a Bradford test.
- The extracts were diluted to reach 15mg/ml.
- 175µl of cell extracts concentrated at 15mg/ml were added in 25µl of protein concentrated at 600µM, which already contained 25µl of ATP 1M and 25µl of D_2O.

After addition of cell extracts to our protein, the sample was let at 303K during 4 hours. Then, as I already explained, the protein was purified again on a nickel column, equilibrated with the buffer of our choice, and after elution, the buffer was changed to remove imidazole. We choose to change the buffer during protein concentration. Then, an ^{1}H-^{15}N HSQC spectrum was recorded, and another one was recorded on the same His-Tag-emerin but that was no modified by cell extracts, in order to superimpose both spectra and observed changes due to modifications in the cell extracts.

V. SOLID STATE NMR

After purification and His-Tag cleavage, proteins were dialyzed against 20mM Tris-HCl pH 8, 30mM NaCl and 5mM DTT and concentrated up to 600µM to initiate self-assembly. Then, filaments were observed by EM before to realize solid-state NMR experiments. The solid-state experiments were performed by Max Zinke, Kitty Hendricks and Prof Adam Lange at FMP Berlin. In the case of EmN and EmN132, solid-state NMR experiments were recorded on a Bruker Avance III HD 800 MHz spectrometer (18.8 T external magnetic field) or on a Bruker Avance III HD 700 MHz spectrometer (16.4 T external magnetic field). All solid-state NMR experiments for EmN delK37 mutant were recorded on a Bruker Avance III HD 700 MHz spectrometer. Temperature was determined using the water-proton signal referenced to DSS and kept at 283 K. Double quantum – single quantum spectra were recorded with magic-angle spinning at 8 kHz and C-C correlation spectra with 11 kHz MAS rate. High-power ^{1}H decoupling was performed during evolution and detection periods using SPINAL-64 at a radio frequency strength of 83 kHz. All spectra were processed using Bruker TopSpin 3.2 and analyzed with CcpNmr.

VI. ELECTRON MICROSCOPY

To observe self-assembled proteins, EmN, EmN49, EmN132, EmC221 and EmN delK37 were dialyzed against 20mM Tris-HCl pH 8.0, 30mM NaCl and 5mM DTT. They were concentrated up to 600 µM and incubated at 293 K during 1 week. Specimens were prepared by negative staining with 2 % uranyl acetate on glow-discharged carbon-coated copper grids. Data collection was performed using a Tecnai Spirit transmission electron microscope (FEI) equipped with a LaB6 filament, operating at 100 kV. Images were recorded on a K2 Base camera (Gatan, 4kx4k) at 15,000 or 4,400 magnification (pixel size at specimen level – 0.25 & 0.83 nm, respectively), at I2BC, by Dr Ana Arteni.

VII. X-RAY CRYSTALLOGRAPHY

1. CRYSTALLIZATION AND DATA COLLECTION

The Emerin/BAF/Igfold complex was purified by gel filtration (Superdex 75, 10/300 GL) in 50mM Tris-HCl pH8, 100mM NaCl and concentrated to 3mg/ml. Crystallization was initiated one week after the gel filtration and in these conditions, EmN was proteolysis and the final complex was composed of LEM/BAF/IgFold. For crystallization, 1 µL of the complex was mixed with 1 µl of the reservoir solution and equilibrated against a 500-µl reservoir by hanging drop at 277K. Crystals were grown in 25% (w/v) polyethylene glycol monomethyl ether (PEG) 3.350, 20mM Bis-Tris pH5.5, 30mM NaCl and 0.2M NH4SO4. Crystals were transferred into a solution composed of 25% PEG 3.350, 20mM Bis-Tris pH5.5 and 0.2M NH4SO4, supplemented with 27% (V/V) ethylene glycol, before being flash-cooled in liquid nitrogen.

2. STRUCTURE DETERMINATION AND DATA REFINEMENT

The 3D structure of the complex was determined by molecular replacement using Molrep in CCP4[277]. The coordinates of BAF dimer with DNA (PDB entry 2BZF), the coordinates of laminA/C globular domain (PDB entry 1IFR) and the coordinates of the emerin LEM domain (PDB entry 2ODC) were used as template. The resulting model was rebuilt using PHENIX[278], manual correction was performed with Coot[279] according to $|F_o| - |F_c|$ and $2|F_o| - |F_c|$ maps, and further refinement was carried out with phenix.refine. All structure figures were prepared with PyMOL (Schrödinger, LLC).

VIII. SELF-ASSEMBLY

Self-assembly was initiated using proteins that were dialyzed in a buffer containing 50mM Tris-HCl pH8 and 30mM NaCl. After dialysis, proteins were concentrated until 600µM, reduced with 5mM DTT and left at room temperature during one week.

IX. THIOFLAVIN KINETICS

After purification and His-Tag cleavage, proteins were dialyzed against 20mM Tris-HCl pH 8.0, 30mM NaCl and 5mM DTT and concentrated until 300µM. Oligomer assembly was followed with time by incubating proteins at 300 µM and 310 K and by regularly taking protein aliquots to be analyzed by fluorescence. These aliquots were diluted in a thioflavin T (ThT) containing buffer so as to obtain 20 µM protein and 2.5 µM ThT in 20mM Tris-HCl (pH 8.0), 30mM NaCl, and 5mM DTT. The fluorescence measurements were carried out in a 60 µl cuvette at 293 K using a fluorimeter JASCO ADP-303T. ThT fluorescence was monitored using an excitation wavelength of 440 nm and fluorescence emission was read at 480 nm.

X. SIZE-EXCLUSION CHROMATOGRAPHY

Size-exclusion chromatography experiments aiming at identifying interactions between emerin, BAF and lamin fragments (WT and mutants) were performed using a Superdex-75 10/300 GL column (GE Healthcare) pre-equilibrated with buffer G (20mM Tris-HCl pH8, 30mM NaCl, 2mM DTT). 500µl of proteins concentrated at 150µM were injected for each experiment at a flow rate of 0.5ml/min at 277K.

XI. ISOTHERMAL TITRATION CALORIMETRY (ITC)

ITC was performed using a high-precision VP-ITC calorimetry system. To characterize interactions between BAF dimer and the different lamin A/C fragments (Igfold WT and mutants, IgfoldProgerin), all proteins were dialyzed against 50mM Tris-HCl pH8, 150mM NaCl, 10mM β-mercaptoethanol and protease inhibitor (Roche). For each experiment using BAF, BAF dimer in the calorimetric cell at 288K was titrated with the lamin A/C fragment in the injection syringe.

To characterize interactions between BAF dimer and EmN, proteins were dialyzed against 50mM Tris-HCl pH8, 100mM NaCl, 10mM β-mercaptoethanol and protease inhibitor (Roche). BAF dimer (40µM) in the calorimetric cell at 288K was titrated with EmN187 (at a concentration of 200µM in the injection syringe).

Analysis of the data were performed using the Origin software provided with the instrument.

XII. ANALYTICAL ULTRACENTRIFUGATION

Sedimentation velocity experiments were realized by measuring the absorbance at 280nm, every 8 minutes, with an optical path of 1.2cm, at 42000 rpm (128 297g), at 293K. An analytical ultracentrifuge XL470 (Beckman Coulter, Palo Alto, USA) and an An-50Ti rotor were used. 400µl of protein sample were useful for each experiment and 410µl of our dialysis buffer was used as reference. Results were analyzed using Sedfit software[255].

XIII. FLUORESCENCE-BASED THERMAL SHIFT ASSAY

The thermal stability of LamIgF proteins was monitored by a fluorescence-based thermal shift assay performed with a QuantStudio 12K Flex instrument (LifeTechnologies). 10 µg of purified protein was mixed with the SYPRO Orange dye (diluted 800-fold from a 5000-fold stock solution, Invitrogen) in 50mM Tris pH 8.0, 150mM NaCl, 10% glycerol. Reactions were carried out in duplicate in a 96-well fast PCR plate at a final volume of 20 µl and each experiment was repeated at least twice independently. The samples were submitted to a denaturation kinetic from 283 to 368K at a rate of 276K/min and fluorescence of Sypro Orange dye was recorded in real time. The protein denaturation temperature (Td +/- s.e.m) was calculated using the Protein Thermal Shift software v1.3 (LifeTechnologies) as the maximum of the derivative of the resulting fluorescence curves.

REFERENCES

1. Hort, W. Quantitative histologische Untersuchungen an wachsenden Herzen. *Virchows Arch. Für Pathol. Anat. Physiol. Für Klin. Med.* **323,** 223–242 (1953).
2. Schneider, R. & Pfitzer, P. [Number of nuclei in isolated human myocardial cells]. *Virchows Arch. B Cell Pathol.* **12,** 238–258 (1973).
3. Bao, G. & Suresh, S. Cell and molecular mechanics of biological materials. *Nat. Mater.* **2,** 715–725 (2003).
4. Herrada, I. *et al.* Muscular Dystrophy Mutations Impair the Nuclear Envelope Emerin Self-assembly Properties. *ACS Chem. Biol.* **10,** 2733–2742 (2015).
5. Chi, Y.-H., Chen, Z.-J. & Jeang, K.-T. The nuclear envelopathies and human diseases. *J. Biomed. Sci.* **16,** 96 (2009).
6. Worman, H. J., Östlund, C. & Wang, Y. Diseases of the Nuclear Envelope. *Cold Spring Harb. Perspect. Biol.* **2,** (2010).
7. Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. *J. Cell Biol.* **110,** 883–894 (1990).
8. Wente, S. R. Gatekeepers of the nucleus. *Science* **288,** 1374–1377 (2000).
9. Maimon, T., Elad, N., Dahan, I. & Medalia, O. The human nuclear pore complex as revealed by cryo-electron tomography. *Struct. Lond. Engl. 1993* **20,** 998–1006 (2012).
10. Lin, D. H. *et al.* Architecture of the symmetric core of the nuclear pore. *Science* **352,** aaf1015 (2016).
11. Higher Nucleoporin-Importinβ Affinity at the Nuclear Basket Increases Nucleocytoplasmic Import - Semantic Scholar. Available at: /paper/Higher-Nucleoporin-Importinβ-Affinity-at-the-Nucle-Azimi-Mofrad/02fdc8d2996f51cc303b0f9094de9c55a38e5568. (Accessed: 29th June 2017)
12. Lim, R. Y. H., Ullman, K. S. & Fahrenkrog, B. Biology and Biophysics of the Nuclear Pore Complex And Its Components. *Int. Rev. Cell Mol. Biol.* **267,** 299–342 (2008).
13. Allen, T. D., Cronshaw, J. M., Bagley, S., Kiseleva, E. & Goldberg, M. W. The nuclear pore complex: mediator of translocation between nucleus and cytoplasm. *J Cell Sci* **113,** 1651–1659 (2000).
14. Kabachinski, G. & Schwartz, T. U. The nuclear pore complex – structure and function at a glance. *J Cell Sci* **128,** 423–429 (2015).

15. Franke, W. W., Deumling, B., Ermen, B., Jarasch, E.-D. & Kleinig, H. NUCLEAR MEMBRANES FROM MAMMALIAN LIVER. *J. Cell Biol.* **46,** 379–395 (1970).

16. Scheer U et al. (1976), Experimental disintegration of the nuclear enve... - Xenbase Paper. Available at: http://www.cbrmed.ucalgary.ca/literature/article.do?method=display&articleId=32531. (Accessed: 30th September 2017)

17. Aaronson, R. P. & Blobel, G. On the attachment of the nuclear pore complex. *J. Cell Biol.* **62,** 746–754 (1974).

18. Xie, W. *et al.* A-type Lamins Form Distinct Filamentous Networks with Differential Nuclear Pore Complex Associations. *Curr. Biol. CB* **26,** 2651–2658 (2016).

19. Gruenbaum, Y. & Foisner, R. Lamins: nuclear intermediate filament proteins with fundamental functions in nuclear mechanics and genome regulation. *Annu. Rev. Biochem.* **84,** 131–164 (2015).

20. Kollmar, M. Polyphyly of nuclear lamin genes indicates an early eukaryotic origin of the metazoan-type intermediate filament proteins. *Sci. Rep.* **5,** (2015).

21. Lin, F. & Worman, H. J. Structural organization of the human gene encoding nuclear lamin A and nuclear lamin C. *J. Biol. Chem.* **268,** 16321–16326 (1993).

22. Höger, T. H., Zatloukal, K., Waizenegger, I. & Krohne, G. Characterization of a second highly conserved B-type lamin present in cells previously thought to contain only a single B-type lamin. *Chromosoma* **99,** 379–390 (1990).

23. Dechat, T., Adam, S. A. & Goldman, R. D. Nuclear lamins and chromatin: when structure meets function. *Adv. Enzyme Regul.* **49,** 157–166 (2009).

24. Röber, R. A., Weber, K. & Osborn, M. Differential timing of nuclear lamin A/C expression in the various organs of the mouse embryo and the young animal: a developmental study. *Dev. Camb. Engl.* **105,** 365–378 (1989).

25. Sullivan, T. *et al.* Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy. *J. Cell Biol.* **147,** 913–920 (1999).

26. Ellenberg, J. *et al.* Nuclear membrane dynamics and reassembly in living cells: targeting of an inner nuclear membrane protein in interphase and mitosis. *J. Cell Biol.* **138,** 1193–1206 (1997).

27. Gerace, L. & Blobel, G. The nuclear envelope lamina is reversibly depolymerized during mitosis. *Cell* **19,** 277–287 (1980).

28. Adam, S. A., Butin-Israeli, V., Cleland, M. M., Shimi, T. & Goldman, R. D. Disruption of lamin B1 and lamin B2 processing and localization by farnesyltransferase inhibitors. *Nucl. Austin Tex* **4,** 142–150 (2013).

29. Dechat, T. *et al.* LAP2α and BAF transiently localize to telomeres and specific regions on chromatin during nuclear assembly. *J. Cell Sci.* **117,** 6117–6128 (2004).

30. Moir, R. D., Yoon, M., Khuon, S. & Goldman, R. D. Nuclear Lamins a and B1. *J. Cell Biol.* **151,** 1155–1168 (2000).

31. Haraguchi, T. *et al.* Live cell imaging and electron microscopy reveal dynamic processes of BAF-directed nuclear envelope assembly. *J. Cell Sci.* **121,** 2540–2554 (2008).

32. Dechat, T. *et al.* Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. *Genes Dev.* **22,** 832–853 (2008).

33. Nigg, E. A. Assembly and cell cycle dynamics of the nuclear lamina. *Semin. Cell Biol.* **3,** 245–253 (1992).

34. Wang, Y. *et al.* A mutation abolishing the ZMPSTE24 cleavage site in prelamin A causes a progeroid disorder. *J. Cell Sci.* **129,** 1975–1980 (2016).

35. Moulson, C. L. *et al.* Homozygous and compound heterozygous mutations in ZMPSTE24 cause the laminopathy restrictive dermopathy. *J. Invest. Dermatol.* **125,** 913–919 (2005).

36. Navarro, C. L. *et al.* Loss of ZMPSTE24 (FACE-1) causes autosomal recessive restrictive dermopathy and accumulation of Lamin A precursors. *Hum. Mol. Genet.* **14,** 1503–1513 (2005).

37. Eriksson, M. *et al.* Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature* **423,** 293–298 (2003).

38. Sandre-Giovannoli, A. D. *et al.* Homozygous Defects in LMNA, Encoding Lamin A/C Nuclear-Envelope Proteins, Cause Autosomal Recessive Axonal Neuropathy in Human (Charcot-Marie-Tooth Disorder Type 2) and Mouse. *Am. J. Hum. Genet.* **70,** 726–736 (2002).

39. Davies, B. S. J., Fong, L. G., Yang, S. H., Coffinier, C. & Young, S. G. The posttranslational processing of prelamin A and disease. *Annu. Rev. Genomics Hum. Genet.* **10,** 153–174 (2009).

40. Dobrzynska, A., Askjaer, P. & Gruenbaum, Y. Lamin-Binding Proteins in Caenorhabditis elegans. *Methods Enzymol.* **569,** 455–483 (2016).

41. Strelkov, S. V., Schumacher, J., Burkhard, P., Aebi, U. & Herrmann, H. Crystal structure of the human lamin A coil 2B dimer: implications for the head-to-tail association of nuclear lamins. *J. Mol. Biol.* **343,** 1067–1080 (2004).

42. Krimm, I. *et al.* The Ig-like structure of the C-terminal domain of lamin A/C, mutated in muscular dystrophies, cardiomyopathy, and partial lipodystrophy. *Struct. Lond. Engl. 1993* **10,** 811–823 (2002).

43. Dhe-Paganon, S., Werner, E. D., Chi, Y.-I. & Shoelson, S. E. Structure of the globular tail of nuclear lamin. *J. Biol. Chem.* **277,** 17381–17384 (2002).

44. Mahamid, J. *et al.* Visualizing the molecular sociology at the HeLa cell nuclear periphery. *Science* **351,** 969–972 (2016).

45. Turgay, Y. & Medalia, O. The structure of lamin filaments in somatic cells as revealed by cryo-electron tomography. *Nucl. Austin Tex* 0 (2017). doi:10.1080/19491034.2017.1337622

46. Shimi, T. *et al.* The A- and B-type nuclear lamin networks: microdomains involved in chromatin organization and transcription. *Genes Dev.* **22,** 3409–3421 (2008).

47. Starr, D. A. & Han, M. ANChors away: an actin based mechanism of nuclear positioning. *J. Cell Sci.* **116,** 211–216 (2003).

48. Crisp, M. & Burke, B. The nuclear envelope as an integrator of nuclear and cytoplasmic architecture. *FEBS Lett.* **582,** 2023–2032 (2008).

49. Wolff, N. *et al.* Structural analysis of emerin, an inner nuclear membrane protein mutated in X-linked Emery–Dreifuss muscular dystrophy. *FEBS Lett.* **501,** 171–176 (2001).

50. Laguri, C. *et al.* Structural characterization of the LEM motif common to three human inner nuclear membrane proteins. *Struct. Lond. Engl. 1993* **9,** 503–511 (2001).

51. Cai, M. *et al.* Solution structure of the constant region of nuclear envelope protein LAP2 reveals two LEM-domain structures: one binds BAF and the other binds DNA. *EMBO J.* **20,** 4399–4407 (2001).

52. Cai, M. *et al.* Solution NMR Structure of the Barrier-to-Autointegration Factor-Emerin Complex. *J. Biol. Chem.* **282,** 14525–14535 (2007).

53. Barton, L. J., Soshnev, A. A. & Geyer, P. K. Networking in the nucleus: A spotlight on LEM-domain proteins. *Curr. Opin. Cell Biol.* **34,** 1 (2015).

54. Lin, F. *et al.* MAN1, an Inner Nuclear Membrane Protein That Shares the LEM Domain with Lamina-associated Polypeptide 2 and Emerin. *J. Biol. Chem.* **275,** 4840–4847 (2000).

55. Holaska, J. M., Lee, K. K., Kowalski, A. K. & Wilson, K. L. Transcriptional Repressor Germ Cell-less (GCL) and Barrier to Autointegration Factor (BAF) Compete for Binding to Emerin in Vitro. *J. Biol. Chem.* **278,** 6969–6975 (2003).

56. Furukawa, K. LAP2 binding protein 1 (L2BP1/BAF) is a candidate mediator of LAP2-chromatin interaction. *J. Cell Sci.* **112 (Pt 15),** 2485–2492 (1999).

57. Huang, Y., Cai, M., Clore, G. M. & Craigie, R. No interaction of barrier-to-autointegration factor (BAF) with HIV-1 MA, cone-rod homeobox (Crx) or MAN1-C in absence of DNA. *PloS One* **6,** e25123 (2011).

58. Kondé, E. *et al.* Structural analysis of the Smad2-MAN1 interaction that regulates transforming growth factor-β signaling at the inner nuclear membrane. *Biochemistry (Mosc.)* **49,** 8020–8032 (2010).

59. Gu, M. *et al.* LEM2 recruits CHMP7 for ESCRT-mediated nuclear envelope closure in fission yeast and human cells. *Proc. Natl. Acad. Sci. U. S. A.* **114,** E2166–E2175 (2017).

60. Bourgeois, B. *et al.* Inhibition of TGF-β signaling at the nuclear envelope: characterization of interactions between MAN1, Smad2 and Smad3, and PPM1A. *Sci. Signal.* **6,** ra49 (2013).

61. Brachner, A. *et al.* The endonuclease Ankle1 requires its LEM and GIY-YIG motifs for DNA cleavage in vivo. *J. Cell Sci.* **125,** 1048–1057 (2012).

62. Asencio, C. *et al.* Coordination of kinase and phosphatase activities by Lem4 enables nuclear envelope reassembly during mitosis. *Cell* **150,** 122–135 (2012).

63. Worman, H. J., Yuan, J., Blobel, G. & Georgatos, S. D. A lamin B receptor in the nuclear envelope. *Proc. Natl. Acad. Sci. U. S. A.* **85,** 8531–8534 (1988).

64. Worman, H. J., Evans, C. D. & Blobel, G. The lamin B receptor of the nuclear envelope inner membrane: a polytopic protein with eight potential transmembrane domains. *J. Cell Biol.* **111,** 1535–1542 (1990).

65. Waterham, H. R. *et al.* Autosomal recessive HEM/Greenberg skeletal dysplasia is caused by 3 beta-hydroxysterol delta 14-reductase deficiency due to mutations in the lamin B receptor gene. *Am. J. Hum. Genet.* **72,** 1013–1017 (2003).

66. Shultz, L. D. *et al.* Mutations at the mouse ichthyosis locus are within the lamin B receptor gene: a single gene model for human Pelger-Huët anomaly. *Hum. Mol. Genet.* **12,** 61–69 (2003).

67. Wassif, C. A. *et al.* HEM dysplasia and ichthyosis are likely laminopathies and not due to 3beta-hydroxysterol Delta14-reductase deficiency. *Hum. Mol. Genet.* **16,** 1176–1187 (2007).

68. Tsai, P.-L., Zhao, C., Turner, E. & Schlieker, C. The Lamin B receptor is essential for cholesterol synthesis and perturbed by disease-causing mutations. *eLife* **5,** (2016).

69. Liokatis, S. *et al.* Solution structure and molecular interactions of lamin B receptor Tudor domain. *J. Biol. Chem.* **287,** 1032–1042 (2012).

70. Hirano, T. *et al.* Conformational Dynamics of Wild-type Lys-48-linked Diubiquitin in Solution. *J. Biol. Chem.* **286,** 37496–37502 (2011).

71. Ye, X. S., Xu, G., Fincher, R. R. & Osmani, S. A. Characterization of NIMA protein kinase in Aspergillus nidulans. *Methods Enzymol.* **283,** 520–532 (1997).

72. Lechner, M. S., Schultz, D. C., Negorev, D., Maul, G. G. & Rauscher, F. J. The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that targets the chromoshadow domain. *Biochem. Biophys. Res. Commun.* **331,** 929–937 (2005).

73. Solovei, I. *et al.* LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell* **152,** 584–598 (2013).

74. Olins, A. L., Rhodes, G., Welch, D. B. M., Zwerger, M. & Olins, D. E. Lamin B receptor. *Nucleus* **1,** 53–70 (2010).

75. Razafsky, D. & Hodzic, D. Bringing KASH under the SUN: the many faces of nucleo-cytoskeletal connections. *J. Cell Biol.* **186,** 461–472 (2009).

76. Starr, D. A. & Fischer, J. A. KASH 'n Karry: the KASH domain family of cargo-specific cytoskeletal adaptor proteins. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **27,** 1136–1146 (2005).

77. Fridkin, A., Penkner, A., Jantsch, V. & Gruenbaum, Y. SUN-domain and KASH-domain proteins during development, meiosis and disease. *Cell. Mol. Life Sci. CMLS* **66,** 1518–1533 (2009).

78. Méjat, A. & Misteli, T. LINC complexes in health and disease. *Nucl. Austin Tex* **1,** 40–52 (2010).

79. Cartwright, S. & Karakesisoglou, I. Nesprins in health and disease. *Semin. Cell Dev. Biol.* **29,** 169–179 (2014).

80. Sosa, B. A., Kutay, U. & Schwartz, T. U. Structural Insights into LINC Complexes. *Curr. Opin. Struct. Biol.* **23,** 285–291 (2013).

81. Zhang, Q. *et al.* Nesprins: a novel family of spectrin-repeat-containing proteins that localize to the nuclear membrane in multiple tissues. *J. Cell Sci.* **114,** 4485–4498 (2001).

82. Wilhelmsen, K. *et al.* Nesprin-3, a novel outer nuclear membrane protein, associates with the cytoskeletal linker protein plectin. *J. Cell Biol.* **171,** 799–810 (2005).

83. Roux, K. J. *et al.* Nesprin 4 is an outer nuclear membrane protein that can induce kinesin-mediated cell polarization. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 2194–2199 (2009).

84. Hagan, I. & Yanagida, M. The product of the spindle formation gene sad1+ associates with the fission yeast spindle pole body and is essential for viability. *J. Cell Biol.* **129,** 1033–1047 (1995).

85. Malone, C. J., Fixsen, W. D., Horvitz, H. R. & Han, M. UNC-84 localizes to the nuclear envelope and is required for nuclear migration and anchoring during C. elegans development. *Dev. Camb. Engl.* **126,** 3171–3181 (1999).

86. Hiraoka, Y. & Dernburg, A. F. The SUN rises on meiotic chromosome dynamics. *Dev. Cell* **17,** 598–605 (2009).

87. Xing, X.-W. *et al.* Identification of a novel gene SRG4 expressed at specific stages of mouse spermatogenesis. *Acta Biochim. Biophys. Sin.* **36,** 351–359 (2004).

88. Göb, E., Schmitt, J., Benavente, R. & Alsheimer, M. Mammalian sperm head formation involves different polarization of two novel LINC complexes. *PloS One* **5,** e12072 (2010).

89. Hieda, M. Implications for Diverse Functions of the LINC Complexes Based on the Structure. *Cells* **6,** (2017).

90. Wang, W. *et al.* Structural insights into SUN-KASH complexes across the nuclear envelope. *Cell Res.* **22,** 1440–1452 (2012).

91. Sosa, B. A., Rothballer, A., Kutay, U. & Schwartz, T. U. LINC Complexes Form by Binding of Three KASH Peptides to the Interfaces of Trimeric SUN proteins. *Cell* **149,** 1035–1047 (2012).

92. Zhou, Z. *et al.* Structure of Sad1-UNC84 homology (SUN) domain defines features of molecular bridge in nuclear envelope. *J. Biol. Chem.* **287,** 5317–5326 (2012).

93. Nie, S. *et al.* Coiled-Coil Domains of SUN Proteins as Intrinsic Dynamic Regulators. *Struct. Lond. Engl. 1993* **24,** 80–91 (2016).

94. Swift, J. *et al.* Nuclear Lamin-A Scales with Tissue Stiffness and Enhances Matrix-Directed Differentiation. *Science* **341,** 1240104 (2013).

95. Dahl, K. N., Ribeiro, A. J. S. & Lammerding, J. Nuclear Shape, Mechanics, and Mechanotransduction. *Circ. Res.* **102,** 1307–1318 (2008).

96. Newport, J. W., Wilson, K. L. & Dunphy, W. G. A lamin-independent pathway for nuclear envelope assembly. *J. Cell Biol.* **111,** 2247–2259 (1990).

97. Lammerding, J. *et al.* Lamin A/C deficiency causes defective nuclear mechanics and mechanotransduction. *J. Clin. Invest.* **113,** 370–378 (2004).

98. Physical plasticity of the nucleus in stem cell differentiation. Available at: http://www.pnas.org/content/104/40/15619.full. (Accessed: 9th October 2017)

99. Krause, M. & Wolf, K. Cancer cell migration in 3D tissue: Negotiating space by proteolysis and nuclear deformability. *Cell Adhes. Migr.* **9,** 357–366 (2015).

100. Thiam, H.-R. *et al.* Perinuclear Arp2/3-driven actin polymerization enables nuclear deformation to facilitate cell migration through complex environments. *Nat. Commun.* **7,** (2016).

101. Skau, C. T. *et al.* FMN2 Makes Perinuclear Actin to Protect Nuclei during Confined Migration and Promote Metastasis. *Cell* **167,** 1571–1585.e18 (2016).

102. Raab, M. *et al.* ESCRT III repairs nuclear envelope ruptures during cell migration to limit DNA damage and cell death. *Science* **352,** 359–362 (2016).

103. Gundersen, G. G. & Worman, H. J. Nuclear positioning. *Cell* **152,** 1376–1389 (2013).

104. Luxton, G. W. G., Gomes, E. R., Folker, E. S., Vintinner, E. & Gundersen, G. G. Linear Arrays of Nuclear Envelope Proteins Harness Retrograde Actin Flow for Nuclear Movement. *Science* **329,** 956–959 (2010).

105. Antoku, S., Zhu, R., Kutscheidt, S., Fackler, O. T. & Gundersen, G. G. Reinforcing the LINC complex connection to actin filaments: the role of FHOD1 in TAN line formation and nuclear movement. *Cell Cycle* **14,** 2200–2205 (2015).

106. Lamin A variants that cause striated muscle disease are defective in anchoring transmembrane actin-associated nuclear lines for nuclear movement. Available at: http://www.pnas.org/content/108/1/131.a. (Accessed: 9th October 2017)

107. Guilluy, C. *et al.* Isolated nuclei adapt to force and reveal a mechanotransduction pathway within the nucleus. *Nat. Cell Biol.* **16,** 376–381 (2014).

108. Margalit, A. *et al.* Barrier to autointegration factor blocks premature cell fusion and maintains adult muscle integrity in C. elegans. *J. Cell Biol.* **178,** 661–673 (2007).

109. Jamin, A. & Wiebe, M. S. Barrier to Autointegration Factor (BANF1): interwoven roles in nuclear structure, genome integrity, innate immunity, stress responses and progeria. *Curr. Opin. Cell Biol.* **34,** 61–68 (2015).

110. Cai, M. *et al.* Solution structure of the cellular factor BAF responsible for protecting retroviral DNA from autointegration. *Nat. Struct. Mol. Biol.* **5,** 903–909 (1998).

111. Bradley, C. M., Ronning, D. R., Ghirlando, R., Craigie, R. & Dyda, F. Structural basis for DNA bridging by barrier-to-autointegration factor. *Nat. Struct. Mol. Biol.* **12,** 935–936 (2005).

112. Zheng, R. *et al.* Barrier-to-autointegration factor (BAF) bridges DNA in a discrete, higher-order nucleoprotein complex. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 8997–9002 (2000).

113. A previously unidentified host protein protects retroviral DNA from autointegration. Available at: http://www.pnas.org/content/95/4/1528.full. (Accessed: 9th October 2017)

114. Margalit, A., Segura-Totten, M., Gruenbaum, Y. & Wilson, K. L. Barrier-to-autointegration factor is required to segregate and enclose chromosomes within the nuclear envelope and assemble the nuclear lamina. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 3290–3295 (2005).

115. Furukawa, K. *et al.* Barrier-to-autointegration factor plays crucial roles in cell cycle progression and nuclear organization in Drosophila. *J. Cell Sci.* **116,** 3811–3823 (2003).

116. Gorjánácz, M. *et al.* Caenorhabditis elegans BAF-1 and its kinase VRK-1 participate directly in post-mitotic nuclear envelope assembly. *EMBO J.* **26,** 132–143 (2007).

117. Nichols, R. J., Wiebe, M. S. & Traktman, P. The vaccinia-related kinases phosphorylate the N' terminus of BAF, regulating its interaction with DNA and its retention in the nucleus. *Mol. Biol. Cell* **17,** 2451–2464 (2006).

118. Molitor, T. P. & Traktman, P. Depletion of the protein kinase VRK1 disrupts nuclear envelope morphology and leads to BAF retention on mitotic chromosomes. *Mol. Biol. Cell* **25,** 891–903 (2014).

119. Zhuang, X., Semenova, E., Maric, D. & Craigie, R. Dephosphorylation of barrier-to-autointegration factor by protein phosphatase 4 and its role in cell mitosis. *J. Biol. Chem.* **289,** 1119–1127 (2014).

120. Samwer, M. *et al.* DNA Cross-Bridging Shapes a Single Nucleus from a Set of Mitotic Chromosomes. *Cell* **170,** 956–972.e23 (2017).

121. Peric-Hupkes, D. & Steensel, B. van. Role of the Nuclear Lamina in Genome Organization and Gene Expression. *Cold Spring Harb. Symp. Quant. Biol.* **75,** 517–524 (2010).

122. Harr, J. C. *et al.* Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins. *J. Cell Biol.* **208,** 33–52 (2015).

123. Lange: Junqueira's Basic Histology, 12th Edition. Available at: http://www.langetextbooks.com/0071842705.php?c=home. (Accessed: 9th October 2017)

124. Czapiewski, R., Robson, M. I. & Schirmer, E. C. Anchoring a Leviathan: How the Nuclear Membrane Tethers the Genome. *Front. Genet.* **7,** (2016).

125. Collas, P., Lund, E. G. & Oldenburg, A. R. Closing the (nuclear) envelope on the genome: how nuclear lamins interact with promoters and modulate gene expression. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **36,** 75–83 (2014).

126. Capelson, M. *et al.* Chromatin-bound nuclear pore components regulate gene expression in higher eukaryotes. *Cell* **140,** 372–383 (2010).

127. Kalverda, B. & Fornerod, M. Characterization of genome-nucleoporin interactions in Drosophila links chromatin insulators to the nuclear pore complex. *Cell Cycle Georget. Tex* **9,** 4812–4817 (2010).

128. Akhtar, A. & Gasser, S. M. The nuclear envelope and transcriptional control. *Nat. Rev. Genet.* **8,** 507–517 (2007).

129. Capelson, M. & Hetzer, M. W. The role of nuclear pores in gene regulation, development and disease. *EMBO Rep.* **10,** 697–705 (2009).

130. Schmid, M. *et al.* Nup-PI: the nucleopore-promoter interaction of genes in yeast. *Mol. Cell* **21,** 379–391 (2006).

131. Taddei, A. *et al.* Nuclear pore association confers optimal expression levels for an inducible yeast gene. *Nature* **441,** 774–778 (2006).

132. Menon, B. B. *et al.* Reverse recruitment: The Nup84 nuclear pore subcomplex mediates Rap1/Gcr1/Gcr2 transcriptional activation. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 5749–5754 (2005).

133. Brown, C. R., Kennedy, C. J., Delmar, V. A., Forbes, D. J. & Silver, P. A. Global histone acetylation induces functional genomic reorganization at mammalian nuclear pore complexes. *Genes Dev.* **22,** 627–639 (2008).

134. Guelen, L. *et al.* Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453,** 948–951 (2008).

135. van Steensel, B. & Belmont, A. S. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* **169,** 780–791 (2017).

136. Zullo, J. M. *et al.* DNA sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. *Cell* **149,** 1474–1487 (2012).

137. Pickersgill, H. *et al.* Characterization of the Drosophila melanogaster genome at the nuclear lamina. *Nat. Genet.* **38,** 1005–1014 (2006).

138. Towbin, B. D. *et al.* Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell* **150,** 934–947 (2012).

139. Demmerle, J., Koch, A. J. & Holaska, J. M. The Nuclear Envelope Protein Emerin Binds Directly to Histone Deacetylase 3 (HDAC3) and Activates HDAC3 Activity. *J. Biol. Chem.* **287,** 22080–22088 (2012).

140. Somech, R. *et al.* The nuclear-envelope protein and transcriptional repressor LAP2beta interacts with HDAC3 at the nuclear periphery, and induces histone H4 deacetylation. *J. Cell Sci.* **118,** 4017–4025 (2005).

141. Cox, J. L. *et al.* Banf1 is required to maintain the self-renewal of both mouse and human embryonic stem cells. *J. Cell Sci.* **124,** 2654–2665 (2011).

142. Wang, X. *et al.* Barrier to Autointegration Factor Interacts with the Cone-Rod Homeobox and Represses Its Transactivation Function. *J. Biol. Chem.* **277,** 43288–43300 (2002).

143. Gonzalo, S. DNA Damage and Lamins. *Adv. Exp. Med. Biol.* **773,** 377–399 (2014).

144. Liu, B. *et al.* Genomic instability in laminopathy-based premature aging. *Nat. Med.* **11,** 780–785 (2005).

145. al, R. S., et. The accumulation of un-repairable DNA damage in laminopathy progeria fibroblasts is caused by ROS generation and is prevented by treatment with N-a... - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/21807766. (Accessed: 28th November 2017)

146. Panier, S. & Boulton, S. J. Double-strand break repair: 53BP1 comes into focus. *Nat. Rev. Mol. Cell Biol.* **15,** 7–18 (2014).

147. Pegoraro, E. *et al.* Co-segregation of LMNA and PMP22 gene mutations in the same family. *Neuromuscul. Disord.* **15,** 858–862 (2005).

148. Krishnan, V. *et al.* Histone H4 lysine 16 hypoacetylation is associated with defective DNA repair and premature senescence in Zmpste24-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 12325–12330 (2011).

149. Gonzalez-Suarez, I. *et al.* Novel roles for A-type lamins in telomere biology and the DNA damage response pathway. *EMBO J.* **28,** 2414–2427 (2009).

150. Columbaro, M. *et al.* Rescue of heterochromatin organization in Hutchinson-Gilford progeria by drug treatment. *Cell. Mol. Life Sci.* **62,** 2669–2678 (2005).

151. Scaffidi, P. & Misteli, T. Reversal of the cellular phenotype in the premature aging disease Hutchinson-Gilford progeria syndrome. *Nat. Med.* **11,** 440–445 (2005).

152. Shumaker, D. K. *et al.* Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 8703–8708 (2006).

153. Dimitrova, N., Chen, Y.-C. M., Spector, D. L. & de Lange, T. 53BP1 promotes non-homologous end joining of telomeres by increasing chromatin mobility. *Nature* **456,** 524–528 (2008).

154. Wood, A. M. *et al.* TRF2 and lamin A/C interact to facilitate the functional organization of chromosome ends. *Nat. Commun.* **5,** ncomms6467 (2014).

155. Smogorzewska, A. & de Lange, T. Different telomere damage signaling pathways in human and mouse cells. *EMBO J.* **21,** 4338–4348 (2002).

156. 53BP1 and the LINC complex promote microtubule-dependent DSB mobility and DNA repair. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4636737/. (Accessed: 9th October 2017)

157. Kennedy, B. K., Barbie, D. A., Classon, M., Dyson, N. & Harlow, E. Nuclear organization of DNA replication in primary mammalian cells. *Genes Dev.* **14,** 2855–2868 (2000).

158. Singh, M. *et al.* Lamin A/C Depletion Enhances DNA Damage-Induced Stalled Replication Fork Arrest. *Mol. Cell. Biol.* **33,** 3390 (2013).

159. Bione, S. *et al.* Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular dystrophy. *Nat. Genet.* **8,** 323–327 (1994).

160. Manilal, S., Man, N. thi, Sewry, C. A. & Morris, G. E. The Emery-Dreifuss Muscular Dystrophy Protein, Emerin, is a Nuclear Membrane Protein. *Hum. Mol. Genet.* **5,** 801–808 (1996).

161. Nagano, A. *et al.* Emerin deficiency at the nuclear membrane in patients with Emery-Dreifuss muscular dystrophy. *Nat. Genet.* **12,** 254–259 (1996).

162. Manilal, S. *et al.* Mutations in Emery-Dreifuss muscular dystrophy and their effects on emerin protein expression. *Hum. Mol. Genet.* **7,** 855–864 (1998).

163. Ellis, J. A., Craxton, M., Yates, J. R. & Kendrick-Jones, J. Aberrant intracellular targeting and cell cycle-dependent phosphorylation of emerin contribute to the Emery-Dreifuss muscular dystrophy phenotype. *J. Cell Sci.* **111 (Pt 6),** 781–792 (1998).

164. Ben Yaou, R. *et al.* Multitissular involvement in a family with LMNA and EMD mutations: Role of digenic mechanism? *Neurology* **68,** 1883–1894 (2007).

165. Berk, J. M. *et al.* The molecular basis of emerin–emerin and emerin–BAF interactions. *J Cell Sci* **127,** 3956–3969 (2014).

166. Sakaki, M. *et al.* Interaction between emerin and nuclear lamins. *J. Biochem. (Tokyo)* **129,** 321–327 (2001).

167. Berk, J. M., Tifft, K. E. & Wilson, K. L. The nuclear envelope LEM-domain protein emerin. *Nucleus* **4,** 298–314 (2013).

168. Fairley, E. A., Kendrick-Jones, J. & Ellis, J. A. The Emery-Dreifuss muscular dystrophy phenotype arises from aberrant targeting and binding of emerin at the inner nuclear membrane. *J. Cell Sci.* **112 (Pt 15),** 2571–2582 (1999).

169. Vaughan, A. *et al.* Both emerin and lamin C depend on lamin A for localization at the nuclear envelope. *J. Cell Sci.* **114,** 2577–2590 (2001).

170. Clements, L., Manilal, S., Love, D. R. & Morris, G. E. Direct interaction between emerin and lamin A. *Biochem. Biophys. Res. Commun.* **267,** 709–714 (2000).

171. Shimi, T. *et al.* Dynamic interaction between BAF and emerin revealed by FRAP, FLIP, and FRET analyses in living HeLa cells. *J. Struct. Biol.* **147,** 31–41 (2004).

172. Bengtsson, L. & Wilson, K. L. Barrier-to-Autointegration Factor Phosphorylation on Ser-4 Regulates Emerin Binding to Lamin A In Vitro and Emerin Localization In Vivo. *Mol. Biol. Cell* **17,** 1154–1163 (2006).

173. Pederson, T. & Aebi, U. Actin in the nucleus: what form and what for? *J. Struct. Biol.* **140,** 3–9 (2002).

174. Lattanzi, G. *et al.* Association of emerin with nuclear and cytoplasmic actin is regulated in differentiating myoblasts. *Biochem. Biophys. Res. Commun.* **303,** 764–770 (2003).

175. Holaska, J. M., Kowalski, A. K. & Wilson, K. L. Emerin caps the pointed end of actin filaments: evidence for an actin cortical network at the nuclear inner membrane. *PLoS Biol.* **2,** E231 (2004).

176. Holaska, J. M. & Wilson, K. L. An emerin 'proteome': purification of distinct emerin-containing complexes from HeLa cells suggests molecular basis for diverse roles including gene regulation, mRNA splicing, signaling, mechanosensing, and nuclear architecture. *Biochemistry (Mosc.)* **46,** 8897–8908 (2007).

177. Mislow, J. M. K. *et al.* Nesprin-1alpha self-associates and binds directly to emerin and lamin A in vitro. *FEBS Lett.* **525,** 135–140 (2002).

178. Libotte, T. *et al.* Lamin A/C-dependent localization of Nesprin-2, a giant scaffolder at the nuclear envelope. *Mol. Biol. Cell* **16,** 3411–3424 (2005).

179. Haque, F. *et al.* Mammalian SUN Protein Interaction Networks at the Inner Nuclear Membrane and Their Role in Laminopathy Disease Processes. *J. Biol. Chem.* **285,** 3487–3498 (2010).

180. Wilkinson, F. L. *et al.* Emerin interacts in vitro with the splicing-associated factor, YT521-B. *Eur. J. Biochem.* **270,** 2459–2466 (2003).

181. Haraguchi, T. *et al.* Emerin binding to Btf, a death-promoting transcriptional repressor, is disrupted by a missense mutation that causes Emery-Dreifuss muscular dystrophy. *Eur. J. Biochem.* **271,** 1035–1045 (2004).

182. Markiewicz, E. *et al.* The inner nuclear membrane protein Emerin regulates β-catenin activity by restricting its accumulation in the nucleus. *EMBO J.* **25,** 3275–3285 (2006).

183. Brill, L. M. *et al.* Robust phosphoproteomic profiling of tyrosine phosphorylation sites from human T cells using immobilized metal affinity chromatography and tandem mass spectrometry. *Anal. Chem.* **76,** 2763–2772 (2004).

184. Hirano, Y. *et al.* Dissociation of emerin from barrier-to-autointegration factor is regulated through mitotic phosphorylation of emerin in a xenopus egg cell-free system. *J. Biol. Chem.* **280,** 39925–39933 (2005).

185. Roberts, R. C. *et al.* The Emery-Dreifuss muscular dystrophy associated-protein emerin is phosphorylated on serine 49 by protein kinase A. *FEBS J.* **273,** 4562–4575 (2006).

186. Shu, H., Chen, S., Bi, Q., Mumby, M. & Brekken, D. L. Identification of phosphoproteins and their phosphorylation sites in the WEHI-231 B lymphoma cell line. *Mol. Cell. Proteomics MCP* **3,** 279–286 (2004).

187. Leach, N. *et al.* Emerin is hyperphosphorylated and redistributed in herpes simplex virus type 1-infected cells in a manner dependent on both UL34 and US3. *J. Virol.* **81,** 10792–10803 (2007).

188. Morris, J. B., Hofemeister, H. & O'Hare, P. Herpes simplex virus infection induces phosphorylation and delocalization of emerin, a key inner nuclear membrane protein. *J. Virol.* **81,** 4429–4437 (2007).

189. Huyer, G. *et al.* Mechanism of inhibition of protein-tyrosine phosphatases by vanadate and pervanadate. *J. Biol. Chem.* **272,** 843–851 (1997).

190. Tifft, K. E., Bradbury, K. A. & Wilson, K. L. Tyrosine phosphorylation of nuclear-membrane protein emerin by Src, Abl and other kinases. *J. Cell Sci.* **122,** 3780–3790 (2009).

191. Bukong, T. N., Hall, W. W. & Jacqué, J.-M. Lentivirus-associated MAPK/ERK2 phosphorylates EMD and regulates infectivity. *J. Gen. Virol.* **91,** 2381–2392 (2010).

192. Emery, A. E. & Dreifuss, F. E. Unusual type of benign x-linked muscular dystrophy. *J. Neurol. Neurosurg. Psychiatry* **29,** 338–342 (1966).

193. Wilding, B. R., McGrath, M. J., Bonne, G. & Mitchell, C. A. FHL1 mutants that cause clinically distinct human myopathies form protein aggregates and impair myoblast differentiation. *J. Cell Sci.* **127,** 2269–2281 (2014).

194. Yates, J. R. W. *et al.* Genotype-phenotype analysis in X-linked Emery–Dreifuss muscular dystrophy and identification of a missense mutation associated with a milder phenotype. *Neuromuscul. Disord.* **9,** 159–165 (1999).

195. Mora, M. *et al.* X-linked emery-dreifuss muscular dystrophy can be diagnosed from skin biopsy or blood sample. *Ann. Neurol.* **42,** 249–253 (1997).

196. Berk, J. M., Tifft, K. E. & Wilson, K. L. The nuclear envelope LEM-domain protein emerin. *Nucleus* **4,** 298–314 (2013).

197. Bonne, G. *et al.* Mutations in the gene encoding lamin A/C cause autosomal dominant Emery-Dreifuss muscular dystrophy. *Nat. Genet.* **21,** 285–288 (1999).

198. Morris, G. E. The role of the nuclear envelope in Emery–Dreifuss muscular dystrophy. *Trends Mol. Med.* **7,** 572–577 (2001).

199. Scharner, J., Gnocchi, V. F., Ellis, J. A. & Zammit, P. S. Genotype–phenotype correlations in laminopathies: how does fate translate? *Biochem. Soc. Trans.* **38,** 257–262 (2010).

200. Brodsky, G. L. *et al.* Lamin A/C gene mutation associated with dilated cardiomyopathy with variable skeletal muscle involvement. *Circulation* **101,** 473–476 (2000).

201. Mercuri, E. *et al.* Extreme variability of skeletal and cardiac muscle involvement in patients with mutations in exon 11 of the lamin A/C gene. *Muscle Nerve* **31,** 602–609 (2005).

202. Rankin, J. *et al.* Extreme phenotypic diversity and nonpenetrance in families with the LMNA gene mutation R644C. *Am. J. Med. Genet. A.* **146A,** 1530–1542 (2008).

203. Zirn, B. *et al.* Association of homozygous LMNA mutation R471C with new phenotype: mandibuloacral dysplasia, progeria, and rigid spine muscular dystrophy. *Am. J. Med. Genet. A.* **146A,** 1049–1054 (2008).

204. Szeverenyi, I. *et al.* The Human Intermediate Filament Database: comprehensive information on a gene family involved in many human diseases. *Hum. Mutat.* **29,** 351–360 (2008).

205. Quijano-Roy, S. *et al.* De novo LMNA mutations cause a new form of congenital muscular dystrophy. *Ann. Neurol.* **64,** 177–186 (2008).

206. Fatkin, D. *et al.* Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N. Engl. J. Med.* **341,** 1715–1724 (1999).

207. Taylor, M. R. G. *et al.* Natural history of dilated cardiomyopathy due to lamin A/C gene mutations. *J. Am. Coll. Cardiol.* **41,** 771–780 (2003).

208. Chang, S.-H., Tsai, C.-T., Lai, L.-P. & Lei, M.-H. Identification of a lamin A/C gene mutation in a Taiwanese family with limb girdle muscular dystrophy and cardiomyopathy. *Int. J. Cardiol.* **145,** 598–599 (2010).

209. Bushby, K. Diagnosis and management of the limb girdle muscular dystrophies. *Pract. Neurol.* **9,** 314–323 (2009).

210. Köbberling, J., Willms, B., Kattermann, R. & Creutzfeldt, W. Lipodystrophy of the extremities. A dominantly inherited syndrome associated with lipatrophic diabetes. *Humangenetik* **29,** 111–120 (1975).

211. Shackleton, S. *et al.* LMNA, encoding lamin A/C, is mutated in partial lipodystrophy. *Nat. Genet.* **24,** 153–156 (2000).

212. D'Apice, M., Tenconi, R., Mammi, I., Ende, J. van den & Novelli, G. Paternal origin of LMNA mutations in Hutchinson–Gilford progeria. *Clin. Genet.* **65,** 52–54 (2004).

213. Goldman, R. D. *et al.* Accumulation of mutant lamin A causes progressive changes in nuclear architecture in Hutchinson-Gilford progeria syndrome. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 8963–8968 (2004).

214. McCord, R. P. *et al.* Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson-Gilford progeria syndrome. *Genome Res.* **23,** 260–269 (2013).

215. Barton, R. M. & Worman, H. J. Prenylated prelamin A interacts with Narf, a novel nuclear protein. *J. Biol. Chem.* **274,** 30008–30018 (1999).

216. Mattioli, E. *et al.* Prelamin A-mediated recruitment of SUN1 to the nuclear envelope directs nuclear positioning in human muscle. *Cell Death Differ.* **18,** 1305–1315 (2011).

217. Robin, J. D. & Magdinier, F. Physiological and Pathological Aging Affects Chromatin Dynamics, Structure and Function at the Nuclear Edge. *Front. Genet.* **7,** (2016).

218. Capanni, C. *et al.* Lamin A precursor induces barrier-to-autointegration factor nuclear localization. *Cell Cycle* **9,** 2600–2610 (2010).

219. Loi, M. *et al.* Barrier-to-Autointegration Factor (BAF) involvement in prelamin A-related chromatin organization changes. *Oncotarget* **7,** 15662–15677 (2015).

220. Kubben, N. *et al.* Identification of differential protein interactors of lamin A and progerin. *Nucl. Austin Tex* **1,** 513–525 (2010).

221. Zhong, N., Radu, G., Ju, W. & Brown, W. T. Novel progerin-interactive partner proteins hnRNP E1, EGF, Mel 18, and UBC9 interact with lamin A/C. *Biochem. Biophys. Res. Commun.* **338,** 855–861 (2005).

222. Lee, J. M. *et al.* Modulation of *LMNA* splicing as a strategy to treat prelamin A diseases. *J. Clin. Invest.* **126,** 1592–1602 (2016).

223. Shalev, S. A., De Sandre-Giovannoli, A., Shani, A. A. & Levy, N. An association of Hutchinson–Gilford progeria and malignancy. *Am. J. Med. Genet. A.* **143A,** 1821–1826 (2007).

224. Chen, L. *et al.* LMNA mutations in atypical Werner's syndrome. *Lancet Lond. Engl.* **362,** 440–445 (2003).

225. Soria-Valles, C. *et al.* Novel LMNA mutations cause an aggressive atypical neonatal progeria without progerin accumulation. *J. Med. Genet.* (2016). doi:10.1136/jmedgenet-2015-103695

226. Caux, F. *et al.* A new clinical condition linked to a novel mutation in lamins A and C with generalized lipoatrophy, insulin-resistant diabetes, disseminated leukomelanodermic papules, liver steatosis, and cardiomyopathy. *J. Clin. Endocrinol. Metab.* **88,** 1006–1013 (2003).

227. Capanni, C. *et al.* Altered pre-lamin A processing is a common mechanism leading to lipodystrophy. *Hum. Mol. Genet.* **14,** 1489–1502 (2005).

228. Cao, H. & Hegele, R. A. LMNA is mutated in Hutchinson-Gilford progeria (MIM 176670) but not in Wiedemann-Rautenstrauch progeroid syndrome (MIM 264090). *J. Hum. Genet.* **48,** 271–274 (2003).

229. Verstraeten, V. L. R. M. *et al.* Compound heterozygosity for mutations in LMNA causes a progeria syndrome without prelamin A accumulation. *Hum. Mol. Genet.* **15,** 2509–2522 (2006).

230. Madej-Pilarczyk, A. *et al.* Progeroid syndrome with scleroderma-like skin changes associated with homozygous R435C LMNA mutation. *Am. J. Med. Genet. A.* **149A,** 2387–2392 (2009).

231. Agarwal, A. K., Kazachkova, I., Ten, S. & Garg, A. Severe mandibuloacral dysplasia-associated lipodystrophy and progeria in a young girl with a novel homozygous Arg527Cys LMNA mutation. *J. Clin. Endocrinol. Metab.* **93,** 4617–4623 (2008).

232. Novelli, G. *et al.* Mandibuloacral dysplasia is caused by a mutation in LMNA-encoding lamin A/C. *Am. J. Hum. Genet.* **71,** 426–431 (2002).

233. Simha, V., Agarwal, A. K., Oral, E. A., Fryns, J.-P. & Garg, A. Genetic and phenotypic heterogeneity in patients with mandibuloacral dysplasia-associated lipodystrophy. *J. Clin. Endocrinol. Metab.* **88,** 2821–2824 (2003).

234. Shen, J. J., Brown, C. A., Lupski, J. R. & Potocki, L. Mandibuloacral dysplasia caused by homozygosity for the R527H mutation in lamin A/C. *J. Med. Genet.* **40,** 854–857 (2003).

235. Garavelli, L. *et al.* Mandibuloacral dysplasia type A in childhood. *Am. J. Med. Genet. A.* **149A,** 2258–2264 (2009).

236. Kosho, T. *et al.* Mandibuloacral dysplasia and a novel LMNA mutation in a woman with severe progressive skeletal changes. *Am. J. Med. Genet. A.* **143A,** 2598–2603 (2007).

237. Garg, A., Cogulu, O., Ozkinay, F., Onay, H. & Agarwal, A. K. A novel homozygous Ala529Val LMNA mutation in Turkish patients with mandibuloacral dysplasia. *J. Clin. Endocrinol. Metab.* **90,** 5259–5264 (2005).

238. Plasilova, M. *et al.* Homozygous missense mutation in the lamin A/C gene causes autosomal recessive Hutchinson-Gilford progeria syndrome. *J. Med. Genet.* **41,** 609–614 (2004).

239. Witt, D. R. *et al.* Restrictive dermopathy: a newly recognized autosomal recessive skin dysplasia. *Am. J. Med. Genet.* **24,** 631–648 (1986).

240. Navarro, C. L. *et al.* Lamin A and ZMPSTE24 (FACE-1) defects cause nuclear disorganization and identify restrictive dermopathy as a lethal neonatal laminopathy. *Hum. Mol. Genet.* **13,** 2493–2503 (2004).

241. Youn, G. *et al.* Autosomal recessive LMNA mutation causing Restrictive Dermopathy. *Clin. Genet.* **78,** 199–200 (2010).

242. Liang, L., Zhang, H. & Gu, X. Homozygous LMNA mutation R527C in atypical Hutchinson-Gilford progeria syndrome: evidence for autosomal recessive inheritance. *Acta Paediatr. Oslo Nor. 1992* **98,** 1365–1368 (2009).

243. Paquet, N. *et al.* Néstor-Guillermo Progeria Syndrome: a biochemical insight into Barrier-to-Autointegration Factor 1, alanine 12 threonine mutation. *BMC Mol. Biol.* **15,** (2014).

244. Puente, X. S. *et al.* Exome sequencing and functional analysis identifies BANF1 mutation as the cause of a hereditary progeroid syndrome. *Am. J. Hum. Genet.* **88,** 650–656 (2011).

245. Lee, K. K. *et al.* Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF. *J. Cell Sci.* **114,** 4567–4573 (2001).

246. LeVine, H. Thioflavine T interaction with synthetic Alzheimer's disease beta-amyloid peptides: detection of amyloid aggregation in solution. *Protein Sci. Publ. Protein Soc.* **2,** 404–410 (1993).

247. Ban, T., Hamada, D., Hasegawa, K., Naiki, H. & Goto, Y. Direct Observation of Amyloid Fibril Growth Monitored by Thioflavin T Fluorescence. *J. Biol. Chem.* **278,** 16462–16465 (2003).

248. Biancalana, M. & Koide, S. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochim. Biophys. Acta* **1804,** 1405–1412 (2010).

249. Le, H. Q. *et al.* Mechanical regulation of transcription controls Polycomb-mediated gene silencing during lineage commitment. *Nat. Cell Biol.* **18,** 864–875 (2016).

250. Samson, C., Herrada, I., Celli, F., Theillet, F.-X. & Zinn-Justin, S. 1H, 13C and 15N backbone resonance assignment of the intrinsically disordered region of the nuclear envelope protein emerin. *Biomol. NMR Assign.* **10,** 179–182 (2016).

251. Wong, Y.-H. *et al.* KinasePhos 2.0: a web server for identifying protein kinase-specific phosphorylation sites based on sequences and coupling patterns. *Nucleic Acids Res.* **35,** W588-594 (2007).

252. Gotzmann, J. & Foisner, R. P. *Lamins and Emerin in Muscular Dystrophy: The Nuclear Envelope Connection*. (Landes Bioscience, 2013).

253. Haraguchi, T. *et al.* Nuclear localization of barrier-to-autointegration factor is correlated with progression of S phase in human cells. *J. Cell Sci.* **120,** 1967–1977 (2007).

254. Wilson, K. L. & Foisner, R. Lamin-binding Proteins. *Cold Spring Harb. Perspect. Biol.* **2,** (2010).

255. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78,** 1606–1619 (2000).

256. pisa.pdf. Available at: https://www.ebi.ac.uk/msd-srv/prot_int/papers/pisa.pdf. (Accessed: 22nd October 2017)

257. Garg, A., Cogulu, O., Ozkinay, F., Onay, H. & Agarwal, A. K. A novel homozygous Ala529Val LMNA mutation in Turkish patients with mandibuloacral dysplasia. *J. Clin. Endocrinol. Metab.* **90,** 5259–5264 (2005).

258. Ashkenazy, H. *et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44,** W344–W350 (2016).

259. Samson, C. *et al.* Emerin self-assembly mechanism: role of the LEM domain. *FEBS J.* n/a-n/a (2016). doi:10.1111/febs.13983

260. Liu, J. *et al.* MAN1 and emerin have overlapping function(s) essential for chromosome segregation and cell division in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 4598–4603 (2003).

261. Mansharamani, M. & Wilson, K. L. Direct Binding of Nuclear Membrane Protein MAN1 to Emerin in Vitro and Two Modes of Binding to Barrier-to-Autointegration Factor. *J. Biol. Chem.* **280,** 13863–13870 (2005).

262. Swaisgood, M. & Schindler, M. Lateral diffusion of lectin receptors in fibroblast membranes as a function of cell shape. *Exp. Cell Res.* **180,** 515–528 (1989).

263. Hofmann, W. A., Johnson, T., Klapczynski, M., Fan, J.-L. & de Lanerolle, P. From transcription to transport: emerging roles for nuclear myosin I. *Biochem. Cell Biol. Biochim. Biol. Cell.* **84,** 418–426 (2006).

264. Vreugde, S. *et al.* Nuclear myosin VI enhances RNA polymerase II-dependent transcription. *Mol. Cell* **23,** 749–755 (2006).

265. Cameron, R. S. *et al.* Myosin16b: The COOH-tail region directs localization to the nucleus and overexpression delays S-phase progression. *Cell Motil. Cytoskeleton* **64,** 19–48 (2007).

266. Pranchevicius, M. C. S. *et al.* Myosin Va phosphorylated on Ser1650 is found in nuclear speckles and redistributes to nucleoli upon inhibition of transcription. *Cell Motil. Cytoskeleton* **65,** 441–456 (2008).

267. Holaska, J. M. Emerin and the nuclear lamina in muscle and cardiac disease. *Circ. Res.* **103,** 16–23 (2008).

268. Margalit, A., Segura-Totten, M., Gruenbaum, Y. & Wilson, K. L. Barrier-to-autointegration factor is required to segregate and enclose chromosomes within the nuclear envelope and assemble the nuclear lamina. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 3290–3295 (2005).

269. Lusk, C. P., Blobel, G. & King, M. C. Highway to the inner nuclear membrane: rules for the road. *Nat. Rev. Mol. Cell Biol.* **8,** 414–420 (2007).

270. Korfali, N. *et al.* The nuclear envelope proteome differs notably between tissues. *Nucl. Austin Tex* **3,** 552–564 (2012).

271. Andrew, C. D., Warwicker, J., Jones, G. R. & Doig, A. J. Effect of phosphorylation on alpha-helix stability as a function of position. *Biochemistry (Mosc.)* **41,** 1897–1905 (2002).

272. Oca, R. M. de, Lee, K. K. & Wilson, K. L. Binding of Barrier to Autointegration Factor (BAF) to Histone H3 and Selected Linker Histones Including H1.1. *J. Biol. Chem.* **280,** 42252–42262 (2005).

273. Montes de Oca, R., Andreassen, P. R. & Wilson, K. L. Barrier-to-Autointegration Factor influences specific histone modifications. *Nucleus* **2,** 580–590 (2011).

274. Agarwal, A. K., Fryns, J.-P., Auchus, R. J. & Garg, A. Zinc metalloproteinase, ZMPSTE24, is mutated in mandibuloacral dysplasia. *Hum. Mol. Genet.* **12,** 1995–2001 (2003).

275. Capanni, C. *et al.* Familial partial lipodystrophy, mandibuloacral dysplasia and restrictive dermopathy feature barrier-to-autointegration factor (BAF) nuclear redistribution. *Cell Cycle* **11,** 3568–3577 (2012).

276. The CCPN data model for NMR spectroscopy: development of a software pipeline. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/15815974. (Accessed: 28th November 2017)

277. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 235–242 (2011).

278. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221 (2010).

279. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501 (2010).

APPENDIX 1

Article published in *Biomol NMR Assign*:

Samson *et al.,* ^{1}H, ^{13}C and ^{15}N backbone resonance assignment of the intrinsically disordered region of the nuclear envelope protein emerin.

^{1}H, ^{13}C and ^{15}N resonance assignments of the intrinsically disordered region of the nuclear envelope protein emerin

Camille Samson[1*], Isaline Herrada[1*], Florian Celli[1], Francois-Xavier Theillet[2], Sophie Zinn-Justin[1#]

[1] Laboratoire de Biologie Structurale et Radiobiologie, Institute for Integrative Biology of the Cell (I2BC), CEA Saclay Bât. 144, 91191 Gif-sur-Yvette Cedex, France

[2] Department of Molecular Pharmacology and Cell Biology, Leibniz-Institut für Molecular Pharmakologie (FMP), 13125 Berlin, Germany

[*] These authors have equally contributed to the work.

[#] To whom correspondence should be addressed: Sophie Zinn-Justin, I2BC CNRS / Univ. Paris South & IBITECS CEA, CEA Saclay Bât 144, 91191 Gif-sur-Yvette Cedex, France, Tel.: +33(0)169083026; Fax: +33(0)169084712; Email: sophie.zinn@cea.fr

Abstract

Human emerin is an inner nuclear membrane protein involved in the response of the nucleus to a mechanical stress. It contributes to the physical connection between the cytoskeleton and the nucleoskeleton. It is also involved in chromatin organization. Its N-terminal region is nucleoplasmic and is characterized by the presence of a globular LEM domain from residue 1 to residue 43. The three-dimensional structure of this LEM domain in complex with the chromatin BAF protein was solved from NMR data. It revealed an asymmetric mode of binding in which one LEM domain binds to two BAF molecules. Apart from the LEM domain, the nucleoplasmic region of emerin is predicted as intrinsically disordered, and the molecular details of its recognition properties are unknown. We here report the assignment of the ^{1}H, ^{15}N and ^{13}C NMR signals of the emerin region from residue 67 to residue 170 in 0 and 8 M urea. Chemical shift analysis confirms that this fragment is intrinsically disordered in both conditions.

Keywords: nuclear envelope, emerin, disorder, binding, urea

Biological context

Inner nuclear envelope proteins play an essential role in cell response to mechanical demands. They regulate cell cytoskeletal tension as a function of tissue stiffness and activate mechano-regulated signaling pathways (Ho et al., 2013; Bertrand et al., 2014; Swift et al., 2014). In particular emerin, SUN1/2 and lamin A contribute to nuclear envelope assembly as well as to the nuclear response to a mechanical stress. However, these proteins contain large regions predicted to be intrinsically disordered regions (IDRs), which hinders their study by classical structural biology tools (Herrada et al., 2015, in press). There is a need today to describe the molecular events involved in nuclear envelope assembly and function, and to understand how IDRs contribute to these events.

Human emerin is an inner nuclear membrane protein of 254 amino acids. It is mutated in patients with X-linked Emery-Dreifuss Muscular dystrophy (Bione et al., 1994). Disease-causing mutations generally lead to an absence of emerin. A small number of missense mutations found in the emerin predicted IDR lead to loss-of-function variants (Ellis et al., 1998). Emerin is anchored to the membrane through its C-terminal α-helix. Its N-terminal regions is nucleoplasmic. Emerin contributes to the physical connection between the cytoskeleton and the nucleoskeleton. In particular it binds to the inner nuclear membrane proteins SUN1/2 (Haque et al. 2006) and the nuclear filaments lamins (Clements et al. 2000). It is also involved in chromatin organization. It is present at the interface between the nuclear envelope and chromatin, and it directly interacts with the DNA binding protein BAF. The three-dimensional structures of emerin N-terminal LEM domain alone (residues 1 to 43; Wolff et al., 2001) and in complex with BAF (Cai et al., 2007) were solved from NMR data. Analysis of the LEM-BAF structure revealed an asymmetric mode of binding in which one LEM domain binds to two BAF molecules. Apart from this LEM domain, the nucleoplasmic region of emerin is predicted as intrinsically disordered, and the molecular details of its recognition properties are unknown. In order to understand the molecular bases of emerin binding properties and disease-causing defects, we undertook NMR studies of emerin nucleoplasmic region from residue 67 to residue 170, in both 0 and 8 M urea. We report here resonance assignments in these two conditions.

Methods and experiments

Protein expression and purification

The human emerin fragment from residue 67 to residue 170 was expressed using a pETM13 vector as a N-terminal octa-histidine fusion protein with a TEV site between the histidine tag and the emerin fragment in *Escherichia coli* BL21 DE3 Star (Novagen). The emerin fragment cDNA was optimized for expression in *Escherichia coli* (Genscript). Bacteria were grown in ^{15}N and ^{13}C labelled minimum medium and induced at an optical density of 1 with 0.5 mM isopropyl β-D-1-thiogalactopyranoside at 293K for 18 h. Cells were lysed in 50 mM Tris‐HCl, pH 8, 300 mM NaCl, 40 mM imidazole, 5% glycerol, 1% Triton X‐100 and 1 mM phenylmethanesulfonylfluoride. As the emerin fragment is produced in inclusion bodies, the pellet was resuspended in buffer A (20 mM Tris‐HCl (pH 8.0), 150 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol and 8 M urea) and centrifuged at 20,000g for 20 min. The protein was then purified on a Ni‐NTA column (GE‐Healthcare) equilibrated in buffer A and eluted directly with 100% of buffer B (buffer A + 1 M imidazole). After dialyzing the sample to remove imidazole and urea, the octa-histidine tag was removed by TEV cleavage and the purity of the sample was checked on SDS-PAGE. The sample was finally dialysed against either buffer D (20 mM sodium phosphate buffer (pH 6.5), 30 mM NaCl and 8 M urea, "Denatured sample" or DS) or buffer N (20 mM sodium phosphate buffer (pH 6.5), 30 mM NaCl and 0 M urea, "Native sample" or NS). NS and DS were concentrated up to 500-600 μM for the NMR assignment experiments.

NMR spectroscopy

All NMR experiments were performed on a 5 mm diameter NMR sample tube containing 500-600 μM uniformly ^{15}N and ^{13}C labelled emerin fragment in 20 mM sodium phosphate buffer (pH 6.5), 30 mM NaCl, 10 mM β-mercaptoethanol or 5mM dithiothreitol, 90 %: 10 %, H_2O: D2O and either 0 or 8 M urea. 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) was added for chemical shift referencing. Experiments were recorded at 303 K on a 750 MHz Bruker Avance spectrometer equipped with a triple resonance cryogenic probe. The data were processed using Topspin3.1 (Bruker) and analyzed with CCPNMR (Vranken et al. 2005). ^{1}H, ^{13}C and ^{15}N resonance frequencies were assigned using 3D HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO and HN(CO)(CA)NH experiments.

Extend of assignments and data deposition

The sequence of the studied protein fragment and its ^{1}H-^{15}N 2D HSQC spectrum in 0M urea are shown in Figure 1. Using a combination of 3D experiments recorded on the NS sample, 96% of ^{1}H-^{15}N pairs (95 out of 99), 99% of ^{13}C (104 out of 105) and 99% of ^{13}C (98 out of 99) and 98% of ^{13}CO (103 out of 105) resonances were assigned for the emerin fragment in 0M urea. Almost all the backbone ^{1}H, ^{13}C, and ^{15}N resonances were completely assigned except for residues G66, T67, Q133 and H166, probably because of broadening of their spectral signatures. The expressed fragment (sequence shown in Figure 1A) contains 6 glycine residues: the N-terminal G66 remaining after TEV cleavage (that is not present in the original emerin sequence), G69, G89, G106, G112 and G156. The ^{1}H-^{15}N HSQC signals of the 5 emerin glycines are clearly visible (Figure 1B). The expressed fragment also contains 6 proline residues: P77, P108, P113, P125, P153 and P169. The main conformations of these prolines are *trans*, as deduced from their Cβ chemical shift values, which are comprised between 31.9 and 32.2 ppm. However the 5 prolines P77, P108, P113, P125 and P169 give rise to supplementary sets of weak signals corresponding to minor *cis* conformations characterized by proline Cβ chemical shift values comprised between 32.9 and 34.8 ppm (Shen and Bax, J Bio NMR 2009). In the 2D 1H-15N HSQC spectrum, the signals due to proline *cis* conformations and the resulting different chemical environments in their vicinity are colored as a function of the proline in *cis* (Figure 1B). The limited chemical shift dispersion of the 1H-15N HSQC signals, particularly in the 1H dimension (between 7.6 and 8.7 ppm), indicates that the emerin fragment is not folded in the NMR experimental conditions (20 mM phosphate, 30 mM NaCl, pH 6.5). Comparison of experimental and calculated backbone 15N chemical shifts confirms that the emerin chemical shifts correspond to random coil (Figure 2A; Tamiola et al. 2010). Similarly, the difference between the secondary chemical shifts of Cα and Cβ yields small values (generally under 1 ppm), which is consistent with intrinsic disorder (Figure 2B). In the case of the 8 M urea-denatured emerin fragment (sample DS), chemical shifts of 97% 1H–15N pairs (96 out of 99 non-proline residues, the 3 missing residues being D70, Q133 and H166), 100% 13Cα (105 out of 105) and 98% 13Cβ (97 out of 99) were assigned (Figures 2 C,D).

Here again, analysis of the differences between the Cα and Cβ secondary chemical shifts confirms that the emerin fragment is disordered in 8 M urea (Figure 2D).

The chemical shift data have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu) under accession numbers 26654 for the emerin fragment in 0 M urea and 26655 for the same fragment in 8 M urea.

Acknowledgments

We thank the French Association against Myopathies (AFM) (research grant n°17243 to S.Z.J and fellowship n°18159 to C.S.) and from the Foundation for Medical Research (FRM) (grant FDT20140931008 to I.H.) for providing financial support to this project.

References

Berk, J M, D N Simon, C R Jenkins-Houk, J W Westerbeck, L M Gronning-Wang, C R Carlson, and K L Wilson. 2014. 'The molecular basis of emerin-emerin and emerin-BAF interactions', *Journal of Cell Science*, 127: 3956-69.

Berk, J. M., K. E. Tifft, and K. L. Wilson. 2013. 'The nuclear envelope LEM-domain protein emerin', *Nucleus*, 4: 298-314.

Clements, L, S Manilal, D R Love, and G E Morris. 2000. 'Direct Interaction between Emerin and Lamin A', *Biochemical and Biophysical Research Communications*, 267: 709-14.

Guilluy, Christophe, Lukas D Osborne, Laurianne Van Landeghem, Lisa Sharek, Richard Superfine, Rafael Garcia-Mata, and Keith Burridge. 2014. 'Isolated nuclei adapt to force and reveal a mechanotransduction pathway in the nucleus', *Nat Cell Biol*, 16: 376-81.

Haque, Farhana, David J Lloyd, Dawn T Smallwood, Carolyn L Dent, Catherine M Shanahan, Andrew M Fry, Richard C Trembath, and Sue Shackleton. 2006. 'SUN1 interacts with nuclear lamin A and cytoplasmic nesprins to provide a physical connection between the nuclear lamina and the cytoskeleton.', *Molecular and Cellular Biology*, 26: 3738-51.

Holaska, James M, Amy K Kowalski, and Katherine L Wilson. 2004. 'Emerin Caps the Pointed End of Actin Filaments: Evidence for an Actin Cortical Network at the Nuclear Inner Membrane', *PLoS Biol*, 2: e231.

Lin, F., D. L. Blake, I. Callebaut, I. S. Skerjanc, L. Holmer, M. W. McBurney, M. Paulin-Levasseur, and H. J. Worman. 2000. 'MAN1, an inner nuclear membrane protein that shares the LEM domain with lamina-associated polypeptide 2 and emerin', *J Biol Chem*, 275: 4840-7.

Roberts, R. C., A. J. Sutherland-Smith, M. A. Wheeler, O. N. Jensen, L. J. Emerson, Spiliotis, II, C. G. Tate, J. Kendrick-Jones, and J. A. Ellis. 2006. 'The Emery-Dreifuss muscular dystrophy associated-protein emerin is phosphorylated on serine 49 by protein kinase A', *FEBS J*, 273: 4562-75.

Vranken, Wim F, Wayne Boucher, Tim J Stevens, Rasmus H Fogh, Anne Pajon, Miguel Llinas, Eldon L Ulrich, John L Markley, John Ionides, and Ernest D Laue. 2005. 'The CCPN data model for NMR spectroscopy: Development of a software pipeline', *Proteins*, 59: 687-96.

Legends

Figure 1

A. Amino acid sequence of the emerin fragment produced for this study, comprising an terminal glycine residue (shown in red) that results from the TEV-cleavage of the purification tag, and the emerin residues between 67 and 170. The colored residues are observed as two signals on the ^{1}H-^{15}N HSQC spectrum, the less intense signal corresponding to a proline *cis*-conformation. They are represented in green (residues around P77), blue (P108 and P113), yellow (P125) and light grey (P169).

B. 2D 1H-15N HSQC spectrum of the emerin fragment in 0 M urea at pH 6.5 and 303 K. Assigned backbone amide resonances are labelled. The spectrum was acquired on a 750 MHz spectrometer equipped with a TCI cryoprobe. An enlarged view of the most crowded region of the spectrum is shown in the down-left corner.

Figure 2

A. Experimental 15N chemical shift values measured in 0 M urea plotted against the calculated 15N chemical shift values predicted for a random coil conformation by ncIDP (http://nmr.chem.rug.nl/ncIDP/).

B. Plot of the difference between the secondary 13C chemical shifts of the Cα (noted as $\Delta\delta C\alpha$) and the Cβ (noted as $\Delta\delta C\beta$) nuclei. Secondary chemical shifts $\Delta\delta C\alpha_i$ and $\Delta\delta C\beta_i$ were obtained by subtracting the random coil chemical shifts (predicted by ncIDP) from the observed chemical shifts. Differences between the secondary chemical shift deviations $\Delta\delta C\alpha$ - $\Delta\delta C\beta$ were plotted against the amino acid residue numbers, taking into account next neighbor effects. Thus $\Delta\delta C\alpha$ – $\Delta\delta C\beta$ was calculated as:

$$\Delta\delta C\alpha - \Delta\delta C\beta = (\Delta\delta C\alpha i\text{-}1 + \Delta\delta C\alpha\ i + \Delta\delta C\alpha\ i\text{+}1) - (\Delta\delta C\beta\ i\text{-}1 + \Delta\delta C\beta\ i + \Delta\delta C\beta\ i\text{+}1)$$

C. Superposition of the 2D 1H-15N HSQC spectra of the emerin fragment recorded at pH 6.5 and 303 K in buffers containing 0 (magenta), 5 (beige) and 8 M urea (purple). Zoom on 2 glycine residues.

D. Same as B) but deduced from the NMR analysis of the emerin fragment in 8 M urea.

Figure 1

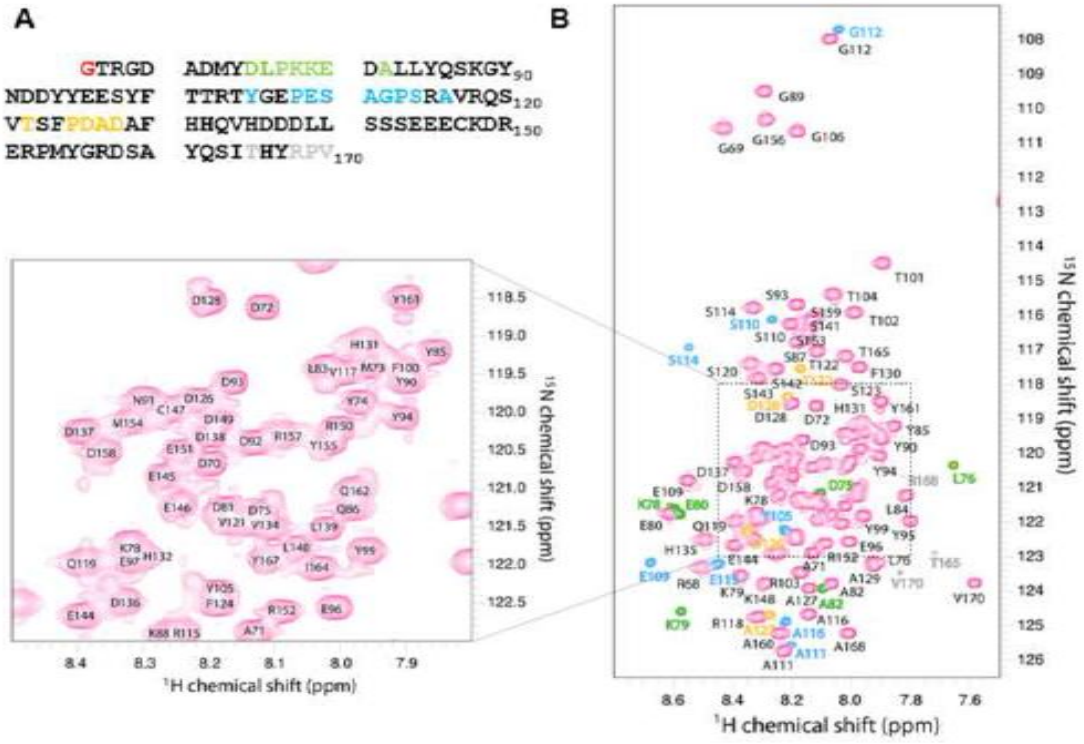
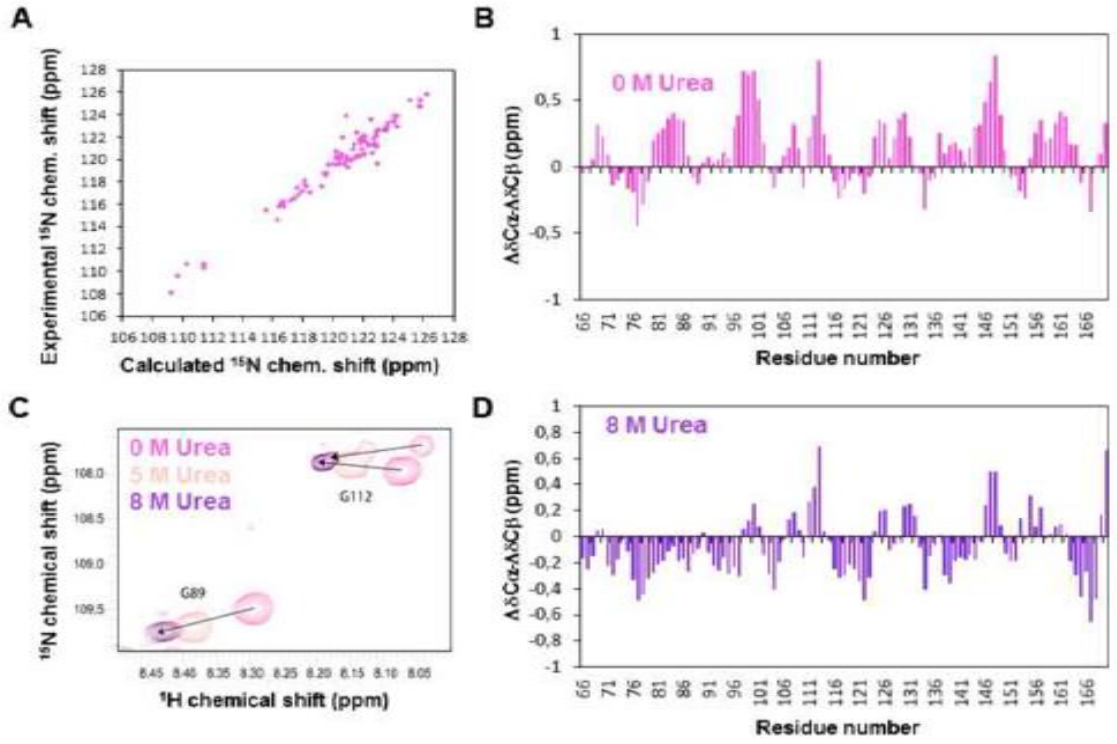


Figure 2



APPENDIX 2

Additional Figures

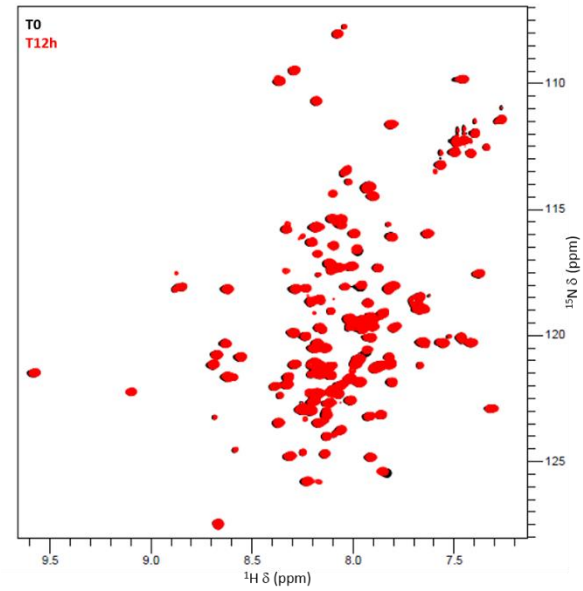


Figure 101 : EmN132 phosphorylation kinetics, in the presence of CK1, by NMR.

EmN132 sample (in black) or an EmN132 sample phosphorylated by CK1 during 12h (in red) at 100µM in 20mM Phosphate pH7, 30mM NaCl, 10mM β-mercaptoethanol, 2mM ATP, 5mM MgCl2, at 303K and 750MHz (FMP Berlin).

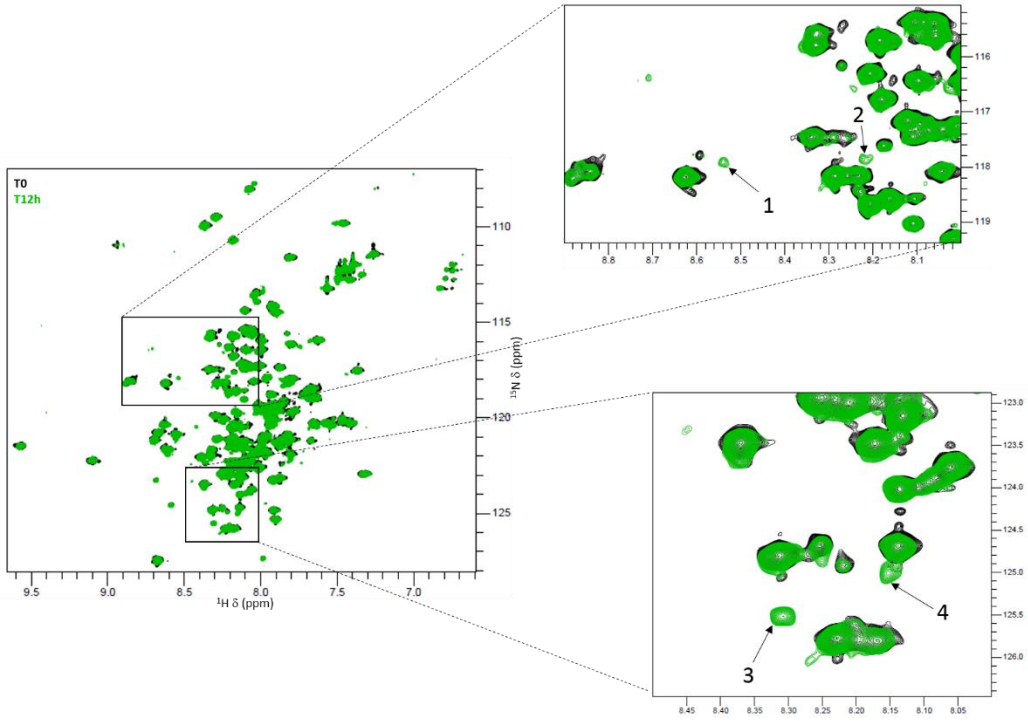


Figure 102 : EmN132 phosphorylation kinetics, in the presence of CK2, by NMR

EmN132 sample (in black) or an EmN132 sample phosphorylated by CK2 during 12h (in green) at 100µM in 20mM Phosphate pH7, 30mM NaCl, 10mM β-mercaptoethanol, 2mM ATP, 5mM MgCl2, at 303K and 750MHz (FMP Berlin).

Experiments done at 15°C	BAF / Igfold		BAF / EmN		BAF / IgfoldProgerin	
	N = 5		N = 3		N = 2	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Ns	0.35	0.09	0.25	0.02	0.22	0.04
Ka (M^{-1})	3.7E5	2.1E5	1.4E6	2.9E5	4.2E5	2.1E5
Kd (µM)	3.15	1.16	0.7	0.17	2.7	1.3
ΔH (cal/mol)	-1.15E4	2.1E3	-4.1E3	9.1E2	-1.6E4	9.5E3
ΔS (cal/mol/deg)	-13.5	9.3	-13.9	3.1	-30.9	34

Table 6 : Table containing the mean and the standard deviation of all ITC experiments done at 288K.

N = number of data points, Ns = number of sites, Ka = the association rate constant, Kd = the dissociation rate constant, ΔH = enthalpy and ΔS = entropy.

Experiments done at 10°C	Igfold WT		Igfold A529V		Igfold R471C	
	N = 2		N = 2		N = 2	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Ns	0.39	0.02	0.38	0.05	0.39	0.16
Ka (M^{-1})	7.2E5	2.6E5	1.5E5	1.4E4	1.4E5	1.6E4
Kd (µM)	1.5	0.54	6.9	1.6	7.2	0.8
ΔH (cal/mol)	-1.3E4	5.9E2	-8E3	1.6E3	-1.1E4	3E3
ΔS (cal/mol/deg)	-20.5	2.75	-4.8	5.9	-14	10

Table 7 : Table containing the mean and the standard deviation of all ITC experiments done at 283K.

N = number of data points, Ns = number of sites, Ka = the association rate constant, Kd = the dissociation rate constant, DH = enthalpy and DS = entropy.

APPENDIX 3

Article submitted in *Nature Structural and Molecular Biology*:

Samson *et al.,* Structural analysis of the ternary complex between lamin A/C, BAF and emerin identifies an interface disrupted in autosomal recessive progeroid diseases

Structural analysis of the ternary complex between lamin A/C, BAF and emerin identifies an interface disrupted in autosomal recessive progeroid diseases

Camille Samson[1], Ambre Petitalot[1], Florian Celli[1], Isaline Herrada[1], Virginie Ropars[1], Marie-Hélène Le Du[1], Naïma Nhiri[2], Eric Jacquet[2], Ana-Andrea Arteni[1], Brigitte Buendia[3] & Sophie Zinn-Justin[1] *

[1] Institut de Biologie Intégrative de la Cellule (I2BC), CEA, CNRS, Université Paris Sud, Université Paris-Saclay, Gif-sur-Yvette, France

[2] Institut de Chimie des Substances Naturelles, Université Paris Sud, Université Paris-Saclay, CNRS UPR 2301, Gif-sur-Yvette, France.

[3] Unité de Biologie Fonctionnelle et Adaptative (BFA), CNRS UMR 8251, Université Paris Diderot, Sorbonne Paris Cité, Paris, France.

* Correspondence should be addressed to S.Z.J. (sophie.zinn@cea.fr)

Abstract

Lamins are the main components of the nucleoskeleton. Whereas their 3D organization was recently described using cryoelectron tomography, no structural data highlights how they interact with their partners at the interface between the inner nuclear envelope and chromatin. A large number of mutations causing rare genetic disorders called laminopathies were identified in the C-terminal globular Igfold domain of lamins A and C. We here present a first structural description of the interaction between the lamin A/C immunoglobulin-like domain and a nuclear envelope protein. We reveal that this lamin A/C domain both directly binds self-assembled emerin and interacts with monomeric emerin LEM domain through the dimeric chromatin-associated Barrier-to-Autointegration Factor (BAF) protein. Mutations causing autosomal recessive progeroid syndromes specifically impair proper binding of lamin A/C domain to BAF, thus destabilizing the interface between lamin A/C and BAF in cells. Our results suggest that these mutations could perturb the BAF-mediated interaction between lamin A/C and chromatin, thus leading to a dysregulation of chromatin organization and gene expression.

Laminopathies are rare genetic disorders caused by mutations in genes encoding lamins or by abnormalities in the processing of lamin A [1]. They display a large variety of clinical symptoms including cardiomyopathy, muscular dystrophy, lipodystrophy, mandibuloacral dysplasia, restrictive dermopathy, peripheral neuropathy, and premature ageing. A major fraction of the disease-causing mutations comprises single point mutations with poorly understood functional implications. Most of them are distributed all along the *LMNA* gene coding for both prelamin A and lamin C. These two proteins are, together with lamins B1 and B2, essential components of the nucleoskeleton. They share an N-terminal region of 566 residues, and whereas lamin C only possesses 6 additional specific C-terminal residues, prelamin A presents 98 additional residues, is farnesylated, carboxymethylated and then cleaved to become mature lamin A. Disease-causing mutations are observed in the large region common to lamins A and C (lamin A/C) as well as in the lamin A specific C-terminal region [2]. At the cellular level, lamins A and C form filaments that are mainly present at the nuclear periphery but are also observed in the nucleoplasm [3,4]. Laminopathies are characterized by nuclear morphological abnormalities and an altered pattern of heterochromatin distribution that is more severe in progeroid syndromes, including Hutchinson-Gilford progeria syndrome (HGPS), mandibuloacral dysplasia (MAD), atypical-Werner syndrome (WS) and restrictive dermopathy (RD) [5,6]. The lamin filament network interacts with chromatin at different stages of the cell cycle: in ana-telophase, lamin A and C dimers are recruited at the core regions of sister chromosomes; in interphase cells, they relocalize both at the nuclear envelope and within the nucleoplasm, where lamins A and C provide anchor points for chromatin [7-10]. The complex molecular organization of the nuclear envelope, nucleoskeleton and their interface with chromatin was recently described using cryo-electron tomography [11,12]. Within the densely packed environment observed close to the nuclear envelope, only nuclear pores, lamin filaments and chromatin could be recognized. Lack of high resolution 3D structures for proteins anchored at the inner nuclear membrane as well as complexes between these proteins, the lamina and chromatin still precludes any further description of the nuclear periphery architecture and of the structural defects caused by mutations in the *LMNA* gene coding for lamins A and C.

We set out to understand how the C-terminal immunoglobulin-like (Igfold) domain of lamin A/C interacts with the inner nuclear membrane protein emerin and the chromatin-associated protein Barrier-to-Autointegration Factor (BAF). Indeed, a large number of disease-causing mutations were identified in this lamin Igfold domain [13,14]. Mutations causing muscle diseases affect residues of the hydrophobic core, suggesting that destabilization of the lamin A/C Igfold domain is responsible for these diseases [14]. Mutations causing either lipodystrophies or progeroid syndromes involve residues localized on two different solvent-exposed sites of the lamin A/C Igfold domain, suggesting that these accessible regions are binding sites for unknown lamin partners and that disruption of these binding events cause the diseases [13-15]. Lamin A/C binds to the inner nuclear membrane protein emerin, as shown by blot overlay, coimmunoprecipitation and yeast 2-hybrid assays [16-18]. Moreover, the tail region of lamin A/C (residues 384 to 566, comprising the Igfold domain) is responsible for emerin binding [18,19]. The central region of emerin (residues 70 to 178) is essential for lamin binding [16,19], and its N-terminal domain (residues 1 to 45, the LEM domain) interacts with BAF, as shown by a panel of techniques going from blot overlay and coimmunoprecipitation assays to X-ray crystallography analyses [16,20].

The structural inter-dependence of lamins, emerin and BAF was revealed by downregulating either lamins A and C or BAF or the LEM domain proteins emerin and LEM2 in *C. elegans* embryos. If any one component was missing, the other two failed to co-assemble, with severe consequences for mitotic spindle assembly and positioning [21], mitotic chromosome segregation and postmitotic nuclear assembly [22,23]. *In vitro* experiments revealed that human lamin A binds weakly (K_d of 1 μM) but directly to BAF and that lamin A and BAF simultaneously bind to emerin [24]. Moreover, BAF enhances binding of prelamin A tails to emerin [25].

We now present the 3D structure of a complex involving lamin A/C, emerin and BAF. We show that emerin oligomerization regulates direct lamin A/C binding. We discuss the role of these lamin-emerin interactions during the cell cycle. We further show that BAF assembly within these complexes is impaired by lamin mutations causing autosomal recessive progeroid syndromes, thus highlighting the essential functional role of interactions between nuclear envelope and chromatin-associated proteins in human cells.

RESULTS

The lamin A/C Igfold domain interacts with the self-assembled emerin fragment EmN

To determine how the lamin A/C tail interacts with emerin, we generated two fragments of the emerin nucleoplasmic region (Figure 1A). We tested binding between the tail region common to lamins A and C (LamIgF comprising aa 411 to aa 566) and these emerin fragments. As previously published, the whole emerin nucleoplasmic region from aa 1 to aa 221 is poorly soluble [26]. We thus produced two overlapping emerin peptides, which we called EmN [27] and EmC (Figure 1A). EmN (aa 1 to aa 187) comprises a LEM domain and a region that is intrinsically disordered [26,28]. We previously showed that, *in vitro*, it can be observed either as a monomer or as an oligomer [27] (Figure 1B, left panel). EmC (aa 67 to aa 221) is entirely unstructured as observed by NMR (Suppl. Fig. 1A), and after purification immediately oligomerizes as observed by fluorescence (Suppl. Fig 1B) and negative-staining electron microscopy (Figure 1B, right panel). In order to identify an interaction between lamin A/C and emerin, we produced a ^{15}N labeled LamIgF peptide, and mixed the labeled peptide with either monomeric EmN, oligomeric EmN or oligomeric EmC (Figure 1C). Nuclear Magnetic Resonance (NMR) analysis of these samples revealed that LamIgF only interacts with oligomeric EmN in these conditions. This interaction is specific to lamin A/C as NMR analysis of a sample containing ^{15}N labeled lamin B1 tail and oligomeric EmN did not demonstrate any binding (Figure 1C, lower and right panel).

The Igfold domain of lamin A/C forms a ternary complex with the LEM domain of emerin and BAF

The chromatin-associated protein BAF binds to both lamin A/C and emerin [29,30]. We examined the role of BAF in mediating an interaction between LamIgF, EmN and BAF by Size-Exclusion Chromatography (SEC). Consistent with our previous NMR results, we observed that LamIgF and monomeric EmN did not co-elute under the conditions used (Figure 2A). However, they co-eluted in the presence of BAF. Similarly, we tested the binding between the lamin B1 tail fragment, EmN and BAF. In this case, no binding could be observed between lamin B1 and BAF, precluding the formation of a ternary complex with EmN. NMR further confirmed that the lamin B1 tail fragment did not interact with BAF (Suppl. Fig. 2).

Isothermal Titration Calorimetry (ITC) experiments showed that LamIgF and EmN bind to BAF with an affinity of 3.2 ± 1.2 μM and 0.7 ± 0.2 μM, respectively (Figure 2B; Suppl. Table 1). Crystal structure of the ternary complex was solved at 1.9 Å resolution (Figure 3A; Suppl. Table 2). One BAF dimer interacts on one side with the LEM domain of emerin (interface 1) and on the other side with the Igfold domain of lamin A/C (interface 2). The emerin fragment outside of the LEM domain is not visible in the electron density map. SDS-PAGE revealed that it was proteolyzed between protein purification and crystallization.

Dimerization of BAF is essential for both lamin and emerin binding

The 3D structure of the ternary complex highlights that BAF dimerization is essential for emerin and lamin A/C binding. First, BAF monomers 1 and 2 both interact with the LEM domain of emerin (interfaces of 395 $Å^2$ and 211 $Å^2$, respectively) (Figure 3B). BAF monomer 1 loop α2α3 (Phe39), helix α3 (Gln48, Val51, Leu52) and helix α4 (Leu58, Glu61, Trp62) contact a LEM surface comprising the β-strand (Gly24, Pro25), loop β1α2 (Ser29, Thr30) and helix α2 (Leu33, Tyr34, Lys36, Lys37). BAF monomer 2 loop α2α3 (Arg37, Phe39) and helix α3 (Glu48) contact an overlapping LEM surface comprising helix α1 (Thr13), the β-strand (Pro25) and loop β1α2 (Val27). A similar 3D structure between BAF dimer and emerin LEM domain was previously modeled based on NMR data (Figure 3C; PDB reference 2ODG [20]). The surface buried at the LEM/BAF interface yields 510 $Å^2$ in the NMR structure compared to 606 $Å^2$ in the X-ray structure, showing that the interaction is significantly tighter in the crystal structure.

Both BAF monomers also contribute to lamin A/C binding, through interfaces of 270 $Å^2$ and 192 $Å^2$, respectively (Figure 4A). BAF monomer 1 helix α1 (Val11, Ala12), loop α1α2 (Pro14) and helix α5 (Glu83, Ala87, Phe88) contact lamin A/C strand β1 (His433, Ala434, Arg435) and strand β9 (Met540, Lys542), whereas BAF monomer 2 helix α1 (Ala12), loop α1α2 (Pro14) and helix α5 (Ala87, Phe88) contact lamin A/C strand β1 (Arg435) and loop β8β9 (Glu537). It was reported that BAF mutant G47E lost its ability to bind to lamin A as well as prelamin A in cells [6]. Gly47 is located at the center of the dimerization interface (in magenta on Figure 4A). Mutation G47E might destabilize the BAF dimer, therefore hindering BAF interaction with both emerin LEM domain and lamin A/C Igfold.

The interface between lamin A/C and BAF involves residues mutated in progeroid syndromes

By solving the 3D structure of a ternary complex formed by BAF, EmN and LamIgF, we revealed a yet undescribed interface between the BAF dimer and the lamin A/C Igfold domain. On the BAF side, Ala12, Pro14 and Phe88 of both monomers are more than 50% buried in the complex (Figure 4A). An Ala12 to Thr amino acid substitution occurs in patients with a Nestor-Guillermo progeria syndrome [31]. Our 3D structure predicts that this substitution causes a defect in BAF/lamin binding. Consistently, in cells, it significantly decreases BAF binding to lamin A and prelamin A [6].

On the lamin side, Arg435, Met540 and K542 are more than 50% buried at the lamin/BAF monomer 2 interface (Figure 4A). Arg435 side-chain is hydrogen-bonded to Phe88 backbone oxygens of both BAF monomers, while Lys542 side-chain is bonded to BAF monomer 2 Asp86 backbone oxygen. These lamin residues are mutated in patients with severe autosomal recessive progeroid syndromes (RD and HGPS [15,32-34]). They belong to a hot spot for mutations causing progeroid diseases [15]. In order to verify that in solution, this hot spot surface is contacting BAF, we followed the NMR ^{1}H and ^{15}N signals of LamIgF while adding BAF to the NMR sample. We observed that the intensities of the NMR signals corresponding to several residues of strand $\beta 1$ (as Arg435), loop $\beta 1\beta 2$, strand $\beta 4$ (as Arg471), loop $\beta 8\beta 9$ and strand $\beta 9$ (as Met540 and Lys542) were significantly decreased (Figure 4B). This demonstrated that the β-sheet formed by strands $\beta 1$ and $\beta 9$ is indeed involved in BAF binding in solution.

Lamin A/C Igfold mutations causing progeroid syndromes decrease binding affinity for BAF

To identify the consequences of the LamIgF mutations causing progeroid diseases on the formation of the lamin-BAF complex, we produced a set of LamIgF peptides mutated on sites responsible for severe diseases. We focused our interest on 5 homozygous LamIgF mutations, R435C (strand $\beta 1$) [32,33,35], R471C (strand $\beta 4$) [36], R527H [37-40], A529V [41] (strand $\beta 8$) and K542N (strand $\beta 9$) [34], which were identified in patients with atypical progeroid syndromes. Two of these mutations cause syndromes generally appearing during early childhood as RD (R435C) and HGPS (K542N), characterized by severe premature aging features, whereas mutations R471C, R527H and A529V cause a disease generally appearing later, called MAD, characterized by growth retardation and skeletal abnormalities.

In the case of R471C, the disease was particularly severe and the phenotype combined MAD, progeria, and rigid spine muscular dystrophy [36]. We first verified that the protein variants are well folded and have a thermal stability close to that of WT LamIgF (i.e. ± 5° relatively to WT; Suppl. Fig. 3). We then measured their affinity for BAF by ITC (Figure 5A; Suppl. Table 1). Compared to WT LamIgF and to mutants R453W and R482W causing Emery-Dreifuss Muscular Dystrophy (EDMD) and Dunnigan-type Familial Partial Lipodystrophy (FPLD), respectively, mutant R435C showed a complete loss of detectable binding to BAF in our conditions. Mutants K542N and R527H bind BAF but the corresponding heat release is so weak that no affinity could be measured. R471C and A529V exhibit a 5-fold decreased affinity for BAF. We also tested the binding of LamIgF mutants to BAF by SEC (Figure 5B). We observed that mutants R453W and R482W coeluted together with BAF as WT LamIgF. However, mutant R435C did not induce any elution volume shift of BAF, and mutants R471C, A529V, R527H and K542N only weakly shifted BAF elution volume. We concluded that the 5 tested mutations causing atypical progeroid syndromes decrease the affinity of lamin A/C for BAF. Moreover, the binding defect seems to globally correlate with the disease severity (Figures 5A,B).

Finally, we transfected HeLa cells with plasmids coding for GFP-BAF and FLAG-mature lamin A (aa 1 to aa 646) WT or variants and detected using an *in situ* Proximity Ligation Assay the proximities between BAF and lamin A in these cells. We thus tested the impact of mutations related to RD (R435C), HGPS (K542N) and MAD (R527H, A529V) on these proximities. FLAG-lamin variants were expressed at similar protein levels as FLAG-lamin WT; they localized both at the nuclear periphery and in the nucleoplasm (Figure 6A,B). We observed that all 4 lamin A mutations associated with atypical progeroid syndromes similarly reduced the frequency of the proximity events between BAF and lamin A (Figure 6C, D).

DISCUSSION

The mechanical roles of lamins and their functions in gene regulation are often viewed as independent activities, but recent findings suggest a highly cross-linked and interdependent regulation of these different functions [42-44]. Lamins interact with inner nuclear membrane proteins to mediate mechanical signaling but also contact chromatin-associated proteins that contribute to gene transcription regulation. Most of these interactions depend on the lamin ability to assemble into higher-order structures [45,46]. Therefore, lamin assembly has been studied intensively *in vitro* and the resulting filaments have been observed in cells [12]. In contrast, little is known about the contacts of lamins with their partners *in vivo*. Lamins are extensively modified and interact with a large number of proteins [47,48], the inner nuclear membrane [49] and chromatin [50]. Their tail region is responsible for association with partners. It contains a globular Igfold domain that is mutated in several laminopathies [14]. Until this study, it was not known how this domain interacted with partners anchored at the inner nuclear membrane and established contacts with chromatin.

The lamin A/C Igfold domain directly binds to self-assembled emerin but also recognizes monomeric emerin through BAF

We showed that the lamin A/C Igfold domain interacts with the inner nuclear membrane protein emerin through two mechanisms: it either directly binds to self-assembled EmN or interacts with monomeric EmN through the chromatin-associated protein BAF. It was previously published that direct lamin A tail interaction with emerin is disrupted not by mutations in the LEM domain, but by mutations in emerin region from aa 70 to aa 164: mutations around positions 70, 76, 95, 112, as well as 141 and 164, decreased binding of lamin A to emerin [16,30]. These results are consistent with our analysis. Indeed, LamIgF only binds directly to self-assembled emerin and the disordered region of emerin between aa 50 and aa 132 is necessary for emerin self-assembly [27]. Similarly, GST fusions of an emerin fragment from aa 1 to aa 132 is sufficient to bind mature lamin A tails [19], and we showed that emerin aa 1 to aa 132 is sufficient for self-assembly [27]. Finally, the EDMD-associated emerin deletion mutation del95-99, which impairs EmN self-assembly *in vitro,* causes a decrease of lamin A/C binding in cells [27].

Altogether these results strongly suggest that emerin self-assembly regulates direct lamin A/C binding in cells. On the other hand, we showed that the monomeric LEM domain (emerin aa 1 to aa 45) interacts with the lamin A/C Igfold domain through a dimer of BAF, consistent with previous data showing that BAF enhanced binding of prelamin A tails to emerin [25].

A defect at the interface between lamin A/C and the chromatin-associated BAF protein is observed in atypical progeroid syndromes

By analyzing the 3D structure of our ternary complex comprising domains of lamin A/C, BAF and emerin, we noticed that one of the interfaces coincides with a hot spot of lamin A/C residues mutated in progeroid syndromes [15]. We measured *in vitro* and in cells the impact of several homozygous mutations causing autosomal recessive progeroid syndromes and showed that these mutations disrupt the interaction between the lamin A/C Igfold and BAF. In HeLa cells, all tested mutations consistently reduced the frequency of the proximity events between lamin A and BAF. As endogeneous lamin A probably forms filaments together with the FLAG-tagged lamin A, a number of PLA events were expected due to interactions between endogeneous lamin A and BAF that create proximities between associated FLAG-tagged lamin and BAF. Consistently, the mutations reduced the number of PLA signals to about 50% of the signals measured for WT lamin A and BAF. *In vitro*, a range of affinity decreases were observed. In particular, SEC experiments revealed that, in our conditions, LamIgF control mutants R453W and R482W bind BAF as WT LamIgF, mutants R471C, A529V, R527H and K542N weakly bind BAF, and mutant R435C did not bind BAF. ITC experiments provided only estimations for affinity decreases, due to the weak affinities of the mutants. However, an affinity decrease of about 5 fold could be calculated for mutants R471C and A529T, whereas essentially no binding could be detected for R527H, K542N and R435C. Interestingly, this affinity scale matches with the severity of the associated diseases: R471C and A529V cause MAD that generally appears during late childhood [36,41], whereas K542N causes a HGPS with an onset between 18 and 24 months [34] and R435C causes the most severe disease, RD, which appears during the first months after birth [32,33,35]; only mutant R527H, which causes a MAD [37-40] , also shows a strong BAF binding defect under the conditions we used. We conclude that a destabilization of the interface between lamin A/C and BAF might mediate the disease-mechanism in atypical progeroid syndromes due to homozygous mutations in LamIgF.

Another progeroid syndrome called the Nestor-Guillermo Progeria Syndrome occurs in patients older than the average lifespan of progeroid patients. This syndrome is due to the homozygous BAF mutation A12T [31]. The mutated protein is expressed and stable [51]. In cells, it still co-immunoprecipitates with endogeneous emerin, lamin A and histone H3. Our preliminary data show that, *in vitro*, BAF mutant A12T consistently interacts with LamIgF, but with an affinity lower than BAF WT and BAF S4E (Suppl. Fig. 5). Immunoprecipitation experiments performed in co-transfected cells showed a reduced interaction between BAF mutant A12T and lamin A [6]. BAF mutant A12T also exhibits a marked defect in double-stranded DNA binding compared to WT BAF [51]. Altogether, these data strongly suggest that the complex between BAF, lamin A/C, DNA and emerin is destabilized by the mutation A12T, highlighting the similarity between the mechanisms of autosomal recessive progeroid syndromes due to mutations in LamIgF and BAF.

Complexes including lamin A/C and emerin are assembled at different stages of the cell cycle

In cells, formation and localization of complexes including lamin A/C, BAF and emerin depend on the cell cycle, in accord with the disassembly/reformation of the nuclear envelope that delimits the interphase nuclear compartment. While in interphase cells, lamins assemble into polymers within the nuclei, they disassemble at mitosis, leading to cytoplasmic lamin A/C dimers. At that stage, pools of lamin A/C and BAF form stable protein complexes with the LEM domain protein LAP2α; these complexes contribute to microtubule spindle assembly and orientation [21]. Interestingly, knocking down lamin A severely affects binding of LAP2α to GFP-BAF and knocking down BAF disrupts binding of LAP2α to GFP-lamin A, suggesting that formation of a lamin A-BAF complex is essential for the recruitment of the LEM domain LAP2α protein. Later, in ana-telophase, first BAF, and then emerin and lamin A/C accumulate at the core region of chromosomes [52]. On the BAF side, mutation G25E, at the interface with DNA, impairs localization of BAF at the core region of chromosomes, emerin localization during telophase and further lamin A accumulation at reforming nuclear envelopes. On the emerin side, the LEM domain is necessary and sufficient for the core localization of emerin. Superimposition of our 3D structure with that of a complex between BAF dimer and DNA (PDB reference 2BZF [53]) revealed that the BAF dimer can simultaneously bind to emerin, lamin and DNA (Figure 7, left panel).

We propose that BAF mediates the interaction between emerin and lamin A at the core region. The fact that emerin must interact with BAF at the core region of chromosomes before being assembled at the nuclear envelope [54,52] also suggests that post-translational modification events occur at the core that further enable emerin assembly at the nuclear envelope.

In interphase cells, emerin binds to lamin A/C [30]. Moreover, an emerin protein mutated in its LEM domain (residues 24-27 being changed into alanines), which does not interact with BAF [16], localizes normally at the nuclear envelope [54], consistent with its ability to directly bind lamin A/C *in vitro* [16]. Emerin self-assembles at the nuclear envelope [19,55]. From our NMR results, we propose that emerin self-assembly creates a high concentration of binding sites for the lamin filaments at the nuclear periphery of cells (Figure 7, right panel). During *in vitro* self-assembly, emerin exhibits a large LEM domain conformational change, strongly suggesting that self-assembled emerin does not bind BAF [26]. Altogether, previous studies as well as our data support a model in which, in interphase cells, emerin exists in two conformations, a self-assembled conformation that interacts with lamin A/C and a conformation in which the LEM domain is free to bind BAF [19,55]. Whether trimeric emerin-BAF-lamin A/C complexes also form at the nuclear envelope of interphase cells remains an open question.

A defective interaction between lamin A/C and BAF could impact chromatin organization of interphase cells

Fluorescence resonance energy transfer experiments showed that YFP-emerin and CFP-BAF interact at the inner nuclear membrane of interphase HeLa cells [56]. In this case, emerin is still relatively immobile, but BAF is dynamic and mobile: it binds to emerin frequently but transiently using a “touch-and-go” mechanism. We also observed proximities between lamin A and BAF in interphase HeLa cells. These proximities detected in situ using a proximity assay are particularly enriched at the nuclear periphery (Suppl. Fig. 6). They are less frequent in cells expressing lamin A variants that are defective for BAF binding *in vitro*. Interestingly, defects of the nuclear envelope-chromatin interface were revealed using electron microscopy in the case of progeroid syndromes caused by lamin A or BAF mutations.

A loss of nuclear peripheral heterochromatin was observed both in cells of patients with autosomal dominant HGPS due to a mutation G608G in prelamin A and in HEK293 cells overexpressing lamin A and BAF A12T [6,57,58,59]. Moreover, lamin A and BAF were shown to directly bind histones and to influence histone epigenetic marks, as H3K9me3, H3K27me3, H3K4me2 and H3K79me2 [60,61,62]. We conclude that defect in the recruitment of BAF to the lamina via the immunoglobulin-like domain of lamin A/C could be responsible for abnormal heterochromatin distribution at the nuclear envelope of interphase cells. More work is needed to understand how the lamin-BAF complex interacts with DNA and histones, and how a defect in this complex compromises the interactions between lamina and chromatin at the nuclear periphery in autosomal recessive progeroid syndromes.

In conclusion, the Igfold domain of lamin A/C can both directly interact with self-assembled emerin and bind to monomeric emerin through the chromatin-associated protein BAF. A defect in the interaction between lamin A/C and BAF might destabilize ternary complexes with emerin, thus leading to a less efficient targeting of emerin at the reforming nuclear envelope. However, diseases associated with mutations disrupting the lamin A/C-BAF interaction are much more severe than diseases associated with emerin loss of function. This strongly suggests that the impact of these lamin A/C and BAF mutations goes beyond LEM-domain protein mislocalization and loss of function. We propose that the disease-causing mutations perturb the BAF-mediated interaction between lamin A/C and chromatin, thus leading to a dysregulation of chromatin organization and gene expression.

Acknowledgments - We thank Prof. H.J. Worman for providing the plasmid coding for GST-LamIgF and reviewing the manuscript and Prof. G. Lattanzi for providing the plasmid coding for GFP-BAF. We also thank N. Vadrot for her help with cell cultures. This work was supported by CEA, CNRS, University Paris South and University Paris Diderot, by the French Infrastructure for Integrated Structural Biology (https://www.structuralbiology.eu/networks/frisbi, grant number ANR-10-INSB-05-01, Acronym FRISBI) and by the French Association against Myopathies (AFM; research grants no. 17243 and 20018 to S.Z.-J. and PhD fellowship no. 18159 to C.S.).

Author contributions – S.Z.J. conceived and designed new experiments reported in this paper. C.S., A.P., F.C., I.H., V.R., N.N, E.J., A.A.A. and B.B. performed the experiments. C.S., A.P., E.J., B.B. and S.Z.J. analyzed data. M.H.L. supervised the analysis of the X-ray crystallography data. C.S., B.B. and S.Z.J. wrote the manuscript.

Competing financial interests – The authors declare no competing financial interests.

Methods.

Protein Expression and Purification.

Human WT emerin fragments from amino acid 1 to amino acid 187 (EmN), from amino acid 67 to 170 (EmC170), from amino acid 67 to 187 (EmC187) and from amino acid 67 to amino acid 221 (EmC), human BAF with all cysteines mutated in alanine and human lamin B1 tail from amino acid 409 to amino acid 586 (LB1) were expressed using a pETM13 vector as N-terminal octa-histidine fusions in *Escherichia coli* BL21 DE3 Star (Novagen). The emerin and BAF cDNAs were optimized for expression in *Escherichia coli* (Genscript). Human wild-type lamin fragment from amino acid 411 to amino acid 566 (LamIgF) was expressed using a pGEX vector as an N-terminal GST fusion in *Escherichia coli* BL21 DE3 Star (Novagen). All LamIgF mutant expression vectors (R435C, R453W, R471C, R482W, R527H, A529V, and K542N) were obtained by mutagenesis using the Quikchange (Agilent) kit from the LamIgF expression vector. Bacteria were cultured in rich medium (*lysogeny broth, LB*) or ^{15}N-labeled minimum medium, induced at an optimal density of 1 with 0.5mM isopropyl β-D-1-thiogalactopyranoside, grown overnight at 293K, and lysed in 50mM Tris-HCl pH 8.0, 300mM NaCl, 40mM imidazole, 5% glycerol, 1% Triton X-100, and 1mM phenylmethanesulfonylfluoride.

For EmN, EmC170, EmC187, EmC and BAF, after sonication at room temperature and centrifugation at 20 000 g, for 20 min, at 277K, the pellet was resuspended in buffer C8 (50mM Tris-HCl pH8.0, 150mM NaCl, 20mM imidazole, 8M urea). A second centrifugation step was performed at 20 000 g, for 20 min, at 293K. The soluble extract was then loaded onto Ni-NTA beads (GE-Healthcare) equilibrated with buffer C8. Proteins were eluted directly with buffer E8 (50mM Tris-HCl pH 8.0, 150mM NaCl, 1M imidazole, 8M urea). Then, proteins were refolded by dialysis in buffer D1 (20mM Tris-HCl pH8.0, 30mM NaCl, 2mM DTT) for EmN, EmC170, EmC187 and EmC and buffer D2 (50mM Tris-HCl pH8.0, 150mM NaCl, 2mM DTT) for BAF. After refolding, they were incubated with the TEV protease during 3 h at room temperature and finally dialyzed into the selected buffer. The final yield was typically 10 mg of purified protein per liter of bacterial culture [27].

Self-assembly was initiated using proteins that were dialyzed in a buffer containing 20mM Tris-HCl pH8 and 30mM NaCl. After dialysis, proteins were concentrated until 600µM, reduced with 5mM DTT and stored at room temperature during one week.

For LB1, after sonication at 277K, benzonase (SigmaAldrich) addition and centrifugation at 20 000 g, for 30 min, at 277K, the soluble extract was loaded onto Ni-NTA beads (GE-Healthcare) equilibrated with buffer C (50mM Tris-HCl pH8.0, 150mM NaCl and 20mM imidazole). The protein was then eluted directly with buffer E (50mM Tris-HCl pH 8.0, 150mM NaCl and 1M imidazole) and dialyzed in buffer C. It was incubated with the TEV protease during 3h at room temperature, loaded again onto Ni-NTA beads and finally, after flow-through recovery, dialyzed into the selected buffer. The final yield was typically 20 mg of purified protein per liter of bacterial culture.

For LamIgF (WT and mutants), after sonication at 277K, benzonase (SigmaAldrich) addition and centrifugation at 20 000 g, for 30 min, at 277K, the soluble extract was loaded onto GST-beads (GE-Healthcare) equilibrated with buffer T2 (50mM Tris-HCl pH7.5, 150mM NaCl, 2mM DTT). After 2h of incubation at 277K, GST-beads were washed with buffer T2 and 1M NaCl, and then thrombin was added and incubated with the beads overnight. The protein was recovered in the flow-through and finally dialyzed into the selected buffer. The final yield was typically 2 mg of purified protein per liter of bacterial culture.

Liquid-State NMR Spectroscopy.

NMR experiments were performed on a 700 MHz Bruker Avance spectrometer equipped with a triple resonance cryogenic probe. To study interaction between the lamin fragments and BAF, two-dimensional ^{1}H-^{15}N HSQC spectra were acquired at 288 K, on a 3-mm-diameter NMR sample tube containing 100-200µM uniformly ^{15}N-labeled lamin peptides and non-labeled BAF, in 20mM sodium phosphate pH 7, 100mM NaCl and 80:20% H_2O/D_2O. To study interaction between lamin fragments and EmN or EmC, two-dimensional ^{1}H-^{15}N HSQC spectra were acquired at 303 K, on a 3-mm-diameter NMR sample tube containing 150 µM uniformly ^{15}N-labeled lamin peptides and non-labeled EmN or EmC, in 20mM Tris-HCl pH8, 30mM NaCl and 80:20% H_2O/D_2O.

In order to reassign LamIgF NMR signals in our conditions, three-dimensional ^{1}H-^{15}N-^{13}C correlation spectra were acquired using 3D HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO and HN(CO)(CA)NH pulse sequences at 303 K, on a 3-mm-diameter NMR sample tube containing 500 µM uniformly $^{15}N/^{13}C$-labeled LamIgF, in 20mM sodium phosphate pH 7, 30mM NaCl and 80:20% H_2O/D_2O. The data were processed using Topspin3.1 (Bruker) and analyzed using CCPNMR [63].

Size-exclusion chromatography.

Size-exclusion chromatography experiments aiming at identifying interactions between emerin, BAF and lamin fragments (LamIgF WT and mutants, LB1) were performed using a Superdex-75 10/300 GL column (GE Healthcare) pre-equilibrated with buffer G (20mM Tris-HCl pH8, 30mM NaCl, 2mM DTT). 500µl of proteins concentrated at 150µM were injected for each experiment at a flow rate of 0.5ml/min at 277K.

Isothermal Titration Calorimetry.

ITC was performed using a high-precision VP-ITC calorimetry system. To characterize interactions between the BAF dimer and LamIgF (WT and mutants), all proteins were dialyzed against 50mM Tris-HCl pH8, 150mM NaCl, 10mM β-mercaptoethanol and protease inhibitors (Roche). BAF (20µM) in the calorimetric cell at 288 K was titrated with LamIgF WT (at a concentration of 100 µM in the injection syringe) or BAF in the calorimetric cell at 283 K was titrated with LamIgF mutants (at a concentration of 100µM in the injection syringe). To characterize interactions between the BAF dimer and EmN, all proteins were dialyzed against 50mM Tris-HCl pH8, 150mM NaCl, 10mM β-mercaptoethanol and protease inhibitors (Roche). BAF (30µM) in the calorimetric cell at 288 K was titrated with EmN (at a concentration of 150µM in the injection syringe). Analyses of the data were performed using the Origin software provided with the instrument.

Negative-staining Electron Microscopy.

To observe self-assembled EmN and EmC proteins, with or without LamIgF WT or LB1, protein samples containing 2 % uranyl acetate were deposited on glow-discharged carbon-coated copper grids. Data collection was performed using a Tecnai Spirit transmission electron microscope (FEI) equipped with a LaB6 filament, operating at 100 keV. Images were recorded using a K2 Base camera (Gatan, 4kx4k) at 15,000 magnification (at the level of the microscope) with a pixel size of 0.25 nm at the specimen level.

X-ray crystallography.

Crystallization and data collection - The EmN-Baf-LamIgF complex was purified by gel filtration (Superdex 75, 10/300 GL) in 50mM Tris-HCl pH8, 100mM NaCl and concentrated to 3mg/ml. Crystallization was initiated one week after the gel filtration and in these conditions, EmN was proteolyzed and the final complex was composed of LEM-BAF-LamIgF. For crystallization, 1 µL of the complex was mixed with 1 µl of the reservoir solution and equilibrated against a 500-µl reservoir by hanging drop at 277K. Crystals were grown in 25% (w/v) polyethylene glycol monomethyl ether (PEG) 3.350, 20mM Bis-Tris pH5.5, 30mM NaCl and 0.2M NH4SO4. They were flash-cooled in liquid nitrogen, using a cryo-protection solution of 25% PEG 3.350, 20mM Bis-Tris pH5.5 and 0.2M NH4SO4, supplemented with 27% (V/V) ethylene glycol.

Structure determination and refinement - The 3D structure of the complex was solved by molecular replacement using Molrep in CCP4 [64]. The coordinates of the BAF dimer with DNA (PDB entry 2BZF), the coordinates of lamin A/C globular domain (PDB entry 1IFR) and the coordinates of the emerin LEM domain (PDB entry 2ODC) were used as templates. The resulting model was rebuilt using PHENIX [65], manual correction was performed with Coot [66] according to $|F_o| - |F_c|$ and $2|F_o| - |F_c|$ maps, and further refinement was carried out with phenix.refine. All structure figures were prepared using PyMOL (Schrödinger, LLC).

Fluorescence-based thermal shift assay.

The thermal stability of LamIgF proteins was monitored by a fluorescence-based thermal shift assay performed with a QuantStudio 12K Flex instrument (LifeTechnologies). 10 µg of purified protein was mixed with the SYPRO Orange dye (diluted 800-fold from a 5000-fold stock solution, Invitrogen) in 50mM Tris pH 8.0, 150mM NaCl, 10% glycerol. Reactions were carried out in duplicate in a 96-well fast PCR plate at a final volume of 20 µl and each experiment was repeated at least twice independently. The samples were submitted to a denaturation kinetic from 10 to 95°C at a rate of 3°C/min and fluorescence of Sypro Orange dye was recorded in real time. The protein denaturation temperature (Td +/- s.e.m) was calculated using the Protein Thermal Shift software v1.3 (LifeTechnologies) as the maximum of the derivative of the resulting fluorescence curves.

Proximity Ligation Assays.

Transfection - HeLa cells were obtained from American Type Culture Collection and cultured in Minimum Essential Medium containing Glutamax (Gibco), 1% non-essential amino acids and 10% fetal bovine serum. HeLa cells were transfected using XtremeGene 9 (Roche). After 24 h, cells were processed for either immunoblotting, immunofluorescence or Proximity ligation assay.

Immunoblotting - Whole cell protein extracts were suspended in Laemmli sample buffer, separated by SDS-PAGE and transferred to nitrocellulose membranes. Membranes were blocked for 1.5 h in TBST (10 mM Tris pH 8.0, 150 mM NaCl, 0.05% Tween-20) containing 5% dry milk, incubated with mouse anti FLAG antibody (Sigma; 1:600) and rabbit anti-BAF (Santacruz, 1:200) for 1 h in TBST/1% milk, washed 4 times and incubated with HRP-conjugated secondary antibodies. After 4 washes in TBST, proteins were detected by enhanced chemiluminescence.

Immunofluorescence microscopy - Cells were fixed with 3% paraformaldehyde for 12 min at R.T, permeabilized with phosphate buffered saline containing 0.5 % Triton for 5 min at R.T, and quenched with 2% bovine serum albumin diluted in phosphate buffered saline containing 0.1% Triton. Primary antibodies were rabbit anti-FLAG (Sigma, 1:300) and mouse anti-GFP (Roche, 1:200). Fluorescent labeled secondary antibodies (donkey anti-mouse Cy2 1:60 and donkey anti-rabbit 594 1:200) were from Jackson ImmunoResearch. DNA was stained with Hoechst 33258 (1 µg/ml).

Proximity ligation assay - PLA was used to detect interactions between GFP-BAF and FLAG-Lamin A, either WT or mutant (R435C, R527H, A529V, K542N), based on proximity (< 40 nm) of two secondary antibodies directed against these tagged proteins. After cell fixation, cell permeabilisation and quenching (as above), pairs of primary antibodies, rabbit anti-FLAG 1:300 and mouse anti-GFP 1:200, were added on HeLa cells expressing GFP-BAF together with FLAG-Lamin A, for 40 min at R.T. Next, Duolink PLA probe anti-rabbit plus, Duolink PLA probe anti-mouse minus and Duolink detection reagents orange (detected with a Cy3 filter) were used according to manufacturer's instructions (Olink, Bioscience). At last, cells were mounted in Duolink mounting medium with Dapi (Olink, Bioscience). Confocal microscopy image acquisition was performed using a LSM 700 Laser scanning microscope (Zeiss) at the imaging facility of the BFA institute. Quantitative analysis of PLA signals was done on images using Image J. Data were then analyzed by comparing median values for the total number of pixels showing a signal (Cy3) per nucleus. Data of the three independent experiments were normalized to 1 (100%) for cells expressing FLAG-lamin A WT. Statistical analysis were performed using Kruskall-Wallis tests.

Legends

Figure 1. The lamin A/C Igfold domain directly binds to emerin self-assembled nucleoplasmic region. (A) Architecture of the emerin protein, highlighting its LEM domain (in orange) and its transmembrane domain (in dark grey), as well as the 2 fragments of the nucleoplasmic region that could be obtained as soluble peptides (EmN and EmC). **(B)** Negative staining EM images of EmN and EmC fragments obtained after purification, concentration and incubation at 293K (white bar: 100 nm). EmN filaments were observed at 600 μM after 1 week [27], whereas EmC filaments were already visible at 300 μM after 1 day. The fragment corresponding to the region common to EmN and EmC did not self-assemble in these conditions (Suppl. Figure 1B). **(C)** Superimposition of the solution-state NMR ^{1}H-^{15}N HSQC spectra of the lamin Igfold domain (lamin A/C in grey or lamin B1 in green) recorded in the absence or presence of emerin fragments EmN or EmC (EmN in red, EmC in purple; 1:1 ratio). The emerin fragments were EmN filaments (upper left and lower right views), EmN monomers (lower left view) and EmC filaments (upper right view). All these spectra were recorded at 303K on a 700 MHz spectrometer.

Figure 2. The lamin A/C Igfold domain binds to emerin monomeric LEM domain through BAF. (A) Detection of a ternary complex including LamIgF, EmN and BAF by size exclusion chromatography using a Superdex 75 10/300GL column. Each gel corresponds to the elution of a different set of proteins, using the same column, the same protein concentrations and the same buffer: from the bottom EmN, LamIgF, BAF, Lamin B1 tail, BAF and lamin B1 tail, EmN and LamIgF, EmN and BAF, LamIgF and BAF and finally EmN, LamIgF and BAF. The bands corresponding to the different complexes are boxed: EmN and BAF in orange, LamIgF and BAF in green and the ternary complex in blue. **(B)** ITC titration of BAF onto LamIgF (upper panel) and BAF onto EmN (lower panel). 200 μM LamIgF or EmN were injected in a cell containing 40 μM of BAF at a temperature of 288K. The affinities deduced from the curves are 3.2 ± 1.2 μM and 0.7 ± 0.2 μM, respectively. The LamIgF / BAF interaction is enthalpy driven whereas the EmM/BAF interaction is both enthalpy and entropy driven (Suppl. Table 1).

Figure 3. Three-dimensional structure of the complex between LamIgF, emerin LEM domain and BAF. (A) Cartoon representation of the complex, with emerin LEM domain (residues 3 to 44) in orange, BAF dimer in yellow-green and green (residues 4 to 89), and lamin A/C Igfold domain (residues 432 to 544) in grey. The interfaces corresponding to the emerin / BAF and BAF / lamin interactions are indicated by red circles and numbered as 1 and 2, respectively. (**B)** Superimposition of the 3D structure of the BAF dimer bound to the emerin LEM domain, as determined in this work (same colors as in (A)), and as revealed using NMR by Clore and co-workers (PDB reference 2ODG [20]; in cyan). **(C)** Zoom on the EmN/BAF interface, with residues more than 30% buried within the interface displayed in sticks. On the emerin side, the interface is mainly formed by residues Gly24 to Lys37 (labeled residues on loop $\alpha1\alpha2$ and helix $\alpha2$). On the BAF side, it is mainly formed by helices $\alpha3$ and $\alpha4$ of one monomer and loop $\alpha2\alpha3$ and helix $\alpha3$ of the other monomer.

Figure 4. Lamin A/C Igfold domain binds BAF through a β-sheet including residues mutated in recessive progeroid syndromes. (A) Cartoon representation of the complex between the lamin A/C Igfold domain and the BAF dimer (main chain colors as in Fig. 3). On the lamin side, the interface is mainly polar: it is formed by residues His433, Ala434, Arg435, Glu537, Met540 and Lys542 (in red sticks). On the BAF side, it is formed by residues from helix $\alpha1$, loop $\alpha1\alpha2$ and helix $\alpha5$ from both monomers. At the BAF monomer-monomer interface, Gly47 is displayed in magenta. **(B)** ^{1}H-^{15}N HSQC spectrum of LamIgF in the absence (grey) and presence (green) of BAF (1:1 ratio). The labelled peaks exhibit a significant intensity decrease after addition of BAF. The peaks labelled in italics correspond to residues mutated in patients with progeroid diseases (R435[32,33,35], R471C [36], M540 [67,15], K542 [34]).

Figure 5. Impact of mutants causing recessive progeroid syndromes on BAF binding depends on the disease severity. (A) ITC titration of BAF onto LamIgF, as measured at 283K. Whereas variants R453W and R482W show WT binding, variants A529V and R471C exhibit a significantly lower binding enthalpy and affinity as compared to LamIgF WT. No affinity could be measured for R527H and K542N and no binding could be detected for R435C. In the lower right panel, the spatial distribution of residues mutated in (A) is displayed on the lamin A/C Igfold structure (PDB reference 1IFR [13]).

The 2 residues mutated in control variants, causing muscle and adipose tissue diseases respectively, are colored in green, whereas the 5 residues mutated in variants with a lower affinity for BAF are colored in cyan, purple and red (for weak binding, no measurable affinity and no detected binding, respectively). **(B)** Size exclusion chromatography of LamIgF mutants mixed to BAF (Superdex 75 10/300GL). Each gel corresponds to the elution of a different set of proteins, using the same column, the same concentrations and the same buffer: from bottom BAF, BAF and R435C, BAF and K542N, BAF and R527H, BAF and R471C, BAF and A529V, BAF and R482W, BAF and R453W and BAF and WT LamIgF. The bands corresponding to the BAF-LamIgF complex are boxed in red.

Figure 6. Mature lamin mutants with mutations causing recessive progeroid syndromes impair BAF binding in cells. (A) HeLa cells expressing GFP-BAF (WT) together with either WT, R435C, R527H, A529V or K542N FLAG-LAm were fixed, and processed for immunofluorescence using as primary antibodies, a mix of anti GFP and anti-FLAG antibodies. Observation was done at the confocal microscope. Are shown in the left panels, merges for GFP-BAF (green), FLAG-LAm (red) and DNA (blue), and in the right panels, the FLAG-LAm alone. Scale bar: 20 μM. **(B)** Whole cell protein extracts prepared from HeLa cells control (ctrl) or expressing GFP-BAF (WT) together with either WT, R435C, R527H, A529V or K542N FLAG-LAm were analyzed by western blot using anti FLAG and anti BAF antibodies. **(C)** HeLa cells, expressing GFP-BAF (WT) together with either WT, R435C, R527H, A529V or K542N FLAG-LAm were fixed, labeled with a mix of anti GFP and anti-FLAG antibodies and further processed for PLA before observation at the confocal microscope. The PLA signals (red) are shown merged with DNA (blue) and GFP-BAF (green) or alone (PLAssay). Signals for GFP-BAF and DNA are shown alone in the two right panels, respectively. Scale bar: 15 μM. **(D)** Quantification of PLA signals per nucleus as shown in C. The graph represents the mean values +/- stdevpa (n = 3 experiments; * $p<0.05$ with Kruskal and Wallis test).

Figure 7. A first model of the interface between the LEM-domain proteins, the nucleoskeleton and the chromatin-associated protein BAF. (Left) The lamin A/C Igfold is able to recognize emerin through BAF. Superimposition of the X-ray structure of the LamIgF-LEM-BAF complex onto the X-ray structure of the BAF protein complexed to DNA (PDB reference 2BZF [53]) suggests that BAF can simultaneously interact with lamin, emerin and DNA. The lamin-BAF interface is disrupted by mutations causing autosomal recessive progeroid syndromes, as symbolized by the red cross. **(Right)** In interphase cells, the lamin A/C Igfold domain recognizes self-assembled emerin at the nuclear envelope. It may also binds to monomeric emerin through BAF. The lamin A/C Igfold domain is able to recognize BAF and histones, thus creating an interface between the lamina and chromatin. This interface is disrupted by mutations causing autosomal recessive progeroid syndromes, as symbolized by the red cross.

Figure 1.

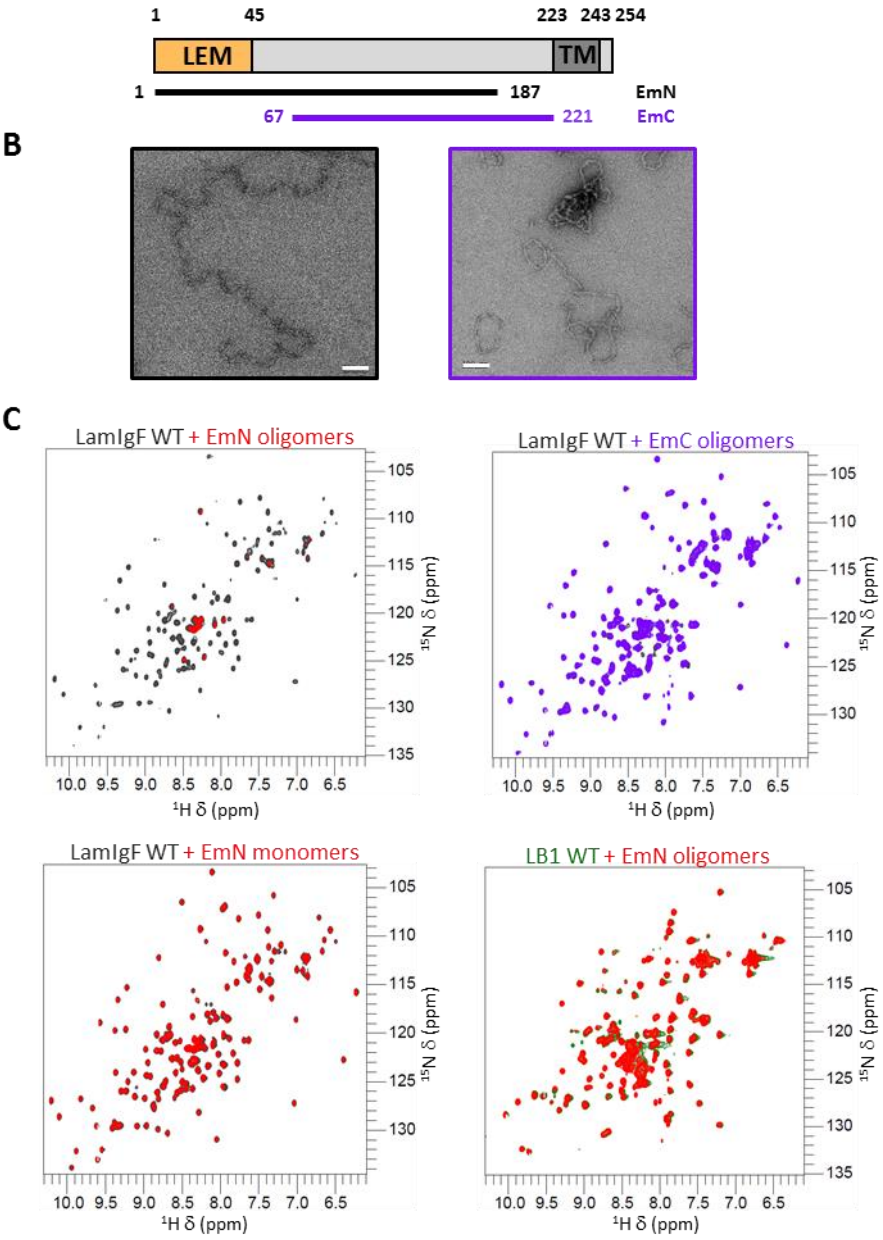

1
45
223
243 254
LEM
TM
1
187
EmN
67
221
EmC
B
C
LamIgF WT + EmN oligomers
LamIgF WT + EmC oligomers
LamIgF WT + EmN monomers
LB1 WT + EmN oligomers
105
110
115
120
125
130
135
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5
¹H δ (ppm)
¹⁵N δ (ppm)

Figure 2.

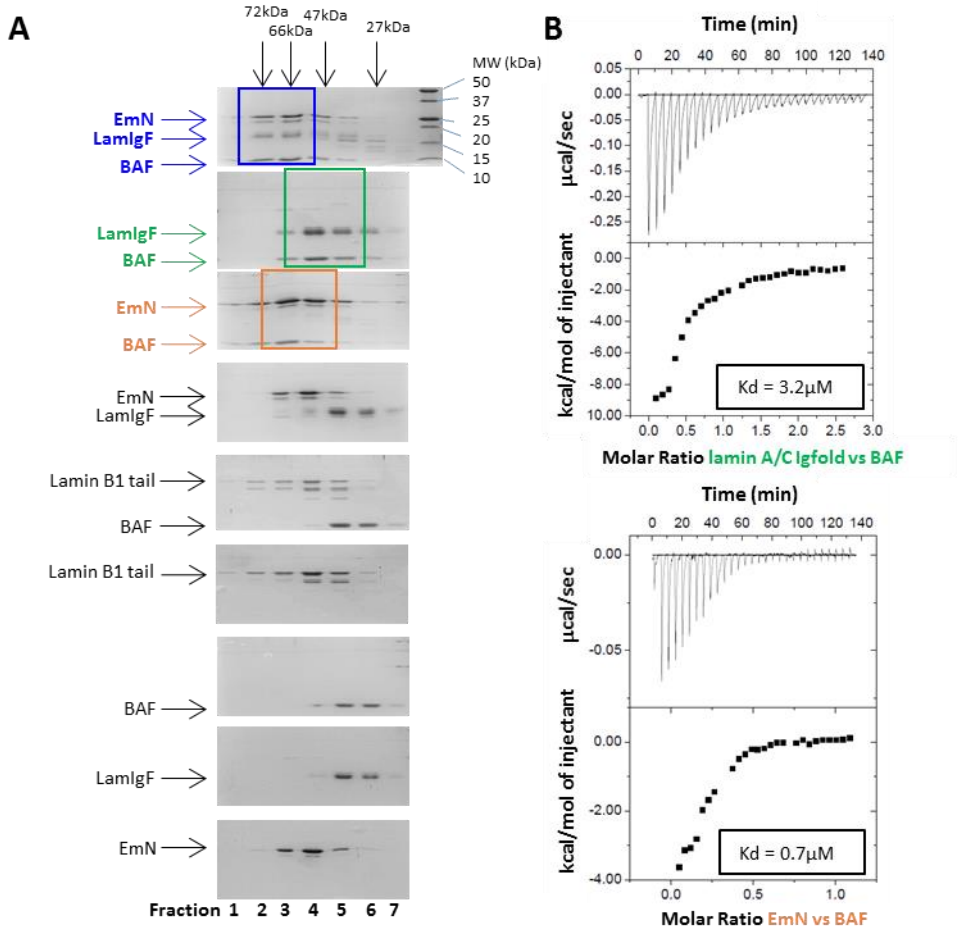

A
72kDa
66kDa
47kDa
27kDa
MW (kDa)
50
37
25
20
15
10
EmN
LamIgF
BAF
LamIgF
BAF
EmN
BAF
EmN
LamIgF
Lamin B1 tail
BAF
Lamin B1 tail
BAF
LamIgF
EmN
Fraction 1 2 3 4 5 6 7
B
Time (min)
0 20 40 60 80 100 120 140
µcal/sec
kcal/mol of injectant
Kd = 3.2µM
Molar Ratio lamin A/C Igfold vs BAF
Time (min)
0 20 40 60 80 100 120 140
µcal/sec
kcal/mol of injectant
Kd = 0.7µM
Molar Ratio EmN vs BAF

Figure 3.

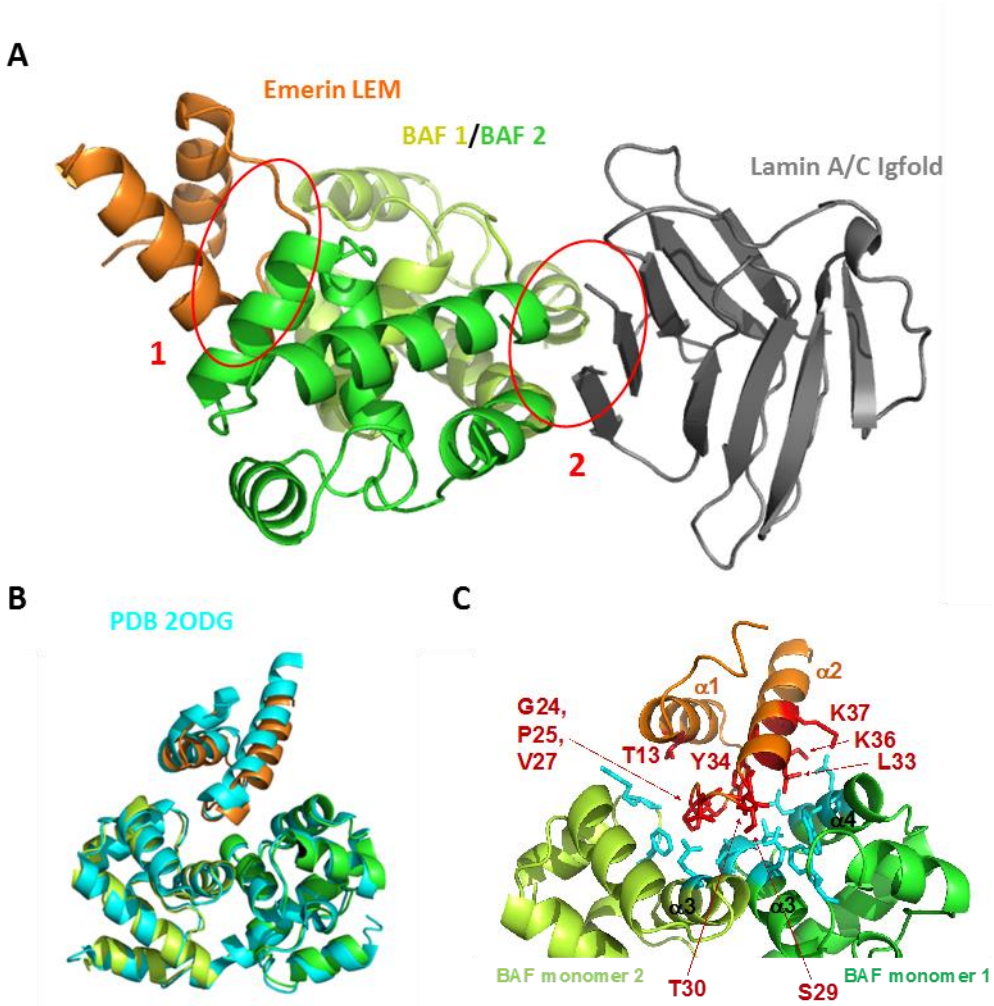
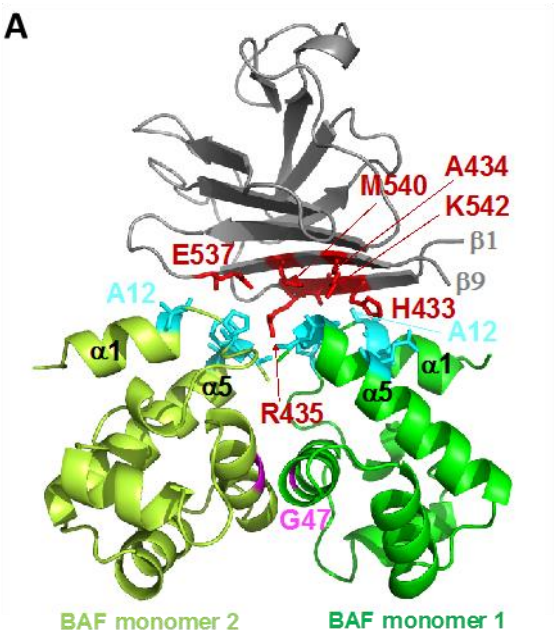

A
Emerin LEM
BAF 1/BAF 2
Lamin A/C Igfold
1
2
B
PDB 2ODG
C
α1
α2
G24,
P25,
V27
T13
Y34
K37
K36
L33
α4
α3
α3
BAF monomer 2
T30
S29
BAF monomer 1

Figure 4.


A
A434
M540
K542
β1
β9
E537
A12
H433
A12
α1
α1
α5
α5
R435
G47
BAF monomer 2
BAF monomer 1

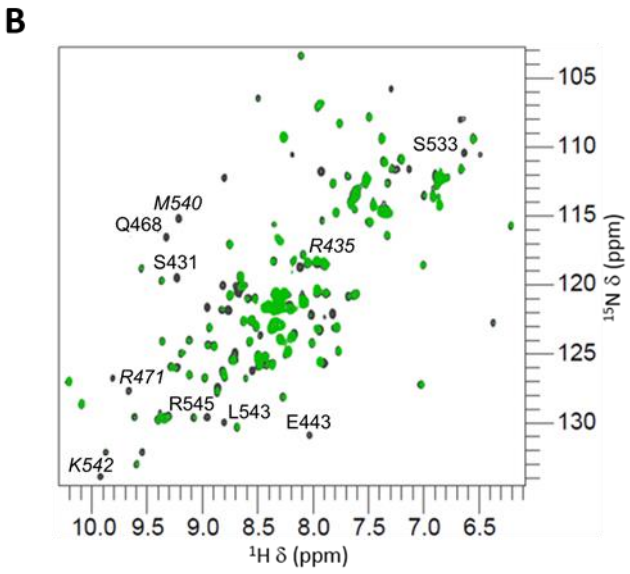

B
105
110
115
120
125
130
15N δ (ppm)
S533
M540
Q468
R435
S431
R471
R545
L543
E443
K542
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5
1H δ (ppm)

Figure 5.

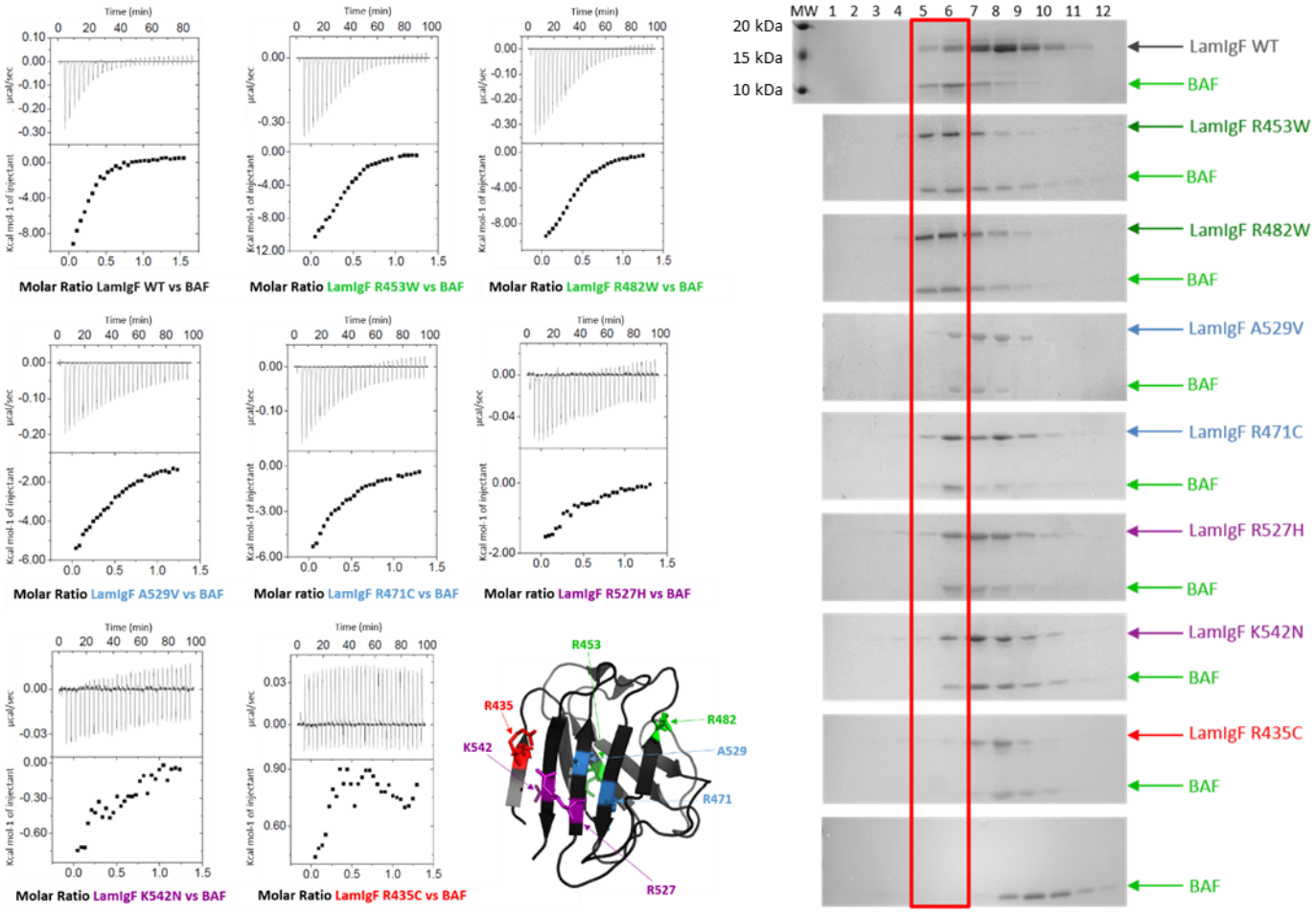

Time (min)
µcal/sec
Kcal mol-1 of injectant
Molar Ratio LamIgF WT vs BAF
Molar Ratio LamIgF R453W vs BAF
Molar Ratio LamIgF R482W vs BAF
Molar Ratio LamIgF A529V vs BAF
Molar ratio LamIgF R471C vs BAF
Molar ratio LamIgF R527H vs BAF
Molar Ratio LamIgF K542N vs BAF
Molar Ratio LamIgF R435C vs BAF
R453
R435
R482
K542
A529
R471
R527
MW 1 2 3 4 5 6 7 8 9 10 11 12
20 kDa
15 kDa
10 kDa
LamIgF WT
BAF
LamIgF R453W
BAF
LamIgF R482W
BAF
LamIgF A529V
BAF
LamIgF R471C
BAF
LamIgF R527H
BAF
LamIgF K542N
BAF
LamIgF R435C
BAF
BAF

Figure 6.

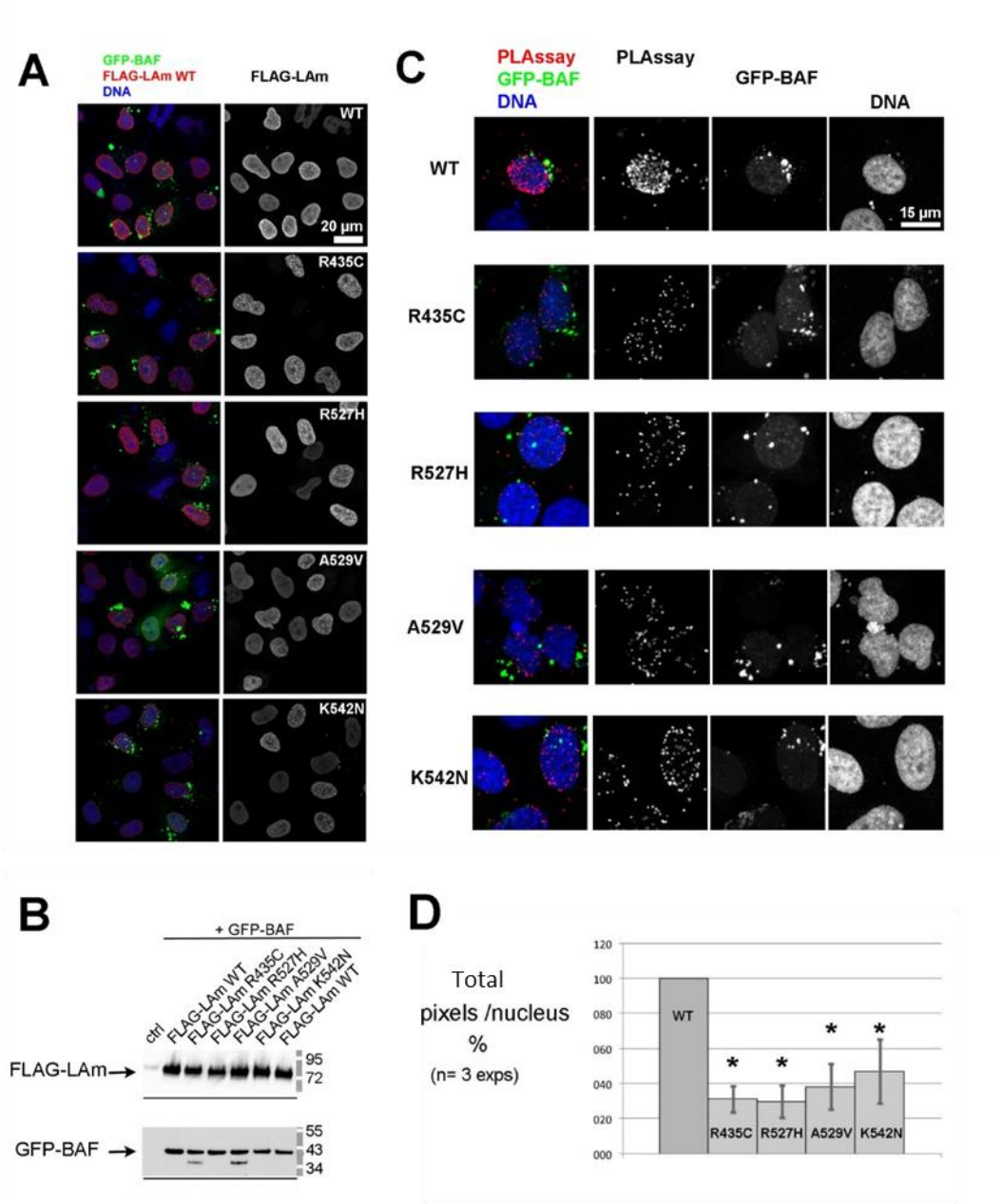
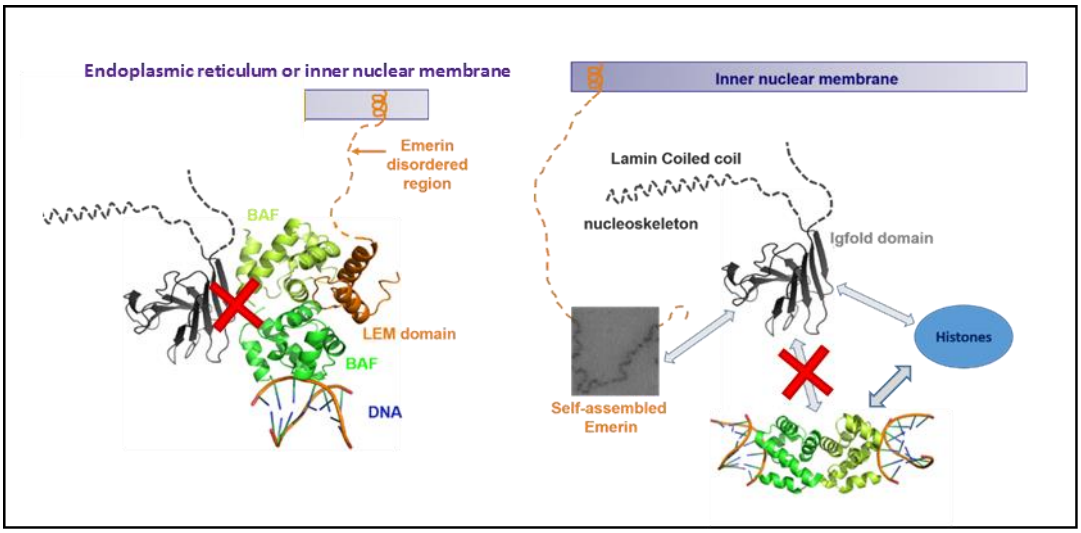

A
GFP-BAF
FLAG-LAm WT
DNA
FLAG-LAm
WT
20 µm
R435C
R527H
A529V
K542N
C
PLAssay
GFP-BAF
DNA
PLAssay
GFP-BAF
DNA
WT
15 µm
R435C
R527H
A529V
K542N
B
+ GFP-BAF
ctrl
FLAG-LAm WT
FLAG-LAm R435C
FLAG-LAm R527H
FLAG-LAm A529V
FLAG-LAm K542N
FLAG-LAm WT
FLAG-LAm
95
72
GFP-BAF
55
43
34
D
Total
pixels /nucleus
%
(n= 3 exps)
120
100
080
060
040
020
000
WT
R435C
R527H
A529V
K542N
*
*
*
*

Figure 7.


Endoplasmic reticulum or inner nuclear membrane
Emerin disordered region
BAF
LEM domain
BAF
DNA
Inner nuclear membrane
Lamin Coiled coil
nucleoskeleton
Igfold domain
Histones
Self-assembled Emerin

Supplementary information

Suppl. Fig. 1. Monomeric structure of fragments EmN49, EmC170, EmC187 and EmC, as observed by NMR. **(A)** Superimposition of the ^{1}H-^{15}N HSQC spectra of different emerin fragments is shown. This superimposition reveals that, whereas the EmN49 fragment, corresponding to the LEM domain (3D structure in the upper left corner; PDB reference 1ODG [68]), is well folded, the 3 other fragments are entirely unstructured. **(B)** ThT fluorescence as a function of the incubation time at 310 K, measured on 3 emerin fragments (EmN, Em67C and EmC) concentrated at 300µM (n=2; error bars correspond to standard deviations). Only EmN and EmC are able to self-assemble. Em67C does not form filaments.

Suppl. Fig. 2. Structure of the LamIgF variants, as observed by NMR and fluorescence. (A) ^{1}H-^{15}N HSQC spectra were recorded at 303K on different LamIgF variants (LamIgF WT in dark grey, LamIgF R453W and R482W in green, LamIgF A529V, R471C and R527H in blue and LamIgF K542N and R435C in red). All proteins were concentrated to 200µM, dialyzed in 20mM Tris-HCl pH 8.0, 30mM NaCl, 2mM DTT and NMR spectra were recorded on a 700MHz spectrometer. **(B)** Thermal stability of the different LamIgF variants was measured using a fluorescence thermal shift assay in 50mM Tris pH 8.0, 150mM NaCl, 10% glycerol. Error bars correspond to the standard deviation calculated from at least 4 measurements.

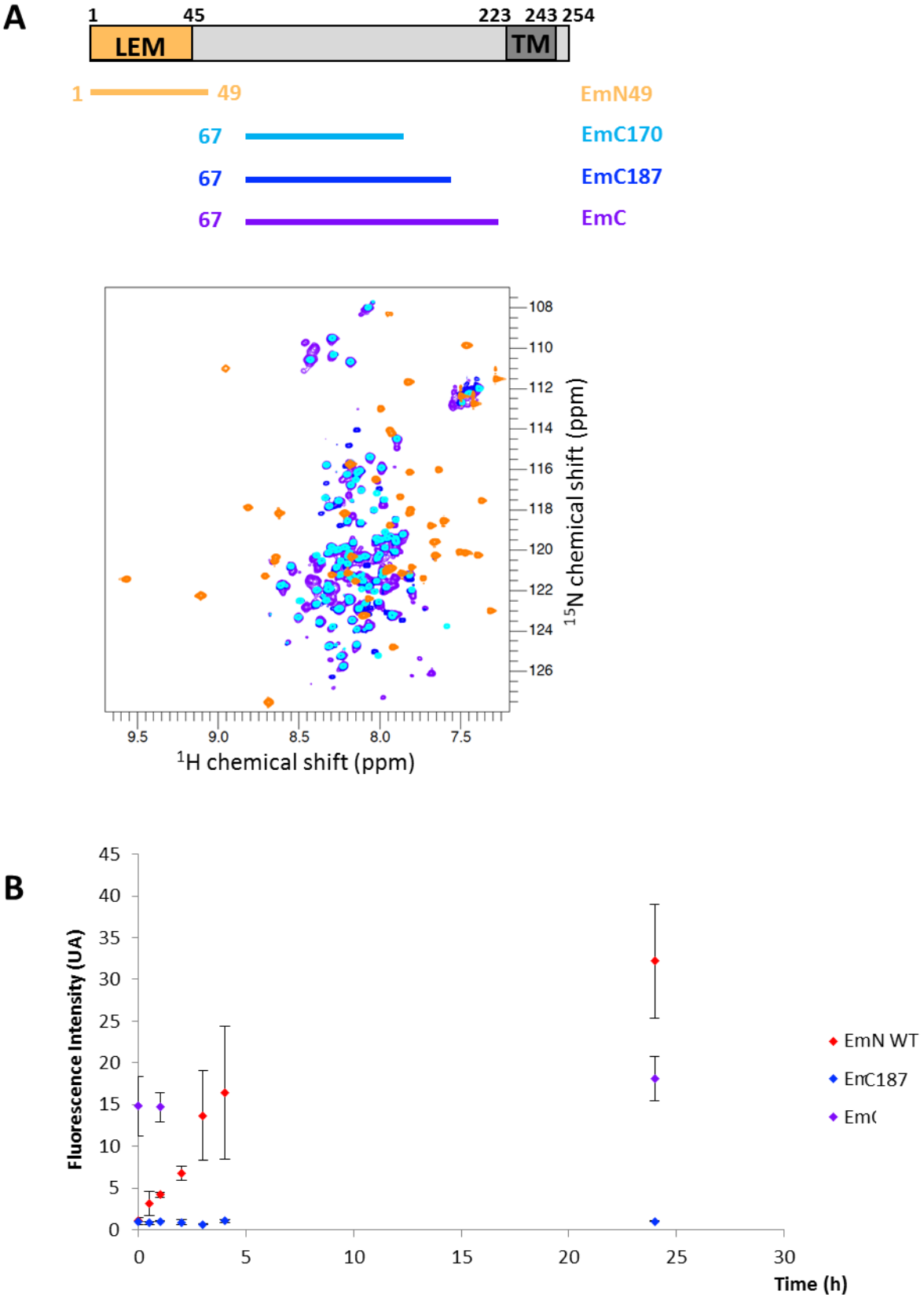
Suppl. Fig. 3. The lamin B1 tail domain does not bind to BAF, as observed by NMR. Addition of BAF onto the ^{15}N labeled lamin B1 tail (residues 395 to 586) does not affect the ^{1}H-^{15}N HSQC spectrum of the lamin B1 tail. The spectra were recorded at 283K, on a 600MHz, in phosphate pH 6.5, 20mM, 150mM NaCl, 1mM DTT.

Suppl. Fig. 4. Quantification of the PLA signals per nucleus (related to Figure 6C). Graphs represent the median values for the total signal expressed in number of pixels per nucleus for three independent experiments (Exp1, Exp2, Exp3; n= 105 to 230 nuclei per sample). Boxes show first and third quartiles (* $p<0.005$ with Kruskal and Wallis test).

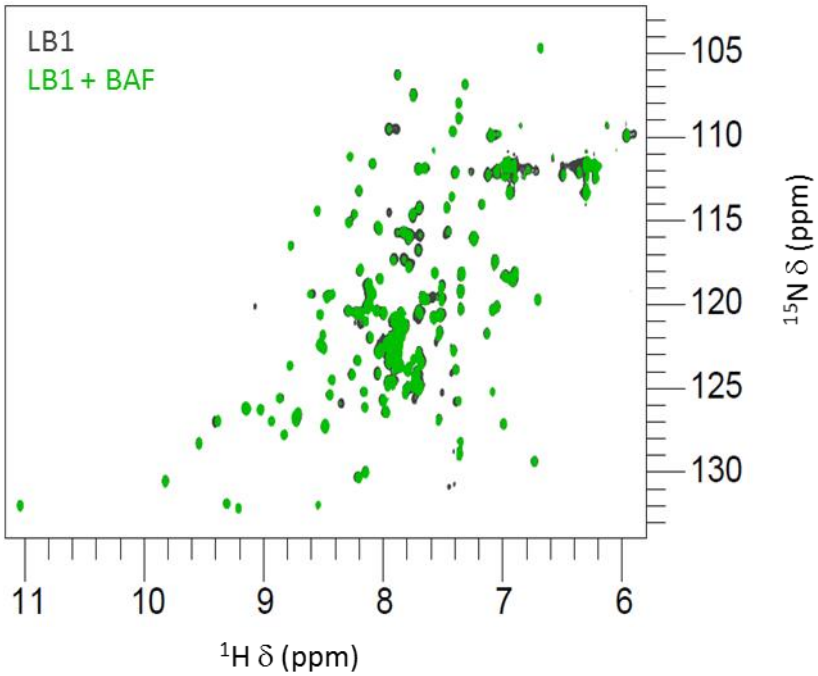
Suppl. Fig. 5. The WT LamIgF still binds to BAF A12T but with a lower affinity. Size exclusion chromatography experiments revealed an interaction between LamIgF and either BAF WT, BAF S4E or BAF A12T (Superdex 75 10/300GL). Each gel corresponds to the elution of a different set of proteins, using the same column, the same concentrations and the same buffer: from the bottom BAF A12T, LamIgF and BAF A12T, BAF S4E, LamIgF and BAF S4E, BAF WT, LamIgF, LamIgF and BAF WT. The bands corresponding to the ternary complex are boxed in red. Whereas BAF S4E binds as BAF WT to LamIgF, BAF A12T shows a weaker affinity for LamIgF.

Suppl. Fig. 6. Lamin A – BAF proximities are enriched at the nuclear periphery of interphase HeLa cells. HeLa cells transfected to express FLAG-LAm together with GFP-BAF were fixed, labeled with a mix of anti-GFP and anti-FLAG antibodies and further processed for PLA before observation at the confocal microscope. Are shown in the left panels the PLA signals (red) merged with GFP-BAF (green) and DNA (blue). Are shown in the right panels the PLA signals alone. Scale bar: 10 μM.

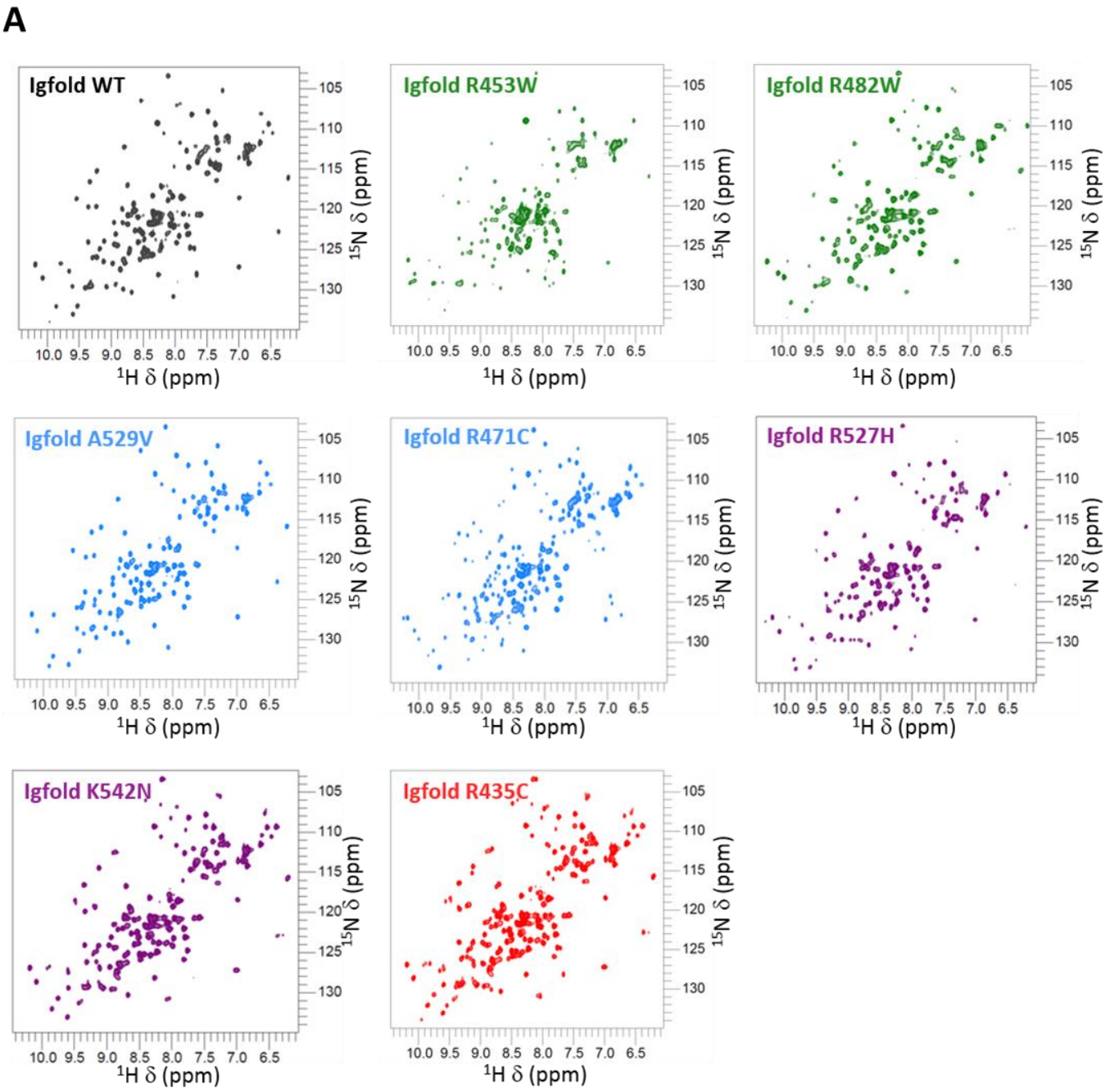
Suppl. Fig. 1.


A
1
45
223
243
254
LEM
TM
1
49
EmN49
67
EmC170
67
EmC187
67
EmC
108
110
112
114
116
118
120
122
124
126
^{15}N chemical shift (ppm)
9.5
9.0
8.5
8.0
7.5
^{1}H chemical shift (ppm)
B
Fluorescence Intensity (UA)
45
40
35
30
25
20
15
10
5
0
0
5
10
15
20
25
30
Time (h)
EmN WT
EmC187
EmC

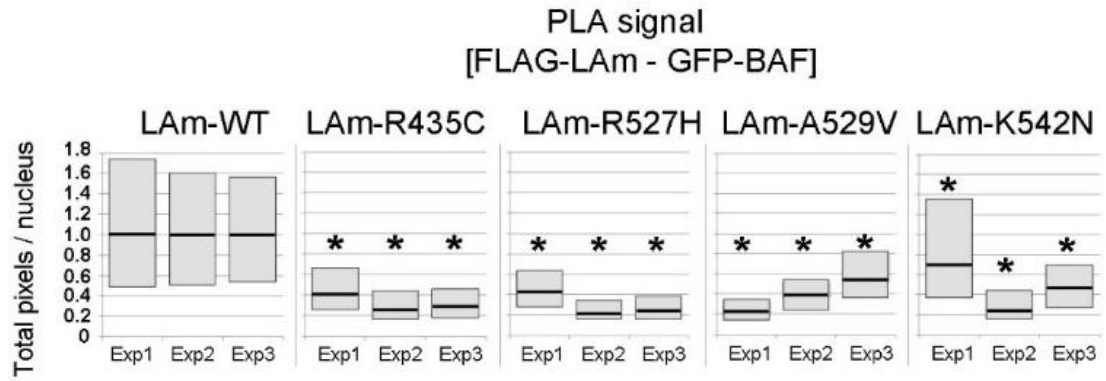
Suppl. Fig. 2.



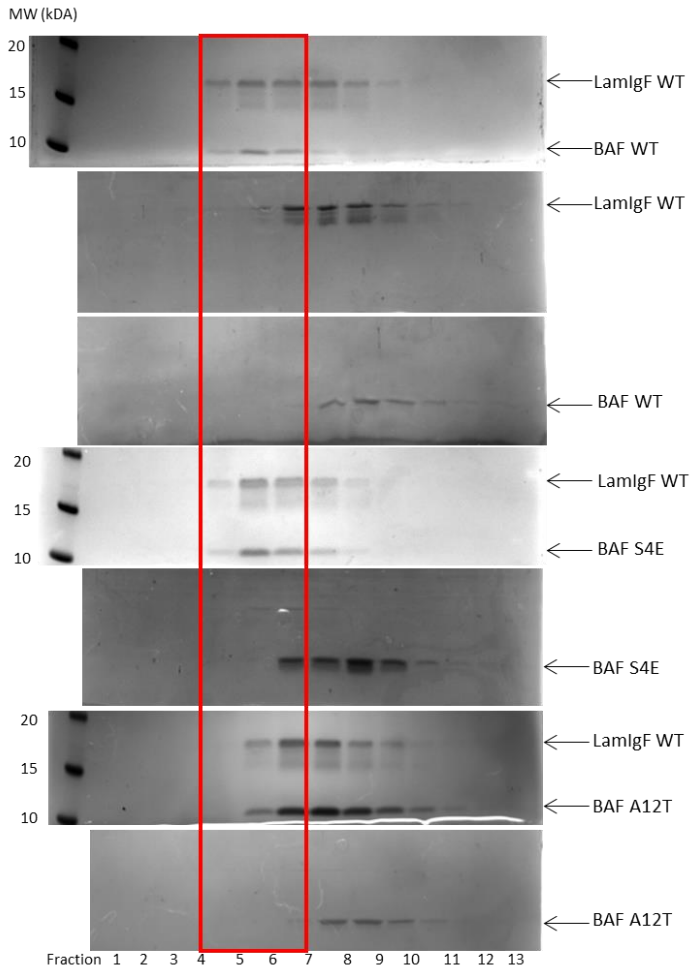
Suppl. Fig. 3.


A
Igfold WT
Igfold R453W
Igfold R482W
Igfold A529V
Igfold R471C
Igfold R527H
Igfold K542N
Igfold R435C
^{15}N δ (ppm)
^{1}H δ (ppm)
B
Denaturation Temperature (°C)
WT
R453W
R482W
A529V
R471C
R527H
K542N
R435C

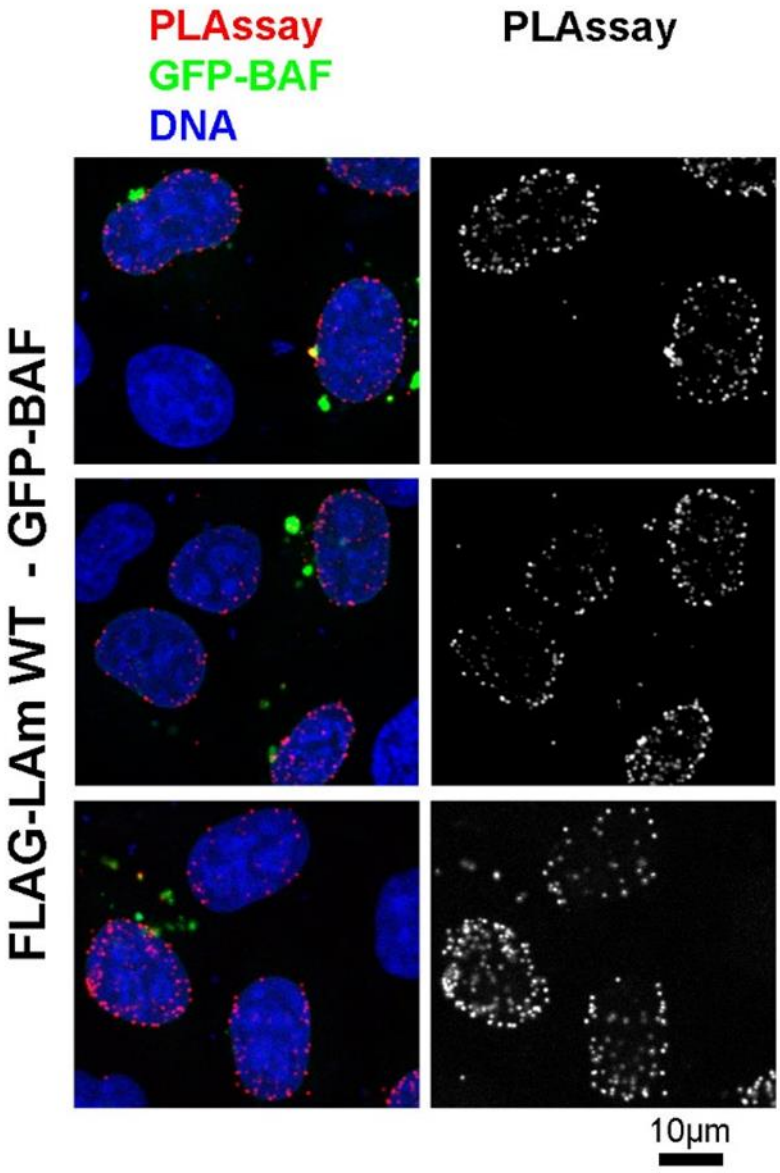
Suppl. Fig. 4.



Suppl. Fig. 5.



Suppl. Fig. 6.



<u>Suppl. Table 1.</u> Summary of Isothermal Titration Calorimetry results. (A) Table containing parameters deduced from all ITC experiments performed at 288K. **(B)** Table containing parameters deduced from all ITC experiments performed at 283K. N is the number of experiments, Ns the stoichiometry, K_a the association rate constant, K_d the dissociation rate constant, ΔH the measured enthalpy and ΔS the resulting entropy.

Experiments done at 15°C	BAF / Igfold		BAF / EmN	
	N = 5		N = 3	
	Mean	Standard deviation	Mean	Standard deviation
Ns	0.35	0.09	0.25	0.02
Ka (M^{-1})	3.7E5	2.1E5	1.4E6	2.9E5
Kd (µM)	3.15	1.16	0.7	0.17
ΔH (cal/mol)	-1.15E4	2.1E3	-4.1E3	9.1E2
ΔS (cal/mol/deg)	-13.5	9.3	-13.9	3.1

Experiments done at 10°C	Igfold WT		Igfold R453W		Igfold R482W		Igfold A529V		Igfold R471C	
	N = 2		N = 2		N = 2		N = 2		N = 2	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Ns	0.39	0.02	0.39	0.03	0.43	0.04	0.38	0.05	0.39	0.16
Ka (M^{-1})	7.2E5	2.6E5	6.5E5	5.3E4	5.6E5	2.6E4	1.5E5	1.4E4	1.4E5	1.6E4
Kd (µM)	1.5	0.54	1.5	0.12	1.7	0.08	6.9	1.6	7.2	0.8
ΔH (cal/mol)	-1.3E4	5.9E2	-1.3E4	6.4E2	-1E4	2E3	-8E3	1.6E3	-1.1E4	3E3
ΔS (cal/mol/deg)	-20.5	2.75	-18.8	2.12	-10.6	7	-4.8	5.9	-14	10

Suppl. Table 2. Data collection and refinement statistics.

Data collection	
Space group	P 1
Cell dimensions	
a, b, c (Å)	57.77 62.75 64.34
a, b, γ (°)	66.44 66.55 88.05
Resolution (Å)	49.152 – 2.1
Anisotropy direction	
Resolution where $CC_{1/2} > 0.3$	2.1
R_{meas} (%)	15.9
I/s(I)	5.9
$CC_{1/2}$	99.1
Completeness (%)	95.67
Redundancy	3.52
B Wilson (Å²)	34.74
Refinement	
No. reflections	83937
R_{work} / R_{free}	0.1879/0.2444
No. atoms	5666
Protein	675
DNA	0
SO_4^{2-}	8
Water	279
B factors	43.38
Protein	42.97
Ligand/ion	80.03
Water	45.37
R.m.s. deviations	1.953
Bond lengths (Å)	0.008
Bond angles (°)	1.16
One crystal per structure.	

References

1. Worman, H.J. Nuclear lamins and laminopathies. *J Pathol* **226**, 316-25 (2012).
2. Worman, H.J. & Bonne, G. "Laminopathies": a wide spectrum of human diseases. *Exp Cell Res* **313**, 2121-33 (2007).
3. Gruenbaum, Y. & Foisner, R. Lamins: nuclear intermediate filament proteins with fundamental functions in nuclear mechanics and genome regulation. *Annu Rev Biochem* **84**, 131-64 (2015).
4. Gruenbaum, Y. & Medalia, O. Lamins: the structure and protein complexes. *Curr Opin Cell Biol* **32**, 7-12 (2015).
5. Robin, J.D. & Magdinier, F. Physiological and Pathological Aging Affects Chromatin Dynamics, Structure and Function at the Nuclear Edge. *Front Genet* **7**, 153 (2016).
6. Loi, M. et al. Barrier-to-autointegration factor (BAF) involvement in prelamin A-related chromatin organization changes. *Oncotarget* **7**, 15662-77 (2016).
7. Burke, B. & Gerace, L. A cell free system to study reassembly of the nuclear envelope at the end of mitosis. *Cell* **44**, 639-52 (1986).
8. Glass, J.R. & Gerace, L. Lamins A and C bind and assemble at the surface of mitotic chromosomes. *J Cell Biol* **111**, 1047-57 (1990).
9. Taniura, H., Glass, C. & Gerace, L. A chromatin binding site in the tail domain of nuclear lamins that interacts with core histones. *J Cell Biol* **131**, 33-44 (1995).
10. Lopez-Soler, R.I., Moir, R.D., Spann, T.P., Stick, R. & Goldman, R.D. A role for nuclear lamins in nuclear envelope assembly. *J Cell Biol* **154**, 61-70 (2001).
11. Mahamid, J. et al. Visualizing the molecular sociology at the HeLa cell nuclear periphery. *Science* **351**, 969-72 (2016).
12. Turgay, Y. et al. The molecular architecture of lamins in somatic cells. *Nature* **543**, 261-264 (2017).
13. Dhe-Paganon, S., Werner, E.D., Chi, Y.I. & Shoelson, S.E. Structure of the globular tail of nuclear lamin. *J Biol Chem* **277**, 17381-4 (2002).
14. Krimm, I. et al. The Ig-like structure of the C-terminal domain of lamin A/C, mutated in muscular dystrophies, cardiomyopathy, and partial lipodystrophy. *Structure* **10**, 811-23 (2002).
15. Verstraeten, V.L. et al. Compound heterozygosity for mutations in LMNA causes a progeria syndrome without prelamin A accumulation. *Hum Mol Genet* **15**, 2509-22 (2006).
16. Lee, K.K. et al. Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF. *J Cell Sci* **114**, 4567-73 (2001).
17. Vaughan, A. et al. Both emerin and lamin C depend on lamin A for localization at the nuclear envelope. *J Cell Sci* **114**, 2577-90 (2001).
18. Sakaki, M. et al. Interaction between emerin and nuclear lamins. *J Biochem* **129**, 321-7 (2001).
19. Berk, J.M. et al. The molecular basis of emerin-emerin and emerin-BAF interactions. *J Cell Sci* **127**, 3956-69 (2014).
20. Cai, M. et al. Solution NMR structure of the barrier-to-autointegration factor-Emerin complex. *J Biol Chem* **282**, 14525-35 (2007).
21. Qi, R. et al. The lamin-A/C-LAP2alpha-BAF1 protein complex regulates mitotic spindle assembly and positioning. *J Cell Sci* **128**, 2830-41 (2015).
22. Liu, J. et al. MAN1 and emerin have overlapping function(s) essential for chromosome segregation and cell division in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **100**, 4598-603 (2003).
23. Margalit, A., Segura-Totten, M., Gruenbaum, Y. & Wilson, K.L. Barrier-to-autointegration factor is required to segregate and enclose chromosomes within the nuclear envelope and assemble the nuclear lamina. *Proc Natl Acad Sci U S A* **102**, 3290-5 (2005).
24. Holaska, J.M., Lee, K.K., Kowalski, A.K. & Wilson, K.L. Transcriptional repressor germ cell-less (GCL) and barrier to autointegration factor (BAF) compete for binding to emerin in vitro. *J Biol Chem* **278**, 6969-75 (2003).
25. Bengtsson, L. & Wilson, K.L. Barrier-to-autointegration factor phosphorylation on Ser-4 regulates emerin binding to lamin A in vitro and emerin localization in vivo. *Mol Biol Cell* **17**, 1154-63 (2006).
26. Samson, C. et al. Emerin self-assembly mechanism: role of the LEM domain. *FEBS J* **284**, 338-352 (2017).
27. Herrada, I. et al. Muscular Dystrophy Mutations Impair the Nuclear Envelope Emerin Self-assembly Properties. *ACS Chem Biol* **10**, 2733-42 (2015).
28. Samson, C., Herrada, I., Celli, F., Theillet, F.X. & Zinn-Justin, S. 1H, 13C and 15N backbone resonance assignment of the intrinsically disordered region of the nuclear envelope protein emerin. *Biomol NMR Assign* **10**, 179-82 (2016).
29. Margalit, A., Brachner, A., Gotzmann, J., Foisner, R. & Gruenbaum, Y. Barrier-to-autointegration factor--a BAFfling little protein. *Trends Cell Biol* **17**, 202-8 (2007).
30. Berk, J.M., Tifft, K.E. & Wilson, K.L. The nuclear envelope LEM-domain protein emerin. *Nucleus* **4**, 298-314 (2013).
31. Puente, X.S. et al. Exome sequencing and functional analysis identifies BANF1 mutation as the cause of a hereditary progeroid syndrome. *Am J Hum Genet* **88**, 650-6 (2011).
32. Madej-Pilarczyk, A. et al. Progeroid syndrome with scleroderma-like skin changes associated with homozygous R435C LMNA mutation. *Am J Med Genet A* **149A**, 2387-92 (2009).
33. Youn, G.J. et al. Autosomal recessive LMNA mutation causing restrictive dermopathy. *Clin Genet* **78**, 199-200 (2010).
34. Plasilova, M. et al. Homozygous missense mutation in the lamin A/C gene causes autosomal recessive Hutchinson-Gilford progeria syndrome. *J Med Genet* **41**, 609-14 (2004).
35. Starke, S. et al. Progeroid laminopathy with restrictive dermopathy-like features caused by an isodisomic LMNA mutation p.R435C. *Aging (Albany NY)* **5**, 445-59 (2013).
36. Zirn, B. et al. Association of homozygous LMNA mutation R471C with new phenotype: mandibuloacral dysplasia, progeria, and rigid spine muscular dystrophy. *Am J Med Genet A* **146A**, 1049-54 (2008).
37. Novelli, G. et al. Mandibuloacral dysplasia is caused by a mutation in LMNA-encoding lamin A/C. *Am J Hum Genet*

71, 426-31 (2002).
38. Simha, V., Agarwal, A.K., Oral, E.A., Fryns, J.P. & Garg, A. Genetic and phenotypic heterogeneity in patients with mandibuloacral dysplasia-associated lipodystrophy. *J Clin Endocrinol Metab* **88**, 2821-4 (2003).
39. Shen, J.J., Brown, C.A., Lupski, J.R. & Potocki, L. Mandibuloacral dysplasia caused by homozygosity for the R527H mutation in lamin A/C. *J Med Genet* **40**, 854-7 (2003).
40. Garavelli, L. et al. Mandibuloacral dysplasia type A in childhood. *Am J Med Genet A* **149A**, 2258-64 (2009).
41. Garg, A., Cogulu, O., Ozkinay, F., Onay, H. & Agarwal, A.K. A novel homozygous Ala529Val LMNA mutation in Turkish patients with mandibuloacral dysplasia. *J Clin Endocrinol Metab* **90**, 5259-64 (2005).
42. Ho, C.Y., Jaalouk, D.E., Vartiainen, M.K. & Lammerding, J. Lamin A/C and emerin regulate MKL1-SRF activity by modulating actin dynamics. *Nature* **497**, 507-11 (2013).
43. Cho, S., Irianto, J. & Discher, D.E. Mechanosensing by the nucleus: From pathways to scaling relationships. *J Cell Biol* **216**, 305-315 (2017).
44. Osmanagic-Myers, S., Dechat, T. & Foisner, R. Lamins at the crossroads of mechanosignaling. *Genes Dev* **29**, 225-37 (2015).
45. Strelkov, S.V. et al. Divide-and-conquer crystallographic approach towards an atomic structure of intermediate filaments. *J Mol Biol* **306**, 773-81 (2001).
46. Herrmann, H., Strelkov, S.V., Burkhard, P. & Aebi, U. Intermediate filaments: primary determinants of cell architecture and plasticity. *J Clin Invest* **119**, 1772-83 (2009).
47. Dorner, D., Gotzmann, J. & Foisner, R. Nucleoplasmic lamins and their interaction partners, LAP2alpha, Rb, and BAF, in transcriptional regulation. *FEBS J* **274**, 1362-73 (2007).
48. Simon, D.N. & Wilson, K.L. Partners and post-translational modifications of nuclear lamins. *Chromosoma* **122**, 13-31 (2013).
49. King, M.C. & Lusk, C.P. A model for coordinating nuclear mechanics and membrane remodeling to support nuclear integrity. *Curr Opin Cell Biol* **41**, 9-17 (2016).
50. Camozzi, D. et al. Diverse lamin-dependent mechanisms interact to control chromatin dynamics. Focus on laminopathies. *Nucleus* **5**, 427-40 (2014).
51. Paquet, N. et al. Nestor-Guillermo Progeria Syndrome: a biochemical insight into Barrier-to-Autointegration Factor 1, alanine 12 threonine mutation. *BMC Mol Biol* **15**, 27 (2014).
52. Haraguchi, T. et al. Live cell imaging and electron microscopy reveal dynamic processes of BAF-directed nuclear envelope assembly. *J Cell Sci* **121**, 2540-54 (2008).
53. Bradley, C.M., Ronning, D.R., Ghirlando, R., Craigie, R. & Dyda, F. Structural basis for DNA bridging by barrier-to-autointegration factor. *Nat Struct Mol Biol* **12**, 935-6 (2005).
54. Haraguchi, T. et al. BAF is required for emerin assembly into the reforming nuclear envelope. *J Cell Sci* **114**, 4575-85 (2001).
55. Barton, L.J., Soshnev, A.A. & Geyer, P.K. Networking in the nucleus: a spotlight on LEM-domain proteins. *Curr Opin Cell Biol* **34**, 1-8 (2015).
56. Shimi, T. et al. Dynamic interaction between BAF and emerin revealed by FRAP, FLIP, and FRET analyses in living HeLa cells. *J Struct Biol* **147**, 31-41 (2004).
57. Columbaro, M. et al. Rescue of heterochromatin organization in Hutchinson-Gilford progeria by drug treatment. *Cell Mol Life Sci* **62**, 2669-78 (2005).
58. Shumaker, D.K. et al. Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proc Natl Acad Sci U S A* **103**, 8703-8 (2006).
59. Scaffidi, P. & Misteli, T. Reversal of the cellular phenotype in the premature aging disease Hutchinson-Gilford progeria syndrome. *Nat Med* **11**, 440-5 (2005).
60. Montes de Oca, R., Lee, K.K. & Wilson, K.L. Binding of barrier to autointegration factor (BAF) to histone H3 and selected linker histones including H1.1. *J Biol Chem* **280**, 42252-62 (2005).
61. Montes de Oca, R., Andreassen, P.R. & Wilson, K.L. Barrier-to-Autointegration Factor influences specific histone modifications. *Nucleus* **2**, 580-90 (2011).
62. Bruston, F. et al. Loss of a DNA binding site within the tail of prelamin A contributes to altered heterochromatin anchorage by progerin. *FEBS Lett* **584**, 2999-3004 (2010).
63. Vranken, W.F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-96 (2005).
64. Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-42 (2011).
65. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).
66. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
67. Bai, S. et al. Mandibuloacral Dysplasia Caused by LMNA Mutations and Uniparental Disomy. *Case Rep Genet* **2014**, 508231 (2014).
68. Wolff, N. et al. Structural analysis of emerin, an inner nuclear membrane protein mutated in X-linked Emery-Dreifuss muscular dystrophy. *FEBS Lett* **501**, 171-6 (2001).

FRENCH SYNTHESIS

Ma thèse portait sur la caractérisation structurale et fonctionnelle de l'enveloppe nucléaire. Cette enveloppe est formée d'une bicouche lipidique (membranes nucléaire interne et externe) et de différentes protéines associées à ces membranes. Elle est impliquée dans:

- la forme et la position du noyau
- la structure et la mobilité cellulaire car elle est reliée au cytosquelette via le complexe LINC
- l'organisation du génome, en interagissant avec la chromatine
- la régulation de voies de signalisation en interagissant avec des facteurs de transcriptions et différentes enzymes.

Des mutations au niveau des gènes codant pour les protéines de l'enveloppe nucléaire interne sont à l'origine d'un grand nombre de maladies humaines, appelées les envelopathies, incluant des dystrophies musculaires, des lipodystrophies et des syndromes de vieillissement prématuré. Toutefois, les données structurales sont rares pour ces protéines, entravant ainsi la description de leurs mécanismes fonctionnels normaux et pathogènes. Dans mon projet de thèse, nous avons proposé de caractériser structuralement le complexe émerine-lamine et d'observer les conséquences de 6 mutations pathogènes de l'émerine, à l'origine de la dystrophie d'Emery-Dreifuss. qui n'empèchent pas l'expression et la bonne localisation de la protéine (delK37, S54F, Q133H, del95-99, P183T, P183H). L'une des nouveautés de ce projet était de réussir à stabiliser des oligomères de l'émerine et de la lamine, afin de reconstituer le complexe fonctionnel impliqué dans l'architecture de l'enveloppe nucléaire.

Lorsque j'ai débuté ma thèse, nous avions des preuves préliminaires que l'émerine dimérise in vitro, forme des oligomères in vitro et dans les cellules, et que cette oligomérisation a une influence sur la liaison de cette protéine à la lamine. Des publications récentes avaient également montré que l'émerine est phosphorylée par Src après l'application d'une force mécanique sur le noyau et que le complexe émerine-lamine est au centre de la réponse nucléaire face à une force mécanique (Ho et al, Nature 2013; Guilluy et al., Nat Cell Biol 2014).

Au cours de ma première année, j'ai caractérisé les oligomères formés par le fragment de l'émerine comprenant les résidus 1 à 187 (protocole de production déjà établi au laboratoire) à l'aide de différentes techniques telles que la microscopie électronique, la fluorescence et la spectroscopie infrarouge. J'ai observé que l'émerine 1-187 est capable de former des filaments in vitro et que ces filaments sont formés de feuillets β. J'ai également mis au point un protocole robuste de production de ces filaments. J'ai caractérisé l'état oligomérique de 6 mutants de l'émerine à l'origine de la dystrophie musculaire d'Emery-Dreifuss (S54F, del95-99, Q133H, P183T et P183H). Pendant une deuxième partie de ma première année de thèse, j'ai également commencé une étude des conséquences structurales de la phosphorylation de l'émerine par la kinase Src (voir introduction) et en particulier l'analyse de l'impact de la phosphorylation sur la formation des filaments. Pour mener à bien ce projet, notre équipe s'est rendue pendant 9 mois (d'Avril à Décembre 2015) dans le laboratoire du Dr Phillip Selenko, au FMP de Berlin-Buch.

Je me suis ensuite intéressée au mécanisme d'assemblage des filaments in vitro et j'ai déterminé quelles parties de l'émerine étaient indispensables à cet assemblage. Pour ce faire, la première approche fut de produire différentes parties de l'émerine et de tester leur capacité à former des filaments. J'ai d'abord produit les fragments 67-170 puis 67-187, afin de savoir si le domaine LEM était indispensable à la formation des filaments et j'ai pu constater que ces deux fragments ne formaient pas ce type d'oligomères. J'ai ensuite voulu savoir si les régions 1-146 puis 1-132 étaient suffisantes pour former des filaments. J'ai choisi ces régions car elles ne comportaient pas l'unique cystéine présente dans l'émerine (en position 147), sachant que nous avions prouvé auparavant que les filaments étaient formés à partir du monomère et non du dimère d'émerine, et car elles comportaient la région 95-99 qui semblait être importante pour la formation des filaments. J'ai pu observer la formation de nos oligomères dans les deux cas. J'ai également produit la région du domaine LEM 1-50 seule et observé l'absence de filaments. Donc la région minimale suffisante pour former des filaments est 1-132. Enfin, j'ai étudié la cinétique de formation d'un mutant du domaine LEM de l'émerine, le mutant delK37, et observé que ce dernier possède un domaine LEM destructuré et est capable de former des filaments plus rapidement que l'émerine 1-187 sauvage.

De manière consistente, j'ai fait la protéolyse ménagée des filaments d'émerine 1-187 sauvage et delK37, et par spectrométrie de masse (collaboration avec l'équipe du Dr Jean Armengaud, CEA Marcoule), j'ai observé que les régions restantes après coupure par la chymotrypsine, la trypsine et l'endoprotéinase GluC correspondaient en majeur partie au domaine LEM, mais également à la région autour des résidus 95-99.

Enfin nous avons étudié nos filaments à l'aide de la technique de la RMN du solide, à travers une collaboration avec l'équipe du Prof Adam Lange, au FMP de Berlin-Buch. Nous avons avons obtenu des premiers spectres des filaments de l'émerine 1-187 et des filaments de l'émerine 1-132. Ces spectres montrent que la région structurée est la même dans les deux types de filaments et comprend une vingtaine de résidus. Puis, nous avons également obtenu des spectres du mutant 1-187 delK37 de l'émerine et observé que les filaments formés par ce mutant étaient identiques à ceux de l'émerine 1-187 WT. Cette analyse par RMN du solide ne nous a pas permis d'identifier la structure 3D du cœur des filaments à cause de l'hétérogénéité conformationelle présente dans ces filaments, mais nous a permis de conclure, avec les données obtenues en spectrométrie de masse, que l'interaction entre le domaine LEM et la région autour des résidus 95-99 de l'émerine est à l'origine de la formation des filaments.

La dernière grande partie de mon projet de thèse concernait l'analyse structurale des interactions entre l'émerine et ses partenaires. Dans un premier temps, nous nous sommes intéressés à l'interaction entre l'émerine et la lamine A, notamment son domaine Igfold (domaine globulaire de la lamine). Lorsque que je suis arrivée au laboratoire, plusieurs expériences avaient déjà été menées sur ce projet, mais l'interaction entre ces deux protéines, connue dans la littérature, n'avait été observée qu'une fois par RMN et n'avait pas été reproduite. Grâce aux différentes études menées pendant ces deux années, j'ai amélioré notre connaissance des déterminants de l'état oligomérique de l'émerine et j'ai pu montrer l'interaction par RMN, ITC et microscopie électronique, entre le monomère du domaine Igfold de la lamine A et les filaments d'émerine 1-187 sauvage et delK37.

Nous avons étudié d'autres interactions lors de ma deuxième année de thèse. Tout d'abord l'interaction entre l'émerine et la queue de la myosine 1B (M1BO1). Cette étude a été possible grâce à une collaboration avec le groupe du Dr Anne Houdusse, de l'institut Curie. Nous avons d'abord caractérisé l'interaction par résonance magnétique nucléaire (RMN), puis nous avons fait des études préliminaires par titration calorimétrique isotherme (ITC) et par thermophorèse à micro-échelle (MST). Après avoir observé que l'interaction de l'émerine 1-187 WT avec M1BO1 semblait induire la formation des filaments d'émerine, nous avons fait des premières études par fluorescence, soit des cinétiques de formation de filaments à l'aide de la thioflavine et avons eu des premières confirmations de notre hypothèse.

Nous avons aussi étudié le complexe LEM/BAF/IgFold. En effet, la liaison entre le domaine globulaire LEM de l'émerine et la protéine BAF avait déjà été bien caractérisée (Cai M et al., 2007) et l'interaction entre l'IgFold de la lamine et la protéine BAF avait été plusieurs fois suggérée. Nous avons donc caractérisé et certifié l'interaction entre la protéine BAF et l'Igfold par RMN, ITC et Gel Filtration. Puis nous avons caractérisé l'interaction entre ces deux protéines et le domaine LEM de l'émerine, démontrant ainsi que les trois protéines peuvent former un complexe ternaire. Enfin nous avons cristallisé le complexe ternaire, obtenu des cristaux et une première structure de ce complexe, résolue à 2.8 Å. Finalement, nous avons étudié l'impact sur l'affinité entre l'Igfold et BAF de mutations causant des pathologies et impliquant des résidus se trouvant à l'interface IgFold/BAF.

Titre : Représentation en trois dimensions de l'interface entre l'enveloppe nucléaire et la chromatine

Mots clés : Interactions protéine-protéine, enveloppe nucléaire, Emery-Dreifuss Muscular Dystrophy

Résumé: Le noyau est un organite caractéristique des cellules eucaryotes et les propriétés mécaniques de ce dernier jouent un rôle essentiel dans le comportement de la cellule, notamment sa motilité, sa polarité et sa survie. Le noyau est entouré par une enveloppe comprenant une membrane interne et une membrane externe, ainsi que de nombreuses protéines. Mes objectifs de thèse étaient de comprendre des mécanismes moléculaires déficients dans deux types de maladies génétiques causées par des mutations dans les lamines: la dystrophie musculaire d'Emery-Dreifuss et les syndromes de type progéroïde.

Dans un premier temps, nous avons montré que l'émerine s'auto-associe *in vitro* et en cellules (Herrada et al. *ACS Chem. Biol.* 2015). J'ai ensuite étudié la structure des oligomères d'émerine, déterminé le fragment protéique minimal nécessaire à la formation de ces oligomères et décrit l'impact de mutations de l'émerine, causant une dystrophie musculaire d'Emery-Dreifuss, sur son auto-assemblage (Samson et al. *Biomol NMR Assign.* 2016 ; Samson et al. *FEBS J.* 2016).

Puis, j'ai montré que seule cette forme auto-assemblée de l'émerine est capable d'interagir avec la lamine A et que la phosphorylation de l'émerine par la kinase Src, observée suite à un stress mécanique, régule cette interaction entre l'enveloppe nucléaire et le nucléosquelette.

Pour finir, j'ai montré que la forme monomérique de l'émerine est capable de former un complexe ternaire avec BAF et la lamine A. Après avoir mesuré les affinités protéine-protéine au sein de ce complexe, identifié les fragments minimaux des différentes protéines permettant de former ce complexe et mis au point un protocole robuste de purification de ce complexe, j'ai pu obtenir des cristaux de ce complexe dans plusieurs conditions. Par la suite, nous avons pu résoudre la structure de ce complexe par remplacement moléculaire avec une résolution de 2 Å. Enfin, j'ai montré que les mutations dans les lamines de type A provoquant des syndromes de type progéroïde pouvaient altérer l'interaction avec BAF in vitro, et nos collaborateurs, l'équipe du Dr B. Buendia (Paris Diderot), ont montré que ces mêmes mutations induisaient une diminution significative de la proximité entre la lamine A et BAF dans les cellules HeLa.

Title : A three-dimensional view of the interface between nuclear envelope and chromatin

Keywords : Protein-protein interactions, nuclear envelope, Emery-Dreifuss Muscular Dystrophy

Abstract: The nucleus is an organelle characteristic of eukaryotic cells and its mechanical properties play an essential role in the behavior of the cell, in particular its motility, polarity and survival. It is surrounded by an envelope comprising an inner membrane and an outer membrane, as well as a large number of proteins. These proteins are either anchored at the nuclear membrane, as emerin, or form a filament meshwork lining the inner nuclear membrane, as lamins. My thesis objectives were to understand molecular mechanisms deficient in two types of genetic diseases caused by mutations in inner nuclear envelope proteins: Emery-Dreifuss muscular dystrophy, associated to mutations in emerin and A-type lamins, and progeroid syndromes caused by mutations in A-type lamins.

First, we showed that the emerin protein self-assembles *in vitro* and in cells (Herrada, Samson et al., *ACS Chem. Biol.*, 2015). I then studied the structure of emerin oligomers, determined the minimal protein fragment necessary for the formation of these oligomers, identify residues forming the structural core of these oligomers by solid-state NMR in collaboration with the group of Prof A. Lange (FMP Berlin).

And described the impact of emerin mutations causing Emery-Dreifuss muscular dystrophy on emerin self-assembly (Samson et al., *Biomol. NMR Assign*. 2016, Samson et al., *FEBS J.* 2017). Then, I observed, mainly using solution-state NMR, that only the self-assembled form of emerin is able to interact with A-type lamin tail, and that mutants causing Emery-Dreifuss muscular dystrophy and unable to self-assemble are also defective in A-type lamin binding. I also obtained preliminary data showing that phosphorylation of emerin by the Src kinase, observed after a mechanical stress in purified nuclei, regulates the interaction between self-assembled emerin and A-type lamins. Finally, I showed that the monomeric form of emerin is able to form a ternary complex with A-type lamin tail through the chromatin-associated protein Barrier-to-Autointegration Factor (BAF). After having measured the protein-protein affinities within this complex, identified the minimal protein fragments involved in the complex and developed a robust protocol for purification of this complex, I was able to obtain crystals under several conditions. Subsequently, I solved the 3D structure of this complex by molecular replacement at a resolution of 2 Å. Finally, I showed that mutations in A-type lamins causing autosomal recessive progeroid syndromes impair interaction with BAF *in vitro*, and our collaborators at Univ. Paris Diderot, the team of Dr B. Buendia, showed that these same mutations induce a significant decrease in the proximity between lamin A and BAF in HeLa cells.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190
Saint-Aubin, France