We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

5,800
Open access books available

142,000
International authors and editors

180M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

NMR Spectroscopy as a Tool to Provide Mechanistic Clues About Protein Function and Disease Pathogenesis

Benjamin Bourgeois[1], Howard J. Worman[2] and Sophie Zinn-Justin[1]
[1]Laboratoire de Biologie Structurale et Radiobiologie,
CEA Saclay & URA CNRS 2096, Gif-sur-Yvette
[2]Department of Medicine and Department of Pathology and Cell Biology,
College of Physicians and Surgeons, Columbia University, New-York
[1]France
[2]USA

1. Introduction

Nuclear magnetic resonance (NMR) has become an important technique for determining the three-dimensional (3D) structure of biological macromolecules. Since 1985, it has been used to determine the structures of approximately 8,000 proteins, 1,000 DNA/RNA complexes and 180 protein/nucleic acid complexes (http://www.pdb.org/pdb/ statistics/holdings.do). At first, NMR was limited to relatively small and soluble proteins or protein domains. The study of large proteins was hindered by the presence of overlapping large peaks in the NMR spectra. This has been in part alleviated by the introduction of isotope (^{2}H, ^{15}N, ^{13}C) labeling and multidimensional (3D, 4D) experiments (Sattler et al., 1999). By using these techniques, it is now possible to study proteins up to 40 kDa. For instance, the 3D solution structure of the maltodextrin-binding protein (41kDa) has been recently solved using NMR (Madl et al., 2009; Figure 1A). Comprehensive NMR studies of integral membrane proteins in solution have long been impaired by substantial problems of sample preparation, including the inability to produce sufficient quantities of isotopically labelled protein as well as the difficulties associated with the limited thermal stability, sample heterogeneity and short lifetimes of such proteins. However, NMR allows investigation of the very conformational mobility that to a large extent interferes with the process of crystallization of membrane proteins. Thus, by focusing on proteins with a sufficient expression yield and screening sample and detergent conditions in a microtiter-plate format, it has been possible to determine 71 structures of integral membrane proteins corresponding to 49 unique proteins (http://www.drorlist.com/ nmr/MPNMR.html). The first NMR structure determination of a detergent-solubilized seven-helix transmembrane (7TM) protein, the phototaxis receptor sensory rhodopsin II (pSRII) from *Natronomonas pharaonis*, was recently reported to illustrate that NMR can provide structures of large membrane proteins (Gautier et al., 2010; Figure 1B). The challenge is now to apply similar techniques to the study of other 7TM proteins, including G-protein coupled receptors (GPCRs) that represent the most important class of targets for current therapeutic agents. These proteins can only be

expressed in small quantities and are of limited stability, particularly in their unbound state. Samples must be conformationally and chemically homogeneous and stable at high concentrations. Despite these problems, there are encouraging indications suggesting that, in the near future, heterologous expression systems will provide labelled GPCRs in sufficient quantities for NMR analysis to characterize the details of the interactions between GPCRs and their ligands.

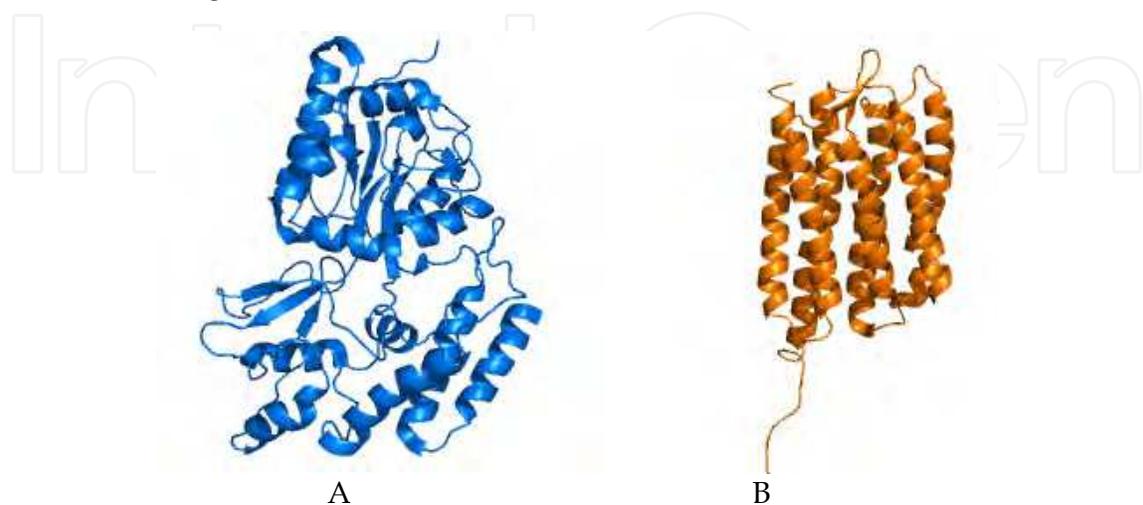

A   B

Fig. 1. Ribbon representation of the solution NMR structures of **(A)** the maltodextrin-binding protein (PDB ID 2KLF; Madl et al., 2009) and **(B)** the phototaxis receptor sensory rhodopsin II (pSRII) from *Natronomonas pharaonis* (PDB ID 2KSY; Gautier et al., 2010).

Most NMR studies of biological macromolecules have been carried out in solution. Since about 10 years ago, developments in NMR technology, sample preparation, pulse-sequence methodology and structure calculation protocols have allowed the structure determination of several proteins by solid-state NMR spectroscopy. Thus, 17 structures corresponding to 11 unique proteins have been solved using oriented samples of protein/lipids mixture (http://www.drorlist.com/ nmr/SPNMR.html). In particular, the M2 protein, a small membrane protein that enables hydrogen ions to enter the influenza A viral particle, was characterized by NMR to better understand its 3D structure and its interaction with drugs used against flu (Hu et al., 2010; Sharma et al., 2010; Figure 2A). Solid-state NMR was also used to solve 18 structures corresponding to 5 proteins forming fibrils (Van Melckebeke et al., 2010; Figure 2B) and 8 proteins yielding microcrystalline samples (Jehle et al., 2010; Figure 2C). Like liquid-state NMR, solid-state NMR is based on local structure information. As a consequence, large amounts of data on several differently labelled samples are typically required to obtain sufficient long-range information. However, NMR is particularly useful to study fibrils and microcrystals because these preparations are highly ordered on a local scale and consequently show narrow resonance lines. In recent years, considerable progress has been made in using solid-state NMR spectroscopy to determine atomic-resolution structures of amyloid fibrils associated with serious disorders such as Alzheimer disease, Parkinson disease, prion diseases and type 2 diabetes mellitus (for review: Tycko, 2011). Although the architecture of fibrils is considered to be a continuous stack of β-sheet ladders termed a cross-β structure, there may be significant variations in the supramolecular organization of the peptides within the fibrils. NMR provides detailed insight on the structural organization of these fibrils, highlighting their assembly process and inspiring the design of fibril-binding compounds and inhibitors.
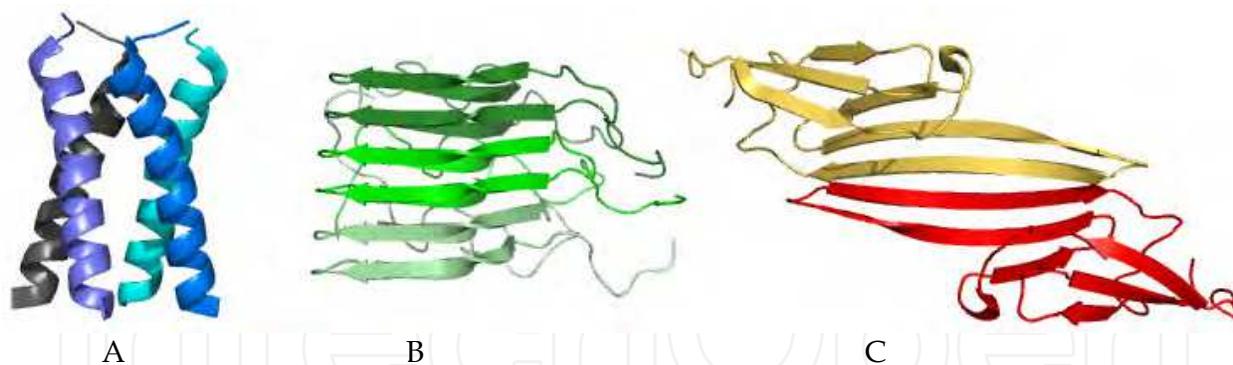
Fig. 2. Ribbon representation of the solid-state NMR structures of **(A)** the M2
transmembrane peptide of the influenza A virus in DMPC lipid bilayers (PDB ID 2KQT;
Cady et al., 2010), (B) the HET-s(218-289) prion in its amyloid form (PDB ID 2KJ2; Van
Melckebeke et al., 2010) and (C) the α-crystallin domain in αB-crystallin oligomers (PDB ID
2KLR; Jehle et al., 2010).

NMR can be used to study a large panel of molecular complexes, including those
characterized by a low affinity binding interaction. At the present time, NMR has been used
to solve about 1,500 3D structures of protein complexes. The majority of these structural
studies have been performed on complexes characterized by submicromolar affinities. In
cases of weak binding, characterized by a millimolar to micromolar affinity, it is possible to
map and characterize details of the interface. The simplest and most popular approach is the
so-called "chemical shift mapping" technique wherein changes in backbone chemical shifts
of one molecule are monitored during titration with another molecule by recording a series
of heteronuclear single-quantum correlation (HSQC) spectra (Figure 3). Upon addition of a
ligand, chemical shift, line width and intensity changes of specific HSQC peaks indicate that
the corresponding residues experience a modification of their chemical environment. This
modification is due to the complex formation, and may result either from the proximity
between the observed residue and the ligand or from binding-induced conformational
changes in the first molecule close to the observed residue. The exact effects of the complex
formation on the HSQC spectrum depend on the relative values of the NMR time scale and
the exchange rate of the complex. The NMR time scale refers to the chemical shift scale:
differences in chemical shifts between species in exchange. If this difference is larger than
the exchange rate, then the exchange is classified as SLOW. If this difference is equal to the
exchange rate, then the exchange is INTERMEDIATE. If this difference is smaller than the
exchange rate, the exchange is FAST. Protein-protein interactions with Kd < 1μM (tight
binding) are generally in slow exchange. In this case, two NMR signals are seen for the free
and bound forms. During the titration, the free form signal gradually disappears and the
bound form signal appears. Protein-protein interactions with Kd > 1mM (weak binding) are
generally in fast exchange. In such cases, only one averaged signal is seen with a chemical
shift fractionally weighted according to the populations and chemical shifts of the free and
bound forms. Protein-protein interactions with a Kd between 1 μM and 1 mM are generally
in intermediate exchange. This gives rise to broadened signals from both the free and bound
forms. During addition of a ligand, identification of the NMR peak changes and mapping of
these changes onto the protein structure provides information on the binding interface and
the conformational changes of the protein due to the interaction (Figure 3). Analysis of these
data opens up great opportunities for understanding the dynamics of interaction networks

(Mackereth et al., 2011). Which molecules are in competition for binding to the studied protein? Which molecules bind simultaneously? Is a conformational change necessary for the complex formation? Is this conformation change involved in the regulation of other interactions? How multiple domain proteins exploit their flexibility in order to recognize their targets?
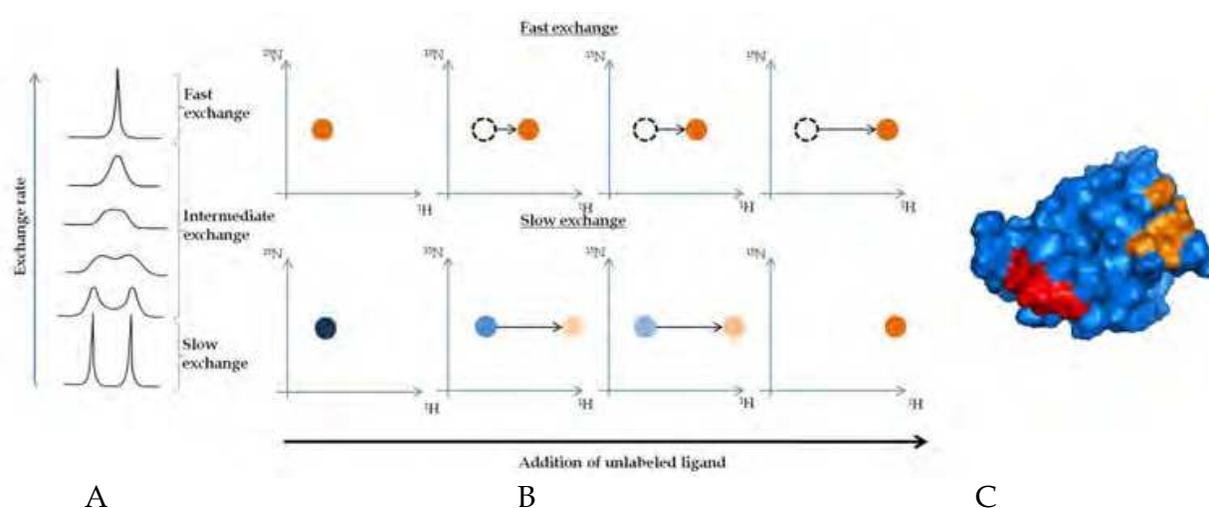

Fig. 3. Mapping of binding sites by NMR. **(A)** In the presence of free and ligand-bound proteins, the NMR signal varies as a function of the exchange rate. Its chemical shifts and linewidths evolve from a FAST exchange to a SLOW exchange condition**.** When the exchange is FAST, only one averaged chemical shift is observed. When it is SLOW, the chemical shifts corresponding to the free and bound forms are observed. **(B)** The chemical shift changes measured during the titration of a labelled protein with its binding partner are displayed in the case of FAST exchange (upper panel) and SLOW exchange (lower panel). **(C)** Residues whose NMR chemical shifts and/or linewidths change during the titration with ligands A and B are colored in red and orange, respectively, on the surface representation of the 3D structure of the labelled protein. This "chemical shift mapping" study shows that the labelled protein recognizes its two ligands by different surfaces.

NMR allows for direct observations of NMR-active nuclei (^{1}H, ^{15}N, ^{13}C, ^{31}P) within any NMR-inactive environment and can therefore be employed to investigate the structures of appropriately labelled biomolecules either in cell extracts or inside live cells. It has been used to structurally characterize biomolecules in the presence of cell extracts, in order to investigate post-translational modifications, conformational changes and binding events at the residue level within macromolecules in conditions mimicking the cellular environment (Liotakis et al., 2010). To accomplish this, NMR-active labelled recombinant proteins are injected into cell extracts and their chemical shift evolution is followed in the extracts after addition of specific enzyme co-factors or binding partners. These experiments are particularly powerful at describing the molecular binding events regulated by post-translational modifications. Furthermore, NMR has been applied to the observation of protein-protein interactions and the determination of protein structures in *E. coli*. The so-called STINT-NMR approach entails sequentially expressing two (or more) proteins within a single bacterial cell in a time-controlled manner and monitoring the protein interactions using in-cell NMR spectroscopy (Burz et al., 2006). NMR has also been used to follow the behavior of intrinsically disordered proteins in the crowded environment of bacterial cells

(Dedmon et al., 2002). The 3D structure of the putative heavy-metal binding protein TTHA1718 from *Thermus thermophilus HB8* was solved from data acquired on living *E. coli* samples (Sakakibara et al., 2009). In eukaryotic cells, the first in-cell NMR studies were performed by injecting proteins into *Xenopus laevis* oocytes (Selenko et al., 2006). Cell-penetrating peptides have also been used to deliver labelled proteins that can be observed in living cells (Inomata et al., 2009). However, the low sensitivity of the method and the short lifetime of the samples have so far prevented the acquisition of structural data in most of these eukaryotic systems. Development of ultra-fast methods for multidimensional experiments has already contributed to the first successes of in-cell NMR and should be further exploited to follow biological events.

Solid-state and liquid-state NMR techniques are now essential tools for understanding biological processes from a mechanistic point of view at the atomic level. They can describe critical steps of a disease mechanism, highlight the mechanism of action of a particular drug and inspire the design of binding compounds and complex assembly inhibitors. We will now focus on the use of NMR for the elucidation of the molecular mechanisms of specific genetic diseases caused by either missense mutations or amino acid deletions/insertions in several proteins located at the nuclear envelope.

2. NMR and genetic diseases involving the nuclear envelope

We have used NMR to study the molecular mechanisms of genetic diseases caused by mutations in genes encoding inner nuclear membrane proteins (see for examples Krimm et al., 2002 and Caputo et al., 2006). These proteins play critical roles in nuclear structure and positioning (Figure 4). They also influence genome spatial and temporal organization and genome functional properties. Because of their many biological functions, the pathological mechanisms resulting from mutations are complex and not well understood (Dauer & Worman, 2009). A wide range of diseases affecting different organs system have been linked to mutations in the *LMNA* gene, encoding the A-type nuclear lamins, which are intermediate filament proteins mainly located at the inner nuclear membrane of differentiated somatic cells. These diseases include dilated cardiomyopathy with variable muscular dystrophy, Dunnigan-type familial partial lipodystrophy, a Charcot-Marie-Tooth type 2 disease, mandibuloacral dysplasia and Hutchinson-Gilford progeria syndrome. Adult-onset autosomal dominant leukodystrophy is caused by duplication of the gene encoding lamin B1, which is an intermediate filament protein located at the inner nuclear membrane of all nucleated somatic mammalian cells. In addition, several diseases are linked to mutations in genes encoding integral proteins of the inner nuclear membrane that are associated with nuclear lamins, such as emerin, LBR and MAN1. Structural studies of the proteins involved in these so-called "laminopathies" or "nuclear envelopathies" have provided mechanistic clues about the functions of the nuclear envelope and insights into the disease pathogenesis.

Some mutations causing laminopathies are frameshifts or premature termination codons resulting in truncated protein. In these cases, the protein is generally rapidly degraded and its function is lost. NMR can however be used to structurally characterize the native protein, thus providing molecular details on its lost function (Caputo et al., 2006). Others mutations generate amino acid substitutions that can destabilize the protein leading to its premature degradation, interfere with its degradation pathway or modify its interaction with other

proteins or nucleic acids. In these cases, NMR can provide information on the structural consequences of the mutation (Krimm et al., 2002). Does the mutation affect a hydrophobic residue in the core of the protein and thus destabilize its 3D structure? Does it affect a residue at the surface of the protein that is critical for a specific interaction with a biological partner? Structural studies utilizing NMR, by identifying the structural consequences of the missense mutations, can point out specific disease mechanisms. Such information could be used to develop therapies targeted to the defects.
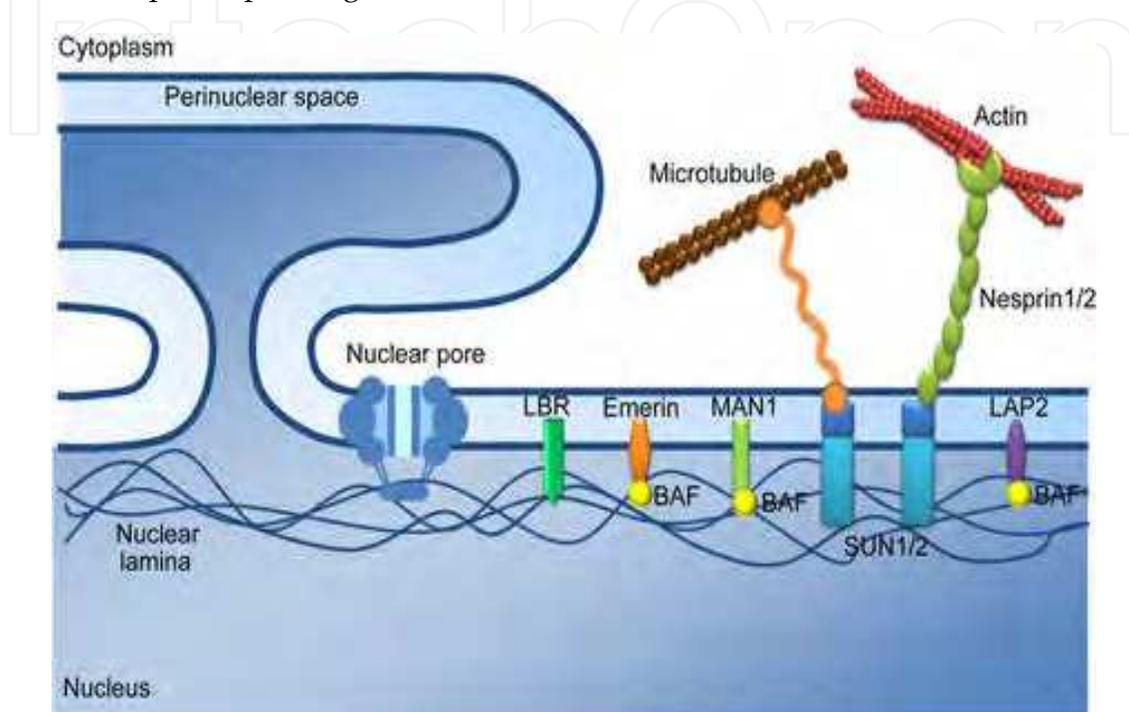

Fig. 4. Architecture of the nuclear envelope (reproduced with permission from Chi et al., 2009). The nuclear lamina, comprising the A- and B-type lamins, interacts with several proteins anchored at the inner nuclear membrane such as LBR, emerin, MAN1 and SUN1/2. SUN1/2 proteins also interact with outer nuclear membrane proteins that in turn bind to microtubules and actin. This network of protein-protein interactions is essential for cell structural organization.

We have used NMR to study A-type lamins, showing how different types of mutations that cause different diseases affect protein structure, hence providing clues about pathogenesis. We have also used NMR to study LEM domain proteins, which mediate interactions between the inner nuclear membrane and DNA and are involved in bone and muscle diseases. We have further analysed in detail MAN1, a LEM domain protein that regulates a key signal transduction pathway and whose heterozygous loss of function causes sclerosing bone dysplasias.

Many proteins associated with cellular control and signalling mechanisms such as A-type lamins and LEM domain proteins are modular in structure in that they contain both well-folded domains and poorly-structured regions (Wright & Dyson, 2009). The presence of poorly-folded regions facilitates protein accessibility to modifying enzymes and interaction with a wide variety of targets. However, this also hinders protein structural characterization, as the poorly-structured regions can adopt a large number of

conformations and fold into different structures on binding to different target proteins. The pathological consequences of mutations located in the poorly-folded regions would then not be observed on the 3D structure of the mutated protein but instead linked to modifications of its binding properties. To tackle the structural description of these proteins, we chose to (1) solve the 3D solution structures of their globular domains and (2) characterize the relative positioning of domains within sub-regions. This "divide-and-conquer" approach is based on the combination of NMR spectroscopy and small-angle X-ray scattering (SAXS) and is a powerful method to characterize the structural ensemble of partially disordered proteins (Madl et al., 2011). From the description of the three-dimensional structures of sub-regions of A-type lamins and LEM domain proteins, we then (1) determined the impact of disease-linked mutations on the structure and stability of the proteins, (2) mapped interaction surfaces of proteins with their biological partners, (3) positioned the mutations relatively to these functionally important surfaces and (4) characterized the 3D structure of complexes involving these proteins.

3. How do mutations in A-type lamins cause different diseases?

3.1 Architecture of the lamin intermediate filaments

Nuclear lamins are classified as either A-type or B-type according to homology in sequence, biochemical properties and localization during the cell cycle. In humans, the *LMNB1* and *LMNB2* genes encode lamin B1 and B2, respectively, which are expressed in most or all nucleated somatic cells. The *LMNA* gene encodes the A-type lamins and the major isoforms, lamin A and lamin C, are expressed in most differentiated somatic cells. Lamins A and C are produced by alternative RNA splicing. They share a common region from amino acids 1-566 and differ at their C-terminus, from amino acids 567-664 for prelamin A, the lamin A precursor that is processed to lamin A, and from amino acids 567-572 for lamin C.

Like all intermediate filament proteins, lamins are fibrous in nature and share a tripartite structural organization. A non-α-helical N-terminal region (head) and a C-terminal region (tail) flank a central α-helical coiled-coil domain (rod). Lamins differ from the cytoplasmic intermediate filament proteins in that they have an extended rod domain (42 amino acid longer), that they have a nuclear localization signal (NLS) between their rod and tail regions and that they display a typical tertiary structure in this tail region (Herrmann et al., 2009; Figure 5). In contrast to most cytoplasmic intermediate filament proteins, lamins also contain sites for phosphorylation by mitotic kinases and most contain carboxy-terminal CAAX motifs that signal posttranslational modification by farnesylation. The lamin α-helical central rod domain is divided into three coiled regions (termed 1A, 1B and 2), which mediate lamin assembly into filaments. *In vitro* experiments have revealed that elementary lamin dimers first assemble through longitudinal N–C interactions and then these assemblies associate through lateral interactions to form intermediate filaments.

X-ray crystallography has been used to determine the atomic structure of various fragments of the lamin intermediate filament rod domain (Herrmann et al., 2009). In particular, it has been used to solve the crystal structure of a human lamin A fragment comprising residues 305 – 387, which corresponds to the second half of coil2 (PDB ID 1X8Y; Strelkov et al., 2004). This structure revealed a left-handed parallel coiled-coil extending to the predicted end of the rod domain (Figure 5). Essentially the same fragment of lamin B1 has also been resolved and a similar structure was observed (PDB ID 3MOV). Interestingly, the crystal structure of

another human lamin A fragment (residues 328–398), which is largely overlapping with fragment 305–387 but harbours a short segment of the tail domain, was also solved. Unexpectedly, there is no parallel coiled-coil form within the crystal. Instead, the α-helices are arranged such that two antiparallel coiled-coil interfaces are formed (PDB ID 2XV5). The most significant interface has a right-handed geometry, which results from a characteristic 15-residue repeat pattern that overlays the canonical heptad repeat pattern (Figure 5). Analysis of these different modes of coil association gives some clues about the mechanisms of lamin polymerization in cells (Kapinos et al., 2011). Several polymerization mechanisms have been proposed that could co-exist and be modulated by specific chaperone-type molecules as well as post-translational modifications.
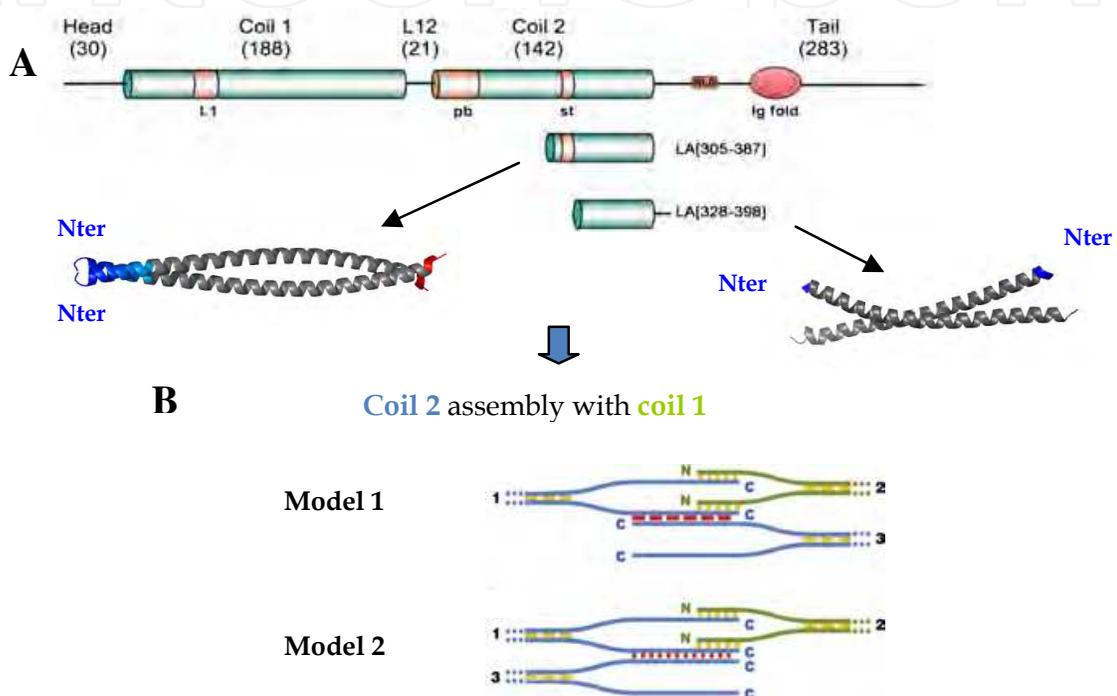

Fig. 5. Architecture of lamins, adapted with permission from Kapinos et al., 2011. **(A)** General structural organization of lamin monomers. Two A-type lamin coil 2 fragments were studied at atomic resolution (PDB IDs 1X8Y and 2XV5). Their X-ray structures are displayed in the cartoon, in which the sequence of amino acids 328-382, common to both fragments, is colored in gray and the remaining residues colored blue (N-terminus) or red (C-terminus). **(B)** Two models of lamin filament structural organization proposed from the analysis of the rod fragment X-ray structures. The C-terminal rod sections of two dimers labelled 1 and 3 are colored blue and the N-terminal rod section of dimer 2 is colored green. Broken lines represent coiled-coil interfaces that have been observed in crystal structures and dotted lines indicate other theoretically possible coiled coil interactions.

We have solved the 3D solution structure of the C-terminal globular domain of A-type lamins, which participates to the recognition of various lamin binding proteins (PDB ID 1IVT; Krimm et al., 2002). Determination of this structure was made from experimental measured values of short inter-proton distances (<6 Å) obtained by ^{1}H, ^{15}N and ^{13}C NMR on a ^{13}C and ^{15}N labelled sample of the domain. Figure 6A shows the well-dispersed ^{1}H-^{15}N HSQC spectrum of the C-terminal globular domain of A-type lamins. The labeling was

essential to assign the inter-proton distances obtained by ^{1}H NMR to specific proton pairs. Molecular modeling calculations were used to calculate a family of 3D structures, all very close one to the other, compatible with the distances measured by NMR (Figure 6B). These structures, which reflect the structure of the domain in solution, have an immunoglobulin fold between residues 430 and 545. Crystal structures of this C-terminal domain of A-type lamins have also been solved and show similar structural organization (Dhe-Paganon et al., 2002; PDB IDs 1IFR and 1UFG).
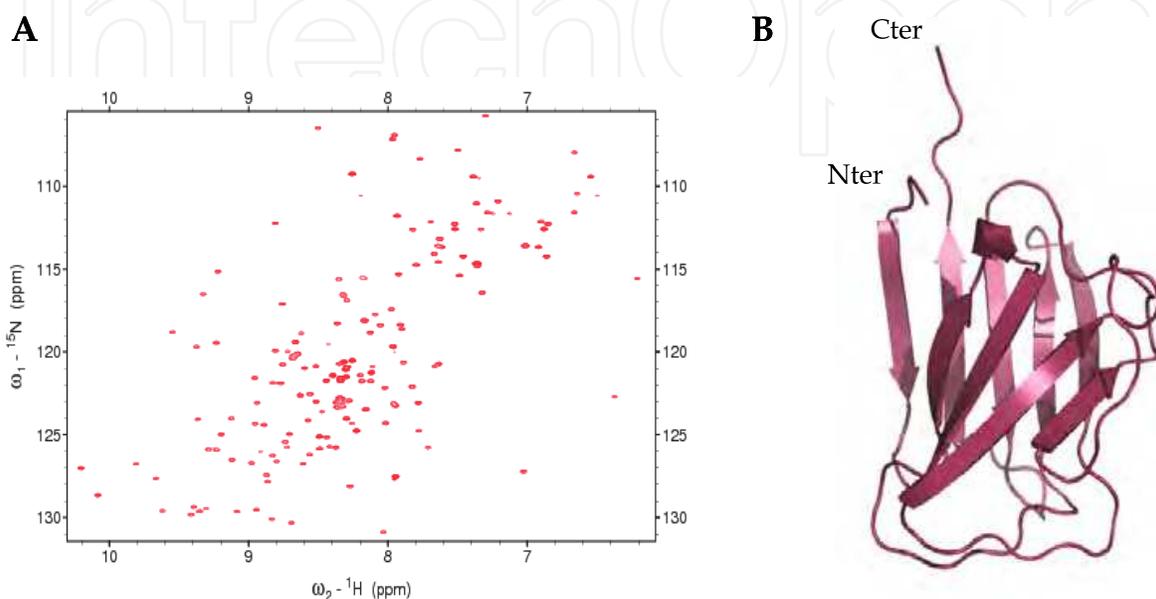

Fig. 6. Structural characterization of the globular domain found in the tail region of lamins. (A) ^{1}H-^{15}N HSQC spectrum of the domain from human A-type lamin. This spectrum was acquired on a 40 μM protein sample in 250 μl phosphate buffer pH 7.4 over 1 hour using a 700 MHz Bruker spectrometer. Each residue of the domain is revealed by a peak whose position depends on the residue chemical environment. Well-folded domains exhibit well-dispersed peaks, as shown for the lamin tail domain. (B) Cartoon representation of the 3D structure of the fragment of amino acids 428-549 from human A-type lamin, as calculated from NMR experimental data.

3.2 Post-translational modification of A-type lamins

Lamins undergo post-translational modifications that alter their localization, assembly and binding properties. Prelamin A is converted into lamin A by four enzymatic post-translational processing steps: farnesylation of a cysteine in a CAAX motif at the carboxyl terminus of the protein, endoproteolytic cleavage of the last three amino acids of the protein, carboxyl methylation of the newly exposed farnesylcysteine and subsequent endoproteolytic release of the last 15 amino acids, including the farnesylcysteine methyl ester. B-type lamins also contain a CAAX motif and undergo the first three of these enzymatic reactions but not the last. A-type and B-type lamins are also phosphorylated by cyclin dependent kinases, which regulate nuclear lamina assembly and disassembly during mitosis. Other phosphorylation, acetylation and sumoylation events have also been reported whose physiological importance is largely unknown. According to the PhosphoSite database, more than 60 phosphorylation sites are present in human A-type

lamins. Most of these are located in unstructured regions of the proteins (Figure 7). These phosphoresidues could be important for lamin functions by regulating interactions with other nuclear proteins.
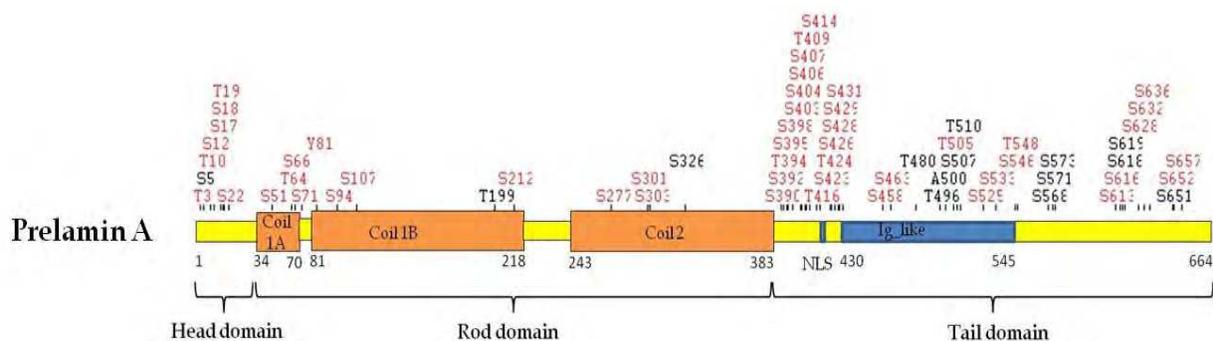

Fig. 7. Phosphorylation sites in human A-type lamins that have been reported in the PhosphoSite database (http://www.phosphosite.org/).

Post-translational modifications can be followed by NMR, either by adding the appropriate enzyme to the labelled protein sample or by diluting the labelled protein into cell extracts and adding the appropriate co-factor (Liotakis et al., JACS 2010). Series of NMR spectra are then recorded to follow the protein chemical shift evolution with time. Residues whose chemical shifts are modified because of the presence of enzymes in the sample are identified. These residues are either the targets of the enzyme or are close to the modified residues. NMR is particularly efficient in dissecting the modification kinetics of protein regions rich in inter-dependant modification sites.

3.3 Localization and structural impact of disease-linked missense mutations identified in A-type lamins

Several different diseases are caused by mutations in the *LMNA* gene encoding lamin A and C (Worman et al., 2010). More than 1,500 mutations have been identified within this gene. Amino acid substitutions in various regions of A-type lamins cause striated muscle diseases, partial lipodystrophy or progeroid disorders but in most cases a correlation between the localization of the mutation and the disease phenotype is not readily apparent. We searched for a correlation between the localization of the mutated residue in the lamin 3D structure and the disease phenotype. We focused on the lamin tail globular domain, because amino acid substitutions in this region cause either striated muscle diseases, lipodystrophy or progeroid disorders. Analysis of the position of the mutated residues in the NMR structure of the immunoglobulin-like domain and experimental measurement of the impact of several selected mutations on this structure were carried out (Krimm et al., 2002). We observed that most dominantly inherited mutations causing diseases of striated muscle affect residues of the hydrophobic core of the immunoglobulin-like domain (Figure 8A). These mutations probably destabilize the 3D structure of this domain, leading to an overall loss of protein function. Such destabilization has been experimentally verified for the Arg453Trp mutant. In contrast, dominantly inherited mutations causing Dunnigan-type familial partial lipodystrophy, a disease affecting adipose tissues, affect residues located at the surface of the immunoglobulin-like domain and cluster at a specific positively charged site (Figure 8B). This site is conserved in A-type and B-type lamins. This suggests that mutations causing

adipose tissue diseases do not destabilize the domain structure but may hinder the interaction of A-type lamin with a binding partner that plays a critical role in adipose tissue. Recessive mutations in the immunoglobulin-like domain of A-type lamins causing progeroid disorders also affect residues located at the surface of this domain and cluster at a specific site (Figure 8C; Verstraeten et al., 2006). This hot spot is distinct from the surface affected in partial lipodystrophy and is conserved in A-type lamins. It may correspond to another protein binding site specific to A-type lamins, with mutations causing progeroid disorders hindering this putative interaction. Experiments to identify the proteins interacting with A-type lamins at the different identified hot spots are now underway.
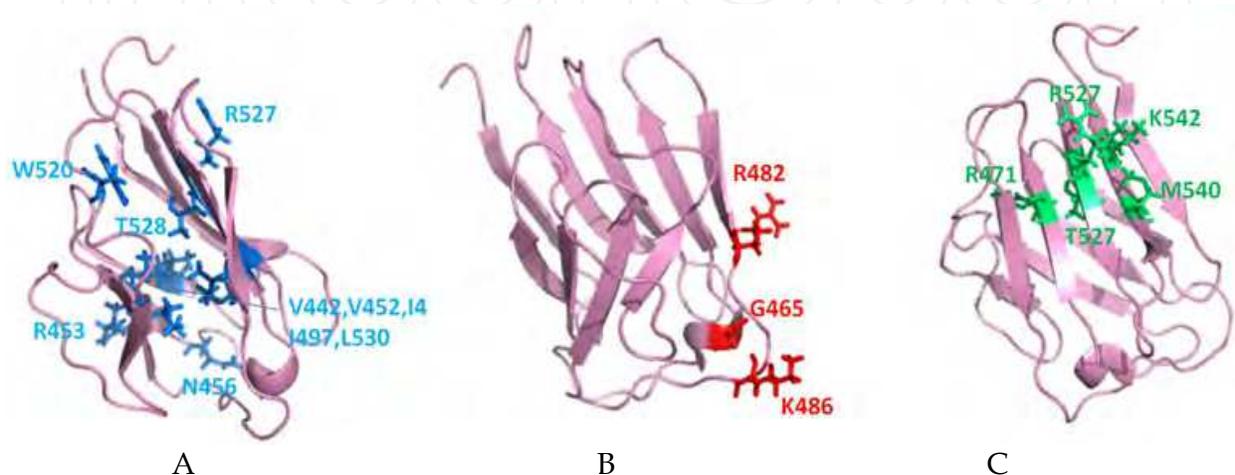

Fig. 8. Localization of the residues identified as mutated in laminopathies in the 3D structure of the tail globular domain of A-type lamins (PDB ID 1IVT; Krimm et al., 2002; Verstraeten et al., 2006). In **(A)**, residues mutated in striated muscle diseases are displayed in blue sticks. In **(B)**, residues mutated in adipose tissue diseases are displayed in red sticks. In **(C)**, residues mutated in progeria-like diseases are displayed in green sticks.

3.4 Perspectives: NMR study of how lamin post-translational modifications affect protein structure and interactions

Several studies have shown that the lamin tail globular domain and the flexible lamin regions are involved in binding to other proteins. We know that flexible regions of lamins are highly modified in cells (Figure 7). This suggests that post-translational modifications could be involved in regulating lamin recognition. In fact, post-translational modification of partners has already been shown to be involved in regulating lamin binding. For example, lamin A is involved in the sequestration of c-Fos, a member of the dimeric activator protein 1 (AP-1) transcription factor family that regulates different cellular pathways including cell proliferation, death survival differentiation and oncogenic transformation. Phosphorylation of c-Fos by ERK1/2 protein kinase reduces lamin A/c-Fos interaction, thus activating AP1 complexes (Gonzales et al., 2008).

In-cell NMR is a useful way to analyse these regulation pathways involving post-translational modifications of lamins or lamin partners. Using NMR chemical shift mapping as presented in part 1, modified residues can be identified onto the lamin tail or onto its binding partners. Next, interaction between modified or unmodified lamins and their binding partners, and conversely between lamins and modified or unmodified partners, can

be tested. Our strategy is to produce recombinant lamins and partners and inject them into human cell extracts with or without addition of cofactors that are essential for post-translational-modifications. Figure 9 shows our first ^{1}H-^{15}N HSQC spectra of the tail region of lamin C incubated into 293T cell extracts during 1 to 24 hours. This experiment shows that it is possible to follow the fate of a partially unfolded protein in cell extracts during one day.
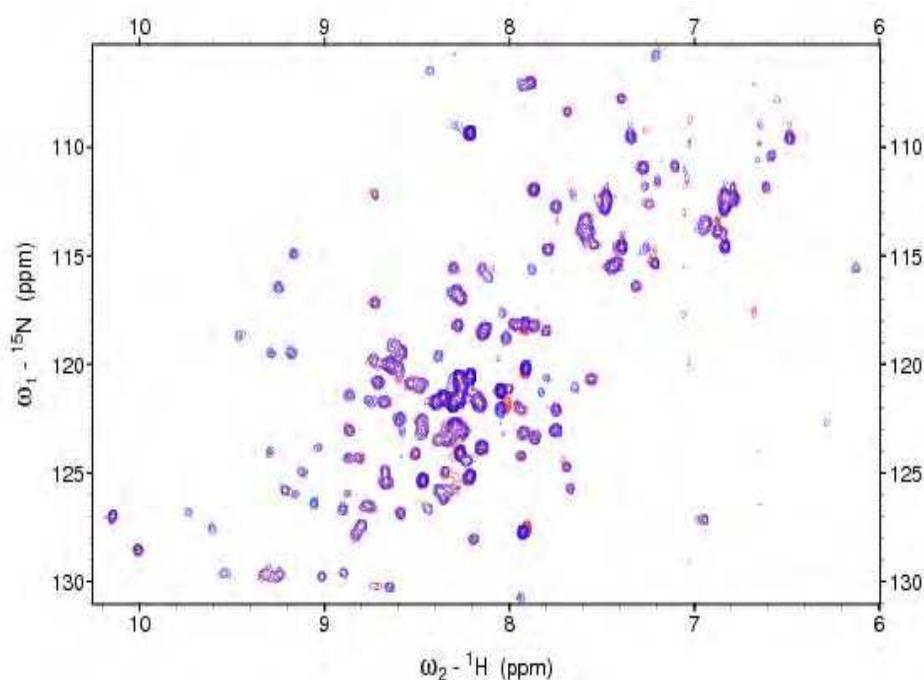

Fig. 9. Following modification and binding events involving the lamin C tail in 293T cell extracts: feasibility experiments. ^{1}H-^{15}N HSQC spectra were recorded on a Bruker 700 MHz spectrometer after 1 hour (red) and 24 hours (blue) of incubation of a 50 μM sample of labelled protein in 250 μl of cell extracts at 5mg/ml. The protein is stable in these conditions. Co-factors should then be added in order to allow modification of the labelled protein.

4. How do mutations in LEM domain proteins affect interactions between the nuclear envelope and chromatin?

4.1 Architecture of LEM domain proteins

LAP2 isoform β (LAP2 β), emerin and MAN1 are human nuclear proteins that share a highly conserved structural motif called the LEM domain (Lin et al., 2000). The LEM domain contains approximately 50 amino acids. It is located at the N-terminus of these proteins and is often separated from the rest of the protein by predicted unstructured regions (Figure 10). Both MAN1 and emerin possess one copy of the LEM domain in their N-terminal nucleoplasmic region whereas LAP2 β contains a LEM-like domain and a LEM domain connected by a predicted unstructured linker.

The 3D solution structures of emerin and LAP2 β LEM domains and LAP2 β LEM-like domain have been solved using NMR (Wolff et al., 2001; Laguri et al., 2001; Cai et al., 2001). These LEM and LEM-like domains share a similar 3D structure, mainly composed of a short N-terminal helix and two large α-helices interacting through a set of conserved amino acids

residues (Figure 10). NMR characterization of the fragment from amino acids 1-168 of LAP2 β containing both the LEM-like and LEM domains has shown that the two domains are structurally independent, non-interacting domains connected by an approximately 60-residue flexible linker (Cai et al., 2001).

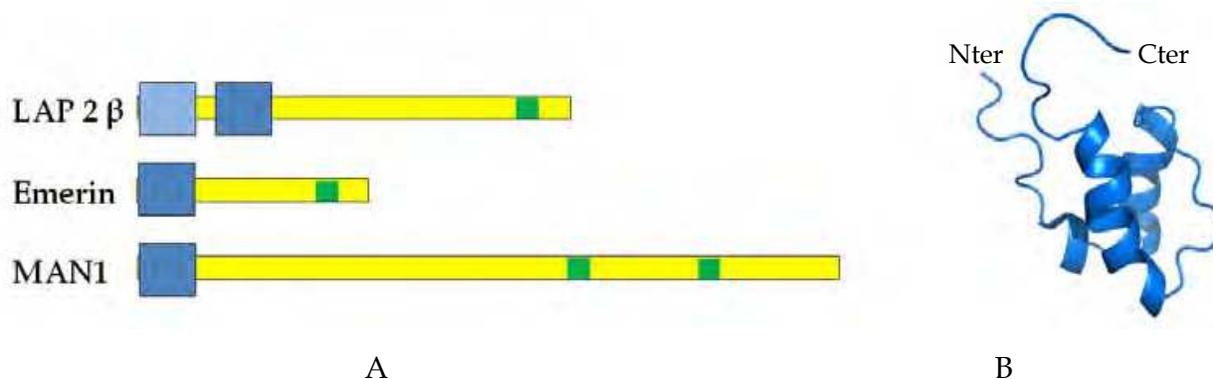A   B

Fig. 10. Schematic organization of nuclear envelope proteins containing a LEM domain. **(A)** Three human inner nuclear membrane proteins are displayed: LAP2 β, MAN1 and emerin. Dark blue boxes represent LEM domains, the light blue box represents a LEM-like domain, green rectangles represent transmembrane segments. **(B)** Ribbon representation of the average solution structure of the LEM domain of LAP2 β (PDB ID 1H9F; Laguri et al., 2001).

4.2 LEM domains recognize the DNA binding BAF protein

Barrier-to-autointegration factor (BAF) is a highly conserved metazoan protein, which functions in nuclear assembly, chromatin organization and gene transcription. BAF dimers bind to double-stranded DNA non-specifically and thereby bridge DNA molecules to form a large, discrete nucleoprotein complex. BAF is able to interact simultaneously with both LAP2 and DNA *in vitro* (Shumaker et al., 2001). This suggested that LEM proteins and BAF mediate chromatin attachment to the nuclear envelope during nuclear assembly or interphase, or both.

The interaction between the fragment containing amino acids 1-168 of LAP2 β and the BAF–DNA complex was characterized by NMR (Cai et al., 2001). It was proposed on the basis of chemical shift mapping experiments that the LEM-like domain recognizes DNA (Figure 11A) while the LEM domain recognizes the DNA binding protein BAF (Figure 11B). However, when the whole BAF-DNA complex was added to the labelled LAP2 β fragment, only the peaks corresponding to the LEM domain shifted, indicating that only the LEM domain interacts with the BAF–DNA nucleoprotein complex (Figure 11C).

Subsequently, the solution structure of the BAF-emerin LEM domain complex was solved on the basis of NMR data (Cai et al., 2007). This structure revealed that one BAF dimer interacts with one emerin LEM domain (Figure 12A). The stoichiometry of the complex prevents association between one molecule of BAF dimer and multiple LEM domain proteins. Moreover, the interaction surface between BAF and emerin does not overlap with the two symmetry related DNA binding sites on the BAF dimer (Figure 12B). Thus, BAF can simultaneously interact with emerin and DNA.
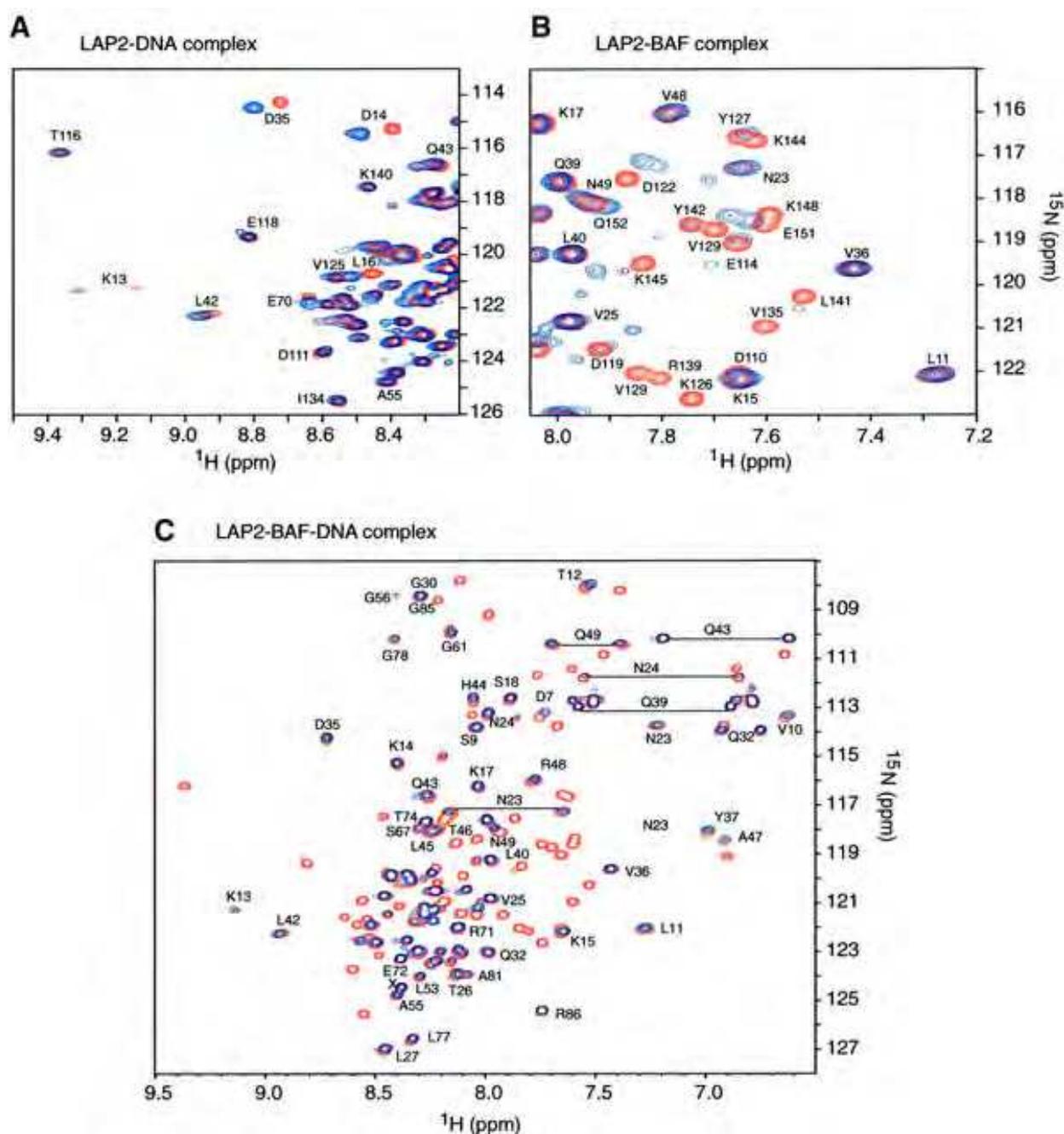
Fig. 11. Mapping the LAP2 β residues involved in BAF–DNA nucleoprotein recognition (reproduced with permission from Cai et al., 2001). (**A**) Comparison of the ^{1}H-^{15}N HSQC spectra of LAP2 β region containing amino acids 1-168 in the absence (red) and presence of 1.5 equivalents (blue) of a duplex DNA dodecamer. Only cross-peaks of the LEM-like domain shift upon DNA binding. (**B**) Comparison of ^{1}H-^{15}N HSQC spectra of LAP2 β region 1-168 in the absence (red) and presence (blue) of one equivalent of BAF. Only the cross-peaks in the LEM domain change significantly upon binding to BAF. (**C**) Comparison of the ^{1}H-^{15}N HSQC spectra of LAP2 β region containing amino acids 1-168 in the absence (red) and presence (blue) of the BAF–DNA nucleoprotein complex. Only the peaks originating from the LEM domain disappear upon binding the BAF–DNA nucleoprotein complex.
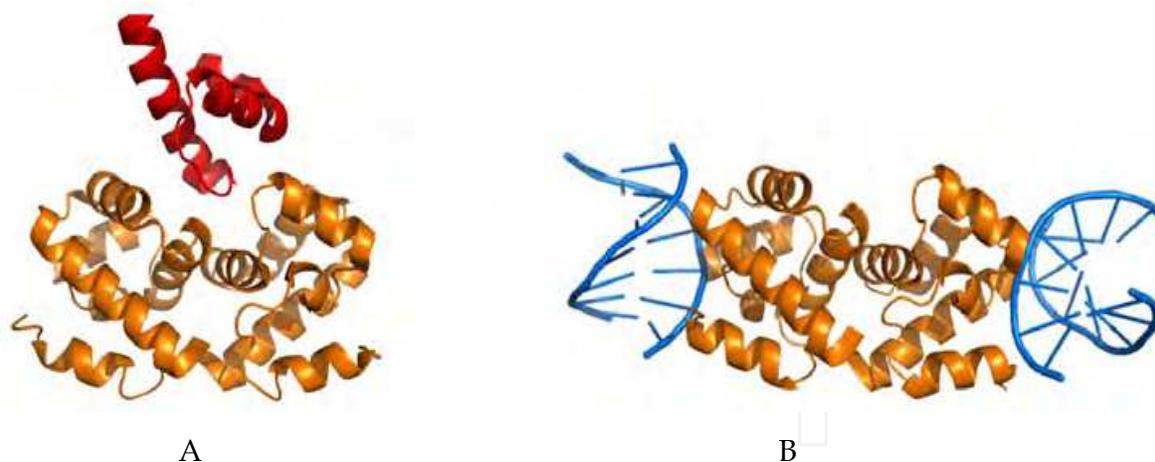
Fig. 12. Ribbon representation of **(A)** the solution structure of the BAF_2 dimer (orange) bound to the emerin LEM domain (red) (PDB ID 2ODG; Cai et al., 2007) and **(B)** the X-ray structure of the BAF_2 dimer (orange) bound to DNA (blue) (PDB ID 2BZF; Bradley et al., 2005).

4.3 Localization and structural impact of disease-linked mutations identified in emerin and BAF

X-linked Emery-Dreyfus muscular dystrophy (EDMD) is characterized by early contractures of the spine, Achilles tendon and elbows, muscle wasting in a humerperoneal distribution and dilated cardiomyopathy with conduction system abnormalities (Muchir & Worman, 2007). It is caused by mutations in the *EMD* gene encoding emerin (Bione et al., 1994), an integral protein of the inner nuclear membrane. At present, 130 mutations homogeneously distributed along the *EMD* gene have been reported (Brown et al., 2011). Approximately 90% of these mutations result in the complete absence of protein as characterized by immunohistochemical staining. The few mutations allowing modified emerin protein expression yield protein that is expressed in reduced amounts and is frequently mislocalized. The majority of patients with residual emerin expression display clinical phenotypes indistinguishable from their null counterparts. However, it is difficult to argue that there is no discernable genotype-phenotype correlation, as clinical data collection on EDMD patients is often incomplete. Moreover, the relationship between emerin absence or decreased expression and muscle abnormalities is poorly understood. As emerin has been hypothesized to regulate muscle specific gene expression and nuclear architecture, some have suggested that absence or reduced levels affect cellular susceptibility to mechanical stress, cellular proliferation and differentiation (Fidziańska & Hausmanowa-Petrusewicz, 2003; Frock el al., 2006).

Recently, a mutation (A12T) in the *BANF1* gene encoding BAF has been shown to induce a progeroid syndrome that shares physiopathological characteristics with Hutchinson-Gilford progeria syndrome (HGPS) (Puente et al., 2011). HPGS has received considerable attention because of its striking premature aging phenotype, including alopecia, diminished subcutaneous fat, premature atherosclerosis, and skeletal abnormalities. The vast majority of HGPS cases are associated with a de novo nucleotide substitution at position 1824 (C->T) in the *LMNA* gene encoding A-type lamins (Eriksson et al., 2003; De Sandre-Giovannoli et al., 2003). This mutation does not affect the coded amino acid (and is thus generally referred to as G608G), but partially activates a cryptic splice donor site in exon 11 of *LMNA*, leading to

the production of a prelamin A mRNA that contains an internal deletion of 150 base pairs. This transcript is then translated into a protein known as progerin, which lacks 50 amino acids near the C terminus. The new progeroid syndrome caused by a mutation in BAF differs from HGPS by the fact that affected patients have not cardiovascular deficiencies and live longer. A 3D model of BAF shows that the mutated residue is located on the surface of the protein (Figure 13). Nevertheless, the A12T mutation is not predicted to affect BAF dimerization or binding to either DNA or emerin, raising the possibility that this amino acid substitution could impair the interaction with other proteins, its subcellular localization or its stability. Examination of the effect of the BAF A12T mutation in the fibroblasts of affected patients showed a reduction in protein levels, indicating that this mutation affects the stability of BAF. Surprisingly, the decrease in BAF appears to be linked to a delocalization of emerin in the endoplasmic reticulum and abnormalities in the nuclear lamina. In HeLa cells that overexpressed a phosphomimetic BAF missense mutant S4E, but not S4A, emerin also mislocalized from the nuclear envelope (Bengtsson & Wilson, 2005). Furthermore, phosphorylation of serine 4 inhibited BAF binding to emerin and lamin A in blot overlay assays. These results suggest that the N-terminal region of BAF is normally involved in emerin localization at the nuclear envelope and that the A12T mutation could destabilize a lamin-emerin-BAF complex. Structural data on this complex are lacking so further structural analysis of the consequences of the BAF A12T mutation is not currently possible.
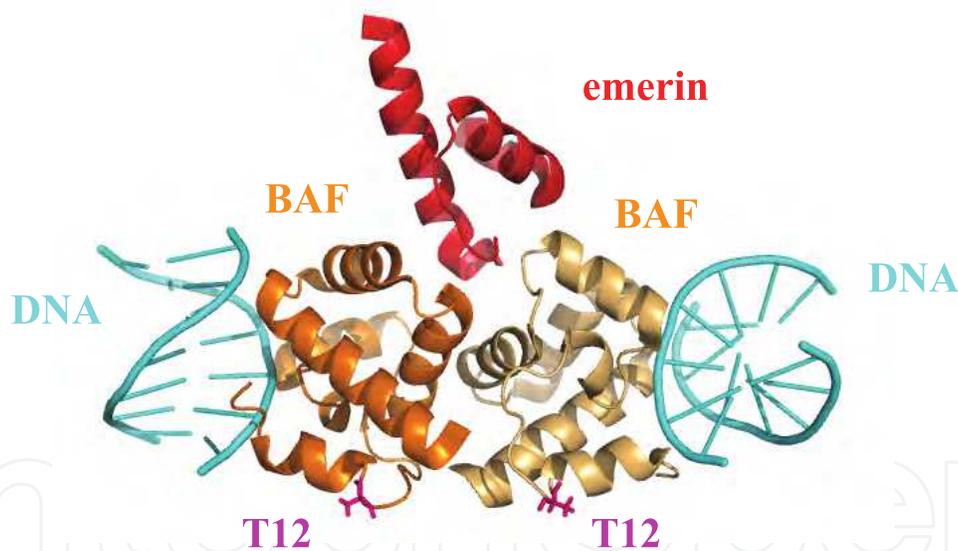

Fig. 13. 3D model of a BAF dimer (orange and light orange in the figure) in complex with emerin (red) and DNA (cyan) built from the PDB files 2ODG and 2BZF. The position of the T12 residue in mutant BAF is indicated in pink sticks (Puente et al., 2011).

5. How do mutations in the inner nuclear envelope MAN1 deregulate TGFβ signaling?

5.1 Architecture of MAN1, a LEM domain protein with an R-Smad binding C-terminal region

The *LEMD3* gene encodes MAN1, an integral protein of the inner nuclear membrane (Lin et al., 2000). MAN1 consists of 911 amino acids and exhibits two transmembrane domains

(Figure 14A). Its N-terminal and the C-terminal regions are located on the nucleoplasmic side of the membrane. The N-terminal nucleoplasmic region of MAN1 (amino acids 1-471) contains an N-terminal LEM (Figure 14B). This domain is essential for emerin, lamin and BAF binding (Liu et al., 2003; Mansharamani & Wilson, 2005). We have used ^{1}H, ^{15}N and ^{13}C NMR and SAXS to study the structural organization of the C-terminal nucleoplasmic region of MAN1 (Caputo et al., 2006; Kondé et al., 2010). This region is composed of a globular winged helix (WH) domain (Figure 14C), a linker region, a globular U2AF Homology Motif (UHM) domain (Figure 14D) and a C-terminal unstructured region. The WH domain presents a fold found in several transcription factors and it recognizes DNA (Caputo et al., 2006). The UHM domain is essential for MAN1 binding to Smad2 and Smad3 that are transcriptional regulators of the TGFβ pathway (Lin et al., 2005; Pan et al., 2005). The C-terminal region also recognizes other transcription regulators as the death-promoting transcriptional repressor Btf and the transcriptional repressor germ-cell-less GCL (Mansharamani & Wilson, 2005).
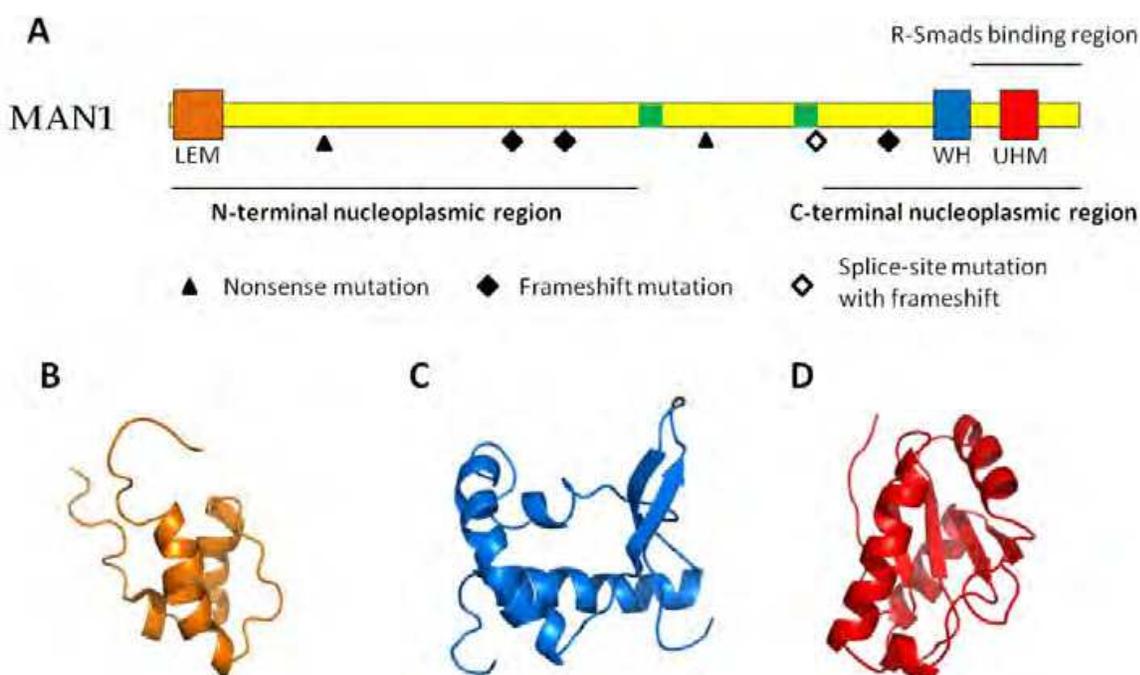

Fig. 14. Structural organization of the MAN1 protein. **(A)** Domain composition of MAN1 (the two transmembrane regions are colored in green) and localization of the mutations causing bone genetics diseases (Hellemans et al., 2004). **(B)** Model of the LEM domain of MAN1, built by homology on the basis of the LAP2 β LEM domain structure (PDB ID 1H9F; Laguri et al., 2001). **(C)** NMR structure of the WH domain (PDB ID 1CH0; Caputo et al., 2006). **(D)** Model calculated for the UHM domain on the basis of the secondary structure information deduced from the NMR chemical shift analysis and the 3D structure of the UHM domain of U2AF65 (PDB ID 1O0P; Kondé et al., 2010).

To understand the mechanisms of regulation of the TGFβ pathway by MAN1, we characterized the molecular determinants of its interaction with Smad2 and Smad3. We showed that the first WH domain is not essential for Smad2 binding. However, the linker region as well as the UHM domain and the C-terminus are all involved in Smad2 binding. NMR and SAXS data were used to reconstitute the fluctuating 3D structure of the Smad2 binding region of MAN1 (Kondé et al., 2010).
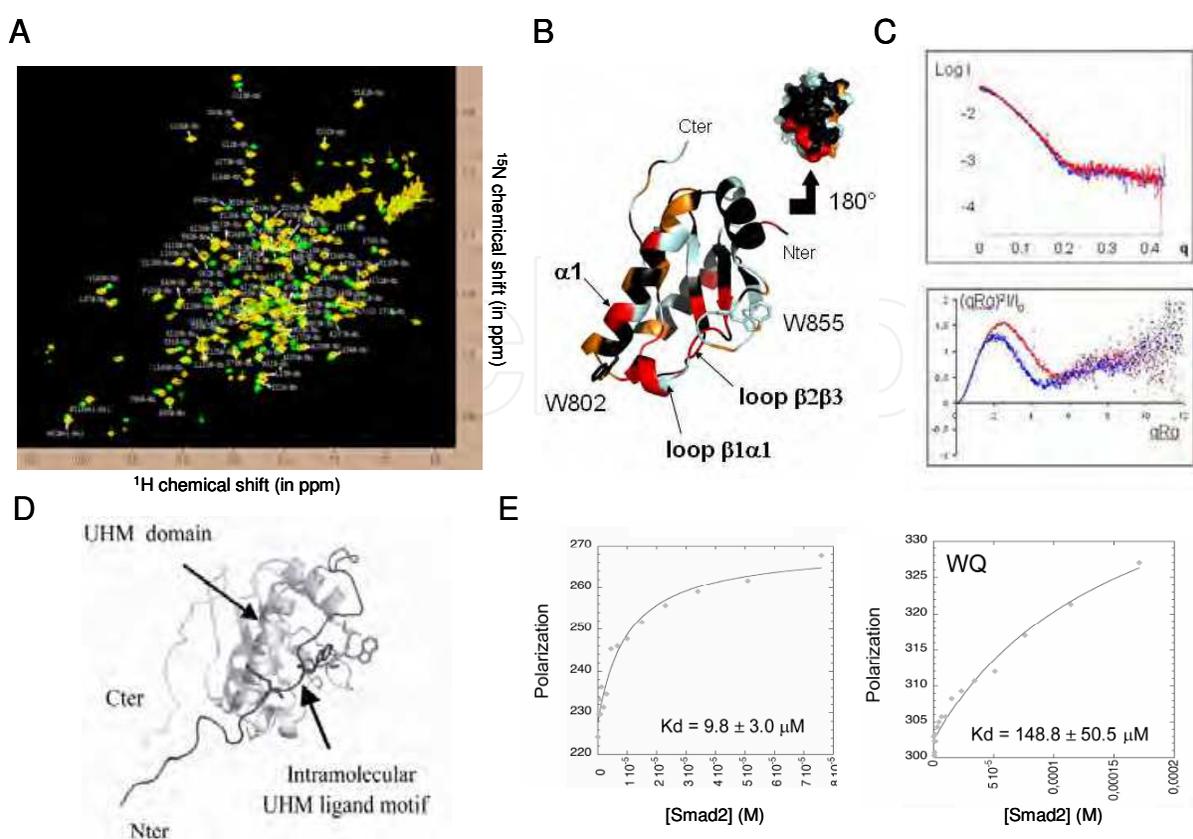
Fig. 15. Structural organization of the Smad2 binding motif of MAN1 (adapted with permission from Kondé et al., 2010). **(A)** NMR ^{1}H-^{15}N HSQC spectra of the wild-type (yellow) and mutated (W765A-Q766A, green) Smad2 binding motif of MAN1, representing the NMR finger print of the fold adopted by these protein fragments. Each NMR peak is assigned to the backbone NH bond of a specific amino acid in MAN1 sequence, and its position in the spectrum reflects its chemical environment in the protein domain. **(B)** NMR chemical shift mapping of the UHM region interacting with the linker. On the black ribbon representation of the MAN1 UHM model, residues whose ^{1}H-^{15}N HSQC peak is largely shifted and could not be assigned due to the mutations W765A and Q766A are colored in red, residues whose HSQC peak is shifted by more than 1.5 times the root mean square deviation of the chemical shift difference distribution but could still be assigned are coloured in orange, and residues whose HSQC peak could not be assigned in the MAN1LuhmWQ spectra are colored in cyan. The surface of the opposite face of the UHM domain is showed as an additional view. **(C)** Comparison of the SAXS data recorded on the wild-type (blue) and mutated (W765A-Q766A, red) Smad2 binding motif of MAN1. The upper view displays the superposition of the logarithm of the SAXS intensities as a function of the amplitude of the diffusion vector q = 4*π*sin(θ)/λ. The lower view corresponds to the superposition of the Kratky plots: the diffusion vector amplitude multiplied by the gyration radius is plotted on the abscissa and the square of the abscissa multiplied by the SAXS intensity divided by the intensity at 0 is plotted on the ordinate. This representation reveals the type of structure: compact, partially folded or unfolded. **(D)** Model of the Smad2 binding domain of MAN1 consistent with the NMR and SAXS data. The backbone is represented as a ribbon; the side chains of W765 and Q766, in the linker, and of several residues lining the hydrophobic cavity of the UHM domain are displayed as sticks. **(E)** Fluorescence polarization curves of the labelled wild-type (left) and mutated (W765A-Q766A, right) Smad2 binding fragment of MAN1 as a function of the concentration in Smad2.

We mutated the highly conserved W765 and Q766 of the linker region and observed the consequences of these mutations on the NMR ^{1}H-^{15}N spectrum of the Smad2 binding region (Figure 15A). Residues with NMR signals modified by the mutations were mapped onto the UHM domain (Figure 15B). They all clustered on one face of the domain and delineated a hydrophobic pocket in which the tryptophane residue binds. We performed SAXS experiments on samples of wild type and mutated Smad2 binding region. SAXS provides accurate information on folding and conformation in solution for both rigid and flexible macromolecules and the indirect Fourier transform of the SAXS intensity gives access to the real space electron pair distance distribution of molecules in solution. Figure 15C (upper panel) shows that the SAXS curves of the wild type and mutated protein fragments are similar, reflecting their common global fold. Calculation of the electron pair distance distributions using GNOM gave access to the radius of gyration Rg, the maximal distance and the I0 value that cannot be measured experimentally and is proportional to the mass. From these values, a Kratky representation was displayed (Figure 15C, lower panel) that highlighted structural differences between wild-type and mutated proteins: the maximum of the curve corresponding to (SRg)2 I/Io is shifted towards higher SRg values, reflecting the more elongated/less structured state of the mutant. Moreover, the wild type protein shows a gyration radius between 20 and 21 Å and a maximal distance between 68 and 72 Å, whereas the mutant shows a gyration radius between 23.3 and 24 Å and a maximal distance is close to 90 Å. This result validates our previous NMR based analysis that strongly suggests a role of W765 and Q766 in anchoring the linker onto the UHM domain. In contrast, the C-terminus of the Smad2 binding region of MAN1 is partially unfolded. From these observations, the positions of the linker and C-terminus regions relatively to the UHM domain were calculated, yielding to a model of the whole Smad2 binding region of MAN1 (Figure 15D). The functional importance of the positioning of the linker region onto the UHM domain was shown using fluorescence binding experiments: mutations of the W765 and Q766 residues into alanine yielded a 15-fold decrease of the affinity of MAN1 for Smad2 (Figure 15E).

5.2 Mechanism of bone diseases caused by the absence of the C-terminal region of MAN1

Heterozygous loss-of-function mutations in *LEMD3* encoding MAN1 cause the sclerosing bone dysplasias osteopoikilosis, Buschke-Ollendorff syndrome and non-sporadic melorheostosis (Hellemans et al., 2004). Several arguments suggest that the bone diseases result from enhanced TGF-ß signalling in cells.

First, all the disease-causing mutations including nonsense or frameshift mutations induce loss of the C-terminal nucleoplasmic fragment of MAN1, which is essential for Smad2 recognition (Figure 14A). Second, the TGF-β pathway is enhanced in cells from patients (Hellemans et al., 2004). Loss of the MAN1-Smad2 interaction may lead to the release of activated Smad2 and Smad3, thus causing an increase of Smad2 dependant transcription (Lin et al., 2005; Pan et al., 2005). It has been proposed that MAN1-Smad2 interaction not only sequestrates Smad2 and Smad3 at the inner nuclear envelope, but also favours their nuclear export (Pan et al., 2005; Figure 16). However, the mechanism by which the nucleoplasmic MAN1 C-terminal region would stimulate this nuclear export remains unclear. MAN1 could simply act as a "nuclear envelope sink" for activated nuclear Smad proteins, sequestering them at the inner nuclear membrane (Figure 16-1) and competitively inhibiting interactions with other nuclear factors involved in gene regulation (Figure 16-2).

Another hypothesis, also proposed for the MAN1 paralog NET25 (Huber et al., 2009), is that MAN1 functions as a scaffold protein recruiting Smad2, Smad3 and kinases/phosphatases to the inner nuclear membrane. Phosphorylation of Smad linker residues by kinases such as MAPKs or dephosphorylation of Smad C-terminal residues by phosphatases would then lead to enhanced nuclear export (Figure 16-3).
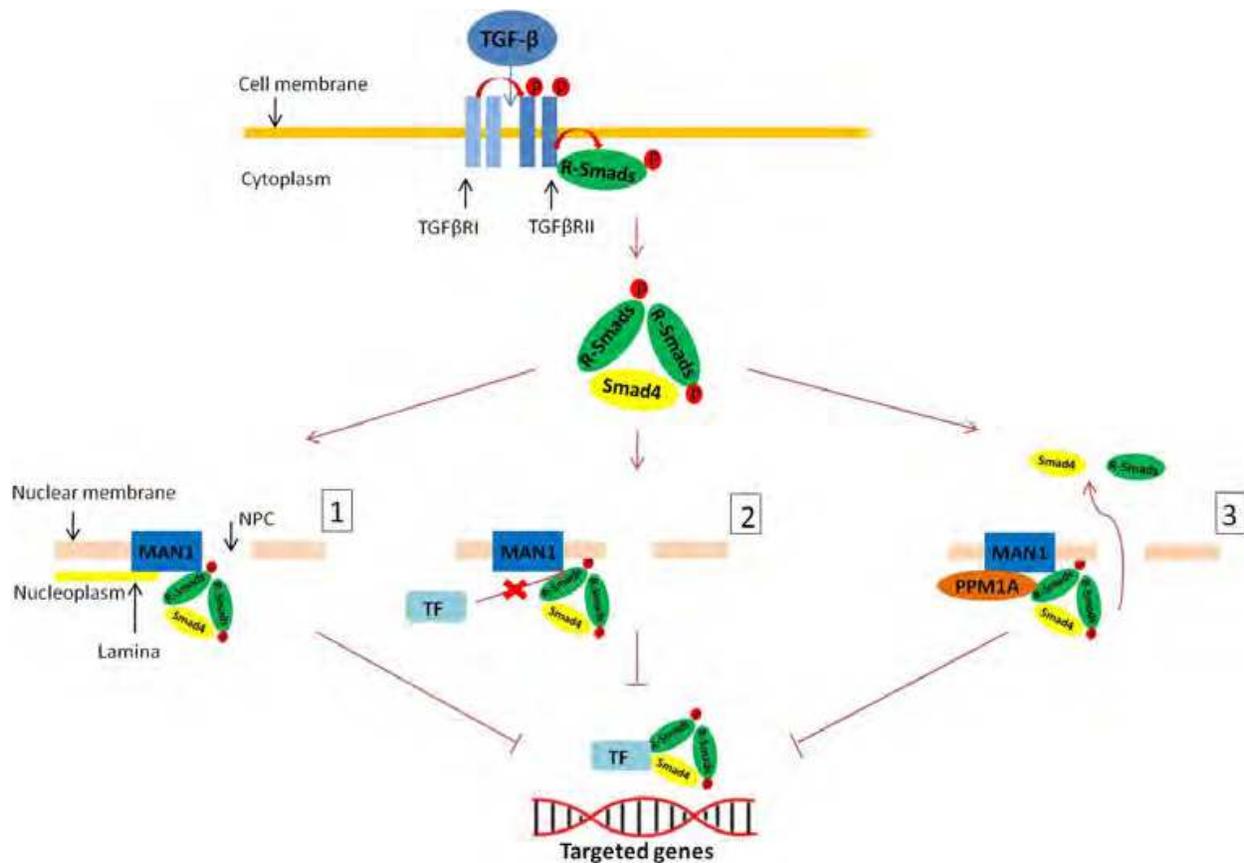

Fig. 16. The role of MAN1 in the TGF-β signaling pathway. The TGF-ß family of cytokines comprises key regulators of metazoan embryonic development and adult tissue homeostasis. In the canonical pathway, ligands of both the TGF-ß and the bone morphogenetic protein members of this family bind to heteromeric serine/threonine kinase receptor complexes, which in turn phosphorylate R-Smad transcription factors at their C-terminal tails. Upon phosphorylation, the R-Smads trimerize. This trimerization allows the R-Smads to interact with the common mediator Smad4. Then, these complexes translocate to the nucleus where they interact with various transcription factors, bind to DNA, and regulate transcription of targeted genes. MAN1 is able to influence the TGF-ß signalling by directly interacting with Smad2 and Smad3 (Hellemans et al., 2004; Lin et al., 2005). This leads to an inhibition of the TGF-ß signalling pathway (Pan et al., 2005). We suggest several putative mechanisms for the TGF-β pathway regulation by MAN1. **(1)** MAN1 could sequestrate the activated complexes R-Smads-Smad4 at the inner nuclear membrane, thus decreasing R-Smads accessibility to its targeted genes. **(2)** MAN1 could also compete with transcription factors for R-Smads binding resulting in a decrease of the transcriptional activity of R-Smads proteins. **(3)** MAN1 could recruit enzymes that modify R-Smads resulting in dissociation of active R-Smads-Smad4 complexes and nuclear export of R-Smads.

We used protein binding assays and isothermal calorimetry experiments to map the Smad2 binding site onto the MAN1 structure (Kondé et al., 2010). These experiments suggested that MAN1 recognizes Smad2 through a motif similar to the Smad Interaction Motif found in transcription factors. MAN1 could thus compete with transcription factors for Smad2 and Smad3 binding. Further experiments are needed to characterize the fate of Smads after MAN1 binding.

6. Conclusion

^{1}H, ^{15}N and ^{13}C NMR spectroscopy provide structural data which, when combined to SAXS data, allow the determination of the 3D structures of partially unfolded proteins. It is now possible to describe not only the structural consequences of disease-linked mutations located in well-folded domains but also their impact on the structural organisation of modular proteins (Mackereth et al., 2011). Such results provide mechanistic clues about how partially folded proteins use conformational switches in order to regulate their functions and highlight the connection between conformational variability, interaction and disease pathogenesis. Regulation of the function of modular proteins is also based on protein modifications. Recent studies propose that NMR characterization of a ^{15}N labelled protein in cell extracts can be a unique tool to examine changes in protein structure and interaction properties upon modification (Liokatis et al., 2010).

Characterization of interaction properties of disease-linked proteins is critical to understanding pathogenic mechanisms and may identify other genes to be screened in similar pathologies. NMR is a powerful tool to map low affinity (mM to μM) interaction surfaces and determine high affinity (nM to pM) complex structures, and must be used in conjunction with other methods such as isothermal calorimetry or fluorescence to deduce the nature, affinity and stochiometry of the interaction. Analysis of these parameters aims at highlighting the competitive or synergic role of the studied interaction within the large cellular protein-protein interaction network, and at identifying how disease-linked mutations perturb this functional network.

All theses different approaches highlight NMR spectroscopy as a technique adapted to the description of the consequences of disease-linked mutations in partially unfolded and multi-partner proteins, as those associated with cellular control and signalling mechanisms. It has become an essential tool for understanding such biological processes from a mechanistic point of view at the atomic level. It can describe critical steps of disease mechanisms, highlight the mechanism of action of a particular drug and inspire the design of binding compounds and complex assembly inhibitors. NMR could be used in the future to develop therapies targeted to the defects.

7. References

Bengtsson L & Wilson KL (2005). Barrier-to-autointegration factor phosphorylation on Ser-4 regulates emerin binding to lamin A in vitro and emerin localization in vivo. *Mol Cell Biol*, Vol. 17, No.3, pp. 1154-63

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. *Nat Genet*, Vol. 8, No.4, pp. 323-7

Bradley CM, Ronning DR, Ghirlando R, Craigie R & Dyda F (2005). Structural basis for DNA bridging by barrier-to-autointegration factor. *Nat Struct Mol Biol*, Vol.12, No.10, pp.935-6

Burz DS, Dutta K, Cowburn D & Shekhtman A (2006). Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR). *Nature Methods*, Vol.3, pp. 91-93

Cady SD, Schmidt-Rohr K, Wang J, Soto CS, Degrado WF & Hong M (2010). Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature*, Vol.443, No.7281, pp. 689-92

Cai M, Huang Y, Ghirlando R, Wilson KL, Craigie R & Clore GM (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*, Vol.20, No.16, pp. 4399-407

Cai M, Huang Y, Suh JY, Louis JM, Ghirlando R, Craigie R & Clore GM (2007). Solution NMR structure of the barrier-to-autointegration factor-Emerin complex. *J Biol Chem*, Vol.282, No.19, pp. 14525-35

Caputo S, Couprie J, Duband-Goulet I, Kondé E, Lin F, Braud S, Gondry M, Gilquin B, Worman HJ & Zinn-Justin S (2006). The carboxyl-terminal nucleoplasmic region of MAN1 exhibits a DNA binding winged helix domain. *J Biol Chem*, Vol.281, No.26, pp. 18208-15

Chi YH, Chen ZJ & Jeang KT (2009). The nuclear envelopathies and human diseases. *J Biomed Sci*, Vol.16, pp. 96

Dauer WT & Worman HJ (2009). The nuclear envelope as a signaling node in development and disease. *Dev Cell*, Vol.17, No.5, pp. 626-38

Dedmon MM, Patel CN, Young GB & Pielak GJ (2002). FlgM gains structure in living cells. *Proc. Natl Acad. Sci. USA*, Vol.99, pp. 12681-12684

De Sandre-Giovannoli A, Bernard R, Cau P, Navarro C, Amiel J, Boccaccio I, Lyonnet S, Stewart CL, Munnich A, Le Merrer M & Lévy N. (2003). Lamin a truncation in Hutchinson-Gilford progeria. *Science*, Vol. 300, No. 5628, pp. 2055.

Dhe-paganon S, Werner ED, Chi YI & Shoelson SE (2002). Structure of the globular tail of nuclear lamin. *J Biol Chem*, Vol.277, No.20, pp. 17381-4

Eriksson M, Brown WT, Gordon LB, Glynn MW, Singer J, Scott L, Erdos MR, Robbins CM, Moses TY, Berglund P, Dutra A, Pak E, Durkin S, Csoka AB, Boehnke M, Glover TW & Collins FS (2003). Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature*, Vol.423, No.6937, pp. 293-8

Fidziańska A & Hausmanowa-Petrusewicz I (2003). Architectural abnormalities in muscle nuclei. Ultrastructural differences between X-linked and autosomal dominant forms of EDMD. *J Neurol Sci*, Vol.210, No.1-2, pp. 47-51

Frock RL, Kudlow BA, Evans AM, Jameson SA, Hauschka SD & Kennedy BK (2006). Lamin A/C and emerin are critical for skeletal muscle satellite cell differentiation. *Genes Dev*, Vol.20, No.4, pp. 486-500

Gautier A, Mott HR, Bostock MJ, Kirkpatrick JP & Nietlispach D (2010). Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. *Nat Struct Mol Biol*, Vol.17, No.6, pp. 768–774

Gonzáles JM, Navarro-Puche A, Casar B, Crespo P & Andrés V (2008). Fast regulation of AP-1 activity throught interaction of lamin A/C, ERK1/2, and c-Fos at the nuclear envelope. *J Cell Biol*, Vol.183, No.4, pp. 653-66

Hellemans J, Preobrazhenska O, Willaert A, Debeer P, Verdonk PC, Costa T, Janssens K, Menten B, Van Roy N, Vermeulen SJ, Savarirayan R, Van Hul W, Vanhoenacker F, Huylebroeck D, De Paepe A, Naeyaert JM, Vandesompele J, Speleman F, Verschueren K, Coucke PJ & Mortier GR (2004). Loss-of-function mutations in LEMD3 result in osteopoikilosis, Buschke-Ollendorff syndrome and melorheostosis. *Nat Genet*, Vol.36, No.11, pp. 1213-8

Herrmann H, Strelkov SV, Burkhard P & Aebi U (2009). Intermediate filaments: primary
        determinants of cell architecture and plasticity. *J Clin Invest*, Vol.119, No.7, pp. 1772-83

Hu F, Luo W & Hong M (2010). Mechanisms of Proton Conduction and Gating in Influenza
        M2 Proton Channels from Solid-State NMR. *Science,* Vol.330, No.6003, pp. 505-508.

Huber MD, Guan T & Gerace L (2009). Overlapping functions of nuclear envelope proteins
        NET25 (Lem2) and emerin in regulation of extracellular signal-regulated kinase
        signalling in myoblast differentiation. *Mol Cell Biol*, Vol.29, No.21, pp. 5715-28

Inomata K, Ohno A, Tochio H, Isogai S, Tenno T, Nakase I, Takeuchi T, Futaki S, Ito Y,
        Hiroaki H & Shirakawa M (2009). High resolution multi-dimensional NMR
        spectroscopy of proteins in human cells. *Nature,* Vol.458, pp. 106-109

Jehle S, Rajagopal P, Bardiaux B, Markovic S, Kühne R, Stout JR, Higman VA, Klevit RE, van
        Rossum BJ & Oschkinat H (2010). Solid-state NMR and SAXS studies provide a
        structural basis for the activation of alphaB-crystallin oligomers. *Nat Struct Mol Biol*,
        Vol.17, No.9, pp. 1037-42

Kapinos LE, Burkhard P, Herrmann H, Aebi U & Strelkov SV (2011). Simultaneous
        formation of right- and left-handed anti-parallel coiled-coil interfaces by a coil2
        fragment of human lamin A. *J Mol Biol*, Vol.408, No.1, pp. 135-46

Kondé E, Bourgeois B, Tellier-Lebegue C, Wu W, Pérez J, Caputo S, Attanda W, Gasparini S,
        Charbonnier JB, Gilquin B, Worman HJ & Zinn-Justin S (2010). Structural analysis
        of the Smad2-MAN1 interaction that regulates transforming growth factor-β
        signaling at the inner nuclear membrane. *Biochemistry,* Vol.49, No.37, pp. 8020-32

Krimm I, Ostlund C, Gilquin B, Couprie J, Hossenlopp P, Mornon JP, Bonne G, Courvalin
        JC, Worman HJ & 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*, Vol.10, No.6, pp. 811-23

Laguri C, Gilquin B, Wolff N, Romi-Lebrun R, Courchay K, Callebault I, Worman HJ &
        Zinn-Justin S (2001). Structural characterization of the LEM motif common to three
        human inner nuclear membrane proteins. *Structure*, Vol.9, No.6, pp. 503-11

Lin F, Blake DL, Callebault I, Skerjanc IS, Holmer L, McBurney MW, Paulin-Levasseur M &
        Worman HJ (2000). MAN1, an inner nuclear membrane protein that shares the
        LEM domain with lamina-associated polypeptide 2 and emerin. *J Biol Chem*,
        Vol.275, No.7, pp. 4840-7

Lin F, Morrison JM, Wu W & Worman HJ (2005). MAN1, an integral protein of the inner
        nuclear membrane, binds Smad2 and Smad3 and antagonizes transforming growth
        factor-beta signaling. *Hum Mol Genet,* Vol.14, No.3, pp. 437-45

Liotakis S, Dose A, Schwarzer D & Selenko P (2010). Simultaneous detection of protein
        phosphorylation and acetylation by high-resolution NMR spectroscopy. *J Am Chem
        Soc*, Vol.132, No.42, pp. 14704-5

Liu J, Lee KK, Segura-Totten M, Neufeld E, Wilson KL & Gruenbaum Y (2003). MAN1 and
        emerin have overlapping function(s) essential for chromosome segregation and cell
        division in Caenorhabditis elegans. *Proc Natl Acad Sci U S A*, Vol.100, No.8, pp. 4598-603

Mackereth C, Madl T, Bonnal S, Simon B, Zanier K, Gasch A, Rybin V, Valcarcel J & Sattler
        M (2011). Multi-domain conformational selection underlies pre-mRNA splicing
        regulation by U2AF. *Nature,* Vol.475, pp. 408-411

Madl T, Bermel W & Zangger K (2009). Use of relaxation enhancements in a paramagnetic
        environment for the structure determination of proteins using NMR spectroscopy.
        Angew Chem Int Ed Engl, Vol.48, No.44, pp. 8259-8262

Madl T, Gabel F & Sattler M (2011). NMR and small-angle scattering-based structural
        analysis of protein complexes in solution. *J Struct Biol,* Vol.173, No.3, pp. 472-82

Mansharamani M & Wilson KL (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*, Vol.280, No.14, pp. 13863-70

Muchir A & Worman HJ (2007). Emery-Dreifuss muscular dystrophy. *Curr Neurol Neurosci Rep*, Vol.7, No.1, pp. 78-83

Pan D, Estévez-Salmerón LD, Stroschein SL, Zhu X, He J, Zhou S & Luo K (2005). The inner nuclear membrane protein MAN1 physically interacts with R-Smad proteins to repress signaling by the transforming growth factor-{beta} superfamily of cytokines. *J Biol Chem*, Vol.280, No.16, pp. 15992-6001

Puente XS, Quesada V, Osorio FG, Cabanillas R, Cadiñanos J, Fraile JM, Ordóñez GR, Puente DA, Gutiérrez-Fernández A, Fanjul- Fernández M, Lévy N, Freije JM & López-Otín C (2011). Exome sequencing and functional analysis identifies BANF1 mutation as the cause of a hereditary progeroid syndrome. *Am J Hum Genet*, Vol.88, No.5, pp. 650-6

Sakakibara D, Sasaki A, Ikeya T, Hamatsu J, Hanashima T, Mishima M, Yoshimasu M, Hayashi N, Mikawa T, Wälchli M, Smith BO, Shirakawa M, Güntert P & Ito Y (2009). Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature,* Vol.458, pp. 102-105

Sattler M, Schleucher J & Griesinger C (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients *Progr Magn Res Spect,* Vol.34, pp. 93-158

Selenko P, Serber Z, Gade B, Ruderman J & Wagner G (2006). Quantitative NMR analysis of the protein G B1 domain in Xenopus laevis egg extracts and intact oocytes. Proc. Natl Acad. Sci. USA, Vol.103, pp. 11904-11909

Sharma M, Yi M, Dong H, Qin H, Peterson E, Busath DD, Zhou HX & Cross TA (2010). Insight into the mechanism of the influenza A proton channel from a structure in a lipid bilayer. *Science,* Vol.330, No.6003, pp. 509-12

Shumaker DK, Lee KK, Tanhehco YC, Craigie R & Wilson KL (2001). LAP2 binds to BAF.DNA complexes: requirement for the LEM domain and modulation by variable regions. *EMBO J*, Vol.20, No.7, pp. 1754-64

Strelkov SV, Schumacher J, Burkhard P, Aebi U & Herrmann H (2004). Crystal structure of the human lamin A coil 2B dimer: implications for the head-to-tail association of nuclear lamins. *J Mol Biol*, Vol.343, No.4, pp. 1067-80

Tycko R (2011). Solid-State NMR studies of amyloid fibril structure. *Annu Rev Phys Chem*, Vol.62, pp. 279-299

Van Melckebeke H, Wasmer C, Lange A, Ab E, Loquet A, Böckmann A & Meier BH (2010). Atomic-resolution three-dimensional structure of HET-s(218-289) amyloid fibrils by solid-state NMR spectroscopy. *J Am Chem Soc*, Vol.132, No.39, pp. 13765-75

Verstraeten VL, Broers JL, van Steensel MA, Zinn-Justin S, Ramaekers FC, Steijlen PM, Kamps M, Kuijpers HJ, Merckx D, Smeets HJ, Hennekam RC, Marcelis CL & Van den Wijngaard A (2006). Compound heterozygosity for mutations in LMNA causes a progeria syndrome without prelamin A accumulation. *Hum Mol Genet,* Vol.15, No.16, pp. 2509-22

Wolff N, Gilquin B, Courchay K, Callebault I, Worman HJ & Zinn-Justin S (2001). Structural analysis of emerin, an inner nuclear membrane protein mutated in X-linked Emery-Dreifuss muscular dystrophy. *FEBS let*, Vol.501, No.2-3, pp. 171-6

Worman HJ, Ostlund C & Wang Y (2010). Diseases of the nuclear envelope. *Cold Spring Harb Perspect Biol*, Vol.2, No.2

Wright PE & Dyson HJ (2009). Linking folding and binding. *Curr Opin Struct Biol,* Vol.19, No.1, pp. 31-8
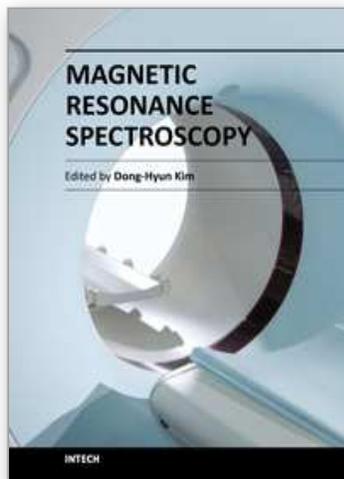
Magnetic Resonance Spectroscopy

Edited by Prof. Dong-Hyun Kim

Magnetic Resonance Spectroscopy (MRS) is a unique tool to probe the biochemistry in vivo providing metabolic information non-invasively. Applications using MRS has been found over a broad spectrum in investigating the underlying structures of compounds as well as in determining disease states. In this book, topics of MRS both relevant to the clinic and also those that are beyond the clinical arena are covered. The book consists of two sections. The first section is entitled 'MRS inside the clinic' and is focused on clinical applications of MRS while the second section is entitled 'MRS beyond the clinic' and discusses applications of MRS in other academic fields. Our hope is that through this book, readers can understand the broad applications that NMR and MRS can offer and also that there are enough references to guide the readers for further study in this important topic.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Benjamin Bourgeois, Howard J. Worman and Sophie Zinn-Justin (2012). NMR Spectroscopy as a Tool to Provide Mechanistic Clues About Protein Function and Disease Pathogenesis, Magnetic Resonance Spectroscopy, Prof. Dong-Hyun Kim (Ed.), ISBN: 978-953-51-0065-2, InTech, Available from: http://www.intechopen.com/books/magnetic-resonance-spectroscopy/nmr-spectroscopy-as-a-tool-to-provide-mechanistic-clues-about-protein-function-and-disease-pathogene

INTECH
open science | open minds