



# Assessing induced folding within the intrinsically disordered C-terminal domain of the Henipavirus nucleoproteins by site directed spin labeling EPR spectroscopy

Marlène Martinho, Johnny Habchi, Zeina El Habre, Léo Nesme, Bruno Guigliarelli, Valérie Belle, Sonia Longhi

## ► To cite this version:

Marlène Martinho, Johnny Habchi, Zeina El Habre, Léo Nesme, Bruno Guigliarelli, et al.. Assessing induced folding within the intrinsically disordered C-terminal domain of the Henipavirus nucleoproteins by site directed spin labeling EPR spectroscopy. *Journal of Biomolecular Structure and Dynamics*, 2013, 31 (5), pp.453-471. 10.1080/07391102.2012.706068 . hal-03555631

**HAL Id: hal-03555631**

**<https://hal.science/hal-03555631>**

Submitted on 3 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Assessing induced folding within the intrinsically disordered C-terminal domain of the  
*Henipavirus* nucleoproteins by site directed spin labeling EPR spectroscopy**

Marlène Martinho<sup>§a</sup> and Johnny Habchi<sup>§b</sup>, Zeina El Habre<sup>b</sup>, Léo Nesme<sup>a</sup>, Bruno Guigliarelli<sup>a</sup>, Valérie Belle<sup>a\*</sup> and Sonia Longhi<sup>b\*</sup>

<sup>a</sup>CNRS, UMR 7281, 13402, Marseille, France, Aix-Marseille Univ, Bioénergétique et Ingénierie des Protéines (BIP), 13402, Marseille, France

<sup>b</sup>CNRS, UMR 7257, 13288, Marseille, France, Aix-Marseille Univ, Architecture et Fonction des Macromolécules Biologiques (AFMB), 13288, Marseille, France

<sup>§</sup>These authors have equally contributed to the work

\* To whom correspondence should be addressed

Sonia Longhi

AFMB, UMR 7257 CNRS and Aix-Marseille University

163, avenue de Luminy, Case 932, 13288 Marseille Cedex 09, France

Tel: (33) 4 91 82 55 80; Fax: (33) 4 91 26 67 20

E-mail [Sonia.Longhi@afmb.univ-mrs.fr](mailto:Sonia.Longhi@afmb.univ-mrs.fr)

Valérie Belle

BIP, UMR 7281 CNRS and Aix-Marseille University

31, Chemin Joseph Aiguier, 13402 Marseille Cedex, France

Tel: (33) 4 91 16 45 57; Fax: (33) 4 91 16 45 78

E-mail [belle@imm.cnrs.fr](mailto:belle@imm.cnrs.fr)

**Running title.** Mapping induced folding by site directed spin labeling EPR spectroscopy

## ABSTRACT

This work aims at characterizing structural transitions within the intrinsically disordered C-terminal domain of the nucleoprotein ( $N_{TAIL}$ ) from the Nipah and Hendra viruses, two recently emerged pathogens gathered within the *Henipavirus* genus. To this end, we used site directed spin labeling combined with EPR spectroscopy to investigate the  $\alpha$ -helical induced folding that *Henipavirus*  $N_{TAIL}$  domains undergo in the presence of the C-terminal X domain of the phosphoprotein ( $P_{XD}$ ). For each  $N_{TAIL}$  protein, six positions located within four previously proposed molecular recognition elements (MoREs) were targeted for spin labeling, with three of these (positions 475, 481, 487) falling within the MoRE responsible for binding to  $P_{XD}$  (Box3). A detailed analysis of the impact of the partner protein on the labeled  $N_{TAIL}$  variants revealed a dramatic modification in the environment of the spin labels grafted within Box3, with the observed modifications supporting the formation of an induced  $\alpha$ -helix within this region. In the free state, the slightly lower mobility of the spin labels grafted within Box3 as compared to the other positions suggests the existence of a transiently populated  $\alpha$ -helix, as already reported for measles virus (MeV)  $N_{TAIL}$ . Comparison with the well-characterized MeV  $N_{TAIL}$ - $P_{XD}$  system, allowed us to validate the structural models of *Henipavirus*  $N_{TAIL}$ - $P_{XD}$  complexes that we previously proposed. In addition, this study highlighted a few notable differences between the Nipah and Hendra viruses. In particular, the observation of composite spectra for the free form of the Nipah virus  $N_{TAIL}$  variants spin labeled in Box3 supports conformational heterogeneity of this partly pre-configured  $\alpha$ -helix, with the pre-existence of stable  $\alpha$ -helical segments. Altogether these results provide insights into the molecular mechanisms of the *Henipavirus*  $N_{TAIL}$ - $P_{XD}$  binding reaction.

## Supplementary Material

Supplementary material dealing with Figures S1 to S5 can be directly downloaded from the JBSD site.

**Key words.** Intrinsic disorder, induced folding, spin labeling, EPR spectroscopy, Nipah, Hendra, nucleoprotein, phosphoprotein.

**<sup>1</sup>Abbreviations.** IDP, intrinsically disordered protein; HeV, Hendra virus; NiV, Nipah virus; MeV, measles virus; N, nucleoprotein; P, phosphoprotein;  $P_{XD}$ , X domain of P; MoRE, Molecular Recognition Element; SDSL, site-directed spin-labeling, EPR, Electron Paramagnetic Resonance; MTSL, (1-oxyl-2,2,5,5-tetramethyl- $\delta$ 3-pyrroline-3-methyl methanethiosulfonate); TFE, 2,2,2-trifluoroethanol; PCR, polymerase chain reaction; SEC-MALLS, size exclusion chromatography coupled to multiple angle laser light scattering; CD, circular dichroism; NMR, nuclear magnetic resonance.

## INTRODUCTION

The Hendra (HeV) and Nipah (NiV) viruses are recently emerged, severe human pathogens within the *Paramyxoviridae* family (Eaton, Mackenzie & Wang 2007). A few distinctive properties set NiV and HeV aside from other paramyxoviruses and led to their classification within the *Henipavirus* genus of the *Paramyxoviridae* family (Wang et al. 2000). The genome of Henipaviruses is encapsidated by the nucleoprotein (N) within a helical nucleocapsid. This helical N:RNA complex, and not naked RNA, is the substrate used by the polymerase complex for both transcription and replication. By analogy with other *Paramyxoviridae* members, the polymerase complex is thought to consist of the L protein and the P protein, with this latter serving as a tether for the recruitment of L onto the nucleocapsid template. *Henipavirus* N and P proteins were shown to interact with each other (Habchi et al. 2011), being able to form both homologous and heterologous N-P complexes (Chan et al. 2004; Omi-Furutani et al. 2010; Blocquel et al. 2012).

So far, the only molecular data available on the *Henipavirus* N and P proteins come from our previous studies (Habchi et al. 2010; Habchi, Blangy et al. 2011; Blocquel, Habchi et al. 2012). Those studies showed that the N-terminal region of P (PNT) and the C-terminal region of N (N<sub>TAIL</sub>, aa 400-532) are intrinsically disordered, although they both contain short order-prone segments (Habchi, Mamelli et al. 2010). The occurrence of some residual, transient secondary and/or tertiary structure within N<sub>TAIL</sub> and PNT led to their classification within the premolten globules subfamily within the family of intrinsically disordered proteins (IDPs) (Dunker et al. 2001; Dunker & Obradovic 2001; Uversky 2002; Uversky 2003; Dyson & Wright 2005; Radivojac et al. 2007). IDPs are ubiquitous, functional proteins that lack highly populated secondary and tertiary structure under physiological conditions and in the absence of a partner/ligand (for recent reviews on IDPs, see (Dunker et al. 2008; Dunker et al. 2008; Turoverov, Kuznetsova & Uversky 2010; Uversky 2010; Uversky & Dunker 2010; Chouard 2011)). Although there are IDPs that carry out their function while remaining permanently disordered (e.g. entropic chains) (Dunker, Lawson et al. 2001), many of them undergo induced folding, i.e. a disorder-to-order transition upon binding to their physiological partner(s) (Kriwacki et al. 1996; Dyson & Wright 2002; Uversky 2002; Fuxreiter et al. 2004; Dunker et al. 2005; Uversky, Oldfield & Dunker 2005).

By contrast, the C-terminal X domain of the *Henipavirus* P proteins (P<sub>XD</sub>, aa 660-709 of NiV P and aa 657-707 of HeV P) was found to be folded, and to adopt an  $\alpha$ -helical conformation (Habchi, Blangy et al. 2011). CD and heteronuclear NMR studies showed that binding to P<sub>XD</sub> triggers  $\alpha$ -helical folding of *Henipavirus* N<sub>TAIL</sub> (Habchi, Blangy et al. 2011). Using bioinformatics approaches, we identified within both HeV and NiV N<sub>TAIL</sub> domains four putative molecular recognition elements (MoREs), where these latter are short order-prone regions within intrinsically disordered regions with a propensity to undergo induced folding upon binding to a partner (Oldfield et al. 2005; Mohan et al. 2006; Fuxreiter, Tompa & Simon 2007; Vacic et al. 2007). The four predicted MoREs are referred to as Box1 (aa 408-422), Box2 (aa 444-464), Box3 (aa 473-493) and Box4 (aa 523-532) (**Fig. 1A**). The behavior of the N<sub>TAIL</sub> resonances in NMR titration studies with unlabeled P<sub>XD</sub> suggested a role for Box3 in binding to P<sub>XD</sub> (Habchi, Blangy et al. 2011). Based on these observations, and also by analogy with the related measles virus (MeV) (Johansson et al. 2003; Kingston et al. 2004), we proposed a

structural model of the *Henipavirus* N<sub>TAIL</sub>-P<sub>XD</sub> interaction where the  $\alpha$ -More located within Box3 adopts an  $\alpha$ -helical conformation and is embedded between helices  $\alpha$ 2 and  $\alpha$ 3 of P<sub>XD</sub> (Habchi, Blangy et al. 2011). Subsequent deletion studies that made use of truncated forms of both HeV and NiV N<sub>TAIL</sub> proteins, showed that Box3 is strictly required for binding to P<sub>XD</sub> and provided direct experimental evidence for the structural state of the four MoREs (Blocquel, Habchi et al. 2012). In particular, CD studies in the presence of trifluoroethanol (TFE) showed that while Box1 and Box4 neither confer  $\alpha$ -helicity nor contribute to  $\alpha$ -helical folding of N<sub>TAIL</sub>, Box3, and to a minor extent Box2, has clear  $\alpha$ -helical propensities. Although all those studies significantly improved our knowledge of *Henipavirus* N and P proteins, as well as our understanding of the interactions they establish (for reviews see (Habchi & Longhi 2012; Habchi, Mamelli & Longhi 2012)), only partial structural information on the *Henipavirus* N<sub>TAIL</sub>-XD complex is presently available. Indeed, NMR studies did not provide full structural information because of the lack of assignment of the N<sub>TAIL</sub> HSQC spectrum, and CD studies only provided global structural information. The objective of the present study was to obtain further insights into the induced folding that *Henipavirus* N<sub>TAIL</sub> undergoes upon binding to P<sub>XD</sub> and to assess the possible contribution of other short order-prone N<sub>TAIL</sub> segments in the establishment of tertiary contacts in both the free and bound form of N<sub>TAIL</sub>. To achieve this goal, we used site directed spin labeling (SDSL) coupled to electron paramagnetic resonance (EPR) spectroscopy.

SDSL EPR spectroscopy has emerged as a valuable tool for studying structural characteristics in a wide range of proteins (Hubbell et al. 1996; Feix & Klug 1998; Biswas et al. 2001; Hubbell et al. 2003; Morin et al. 2006; Belle et al. 2008). In particular, it has been shown that SDSL EPR spectroscopy can significantly contribute to the characterization of IDPs, since their flexibility hinders structure determination by X-ray crystallography (Fanucci & Cafiso 2006; Eralles et al. 2009; Klare & Steinhoff 2009; Belle et al. 2010; Drescher 2011; Longhi et al. 2011; Pirman et al. 2011). SDSL EPR spectroscopy relies on the introduction of a paramagnetic spin label through covalent modification of a unique sulfhydryl group using a selective nitroxide reagent, and on the ensuing monitoring of EPR spectra whose shape reflects the mobility of the spin label. This approach requires the presence of a cysteine residue at a desired position in polypeptide chain but alternative strategy to label other kind of residue are currently in progress (Lorenzi et al. 2011). One of the advantages of this technique is that the label is a sensitive reporter of its local environment thus providing information at the residue level, while introducing minimal perturbation of the system. We have previously reported studies where the induced folding of an intrinsically disordered protein domain, namely the N<sub>TAIL</sub> domain from the closely related MeV, could be mapped using SDSL EPR spectroscopy (Morin, Bourhis et al. 2006; Belle, Rouger et al. 2008; Kavalenka et al. 2010).

In the present study, for each HeV and NiV N<sub>TAIL</sub> protein, we targeted for spin-labeling six different sites, three of which (positions 475, 481 and 487) are located within the  $\alpha$ -MoRE shown to be involved in binding to P<sub>XD</sub> (i.e. Box3) while the other three (positions 421, 458 and 525) are located within the other three predicted MoREs (see **Fig. 1A**). For both viruses, while the mobility of the spin labels grafted within Box1, Box2 and Box4 was **weakly** affected by the addition of P<sub>XD</sub>, that of the spin labels grafted within Box3 was severely affected. A detailed analysis of the impact of the partner protein and a comparison with the well-characterized MeV N<sub>TAIL</sub>-P<sub>XD</sub> complex allowed us to validate

the previously proposed structural models of the *Henipavirus* N<sub>TAIL</sub>-P<sub>XD</sub> complex and to point out some differences between NiV and HeV. Altogether these results provided insights into the molecular mechanisms of the *Henipavirus* N<sub>TAIL</sub>-P<sub>XD</sub> binding.

## MATERIALS AND METHODS

### **Construction of protein expression plasmids**

The HeV and NiV pDest14/N<sub>TAILHN</sub> constructs, encoding residues 400-532 of *Henipavirus* N with a hexahistidine tag fused to their N- termini, and the pDest14/P<sub>XD</sub> constructs, encoding residues 660-709 (NiV) or 657-707 (HeV) of P with a hexahistidine tag fused to their C-termini, have already been described (Habchi, Mamelli et al. 2010; Habchi, Blangy et al. 2011).

All N<sub>TAIL</sub> mutated constructs were obtained by PCR (polymerase chain reaction) using the *Henipavirus* pDest14/N<sub>TAILHN</sub> plasmids as template (Habchi, Mamelli et al. 2010). Each mutant construct was obtained using *Turbo-Pfu* polymerase (Stratagene) and a pair of complementary mutagenic primers of 39 nucleotides in length (Operon) designed to introduce a Cys codon at the desired position. PCR cycles were carried out according to the supplier's instructions. After digestion with *DpnI* to remove the parental, methylated DNA template, transformation of CaCl<sub>2</sub>-competent *E. coli* TAM1 cells (Active Motif) with the amplified PCR product was carried out. The sequence of the coding region of all expression plasmids was verified by sequencing (GATC Biotech) and found to conform to expectations.

### **Expression of N<sub>TAIL</sub> constructs**

The *E. coli* Rosetta [DE3] pLysS strain (Novagen) was used for the expression of the N<sub>TAIL</sub> variants. This strain, which is optimized for the expression of recombinant proteins, also carries the lysozyme gene thus allowing a tight regulation of the expression of the recombinant gene, as well as a facilitated lysis. Cultures were grown overnight to saturation in 2YT medium containing 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. An aliquot of the overnight culture was diluted 1/25 in 2YT medium and grown at 37°C. When the optical density at 600 nm reached 0.6-0.8, isopropyl β-D-thiogalactopyranoside was added to a final concentration of 0.2 mM, and the cells were grown at 37°C for 3 additional hours. The induced cells were harvested, washed and collected by centrifugation (5,000 g, 10 min). The resulting pellets were frozen at -20°C. Expression of tagged HeV and NiV P<sub>XD</sub> was carried out as described in (Habchi, Blangy et al. 2011).

### **Purification of N<sub>TAIL</sub> variants and of P<sub>XD</sub> proteins and characterization of the synthetic NiV P<sub>XD</sub> sample**

The *Henipavirus* N<sub>TAIL</sub> cysteine variants were purified from the soluble fraction of bacterial lysates using the protocols already described for *Henipavirus* parental N<sub>TAIL</sub> proteins (Habchi, Mamelli et al. 2010). The recombinant HeV and NiV P<sub>XD</sub> were purified as described in (Habchi, Blangy et al. 2011).

In the case of the NiV P<sub>XD</sub> sample, we noticed an unexpected variability from lot to lot in terms of impact on the spin label mobility of the various NiV N<sub>TAIL</sub> variants (data not shown). To circumvent this problem, we purchased (China Peptides) a synthetic NiV P<sub>XD</sub> sample (hereafter referred to as NiV P<sub>XD</sub><sup>S</sup>) endowed with exactly the same amino acid sequence as the recombinant sample (referred to as P<sub>XD</sub><sup>R</sup>). The lyophilized peptide (57 residues, 10 mg) was resuspended in 10 mM Tris/HCl pH 7.5, 200 mM NaCl buffer at 10 mg/mL and then stored at –20°C. HPLC analysis carried out by the supplier (data not shown) showed the presence of a single peak, and ESI-MS analysis (data not shown) yielded a molecular mass (6,740 Da) very close to the expected one (6,733 Da) and identical to the value previously found in MALDI-TOF analyses of the recombinant NiV P<sub>XD</sub> sample (Habchi, Blangy et al. 2011). In agreement, SDS-PAGE analysis showed the presence of a single protein band with the expected molecular mass (see **Fig. S1A**). Notably, the purity of the recombinant sample was comparable to that of the NiV P<sub>XD</sub><sup>S</sup> sample (see **Fig. S1A**), thus ruling out the possibility that protein contaminants could be responsible for the lack of reproducibility observed with the NiV P<sub>XD</sub><sup>R</sup> sample in EPR spectroscopy studies.

Strikingly, the NiV P<sub>XD</sub><sup>S</sup> sample was found to be monomeric, as judged from SEC-MALLS analyses that showed that the protein is eluted in a single peak, with a molecular mass of 6,667 ± 333 Da (**Fig. S1B**). The sharpness and symmetry of the peak supports the presence of a well-defined molecular species, whose measured Stokes radius is 12.2 ± 1.8 Å, a value close to the theoretical value (14.4 Å) expected for a folded, monomeric protein (Uversky 1993). For comparative purposes we have compared this profile (obtained upon injection of a NiV P<sub>XD</sub><sup>S</sup> sample at 1.48 mM) with the one obtained by loading the column with a NiV P<sub>XD</sub><sup>R</sup> sample at 1.2 mM. As already observed (Habchi, Blangy et al. 2011), this latter was found to be predominantly trimeric although a minor peak, eluted at 9.6 mL, was also detected (see **Fig. S1B**). Consistent with the trimeric state of NiV P<sub>XD</sub><sup>R</sup>, comparison between the recombinant and synthetic NiV P<sub>XD</sub> sample by native PAGE confirmed that the NiV P<sub>XD</sub><sup>R</sup> sample migrates much more slowly than NiV P<sub>XD</sub><sup>S</sup> (see **Fig. S1A**). Finally, to directly ascertain the overall foldedness of the synthetic NiV P<sub>XD</sub> sample, we recorded its far-UV CD spectrum and compared it to that of the recombinant sample (see **Fig. S1C**). The two spectra are almost superimposable (see **Fig. S1C**), indicating that the two samples possess a quite similar content in secondary structure.

All proteins were concentrated using Centricon Plus-20 (molecular cutoff of 3,000 Da for P<sub>XD</sub>, and of 5,000 Da for N<sub>TAIL</sub>) (Millipore), and were stored at –20°C. All purification steps, except for gel filtrations, were carried out at 4°C. Apparent molecular mass of proteins eluted from SEC columns was deduced from a calibration curve established with Low Molecular Weight (LMW) and High Molecular Weight (HMW) calibration kits (GE Healthcare). The theoretical Stokes radii of native, fully unfolded and premolten globule forms were calculated according to (Uversky 1993).

Protein concentrations were estimated via quantitative analysis of amino acid composition, since estimations based on the theoretical absorption coefficients at 280 nm turned out to be not fully reliable (Habchi, Mamelli et al. 2010; Habchi, Blangy et al. 2011). In those analyses, protein samples were dried and hydrolyzed at 110°C in constant-boiling HCl containing 1% (v/v) phenol for 24 h under reduced pressure and in the absence of oxygen. Amino acids were analyzed on a model Biochrom 30

amino acid analyzer, with the standard sodium citrate eluting buffer system. Calibrations were made with standard solutions that contained all the amino acid except tryptophan. The amino acids concentrations were estimated through reaction with ninhydrine. The internal standard was norleucine. Once the actual protein concentration of stock solutions has been determined in that manner, the corresponding experimental absorption coefficients at 254 nm ( $N_{TAIL}$ ) or 280 nm ( $P_{XD}$ ) could be derived. Indeed, since both  $N_{TAIL}$  proteins are devoid of Trp and Tyr residues, the absorbance was measured at 254 nm.

### ***Spin-labeling***

Before spin-labeling, 20 mM DTT was added to each purified  $N_{TAIL}$  mutated protein (approximately 5 mg). The mixture was incubated for 30 min in an ice bath in order to reduce the unique free cysteine residue. DTT was removed by gel filtration onto a PD-10 column (GE Healthcare) with 10 mM MES, 150 mM NaCl at pH 6.5 as elution buffer. The fractions containing the protein were pooled. The spin label, 1-oxy-2,2,5,5-tetramethyl- $\delta^3$ -pyrroline-3-methyl methanethiosulfonate (MTSL, Toronto Research Chemicals Inc., Toronto, Canada), was immediately added to the sample at a molar excess of 10 using a stock solution at 10 mg/mL in acetonitrile. The reaction was carried out during one hour in the dark and in an ice bath, under gentle stirring and a continuous flow of argon to avoid oxidation. The excess of unbound spin label was removed by gel filtration as described above, except that 10 mM Tris/HCl, 200 mM NaCl at pH 8.0 was used as elution buffer. The fractions containing the labeled protein were pooled and concentrated by ultrafiltration using Centricon Plus-20 (molecular cutoff: 5000 Da) (Millipore). The spin-labeled proteins were stored at  $-80^{\circ}\text{C}$ .

### ***Analytical SEC combined with on-line multi-angle laser light scattering (MALLS) and refractometry (RI)***

Analytical SEC was carried out on a high pressure liquid chromatography (HPLC) system (Alliance 2695, Waters) using silica-based columns (Shodex). A 15-mL KW-802.5 column was used for the characterization of the recombinant and synthetic NiV  $P_{XD}$  samples. Proteins were eluted with a 50 mM Tris/HCl pH 7.3, 200 mM NaCl buffer at a flow of 0.5 mL/min. Separation was performed at room temperature. 30  $\mu\text{L}$  of a NiV  $P_{XD}$  sample either at 1.2 mM (recombinant  $P_{XD}$ ) or at 1.48 mM (synthetic  $P_{XD}$ ) were injected. Detection was performed using a triple-angle light-scattering detector (MiniDAWN<sup>TM</sup> TREOS, Wyatt Technology), a quasi-elastic light-scattering instrument (Dynapro<sup>TM</sup>, Wyatt Technology) and a differential refractometer (Optilab<sup>®</sup> rEX, Wyatt Technology). Molecular mass and Stokes radius determination was performed by the ASTRA V software (Wyatt Technology) using a  $dn/dc$  value of 0.185 mL/g. The column was calibrated with proteins of known Stokes radii and molecular masses.

### ***Far-UV circular dichroism (CD)***

Circular dichroism spectra were recorded on a Jasco 810 dichrograph using 0.1 cm thick quartz cells in 10 mM sodium phosphate pH 7 at  $20^{\circ}\text{C}$ . Protein concentrations of 0.1 mg/mL were used. Structural variations of  $N_{TAIL}$  proteins were measured as a function of changes in the initial CD



spectrum upon addition of 20% TFE. CD spectra were measured between 190 and 260 nm, at 0.2 nm/min and were averaged from three scans. The spectra were corrected for buffer signal and smoothed by using a third-order least square polynomial fit. Finally, mean ellipticity values per residue ( $[\Theta]$ ) were calculated as  $[\Theta] = 3300 \text{ m} \Delta A / (l C n)$ , where  $l$  (path length) = 0.1 cm,  $n$  = number of residues,  $m$  = molecular mass in Daltons and  $C$  = protein concentration expressed in mg/mL. Number of residues ( $n$ ) is 140 for all  $N_{\text{TAIL}}$  proteins and 57 for NiV  $P_{\text{XD}}$ , while  $m$  is approximately 15,000 for all  $N_{\text{TAIL}}$  variants and 6,700 Da for NiV  $P_{\text{XD}}$ .

### ***EPR spectroscopy***

Room temperature ( $296 \pm 1\text{K}$ ) EPR spectra were recorded on an ESP 300E Bruker spectrometer equipped with an ELEXSYS Super High Sensitivity resonator operating at X-band (9.9 GHz). The microwave power was 10 mW, the magnetic field modulation amplitude was optimized at 0.1 or 0.2 mT to avoid over-modulation of the signal and the frequency modulation was 100 kHz. The concentration of spin labels was evaluated by double integration of the EPR signal recorded under non-saturating conditions and comparison with that given by a MTSL standard sample of known concentration. The number of spins *per* protein was estimated by calculating the ratio between the concentration of spin labels and the total protein concentration, where this latter was estimated from the optical density at 254 nm and from the experimental absorption coefficients obtained from the amino acid composition analysis.

EPR spectra of all labeled  $N_{\text{TAIL}}$  variants were recorded both in the absence and in the presence of 20% (v/v) TFE. The possibility that TFE could affect the mobility of the free radical in solution has already been checked and ruled out in our previous studies, based on the comparison of EPR spectra obtained with a 40  $\mu\text{M}$  MTSL solution in the presence or absence of 40% TFE (Belle, Rouger et al. 2008). For all spin labeled variants, EPR spectra were also recorded in the presence of 30% sucrose (v/v), a condition where the contribution of protein rotation to the EPR spectral line shape is cancelled (Timofeev & Tsetlin 1983; McHaourab et al. 1996), with or without the homologous partner. For each labeled variant, we recorded spectra with an increasing excess of the homologous  $P_{\text{XD}}$  so as to ensure that saturation was achieved. Under saturating conditions, the extent of reduction of the mobility of the spin label is maximal, and further increasing the molar excess of  $P_{\text{XD}}$  does not result in any further variation in the radical mobility.

### ***EPR data analysis***

Spectral analysis was carried out on spectra recorded from mixtures containing saturating amounts of either a recombinant (HeV) or a synthetic (NiV)  $P_{\text{XD}}$  sample. As a first step, we described the EPR spectral shape by measuring the peak-to-peak amplitude of the low- and central-field lines, referred to as  $h(+1)/h(0)$ . We previously showed that this latter is a sensitive parameter of spin label mobility in the specific regime of mobility encountered in the study of disordered proteins that are characterized by a fast, anisotropic rotational movement of the spin label (Morin, Bourhis et al. 2006; Belle, Rouger et al. 2008). All the EPR spectra recorded in 30% sucrose were simulated using the EPRSIM-C software program. This program, kindly provided by Dr. J. Strancar (University of Ljubljana,

Slovenia), is based on the so-called motional-restricted fast-motion approximation described in details in (Strancar et al. 2005). Briefly, the partial averaging of the hyperfine and  $g$  tensors is described in this model by the following parameters: an effective rotational correlation time  $\tau_c$  and two angles ( $\theta_0$  and  $\varphi_0$ ) corresponding respectively to the amplitude and the anisotropy of the spin label rotational motion within a cone. The normalization of the two last parameters by  $\Omega = (\theta_0 \varphi_0) / (\pi/2)^2$  represents the free rotational space, which varies from zero (totally restricted movement) to 1 (totally unrestricted movement). Two other parameters are required for simulation: a residual width ( $\omega$ ) and a scalar parameter linked to the polarity of the probe environment ( $p_A$ ) for the adjustment of the hyperfine tensor principle values. The calculation of the spin label rotamers attached at different positions of the crystal structure of the MeV N<sub>TAIL</sub>-P<sub>XD</sub> chimera (pdb code 1T60) (Kingston, Hamel et al. 2004) and of the modeled HeV and NiV N<sub>TAIL</sub>-P<sub>XD</sub> complexes (Habchi, Blangy et al. 2011) was performed using the Matlab package MMM (Multiscale Modeling of Macromolecules <http://www.epr.ethz.ch/software/index>) based on a rotamer library approach (Polyhach, Bordinon & Jeschke 2011).

## RESULTS AND DISCUSSION

### ***Rational design of $N_{TAIL}$ cysteine-substituted variants***

Previous computational studies pointed out the occurrence within *Henipavirus*  $N_{TAIL}$  of four regions enriched in hydrophobic clusters, corresponding to putative MoREs (Habchi, Mamelli et al. 2010), with the MoRE located within Box3 having been shown to be required for binding to  $P_{XD}$  (Blocquel, Habchi et al. 2012).

In order to investigate at the residue level the  $P_{XD}$ -induced folding of Box3 and to check for the possible implication of other boxes in the establishment of long/short-range intramolecular tertiary contacts in either the free or bound form, we herein targeted for spin-labeling six different sites within both HeV and NiV  $N_{TAIL}$  proteins, three located in Box1, Box2 and Box4 respectively (positions 421, 458 and 525) and three located within Box3 (positions 475, 481 and 487) (**Fig. 1A**).

Whenever possible, serine to cysteine substitutions were done. When no native serine residue occurred in the MoRE of interest, as was the case of Box1 and Box4, leucine residues were targeted (**Fig. 1A**). In addition, we targeted for substitution residues that were conserved in the two viruses, so as to introduce similar structural perturbations (if any), thereby allowing meaningful comparisons between the two viruses. Isosteric substitution was not however the sole criterion we used to target the various positions. In fact, in the case of Box3, as no serine residue occurs downstream Ser 481, we targeted for mutagenesis Ala 487 so as to introduce minimal perturbation (**Fig. 1A**).

### ***Expression, purification and spin labeling of $N_{TAIL}$ variants***

All recombinant cysteine  $N_{TAIL}$  variants were expressed in *E. coli* and were recovered from the soluble fraction of the bacterial lysates (data not shown). The  $N_{TAIL}$  mutated proteins were purified to homogeneity (> 95%) in two steps: immobilized metal affinity chromatography and preparative size exclusion chromatography (SEC). For the majority of  $N_{TAIL}$  variants, two distinct peaks were obtained in SEC: one corresponding to the peak obtained with parental *Henipavirus*  $N_{TAIL}$  proteins, and one corresponding to a dimeric form of the protein, as judged from SDS-PAGE analysis under reducing conditions (data not shown).

As shown in **Fig. 1B**, all the  $N_{TAIL}$  mutated proteins migrate in SDS-PAGE with an apparent molecular mass of about 20 kDa (expected molecular mass is approximately 15 kDa). This abnormal migratory behavior has already been documented for both parental (i.e. native) HeV and NiV  $N_{TAIL}$ , where mass spectrometry analysis gave the expected results (Habchi, Mamelli et al. 2010). This anomalous electrophoretic mobility is frequently observed in IDPs and is to be ascribed to a relatively high content in acidic residues (Tomba 2002), and/or to a high degree of protein extension in solution (Blocquel, Habchi et al. 2012). Likewise, the behavior of the mutated  $N_{TAIL}$  proteins can be accounted for by this sequence bias composition. Note that unless 20 mM DTT was added to the samples prior to electrophoretic migration, a protein band of approximately 36 kDa, corresponding to a dimeric, disulfide-bridged form, was also observed in SDS-PAGE for almost all variants (data not shown).

The Stokes radius value of the monomeric form of all *Henipavirus*  $N_{TAIL}$  variants ( $28 \pm 2$  Å), as inferred from the apparent molecular mass observed in SEC, is the same as that observed for both

parental HeV and NiV N<sub>TAIL</sub> proteins (data not shown). This value is consistent with the value expected for IDPs adopting a premolten globule state (Uversky 2002). Thus, the cysteine variants share similar hydrodynamic properties with parental N<sub>TAIL</sub>, being all non-globular while possessing a certain residual compactness typical of the premolten globule state (Uversky 1993).

The purified N<sub>TAIL</sub> variants were successively spin labeled through a two-step procedure consisting of a DTT reduction followed by covalent modification of the sulfhydryl group by the MTSL nitroxide derivative. Labeling yields ranged from 40 to 100%. Given the prevalently unfolded state of both HeV and NiV N<sub>TAIL</sub>, all sites are expected to be equally solvent-accessible. As such, all the variants are expected to give rise to comparable labeling yields. In agreement with this expectation, differences in labeling yields were not position-dependent and likely reflect subtle differences in the experimental conditions rather than different degrees of solvent-exposure of the sulfhydryl group.

### ***Analysis of structural propensities of N<sub>TAIL</sub> variants by CD and EPR spectroscopy***

Although secondary structure predictions using the PSIPRED server (Bryson et al. 2005) pointed out no differences among the various N<sub>TAIL</sub> variants, with the same secondary structure elements predicted for the native proteins being also predicted for the variants (data not shown), a critical point prior to further analysis by EPR spectroscopy was to assess whether the cysteine substitution and the introduction of the covalently bound nitroxide radical affected the overall secondary structure content of the various N<sub>TAIL</sub> variants. To this endeavor, we carried out far-UV circular dichroism (CD) studies. The far-UV CD spectra of spin-labeled N<sub>TAIL</sub> proteins variants at neutral pH quite well superimpose onto that of parental N<sub>TAIL</sub>, and are all typical of unstructured proteins, as shown by their large negative ellipticity at 200 nm and moderate ellipticity at 190 nm (**Figs. 2A, B and S2**). These data indicate that the cysteine substitution and the introduction of the spin label induce little, if any, structural perturbations. We also analyzed the propensities for folding of the spin-labeled N<sub>TAIL</sub> variants in the presence of 20% TFE. The TFE solvent mimics the hydrophobic environment experienced by proteins in protein-protein interactions and is therefore widely used as a probe to unveil disordered regions having a propensity to undergo an induced folding (Hua et al. 1998). We previously reported that the addition of increasing amounts of TFE to *Henipavirus* parental N<sub>TAIL</sub> triggers a gain of  $\alpha$ -helicity (Habchi, Mamelli et al. 2010; Blocquel, Habchi et al. 2012). All N<sub>TAIL</sub> variants show a gain of  $\alpha$ -helicity upon addition of TFE, as judged from the appearance of minima at 208 and 222 nm and from the increased ellipticity at 190 nm (**Figs. 2A, B and S2**). These results therefore indicate that the introduction of the spin label onto a non-native cysteine residue does not impair the ability of *Henipavirus* N<sub>TAIL</sub> to undergo  $\alpha$ -helical folding.

CD experiments however only inform about the overall  $\alpha$ -helical content and do not provide indications on the location of the induced  $\alpha$ -helices. Conversely, SDSL EPR spectroscopy allows a more precise mapping of TFE effects at different, targeted sites, as it is sensitive to structural changes occurring into the close neighborhood around the spin label (Belle, Rouger et al. 2010).

We therefore recorded EPR spectra of spin labeled N<sub>TAIL</sub> proteins at 0% and 20% TFE (**Fig. S3**). When the dynamic of the spin label remains in the rapid regime of mobility, a simple and efficient way to detect mobility change is the measurement of the  $h(+1)/h(0)$  ratio, where  $h(+1)$  and  $h(0)$  are the

peak-to-peak amplitudes of low field and central lines of the EPR spectrum respectively. This semi quantitative parameter decreases with the mobility of the label (Morin, Bourhis et al. 2006; Belle, Rouger et al. 2010; Pirman, Milshteyn et al. 2011). The variations in the radical mobility observed for the spin labeled N<sub>TAIL</sub> proteins in the presence of 20% TFE thus reflect changes in the protein environment in the proximity of the spin label.

The addition of 20% TFE triggers a decrease in the mobility of all HeV and NiV spin labeled N<sub>TAIL</sub> variants, as judged from the drop in the  $h(+1)/h(0)$  ratios (**Fig. 2C, D**). The most pronounced decrease in mobility is observed for the N<sub>TAIL</sub> variants in which the spin label is grafted at positions 475, 481 and 487, with variations in the  $h(+1)/h(0)$  ratio ranging from 0.25 to 0.40 ( $\pm 0.02$ ) (**Fig. 2C, D**). The mobility of the radicals grafted at the other positions (421, 458 and 525) is only moderately affected by the addition of TFE, with the highest variation in the  $h(+1)/h(0)$  being 0.20 ( $\pm 0.02$ ) (**Fig. 2C, D**). This position-dependent TFE-induced gain of rigidity can be accounted for by an  $\alpha$ -helical transition taking place within Box3, in agreement with previous CD studies in the presence of TFE which showed that Box3 is the major determinant of  $\alpha$ -helical folding (Blocquel, Habchi et al. 2012).

A notable difference between the HeV and NiV N<sub>TAIL</sub> proteins resides in the extent of the TFE-induced gain of rigidity within Box3. While the TFE-induced gain of rigidity of the spin label grafted at position 481 is similar in the two viruses, the drop in mobility of spin labels at positions 475 and 487 is more pronounced in the case of NiV (see **Fig. 2C, D**), suggesting a higher extent of  $\alpha$ -helical stabilization by TFE in NiV Box3 as compared to HeV. An additional difference is the higher gain of rigidity of spin labels grafted at positions 458 and 525 in NiV N<sub>TAIL</sub> as compared to HeV. This could reflect differences in Box2 and Box4 between the two viruses, with these two regions having already been shown to differently affect the compaction and folding properties of N<sub>TAIL</sub> in the two viruses (Blocquel, Habchi et al. 2012). We can speculate that for NiV N<sub>TAIL</sub>, the TFE-induced  $\alpha$ -helical folding of Box3 could trigger a more pronounced conformational change in the flanking regions as compared to HeV, causing Box2 and Box4 to establish transient long-range tertiary contacts with Box3. In support of this hypothesis, previous CD studies pointed out a role for Box4 in supporting TFE-induced  $\alpha$ -helical folding of N<sub>TAIL</sub>, although not through inherent  $\alpha$ -helical folding propensities but rather through possible stabilization of the two  $\alpha$ -MoREs located within Box2 and Box3 via long- or short-range tertiary contacts (Blocquel, Habchi et al. 2012).

Incidentally, and in agreement with our previous studies with MeV N<sub>TAIL</sub> (Morin, Bourhis et al. 2006; Belle, Rouger et al. 2008), the observation that TFE has a different impact in the various N<sub>TAIL</sub> positions indicates that TFE does not promote non-native  $\alpha$ -folding. This observation, by showing that TFE acts as a stabilizer (rather than as an inducer) of secondary structures, comes in support of the reliability of structural information derived from studies making use of TFE, a point that is still debated.

In the absence of TFE, for both viruses, slight differences were detected in the EPR spectra according to the positions of the spin labels. The labeled sites can be separated into two groups according to their  $h(+1)/h(0)$  ratios: one with a ratio  $\geq 0.90$  (positions 421, 458, 475 and 525 for HeV and positions 421, 458 and 525 for NiV), and one with a ratio  $< 0.90$  (positions 481 and 487 for HeV and positions 475, 481 and 487 for NiV) (**Fig. 2C, D**). This scenario is comparable to the one previously reported in a similar SDSL study of MeV N<sub>TAIL</sub> (Belle, Rouger et al. 2008), where the slight

decrease in the mobility of radicals located in Box2 has been attributed to a transiently populated  $\alpha$ -helix. This partially pre-configured  $\alpha$ -helix in the free form has been subsequently mapped to the 490–500 region (Kavalenka, Urbancic et al. 2010), which corresponds to the 477-487 region of *Henipavirus* N<sub>TAIL</sub> (Habchi, Blangy et al. 2011). Recent NMR studies on MeV N<sub>TAIL</sub> confirmed the presence of this pre-configuration by showing that N<sub>TAIL</sub> adopts a dynamic equilibrium between a completely unfolded state and several partially helical conformations (Ringkjøbing Jensen et al. 2011).

### ***Analysis of the impact of the partner protein on spin labeled N<sub>TAIL</sub> variants by EPR spectroscopy***

In order to investigate the folding events that both HeV and NiV N<sub>TAIL</sub> undergo in the presence of the partner P<sub>XD</sub>, we monitored EPR spectra of spin-labeled N<sub>TAIL</sub> variants in the presence of 30% sucrose either in the absence or in the presence of a molar excess of the homologous P<sub>XD</sub>. Sucrose is here used as a viscosity agent in order to eliminate the global rotation of the protein that contributes to the partial averaging of the  $g$  and hyperfine tensors anisotropy. Under this condition, the EPR signal reflects the local mobility of the label including the rotation of the side-chain bearing the nitroxide radical and the local backbone fluctuations as well (Timofeev & Tsetlin 1983; McHaourab, Lietzow et al. 1996; Belle, Rouger et al. 2010; Longhi, Belle et al. 2011). In all cases, saturation was achieved with a two- or four-fold molar excess of P<sub>XD</sub>.

All EPR spectra were simulated so as to achieve a more detailed description of the impact of the partner protein on the spin label mobility and to assess the possible presence of stable structural elements in the free form. The simulation allows to decompose an EPR spectrum into possibly different components, each of them being described by two main parameters: the effective rotational correlation time  $\tau_c$  representing a dynamic parameter, and the free rotational space  $\Omega$ , a geometric parameter corresponding to the space angle left for the spin label wobbling, which varies from zero (totally restricted movement) to 1 (totally unrestricted movement) (see Materials and Methods).

We have previously reported a thorough investigation of the  $\alpha$ -helical induced folding that the N<sub>TAIL</sub> domain from the closely related MeV undergoes upon binding to P<sub>XD</sub> (Belle, Rouger et al. 2008; Kavalenka, Urbancic et al. 2010). The availability of as many as 14 spin labeled MeV N<sub>TAIL</sub> variants provides a panel of spectral signatures that have a high predictive value allowing the comparison with the spectral shapes obtained with the *Henipavirus* N<sub>TAIL</sub> variants. Four MeV N<sub>TAIL</sub> variants representative of four different spin label environments in the bound form, namely variants spin labeled at positions 407, 491, 496 and 517, have been chosen for comparative purposes. The spin label grafted at position 407 is located far away from the P<sub>XD</sub> interaction site and is unaffected upon addition of the partner. The spin labels grafted at positions 491 and 496 are both located within the induced  $\alpha$ -helix embedded at the P<sub>XD</sub> interface (with the former pointing towards the partner and the latter being solvent exposed), while the spin label at position 517 is grafted onto a region that does not undergo an  $\alpha$ -helical transition but becomes more rigid upon binding to P<sub>XD</sub>. We have therefore herein simulated the spectra of these MeV spin labeled variants (**Fig. S4** and data not shown). Note that their spectra have been recorded under the same conditions as those used to record the spectra of the HeV and

NiV N<sub>TAIL</sub> variants (i.e. in the presence of 30% sucrose and either with or without saturating amounts of MeV P<sub>XD</sub>) (Belle, Rouger et al. 2008; Kavalenka, Urbancic et al. 2010).

### HeV N<sub>TAIL</sub> proteins

**Figure 3A** shows the EPR spectra of HeV N<sub>TAIL</sub> proteins either with or without saturating amounts of the HeV P<sub>XD</sub> partner protein. In the case of the variants labeled at positions 475, 481 and 487, the EPR spectra obtained in the absence of P<sub>XD</sub> are indicative of a slightly reduced mobility of the spin label as compared to the other positions, as shown by their larger  $\tau_C$  values (**Fig. 3A** and **Table 1**). These spectra were simulated with only one component, corresponding to a single population with similar parameters for all three positions reflecting a high mobility (**Fig. 4** and **Table 1**).

Addition of P<sub>XD</sub> to these three N<sub>TAIL</sub> variants induces a drastic spectral modification indicative of an important decrease in the mobility of the spin labels (**Fig. 3A**). **Figure 4** shows the experimental and simulated spectra of the S475C, S481C and A487C variants. For these variants, the best spectral simulations were obtained with one (S475C) or two (S481C and A487C) broad-shaped components (see spectra 2 and 3 in **Fig. 4**) and one additional narrow component corresponding to a minor fraction of unbound N<sub>TAIL</sub> (see spectra 1 in **Fig. 4**).

**Figure 5A** gives a representation of the results as  $[\tau_C-\Omega]$  plots for the different positions of the spin label either in the absence or in the presence of P<sub>XD</sub>, with the surface of the spheres being proportional to the relative contribution of the various spectral components to the observed, composite spectrum. Addition of the partner triggers a different repartition of the populations, where the label experiences different environments, as illustrated by the distinct values of  $\tau_C$  and  $\Omega$  (**Fig. 5A** and **Table 1**). This likely reflects the occurrence in the bound form of N<sub>TAIL</sub> of conformational sub-ensembles encompassing positions 481 and 487. In these conformational sub-ensembles the spin label exhibits different  $\tau_C$  and  $\Omega$ , possibly corresponding to different orientations of the spin label at the P<sub>XD</sub> interface. It has been previously demonstrated that two-component EPR spectra can result from two rotameric states of the nitroxide side chain, each experiencing a unique environment, a behavior frequently found at helical sites (Langen et al. 2000; Guo et al. 2007; Guo et al. 2008; Bridges, Hideg & Hubbell 2010; Ranaldi et al. 2010). The residual unbound form of N<sub>TAIL</sub> could correspond to a conformational sub-ensemble in which the spin label would adopt an orientation not compatible with binding to P<sub>XD</sub>.

Similar patterns of  $[\tau_C-\Omega]$  plots are observed from the simulation of the EPR spectra of the MeV N<sub>TAIL</sub> S491C, L496C and V517C variants (**Figs. 5B** and **S4**). Spectral simulations in the case of the HeV N<sub>TAIL</sub> variants spin labeled at positions 481 and 487 in the presence of P<sub>XD</sub> yield a pattern comparable to that obtained with the MeV L496C and S491C variants. Aside from the different proportions, the main difference between these HeV and MeV variants concerns the free conformational space parameter  $\Omega$ . For the low  $\Omega$  component, the average omega value of the S481C and A487C HeV variant has an intermediate value (0.38) as compared to the  $\Omega$  values of the S491C MeV variant (0.09) and the L496C MeV variant (0.56) (see **Fig. 5A, B** and **Table 1**). Likewise, for the high  $\Omega$  component, the average  $\Omega$  value of the S481C and A487C HeV variants (0.64) is intermediate between the value of the S491C MeV variant (0.48) and that of the L496C MeV variant

(0.71) (see **Fig. 5A, B** and **Table 1**). We have previously shown that in the MeV N<sub>TAIL</sub> S491C and L496C labeled variants, the spin label is located in the  $\alpha$ -MoRE and points either towards the partner or the solvent, respectively (Belle, Rouger et al. 2008), in agreement with the crystal structure of the chimeric construct consisting of P<sub>XD</sub> and of the 486-504 region of N<sub>TAIL</sub> (Kingston, Hamel et al. 2004) (see **Fig. 6A**). **The spectral similarities between these MeV and HeV variants, together with our previous studies supporting a P<sub>XD</sub>-induced  $\alpha$ -helical transition within HeV N<sub>TAIL</sub> (Habchi, Blangy et al. 2011), allow us to** conclude that in the S481C and A487C HeV N<sub>TAIL</sub> variants, the spin labels are located within the induced  $\alpha$ -helix, with the labels being in an intermediate environment between fully solvent-exposed and in close contact with P<sub>XD</sub>, a situation that is consistent with the structural model of the HeV N<sub>TAIL</sub>-P<sub>XD</sub> complex that we previously proposed (Habchi, Blangy et al. 2011) (see **Fig. 6A**).

To go further, we computed the rotamers populated by each spin label in the S475C, S481C and A487C variants in the structural model of the HeV N<sub>TAIL</sub>-P<sub>XD</sub> complex (Habchi, Blangy et al. 2011) using the MMM software (see Materials and Methods), based on a rotamer library approach (Polyhach, Bordignon et al. 2011). The number of rotamers found for both S481C and A487C N<sub>TAIL</sub> variants (around 60) is comprised between the low number of rotamers (6) of the MeV S491C N<sub>TAIL</sub> variant, where the label is in contact with P<sub>XD</sub>, and the number of rotamers of the MeV L496C variant (116), where the label is in the middle of the induced  $\alpha$ -helix and fully exposed to the solvent (**Fig. 6**). These data are consistent with the EPR analyses (**Fig. 5A**) that showed an environment of the spin labels intermediate between solvent exposed and in contact with P<sub>XD</sub>.

The pattern of the HeV S475C N<sub>TAIL</sub> variant in the bound form is different from that obtained with the S481C and A487C variants, in that it displays only one component of high  $\Omega$  (0.83), corresponding to a rather open conformational space, and a high  $\tau_C$  (4.1 ns), indicating a drop in the spin label mobility upon binding to the partner. This drop in the dynamics of the spin label likely reflects a gain of rigidity arising from the P<sub>XD</sub>-induced  $\alpha$ -helical folding. Using MMM, the number of rotamers populated by the spin label at position 475, as calculated from the modeled HeV N<sub>TAIL</sub>-P<sub>XD</sub> complex, is very high (163 rotamers) (**Fig. 6B**), a result in agreement with the high conformational space deduced from the simulation of the bound form. Interestingly, the pattern obtained with the HeV S475C variant is comparable to that of the MeV V517C N<sub>TAIL</sub> variant (*cf.* **Fig. 5A** and **5B**), where the spectral modification induced by P<sub>XD</sub> is due to a gain of rigidity resulting from the  $\alpha$ -helical folding taking place in the upstream Box2 region (Belle, Rouger et al. 2008). In this latter case, the label is however not part of the  $\alpha$ -helix and the label dynamics is higher than that observed for the HeV S475C variant, as indicated by the lower value of  $\tau_C$  (2.4 ns) of the MeV variant (see **Fig. 5B**). This observation is consistent with a situation where the label of the HeV S475C N<sub>TAIL</sub> variant is located at the extremity of the P<sub>XD</sub>-induced  $\alpha$ -helix in a position fully exposed to the solvent, a result that further validates the proposed structural model of the N<sub>TAIL</sub>-P<sub>XD</sub> complex (**Fig. 6A**).

Taken together, the EPR analyses of the labeled variants within Box3 bring experimental support for the structural model of HeV N<sub>TAIL</sub>-P<sub>XD</sub> complex that we previously proposed (Habchi, Blangy et al. 2011) (**Fig. 6A**).

In the case of positions 421, 458 and 525, upon addition of P<sub>XD</sub>, the spectra are still indicative of a high radical mobility (**Fig. 3A**). Indeed, the EPR spectra obtained with and without P<sub>XD</sub> were found to



be superimposable and the best simulation is obtained with only one component (see **Figs. 4** and **5A** for S458C and **Table 1**). A quite comparable situation is observed for position 407 of MeV N<sub>TAIL</sub> (see **Fig. 5B**), a site that remains disordered in the bound form and does not contact P<sub>XD</sub> (Bourhis et al. 2005; Morin, Bourhis et al. 2006; Belle, Rouger et al. 2008).

The absence of impact of P<sub>XD</sub> on the mobility of the spin label grafted at these positions argues for the lack of involvement of these sites in complex formation, in agreement with previous data that showed that Box1, Box2 and Box4 are dispensable for binding to P<sub>XD</sub> (Blocquel, Habchi et al. 2012).

#### NiV N<sub>TAIL</sub> proteins

In the absence of the partner, the NiV N<sub>TAIL</sub> variants bearing the spin label grafted within Box3 (i.e. positions 475, 481 and 487) show a different behavior with respect to the corresponding HeV variants. Composite spectra were indeed observed, which required several components to obtain a good fit (**Figs. 5C, S5** and **Table 1**). These results indicate that in the absence of the partner, the NiV N<sub>TAIL</sub> variants labeled within Box3 sample in solution multiple conformations, associated to different environments of the label. The broad-shaped components (see **Fig. S5**) are characterized by  $\Omega$  and  $\tau_C$  values ( $\Omega \approx 0.3$  to  $0.8$  and  $\tau_C \approx 0.5$  to  $3.3$  ns) (**Fig. 5C** and **Table 1**), which are reminiscent of those observed for labels in an  $\alpha$ -helical environment (i.e. MeV L496C and S491C, or HeV S481C and S487C) (see **Table 1**). Those populations likely reflect the presence of  $\alpha$ -helical segments within Box3, stable enough to be detected. Taking into account the nanosecond EPR time window, the detection of distinct populations for the NiV N<sub>TAIL</sub> variants labeled within Box3 indicates that the inter-conversion time between these  $\alpha$ -helical segments is longer than the ns. The lack of detection of multiple conformations for the free forms of the corresponding HeV and MeV variants suggests that these pre-structured  $\alpha$ -helical segments within Box3 would be more long-lived in NiV N<sub>TAIL</sub> than in HeV and MeV.

Like in the case of the HeV variants, addition of P<sub>XD</sub> induces an important decrease in the mobility of the spin labels grafted at positions 475, 481 and 487 (**Fig. 3B** and **5C**). In the case of the A487C N<sub>TAIL</sub> variant, three components are observed, two of which are already observed in the free form (**Figs. 5C** and **S5** and **Table 1**). Aside from the occurrence of two distinct forms in the absence of the partner, the pattern obtained with the NiV A487C variant is similar to the one exhibited by the corresponding HeV variant, with close parameters but slightly different proportions of each component. This suggests that the two labels at this position experience a similar environment. For the S481C and S475C variants, addition of P<sub>XD</sub> triggers the appearance of new spectral components, as judged from the modification of the  $\Omega$  and  $\tau_C$  values (**Fig. 5C**). The pattern observed for the spin label grafted at position 481 is reminiscent of that obtained with the HeV variant labeled at the same position, with one major component of low  $\Omega$  (0.36) and low  $\tau_C$  (0.8 ns). This observation indicates that the spin probes grafted at this position experience a similar environment. The two other components with  $\Omega$  around 0.7 and  $\tau_C$  equal to either 1.4 or 2.7 ns could reflect heterogeneities in the bound form. For the variant labeled at position 475, the major spectral population corresponds to components of high  $\Omega$  (around 0.90), indicating a more open conformational space than for positions 481 and 487, although a minor population with a narrower conformational space is also detected. The observation that the dominant

population is characterized by a large open conformational space is in agreement with the solvent orientation of the side-chain of residue 475 proposed in the structural model of the NiV N<sub>TAIL</sub>-P<sub>XD</sub> complex (Habchi, Blangy et al. 2011) (**Fig. 6A**).

Comparing position 475 in the two viruses clearly points out a difference, that was already observed in the presence of TFE, where the decrease in mobility of the label was more pronounced for NiV than HeV (*cf.* **Fig. 2C** and **2D**). For all the spin labels grafted within Box3, the number of possible rotamers built by MMM on the structural model of the NiV N<sub>TAIL</sub>-P<sub>XD</sub> complex (*i.e.* the number of possible rotamers that could be accommodated in the available conformational space) is comparable to that obtained with HeV (**Fig. 6B**). As in the case of HeV, the present results validate the structural model of the NiV N<sub>TAIL</sub>-P<sub>XD</sub> complex (Habchi, Blangy et al. 2011) (**Fig. 6A**).

For the other spin label positions, the EPR spectra of the NiV N<sub>TAIL</sub> proteins bearing the spin label grafted at positions 421, 458 and 525 are indicative of a high radical mobility. Addition of P<sub>XD</sub> induces no change in the EPR spectral shape of the variants labeled at positions 421 and 525 (**Fig. 3B** and **Table 1**) and simulation of the resulting spectra required only one component (**Table 1**). In contrast to HeV, addition of the partner triggers a reduction in the mobility of the spin label grafted at position 458, although this drop is less pronounced than that observed at positions 475, 481 and 487 (**Fig. 3B**). Note that simulation of the spectrum of the NiV S458C variant in the presence of P<sub>XD</sub> required three components including one corresponding to the unbound form (**Figs. 5C** and **S5**). Besides this unbound form, the pattern of the  $[\tau_C - \Omega]$  plot is very similar to the one obtained for the HeV S481C and A487C variants (**Fig. 5A**), corresponding to a position within the induced  $\alpha$ -helix. This observation suggests that in the case of NiV addition of P<sub>XD</sub> would trigger at least partial formation of an  $\alpha$ -helix within Box2.

### ***Similarities and discrepancies between HeV and NiV in terms of molecular mechanisms of binding to P<sub>XD</sub>***

In the absence of the partner, for both HeV and NiV, the EPR spectral analysis pointed out the existence of significant differences in mobility as a function of the spin label position, with the most reduced mobility being observed for the spin labels grafted within Box3, consistent with the existence of a partly pre-configured  $\alpha$ -helix, as already reported for MeV N<sub>TAIL</sub> (Belle, Rouger et al. 2008; Kavalenka, Urbancic et al. 2010). By analogy with MeV, the lower mobility of the spin labels grafted within the Box3 region of HeV N<sub>TAIL</sub> in the absence of the partner as compared to that of the other positions, is indicative of transiently populated structural elements in rapid inter-conversion with the unstructured form of N<sub>TAIL</sub>. For both henipaviruses, the existence of a partly preconfigured  $\alpha$ -MoRE within Box3 prior to binding to P<sub>XD</sub> suggests that the induced folding mechanism would rely, at least partly, on a conformational selection (*i.e.* selection by the partner of a pre-existing conformation) (Tsai et al. 2001; Tsai et al. 2001). It is well established that the occurrence of partly preconfigured MoREs favors the binding process by lowering the entropic penalty associated to the disorder-to-order transition (Tomba 2002; Fuxreiter, Simon et al. 2004; Lacy et al. 2004; Sivakolundu, Bashford & Kriwacki 2005; Mohan, Oldfield et al. 2006; Vacic, Oldfield et al. 2007).

Notably, while in the absence of P<sub>XD</sub> the HeV variants bearing the spin label within Box3 exhibit

spectra with a single component, the corresponding NiV variants have multi-component EPR spectra that argue for a conformational heterogeneity of NiV N<sub>TAIL</sub>. This conformational heterogeneity can reflect the existence within Box3 of different  $\alpha$ -helical segments that would be more long-lived in NiV than in the case of HeV and MeV (i.e. these conformations are stable enough in the case of NiV, at the EPR timescale, to be detected as different EPR components). This difference in the persistence of these helical segments could arise from the establishment of fluctuating short- or long-range tertiary contacts between Box3 and other N<sub>TAIL</sub> regions.

The finding that the spin labels grafted at positions 421 and 525 in both HeV and NiV N<sub>TAIL</sub> have a high mobility in the free forms of the protein is consistent with the lack of a transiently populated  $\alpha$ -helix within Box1 and Box4, in agreement with previous spectroscopic studies (Blocquel, Habchi et al. 2012). On the other hand, this high mobility does not provide any hints about the previously postulated role of Box1 as a determinant of protein compaction and of Box4 as possibly establishing tertiary contacts with Box2 and Box3 (Blocquel, Habchi et al. 2012), and rather supports solvent exposure and lack of involvement of these residues in intra-molecular tertiary contacts.

In the presence of P<sub>XD</sub>, the various sites targeted for spin labeling were found to contribute to different extents. For spin labels at positions 421 and 525, EPR analyses reinforced the previous observations (Blocquel, Habchi et al. 2012), confirming lack of involvement of these two regions in the interaction with P<sub>XD</sub>. This latter observation suggests that Box1 and Box4 retain a considerable flexibility in the bound form, consistent with the fuzzy (Tompa & Fuxreiter 2008) nature of the *Henipavirus* N<sub>TAIL</sub>-P<sub>XD</sub> complexes. Notably, although previously reported experimental data (Habchi, Blangy et al. 2011) support the fuzziness of the *Henipavirus* N<sub>TAIL</sub>-P<sub>XD</sub> complexes (Habchi & Longhi 2012), the use of paramagnetic probes combined with EPR spectroscopy allowed us to precisely identify the N<sub>TAIL</sub> regions remaining flexible in the bound form.

For position 458, which is located in the middle of Box2, the behavior is different between HeV and NiV. While P<sub>XD</sub> induces no spectral variation in the case of HeV, the spectrum is modified for NiV and revealed spectral components characteristic of  $\alpha$ -helical elements (*cf.* **Fig. 5A** and **5C**). The different behavior of the spin label grafted at position 458 in the two viruses could be ascribed to an already previously noticed different role of Box2 in the two viruses, where removal of Box2 led to either a more extended (NiV) or a more compact (HeV) form as compared to parental, full-length N<sub>TAIL</sub> (Blocquel, Habchi et al. 2012). This different role of Box2 may reflect a different extent of involvement in either short- or long-range tertiary contacts with other N<sub>TAIL</sub> regions, resulting in different degrees of protein compaction. One can speculate that in NiV N<sub>TAIL</sub>, binding to P<sub>XD</sub> may be accompanied by a stabilization of these long-range contacts. Alternatively or additionally, binding of P<sub>XD</sub> to Box3 could stabilize the  $\alpha$ -helix within Box2, leading to a reduced mobility of the spin label grafted at position 458.

For the labels within Box3, the comparison of the EPR spectral shapes with those of MeV N<sub>TAIL</sub> variants labeled at positions of well-known behavior and the calculation of spin labels rotamers using MMM allowed us to document the formation of the induced  $\alpha$ -helix and to validate the orientation of this helix with respect to P<sub>XD</sub> as proposed in the structural models of the complexes (Habchi, Blangy et al. 2011), with the side-chains of cysteine 475 being exposed to the solvent and that of cysteines 481 and 487 being closer to the P<sub>XD</sub> surface (**Fig. 6**).

Altogether, the present results provide evidence for the occurrence of a transiently populated  $\alpha$ -helix within the Box3 region of *Henipavirus* N<sub>TAIL</sub> domains prior to binding. They also confirm and extend previous data indicating that P<sub>XD</sub> triggers an  $\alpha$ -helical transition within Box3, while providing information at the residue level and new insights into the molecular mechanisms of the N<sub>TAIL</sub>-P<sub>XD</sub> complex formation. Indeed, the present study, while providing experimental support for the reliability of the previously proposed structural models of the N<sub>TAIL</sub>-P<sub>XD</sub> complexes, pointed out some subtle differences between the two viruses. In particular, the different persistence of the partly pre-configured  $\alpha$ -helical segments of Box3 may have important implications for the design of inhibitors specifically able to block the HeV or the NiV N<sub>TAIL</sub>-P<sub>XD</sub> interaction.

## ACKNOWLEDGEMENTS

We wish to thank Nicolas Armstrong and the mass spectrometry platform of the IFR48 of Marseille for mass spectrometry analyses, and David Blocquel (AFMB) for help with the SEC-MALLS analyses of the synthetic NiV P<sub>XD</sub> sample. We also thank Jean-Pierre Andrieu (Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale, Grenoble, France) and the proteomic platform of the IBS for the amino acid compositions analyses. The authors are also grateful to the EPR facilities of the national TGE RPE and Aix-Marseille University EPR platform, and to Janez Strancar (Jozef Stefan Institut, Ljubljana, Slovenia) for providing access to the EPRSIM-C software.

## FUNDING

This work was carried out with the financial support of the Agence Nationale de la Recherche, specific program "Blanc", ANR-09-BLAN-0100 (B.G.), and with the partial support of the CNRS. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## FIGURE LEGENDS

**Figure 1. Design and purification of cysteine N<sub>TAIL</sub> variants.** (A) Sequence alignment between the HeV and NiV N<sub>TAIL</sub> proteins and structural organization of *Henipavirus* N<sub>TAIL</sub>. Hydrophobic residues (L, I, V, F, W and M) and charged residues are shown in grey. Asterisks and colons below the alignment denote identity and similarity, respectively, while dots indicate weakly similar residues. Lack of symbol below the alignment denotes strongly different residues. The front numbers correspond to the amino acid position in the N sequence. Dots above the alignment indicate intervals of 10 residues. The four boxes corresponding to MoREs (Habchi, Mamelli et al. 2010; Blocquel, Habchi et al. 2012) are shown. The positions targeted for cysteine substitution and spin labeling are shaded in grey in the sequence alignment (top) and highlighted by arrows in the schematic representation of the structural organization of *Henipavirus* N<sub>TAIL</sub> domains (bottom). (B) Coomassie blue staining of a 12% SDS-PAGE loaded with approximately 5 µg of N<sub>TAIL</sub> variants purified from the soluble fraction of *E. coli*. Note that 20 mM DTT was added to the samples before electrophoretic migration.

**Figure 2. Effect of TFE on spin-labeled HeV and NiV N<sub>TAIL</sub> proteins monitored by far-UV CD and EPR spectroscopies.** (A,B) CD spectra of spin labeled HeV A487C (A) and NiV A487C (B) variants representative of the other labeled variants (see Fig. S2) at 0% and 20% TFE (v/v). (C, D)  $h(+1)/h(0)$  ratios, corresponding to the peak-to-peak amplitude of the low-field over the central line of the EPR spectra, of the spin labeled HeV (C) and NiV (D) N<sub>TAIL</sub> variants at 0% and 20% (v/v) TFE.

**Figure 3. Effect of P<sub>XD</sub> on labeled N<sub>TAIL</sub> variants.** Amplitude normalized room temperature EPR spectra of the spin labeled HeV (A) and NiV (B) N<sub>TAIL</sub> variants in the presence of 30% (v/v) sucrose and either in the absence or in the presence of saturating amounts of P<sub>XD</sub>.

**Figure 4. Simulation of the RT EPR spectra of the HeV S458C, S475C, S481C and A487C spin labeled N<sub>TAIL</sub> variants in the presence of sucrose (30% v/v) and in the absence or presence of saturating amounts of P<sub>XD</sub>.** Simulation was performed using the EPRSIM-C software program (Strancar, Koklic et al. 2005). The individual simulated components are given at the top of each panel with the indication of their respective proportion (see Table 1). At the bottom of each panel are shown the simulated spectra (black lines) resulting from the addition of each individual simulated component, superimposed on the experimental spectra (grey lines).

**Figure 5. [ $\tau_C$ - $\Omega$ ] plots of HeV, MeV and NiV spin labeled N<sub>TAIL</sub> variants.** Parameters extracted from the simulation of RT EPR spectra of the HeV (A), MeV (B) and NiV (C) spin labeled N<sub>TAIL</sub> variants in the presence of 30% (v/v) sucrose and in the absence or presence of saturating amounts of P<sub>XD</sub>. The parameters are given as [ $\tau_C$ - $\Omega$ ] plots, with  $\tau_C$  being the effective rotational correlation time (in ns) and  $\Omega$  the normalized free rotational space, in the absence (white spheres) and presence (black spheres) of the partner P<sub>XD</sub>. The surface of the spheres is proportional to the relative contribution of the EPR spectral shape components. For the sake of clarity, when the positions of the black spheres

overlapped those of the white spheres, the former have been slightly shifted (case of HeV S458C, and of NiV S407C and A487C variants, see **Table 1** for parameters).

**Figure 6. Structural models and calculated rotamers. (A)** Crystal structure of the MeV chimera between P<sub>XD</sub> and the N<sub>TAIL</sub> region encompassing residues 486–504 (pdb code 1T60) (Kingston, Hamel et al. 2004), and structural models of the *Henipavirus* complexes between P<sub>XD</sub> and the Box3 region (residues 473-493) of N<sub>TAIL</sub>. P<sub>XD</sub> is shown with a surface representation, whereas the N<sub>TAIL</sub> region is represented as a ribbon. Hydrophobic P<sub>XD</sub> residues are shown in light grey. In the MeV complex, residues 491 and 496 have been replaced by a cysteine and these side chains are shown in sticks. The side chain of cysteine 491 points towards the surface of P<sub>XD</sub>, while that of cysteine 496 is solvent-exposed. In the structural models of the HeV and NiV complexes the native N<sub>TAIL</sub> residues targeted for cysteine mutagenesis and spin labeling have been replaced by a cysteine. the side chains of these cysteines are shown in sticks. Note that in the structural models, the side chain of residue 475 is solvent-exposed, while that of residues 481 and 487 is closer to the P<sub>XD</sub> surface. The pictures were drawn using Pymol (DeLano 2002). **(B)** Calculated rotamers of the spin labels of MeV, HeV and NiV N<sub>TAIL</sub> variants. Ribbon representation of the complexes of MeV (pdb code 1T60) (Kingston, Hamel et al. 2004), HeV and NiV N<sub>TAIL</sub>-P<sub>XD</sub> (Habchi, Blangy et al. 2011) showing, in semi-transparent sticks and balls, the calculated spin label rotamers attached to the engineered cysteines at positions 491 and 496 (MeV) and 475, 481 and 487 (HeV and NiV). The spin label rotamers were computed using the software MMM. The balls correspond to the baricenter of the N-O bond with the size of the ball being proportional to the abundance of the corresponding rotamer.

## REFERENCES

- Belle, V., Rouger, S., Costanzo, S., Liquiere, E., Strancar, J., Guigliarelli, B., Fournel, A. & Longhi, S. (2008). Mapping alpha-helical induced folding within the intrinsically disordered C-terminal domain of the measles virus nucleoprotein by site-directed spin-labeling EPR spectroscopy. *Proteins: Structure, Function and Bioinformatics* 73: 973-988.
- Belle, V., Rouger, S., Costanzo, S., Longhi, S. & Fournel, A. (2010). Site-directed spin labeling EPR spectroscopy. *Instrumental analysis of intrinsically disordered proteins: assessing structure and conformation*. V. N. Uversky and S. Longhi. Hoboken, New Jersey John Wiley and Sons.
- Biswas, R., Kuhne, H., Brudvig, G. W. & Gopalan, V. (2001). Use of EPR spectroscopy to study macromolecular structure and function. *Sci Prog* 84: 45-67.
- Blocquel, D., Habchi, J., Gruet, A., Blangy, S. & Longhi, S. (2012). Compaction and binding properties of the intrinsically disordered C-terminal domain of Henipavirus nucleoprotein as unveiled by deletion studies. *Mol Biosyst* 8: 392-410.
- Bourhis, J. M., Receveur-Bréchet, V., Oglesbee, M., Zhang, X., Buccellato, M., Darbon, H., Canard, B., Finet, S. & Longhi, S. (2005). The intrinsically disordered C-terminal domain of the measles virus nucleoprotein interacts with the C-terminal domain of the phosphoprotein via two distinct sites and remains predominantly unfolded. *Protein Sci* 14: 1975-1992.
- Bridges, M. D., Hideg, K. & Hubbell, W. L. (2010). Resolving Conformational and Rotameric Exchange in Spin-Labeled Proteins Using Saturation Recovery EPR. *Appl Magn Reson* 37: 363.
- Bryson, K., McGuffin, L. J., Marsden, R. L., Ward, J. J., Sodhi, J. S. & Jones, D. T. (2005). Protein structure prediction servers at University College London. *Nucleic Acids Res* 33: W36-38.
- Chan, Y. P., Koh, C. L., Lam, S. K. & Wang, L. F. (2004). Mapping of domains responsible for nucleocapsid protein-phosphoprotein interaction of Henipaviruses. *J Gen Virol* 85: 1675-1684.
- Chouard, T. (2011). Structural biology: Breaking the protein rules. *Nature* 471: 151-153.
- DeLano, W. L. (2002). The PyMOL molecular graphics system *Proteins: Structure, Function and Bioinformatics* 30: 442-454.
- Drescher, M. (2011). EPR in Protein Science : Intrinsically Disordered Proteins. *Top Curr Chem*.
- Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M. & Uversky, V. N. (2005). Flexible nets. *Febs J* 272: 5129-5148.
- Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C. & Obradovic, Z. (2001). Intrinsically disordered protein. *J. Mol. Graph. Model.* 19: 26-59.
- Dunker, A. K. & Obradovic, Z. (2001). The protein trinity--linking function and disorder. *Nat Biotechnol* 19: 805-806.
- Dunker, A. K., Oldfield, C. J., Meng, J., Romero, P., Yang, J. Y., Chen, J. W., Vacic, V., Obradovic, Z. & Uversky, V. N. (2008). The unfoldomics decade: an update on intrinsically disordered proteins. *BMC Genomics* 9 Suppl 2: S1.

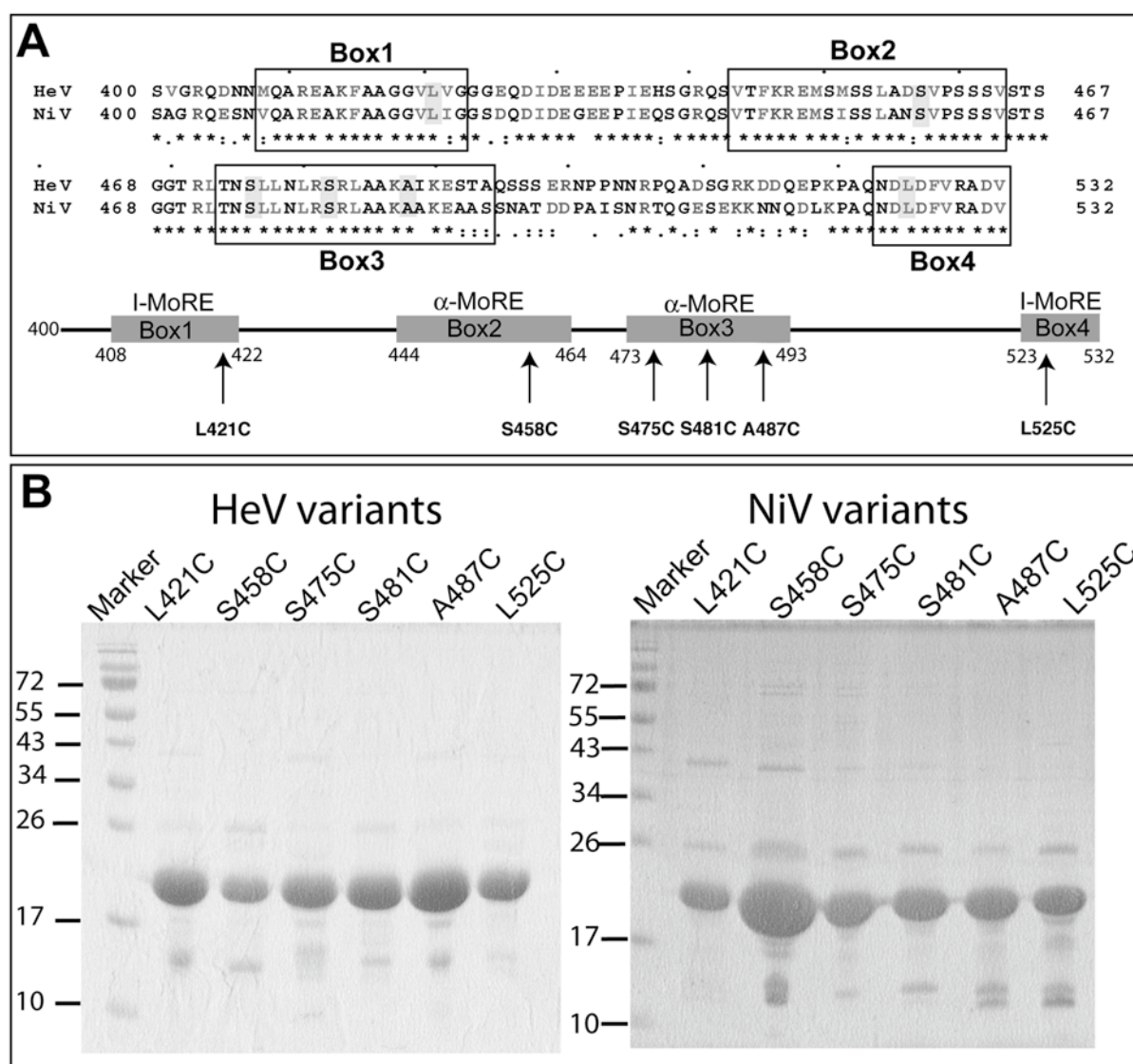
- Dunker, A. K., Silman, I., Uversky, V. N. & Sussman, J. L. (2008). Function and structure of inherently disordered proteins. *Curr Opin Struct Biol* 18: 756-764.
- Dyson, H. J. & Wright, P. E. (2002). Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol* 12: 54-60.
- Dyson, H. J. & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6: 197-208.
- Eaton, B. T., Mackenzie, J. S. & Wang, L. F. (2007). Henipaviruses. *Fields Virology*. B. N. Fields, D. M. Knipe and P. M. Howley. Philadelphia, Lippincott-Raven: 1587-1600.
- Erales, J., Lorenzi, M., Lebrun, R., Fournel, A., Etienne, E., Courcelle, C., Guigliarelli, B., Gontero, B. & Belle, V. (2009). A new function of GAPDH from *Chlamydomonas reinhardtii*: a thiol-disulfide exchange reaction with CP12. *Biochemistry* 48: 6034-6040.
- Fanucci, G. E. & Cafiso, D. S. (2006). Recent advances and applications of site-directed spin labeling. *Current Opinion in Structural Biology* 16: 644-653.
- Feix, J. B. & Klug, C. S. (1998). Site-directed spin-labeling of membrane proteins and peptide-membrane interactions. In *Biological magnetic resonance*. L. Berliner. New York, Plenum Press. *Spin labeling: the next millenium*: 251-281.
- Fuxreiter, M., Simon, I., Friedrich, P. & Tompa, P. (2004). Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J Mol Biol* 338: 1015-1026.
- Fuxreiter, M., Tompa, P. & Simon, I. (2007). Local structural disorder imparts plasticity on linear motifs. *Bioinformatics* 23: 950-956.
- Guo, Z., Cascio, D., Hideg, K. & Hubbell, W. L. (2008). Structural determinants of nitroxide motion in spin-labeled proteins: solvent-exposed sites in helix B of T4 lysozyme. *Protein Sci* 17: 228-239.
- Guo, Z., Cascio, D., Hideg, K., Kalai, T. & Hubbell, W. L. (2007). Structural determinants of nitroxide motion in spin-labeled proteins: tertiary contact and solvent-inaccessible sites in helix G of T4 lysozyme. *Protein Sci* 16: 1069-1086.
- Habchi, J., Blangy, S., Mamelli, L., Ringkjøbing Jensen, M., Blackledge, M., Darbon, H., Oglesbee, M., Shu, Y. & Longhi, S. (2011). Characterization of the interactions between the nucleoprotein and the phosphoprotein of Henipaviruses. *J Biol Chem* 286: 13583-13602.
- Habchi, J. & Longhi, S. (2012). Structural disorder within paramyxovirus nucleoproteins and phosphoproteins. *Mol Biosyst* 8: 69-81.
- Habchi, J., Mamelli, L., Darbon, H. & Longhi, S. (2010). Structural Disorder within Henipavirus Nucleoprotein and Phosphoprotein: From Predictions to Experimental Assessment. *PLoS ONE* 5: e11684.
- Habchi, J., Mamelli, L. & Longhi, S. (2012). Structural disorder within the nucleoprotein and phosphoprotein from measles, Nipah and Hendra viruses. *Flexible viruses: structural disorder in viral proteins*. V. N. Uversky and S. Longhi. Hoboken, New Jersey, John Wiley and Sons.
- Hua, Q. X., Jia, W. H., Bullock, B. P., Habener, J. F. & Weiss, M. A. (1998). Transcriptional activator-coactivator recognition: nascent folding of a kinase-inducible transactivation domain predicts its structure on coactivator binding. *Biochemistry* 37: 5858-5866.



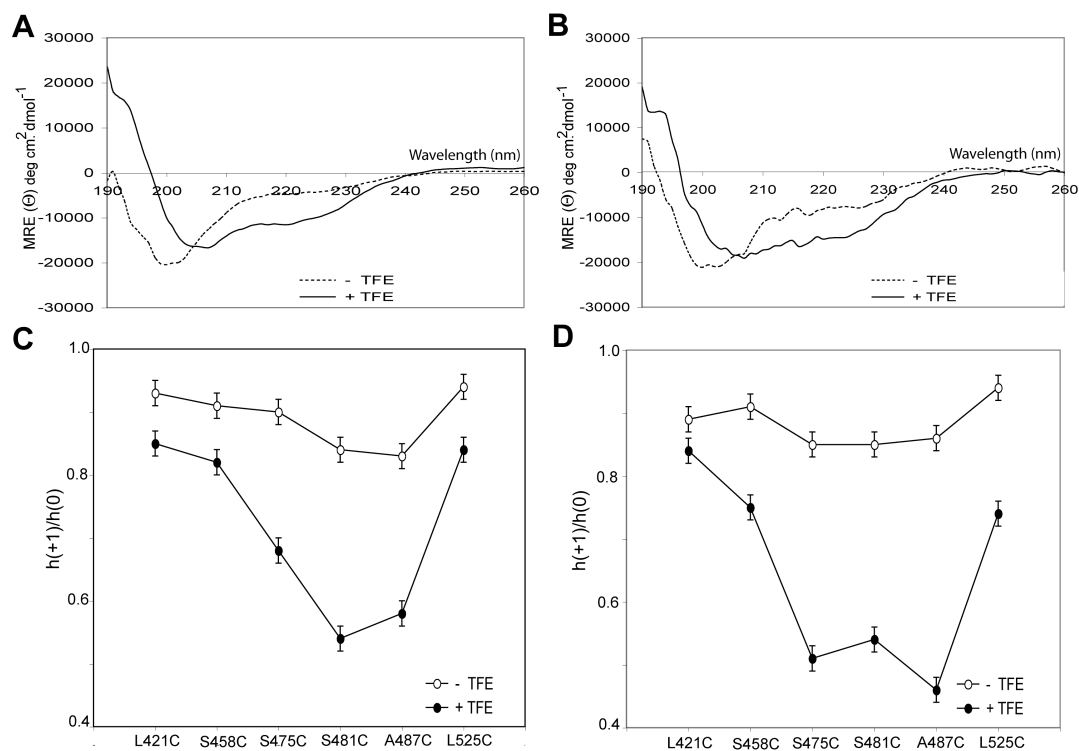
- Hubbell, W. L., Altenbach, C., Hubbell, C. M. & Khorana, H. G. (2003). Rhodopsin structure, dynamics, and activation: a perspective from crystallography, site-directed spin labeling, sulfhydryl reactivity, and disulfide cross-linking. *Adv Protein Chem* 63: 243-290.
- Hubbell, W. L., McHaourab, H. S., Altenbach, C. & Lietzow, M. A. (1996). Watching proteins move using site-directed spin labeling. *Structure* 4: 779-783.
- Johansson, K., Bourhis, J. M., Campanacci, V., Cambillau, C., Canard, B. & Longhi, S. (2003). Crystal structure of the measles virus phosphoprotein domain responsible for the induced folding of the C-terminal domain of the nucleoprotein. *J Biol Chem* 278: 44567-44573.
- Kavalenka, A., Urbancic, I., Belle, V., Rouger, S., Costanzo, S., Kure, S., Fournel, A., Longhi, S., Guigliarelli, B. & Strancar, J. (2010). Conformational analysis of the partially disordered measles virus NTAIL-XD complex by SDSL EPR spectroscopy. *Biophys J* 98: 1055-1064.
- Kingston, R. L., Hamel, D. J., Gay, L. S., Dahlquist, F. W. & Matthews, B. W. (2004). Structural basis for the attachment of a paramyxoviral polymerase to its template. *Proc Natl Acad Sci U S A* 101: 8301-8306.
- Klare, J. P. & Steinhoff, H. J. (2009). Spin labeling EPR. *Photosynthesis Research* 102: 377-390.
- Kriwacki, R. W., Hengst, L., Tennant, L., Reed, S. I. & Wright, P. E. (1996). Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proc Natl Acad Sci U S A* 93: 11504-11509.
- Lacy, E. R., Filippov, I., Lewis, W. S., Otieno, S., Xiao, L., Weiss, S., Hengst, L. & Kriwacki, R. W. (2004). p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nat Struct Mol Biol* 11: 358-364.
- Langen, R., Oh, K. J., Cascio, D. & Hubbell, W. L. (2000). Crystal structures of spin labeled T4 lysozyme mutants: implications for the interpretation of EPR spectra in terms of structure. *Biochemistry* 39: 8396-8405.
- Longhi, S., Belle, V., Fournel, A., Guigliarelli, B. & Carrière, F. (2011). Probing structural transitions in both structured and disordered proteins by site-directed spin-labeling EPR spectroscopy. *J Pept Sci in press*.
- Lorenzi, M., Puppo, C., Lebrun, R., Lignon, S., Roubaud, V., Martinho, M., Mileo, E., Tordo, P., Marque, S. R., Gontero, B., Guigliarelli, B. & Belle, V. (2011). Tyrosine-targeted spin labeling and EPR spectroscopy: an alternative strategy for studying structural transitions in proteins. *Angew Chem Int Ed Engl* 50: 9108-9111.
- McHaourab, H. S., Lietzow, M. A., Hideg, K. & Hubbell, W. L. (1996). Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry* 35: 7692-7704.
- Mohan, A., Oldfield, C. J., Radivojac, P., Vacic, V., Cortese, M. S., Dunker, A. K. & Uversky, V. N. (2006). Analysis of Molecular Recognition Features (MoRFs). *J Mol Biol* 362: 1043-1059.
- Morin, B., Bourhis, J. M., Belle, V., Woudstra, M., Carrière, F., Guigliarelli, B., Fournel, A. & Longhi, S. (2006). Assessing induced folding of an intrinsically disordered protein by site-directed spin-labeling EPR spectroscopy. *J. Phys. Chem. B* 110: 20596-20608.

- Oldfield, C. J., Cheng, Y., Cortese, M. S., Romero, P., Uversky, V. N. & Dunker, A. K. (2005). Coupled Folding and Binding with alpha-Helix-Forming Molecular Recognition Elements. *Biochemistry* 44: 12454-12470.
- Omi-Furutani, M., Yoneda, M., Fujita, K., Ikeda, F. & Kai, C. (2010). Novel phosphoprotein-interacting region in Nipah virus nucleocapsid protein and its involvement in viral replication. *J Virol* 84: 9793-9799.
- Pirman, N. L., Milshteyn, E., Galiano, L., Hewlett, J. C. & Fanucci, G. E. (2011). Characterization of the disordered-to-alpha-helical transition of IA by SDSL-EPR spectroscopy. *Protein Sci* 20: 150-159.
- Polyhach, Y., Bordignon, E. & Jeschke, G. (2011). Rotamer libraries of spin labelled cysteines for protein studies. *Phys Chem Chem Phys* 13: 2356-2366.
- Radivojac, P., Iakoucheva, L. M., Oldfield, C. J., Obradovic, Z., Uversky, V. N. & Dunker, A. K. (2007). Intrinsic disorder and functional proteomics. *Biophys J* 92: 1439-1456.
- Ranaldi, S., Belle, V., Woudstra, M., Bourgeas, R., Guigliarelli, B., Roche, P., Vezin, H., Carriere, F. & Fournel, A. (2010). Amplitude of pancreatic lipase lid opening in solution and identification of spin label conformational subensembles by combining continuous wave and pulsed EPR spectroscopy and molecular dynamics. *Biochemistry* 49: 2140-2149.
- Ringkjøbing Jensen, M., Communie, G., Ribeiro, E. D., Jr., Martinez, N., Desfosses, A., Salmon, L., Mollica, L., Gabel, F., Jamin, M., Longhi, S., Ruigrok, R. W. & Blackledge, M. (2011). Intrinsic disorder in measles virus nucleocapsids. *Proc Natl Acad Sci U S A* 108: 9839-9844.
- Sivakolundu, S. G., Bashford, D. & Kriwacki, R. W. (2005). Disordered p27Kip1 exhibits intrinsic structure resembling the Cdk2/cyclin A-bound conformation. *J Mol Biol* 353: 1118-1128.
- Strancar, J., Koklic, T., Arsov, Z., Filipic, B., Stopar, D. & Hemminga, M. A. (2005). Spin label EPR-based characterization of biosystem complexity. *J Chem Inf Model* 45: 394-406.
- Timofeev, V. P. & Tsetlin, V. I. (1983). Analysis of mobility of protein side chains by spin-label technique. *Biophysics of Structure and Mechanisms* 10: 93-108.
- Tomba, P. (2002). Intrinsically unstructured proteins. *Trends Biochem Sci* 27: 527-533.
- Tomba, P. & Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci* 33: 2-8.
- Tsai, C. D., Ma, B., Kumar, S., Wolfson, H. & Nussinov, R. (2001). Protein folding: binding of conformationally fluctuating building blocks via population selection. *Crit Rev Biochem Mol Biol* 36: 399-433.
- Tsai, C. J., Ma, B., Sham, Y. Y., Kumar, S. & Nussinov, R. (2001). Structured disorder and conformational selection. *Proteins: Structure, Function and Bioinformatics* 44: 418-427.
- Turoverov, K. K., Kuznetsova, I. M. & Uversky, V. N. (2010). The protein kingdom extended: ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. *Prog Biophys Mol Biol* 102: 73-84.
- Uversky, V. N. (1993). Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. *Biochemistry* 32: 13288-13298.

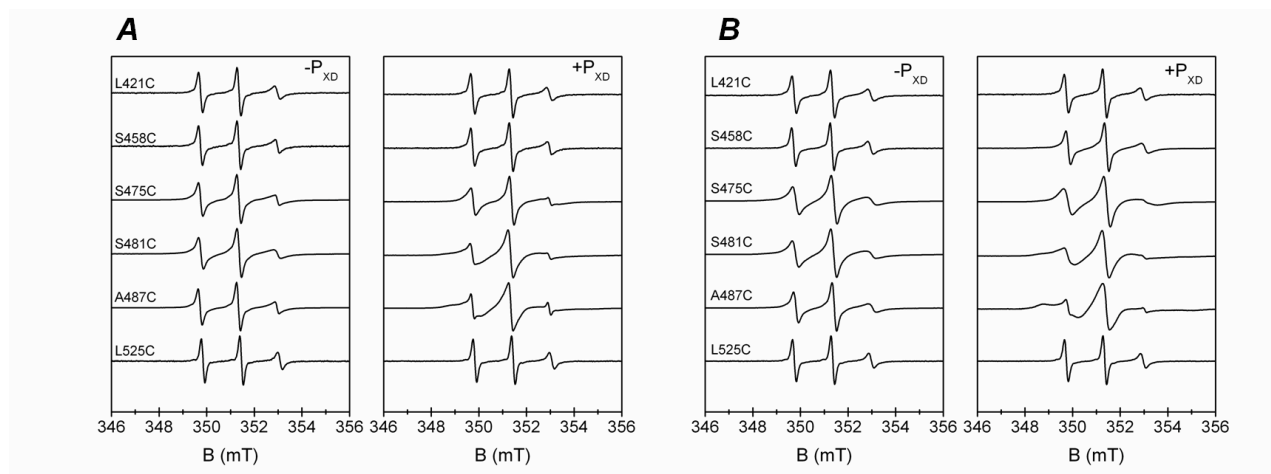
- Uversky, V. N. (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* 11: 739-756.
- Uversky, V. N. (2002). What does it mean to be natively unfolded? *Eur J Biochem* 269: 2-12.
- Uversky, V. N. (2003). Protein folding revisited. A polypeptide chain at the folding-misfolding-nonfolding cross-roads: which way to go? *Cell Mol Life Sci* 60: 1852-1871.
- Uversky, V. N. (2010). The mysterious unfoldome: structureless, underappreciated, yet vital part of any given proteome. *J Biomed Biotechnol* 2010: 568068.
- Uversky, V. N. & Dunker, A. K. (2010). Understanding protein non-folding. *Biochim Biophys Acta* 1804: 1231-1264.
- Uversky, V. N., Oldfield, C. J. & Dunker, A. K. (2005). Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J Mol Recognit* 18: 343-384.
- Vacic, V., Oldfield, C. J., Mohan, A., Radivojac, P., Cortese, M. S., Uversky, V. N. & Dunker, A. K. (2007). Characterization of molecular recognition features, MoRFs, and their binding partners. *J Proteome Res* 6: 2351-2366.
- Wang, L. F., Yu, M., Hansson, E., Pritchard, L. I., Shiell, B., Michalski, W. P. & Eaton, B. T. (2000). The exceptionally large genome of Hendra virus: support for creation of a new genus within the family Paramyxoviridae. *J Virol* 74: 9972-9979.



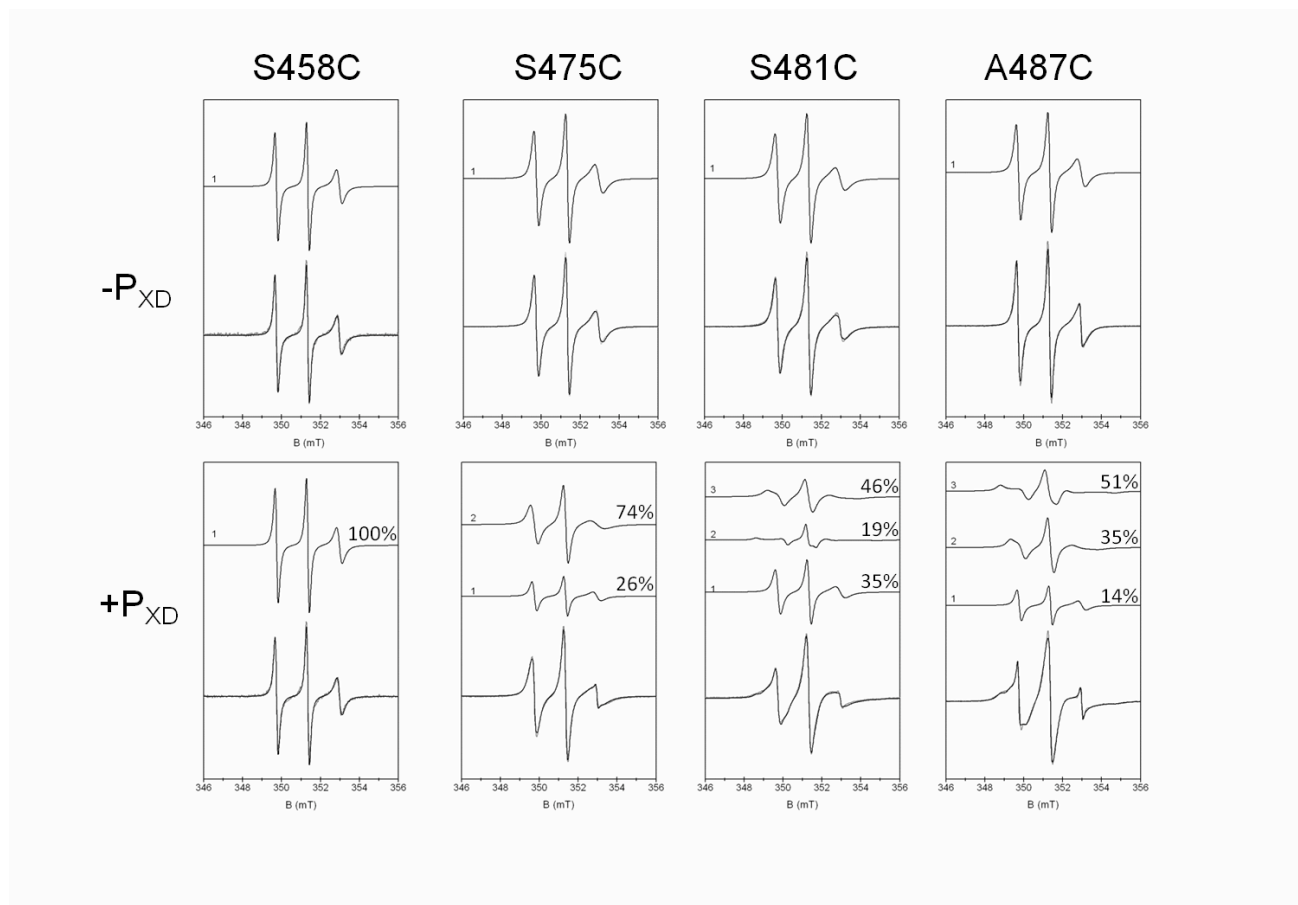
**Figure 1. Design and purification of cysteine N<sub>TAIL</sub> variants.** (A) Sequence alignment between the HeV and NiV N<sub>TAIL</sub> proteins and structural organization of *Henipavirus* N<sub>TAIL</sub>. Hydrophobic residues (L, I, V, F, W and M) and charged residues are shown in grey. Asterisks and colons below the alignment denote identity and similarity, respectively, while dots indicate weakly similar residues. Lack of symbol below the alignment denotes strongly different residues. The front numbers correspond to the amino acid position in the N sequence. Dots above the alignment indicate intervals of 10 residues. The four boxes corresponding to MoREs (Habchi, Mamelli et al. 2010; Blocquel, Habchi et al. 2012) are shown. The positions targeted for cysteine substitution and spin labeling are shaded in grey in the sequence alignment (top) and highlighted by arrows in the schematic representation of the structural organization of *Henipavirus* N<sub>TAIL</sub> domains (bottom). (B) Coomassie blue staining of a 12% SDS-PAGE loaded with approximately 5 µg of N<sub>TAIL</sub> variants purified from the soluble fraction of *E. coli*. Note that 20 mM DTT was added to the samples before electrophoretic migration.



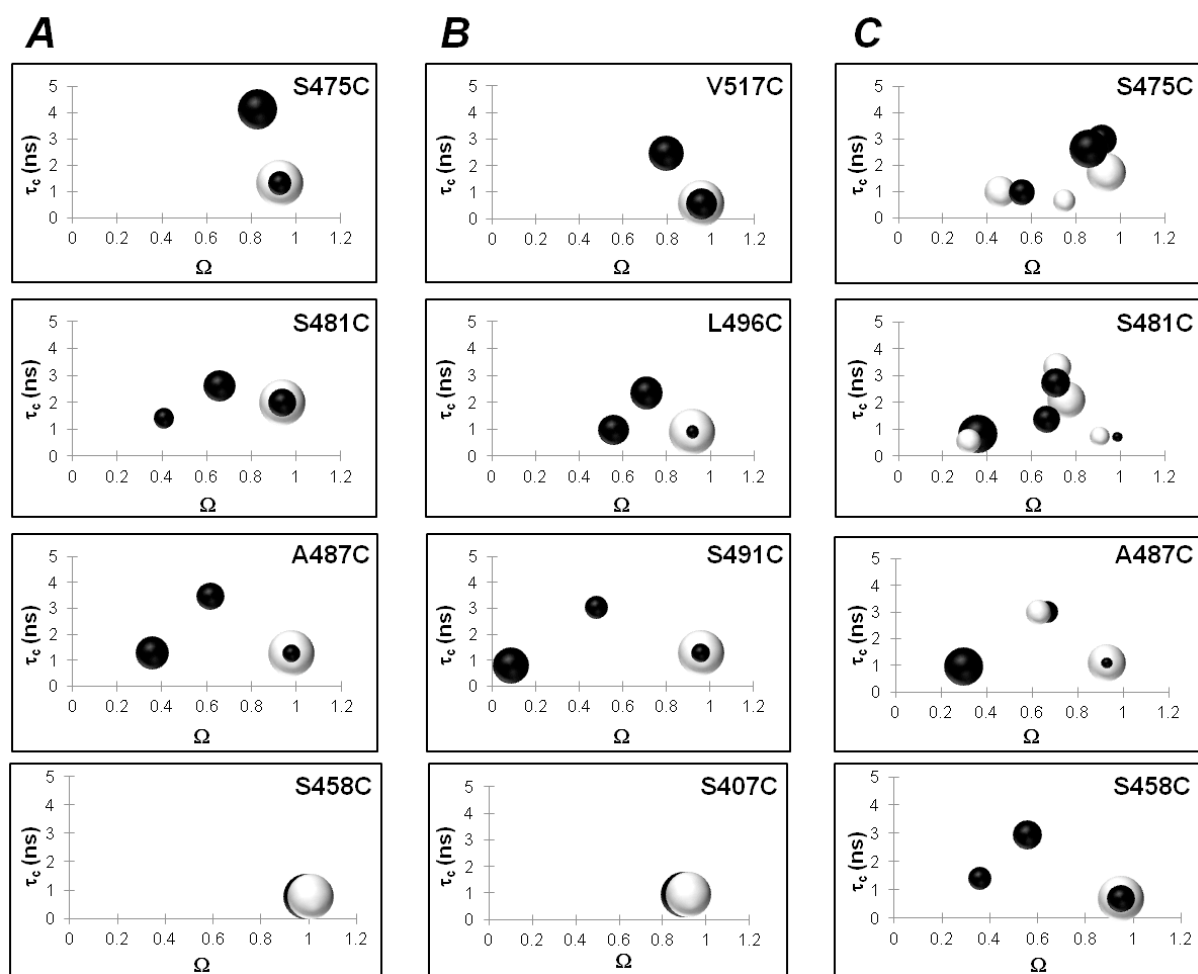
**Figure 2. Effect of TFE on spin-labeled HeV and NiV N<sub>TAIL</sub> proteins monitored by far-UV CD and EPR spectroscopies.** (A,B) CD spectra of spin labeled HeV A487C (A) and NiV A487C (B) variants representative of the other labeled variants (see Fig. S2) at 0% and 20% TFE (v/v). (C, D)  $h(+1)/h(0)$  ratios, corresponding to the peak-to-peak amplitude of the low-field over the central line of the EPR spectra, of the spin labeled HeV (C) and NiV (D) N<sub>TAIL</sub> variants at 0% and 20% (v/v) TFE.



**Figure 3. Effect of  $P_{XD}$  on labeled  $N_{TAIL}$  variants.** Amplitude normalized room temperature EPR spectra of the spin labeled HeV (**A**) and NiV (**B**)  $N_{TAIL}$  variants in the presence of 30% (v/v) sucrose and either in the absence or in the presence of saturating amounts of  $P_{XD}$ .

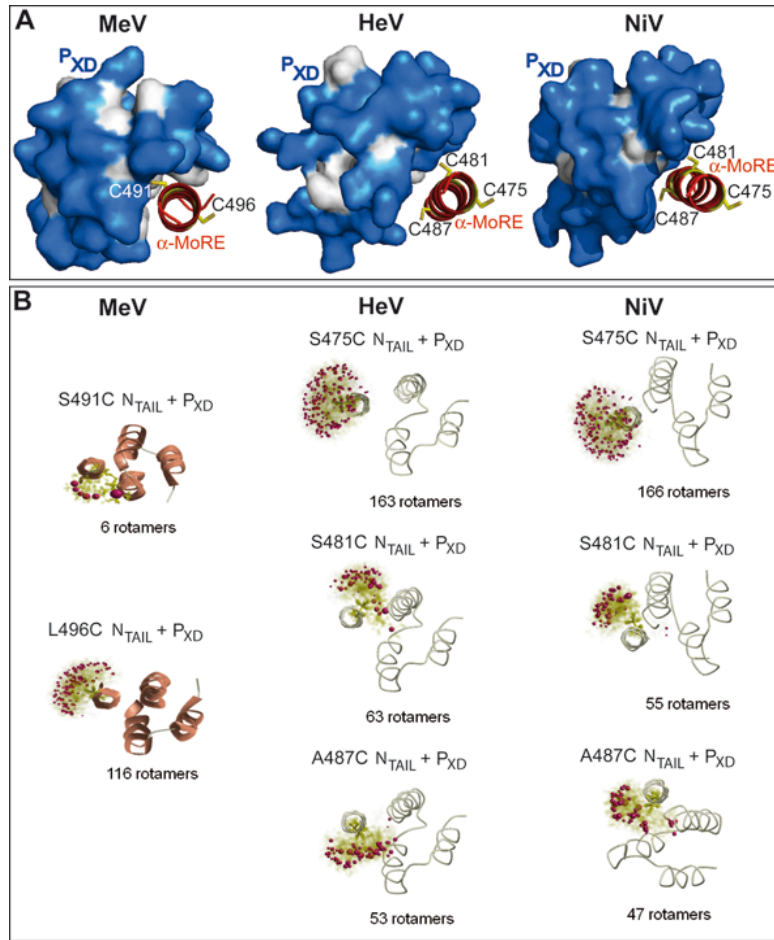


**Figure 4. Simulation of the RT EPR spectra of the HeV S458C, S475C, S481C and A487C spin labeled  $N_{TAIL}$  variants in the presence of sucrose (30% v/v) and in the absence or presence of saturating amounts of  $P_{XD}$ .** Simulation was performed using the EPRSIM-C software program (Strancar, Koklic et al. 2005). The individual simulated components are given at the top of each panel with the indication of their respective proportion (see **Table 1**). **At the bottom of each panel are shown the simulated spectra (black lines) resulting from the addition of each individual simulated component, superimposed on the experimental spectra (grey lines).**



**Figure 5.  $[\tau_c\text{-}\Omega]$  plots of HeV, MeV and NiV spin labeled  $N_{\text{TAIL}}$  variants.** Parameters extracted from the simulation of RT EPR spectra of the HeV (A), MeV (B) and NiV (C) spin labeled  $N_{\text{TAIL}}$  variants in the presence of 30% (v/v) sucrose and in the absence or presence of saturating amounts of  $P_{\text{XD}}$ . The parameters are given as  $[\tau_c\text{-}\Omega]$  plots, with  $\tau_c$  being the effective rotational correlation time (in ns) and  $\Omega$  the normalized free rotational space, in the absence (white spheres) and presence (black spheres) of the partner  $P_{\text{XD}}$ . The surface of the spheres is proportional to the relative contribution of the EPR spectral shape components. For the sake of clarity, when the positions of the black spheres overlapped those of the white spheres, the former have been slightly shifted (case of HeV S458C, and of NiV S407C and A487C variants, see **Table 1** for parameters).





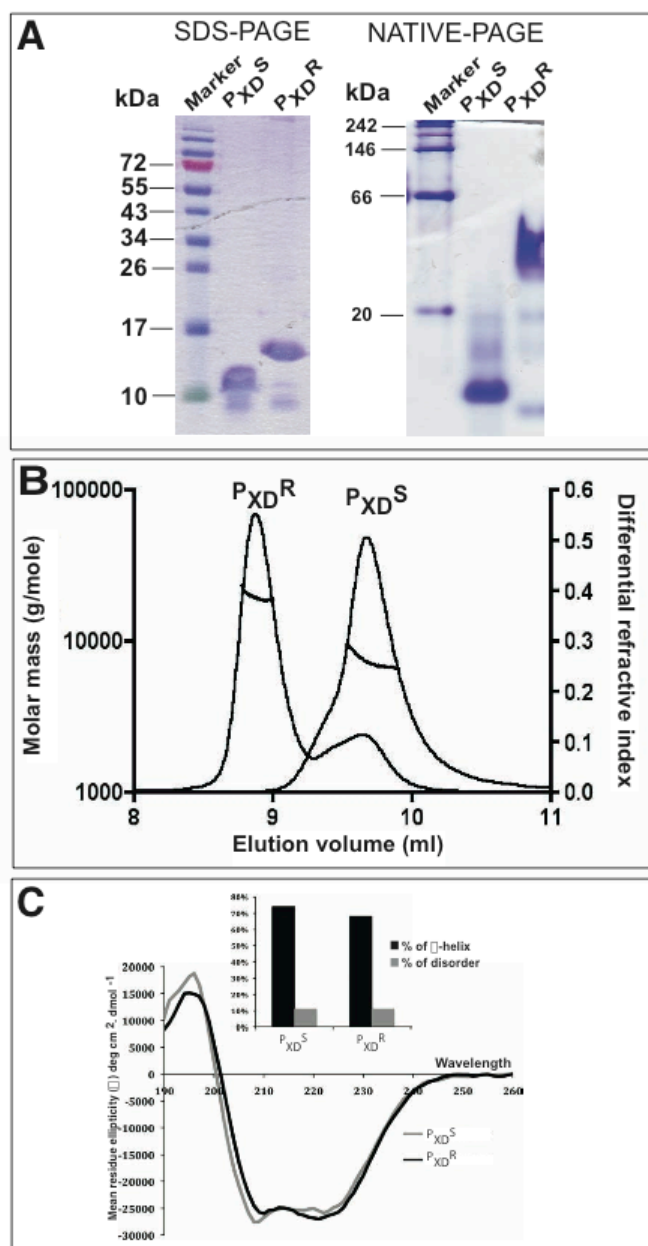
**Figure 6. Structural models and calculated rotamers.** **(A)** Crystal structure of the MeV chimera between P<sub>XD</sub> and the N<sub>TAIL</sub> region encompassing residues 486–504 (pdb code 1T60) (Kingston, Hamel et al. 2004), and structural models of the *Henipavirus* complexes between P<sub>XD</sub> and the Box3 region (residues 473–493) of N<sub>TAIL</sub>. P<sub>XD</sub> is shown with a surface representation, whereas the N<sub>TAIL</sub> region is represented as a ribbon. Hydrophobic P<sub>XD</sub> residues are shown in light grey. In the MeV complex, residues 491 and 496 have been replaced by a cysteine and these side chains are shown in sticks. The side chain of cysteine 491 points towards the surface of P<sub>XD</sub>, while that of cysteine 496 is solvent-exposed. In the structural models of the HeV and NiV complexes the native N<sub>TAIL</sub> residues targeted for cysteine mutagenesis and spin labeling have been replaced by a cysteine. The side chains of these cysteines are shown in sticks. Note that in the structural models, the side chain of residue 475 is solvent-exposed, while that of residues 481 and 487 is closer to the P<sub>XD</sub> surface. The pictures were drawn using Pymol (DeLano 2002). **(B)** Calculated rotamers of the spin labels of MeV, HeV and NiV N<sub>TAIL</sub> variants. Ribbon representation of the complexes of MeV (pdb code 1T60) (Kingston, Hamel et al. 2004), HeV and NiV N<sub>TAIL</sub>-P<sub>XD</sub> (Habchi, Blangy et al. 2011) showing, in semi-transparent sticks and balls, the calculated spin label rotamers attached to the engineered cysteines at positions 491 and 496 (MeV) and 475, 481 and 487 (HeV and NiV). The spin label rotamers were computed using the software MMM. The balls correspond to the baricenter of the N-O bond with the size of the ball being proportional to the abundance of the corresponding rotamer.

**Table 1.** Parameters extracted from the simulation of the RT EPR spectra of MeV, HeV and NiV N<sub>TAIL</sub> variants in the presence of 30% (v/v) sucrose and either in the absence or in the presence of saturating amounts of P<sub>XD</sub>.

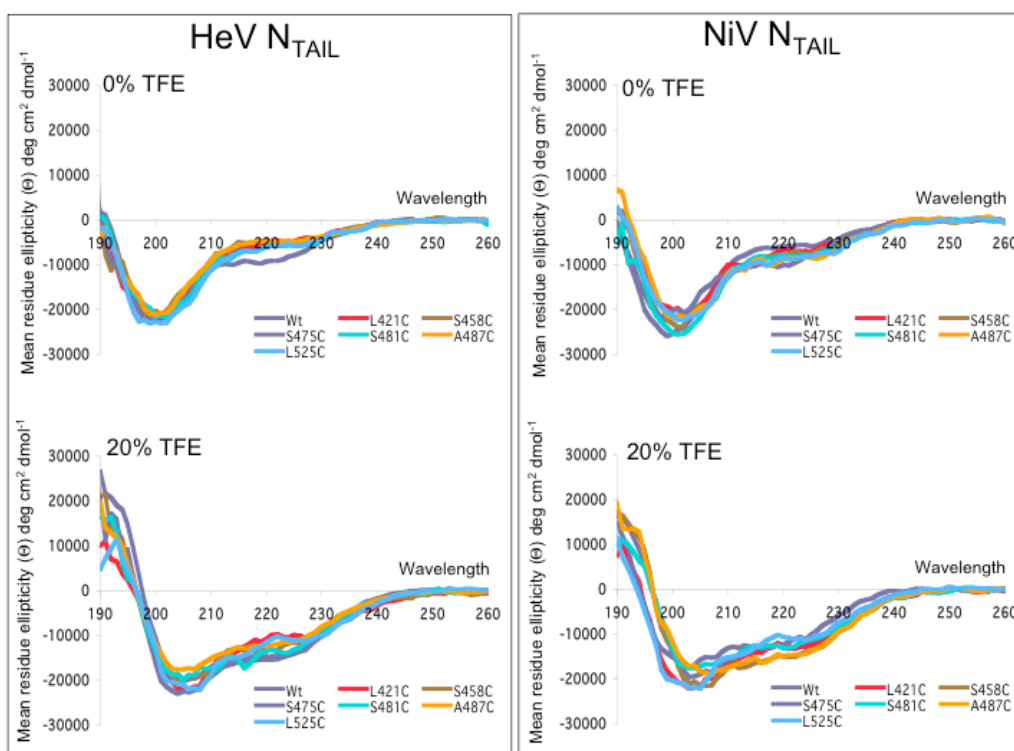
	Variant	Component	- P <sub>XD</sub>			+ P <sub>XD</sub>		
			$\tau_C$ (ns)	$\Omega$	%	$\tau_C$ (ns)	$\Omega$	%
MeV	S407C	1	0.9(0.2)	0.93(0.05)	100	0.9(0.2)	0.93(0.05)	100
	S491C	1	1.3(0.2)	0.96(0.05)	100	1.3(0.2)	0.96(0.05)	15
		2				3.0(0.6)	0.48(0.05)	24
		3				0.8(0.2)	0.09(0.05)	61
	L496C	1	0.9(0.2)	0.92(0.05)	100	0.9(0.2)	0.92(0.05)	7
		2				1.0(0.2)	0.56(0.05)	44
		3				2.3(0.4)	0.71(0.05)	49
	V517C	1	0.6(0.2)	0.93(0.05)	100	0.6(0.2)	0.93(0.05)	44
		2				2.4(0.4)	0.80(0.05)	56
HeV	L421C	1	0.8(0.2)	0.93(0.05)	100	0.8(0.2)	0.93(0.05)	100
	S458C	1	0.8(0.2)	0.99(0.05)	100	0.8(0.2)	0.99(0.05)	100
	S475C	1	1.3(0.2)	0.93(0.05)	100	1.3(0.2)	0.93(0.05)	26
		2				4.1(0.8)	0.83(0.05)	74
	S481C	1	2.0(0.4)	0.94(0.05)	100	2.0(0.4)	0.94(0.05)	35
		2				1.4(0.2)	0.41(0.05)	19
		3				2.6(0.4)	0.66(0.05)	46
	A487C	1	1.2(0.2)	0.98(0.05)	100	1.2(0.2)	0.98(0.05)	14
		2				3.5(0.6)	0.62(0.05)	35
		3				1.3(0.2)	0.36(0.05)	51
	L525C	1	0.5(0.1)	0.98(0.05)	100	0.5(0.2)	0.98(0.05)	100
NiV	L421C	1	0.8(0.2)	0.92(0.05)	100	0.8(0.2)	0.92(0.05)	100
	S458C	1	0.7(0.2)	0.95(0.05)	100	0.7(0.2)	0.95(0.05)	36
		2				1.4(0.2)	0.36(0.05)	25
		3				3.0(0.6)	0.56(0.05)	39
	S475C	1	1.7(0.2)	0.94(0.05)	52	3.0(0.6)	0.92(0.05)	29
		2	0.6(0.1)	0.75(0.05)	16	2.6(0.5)	0.86(0.05)	49
		3	1.0(0.2)	0.46(0.05)	32	1.0(0.2)	0.56(0.05)	22
	S481C	1	0.7(0.2)	0.91(0.05)	12	0.7(0.2)	0.99(0.05)	3
		2	3.3(0.6)	0.72(0.05)	25	2.7(0.5)	0.71(0.05)	26
		3	0.5(0.1)	0.32(0.05)	18	0.8(0.2)	0.36(0.05)	48
		4	2.1(0.2)	0.76(0.05)	45	1.4(0.2)	0.67(0.05)	23
	A487C	1	1.1(0.2)	0.93(0.05)	70	1.1(0.2)	0.93(0.05)	6
		2	3.0(0.6)	0.67(0.05)	30	3.0(0.6)	0.67(0.05)	22
		3				1.0(0.2)	0.30(0.05)	72
	L525C	1	0.6(0.2)	0.96(0.05)	100	0.6(0.2)	0.96(0.05)	100

\* P<sub>XD</sub><sup>S</sup> for NiV N<sub>TAIL</sub> proteins and P<sub>XD</sub><sup>R</sup> for HeV and MeV N<sub>TAIL</sub> proteins.

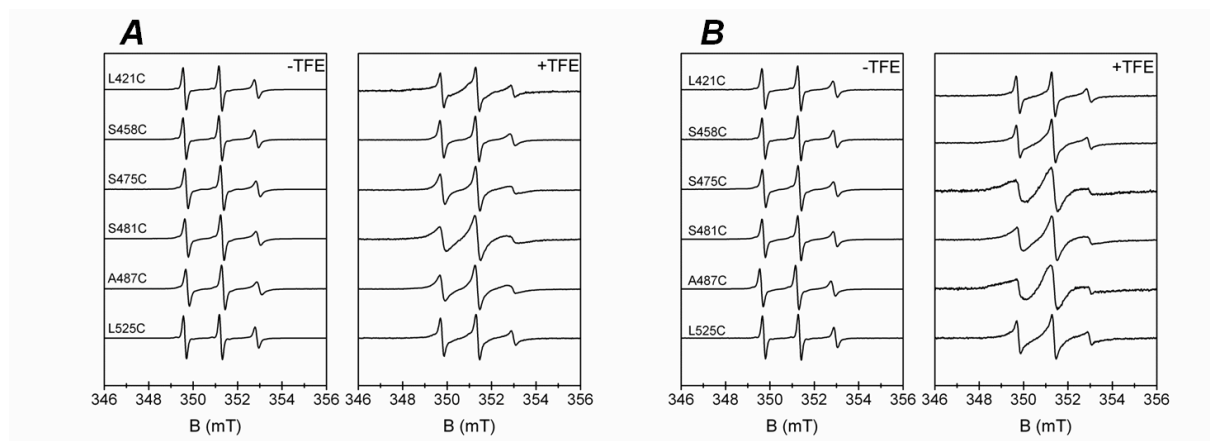
## SUPPLEMENTARY FIGURES



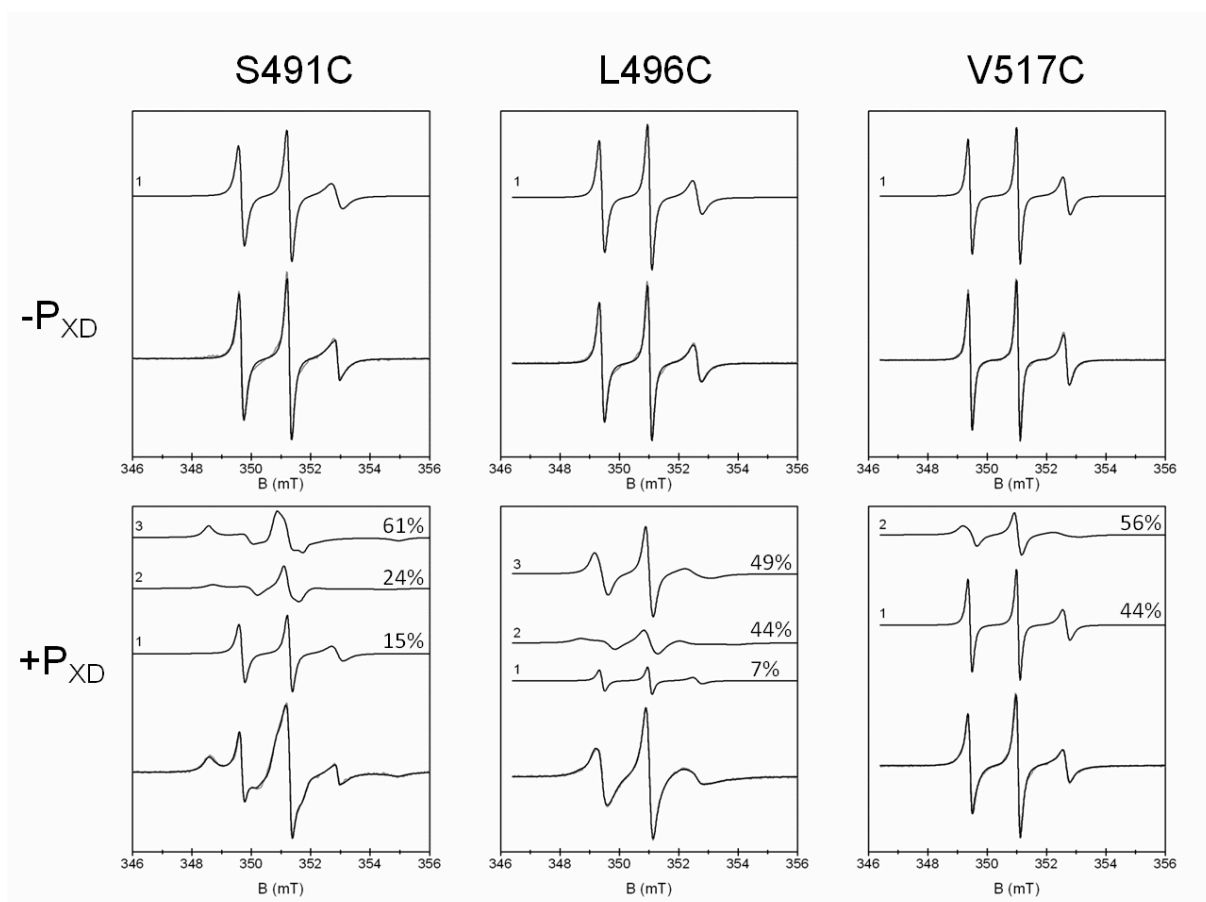
**Figure S1. Comparison between the conformational and spectroscopic properties of recombinant and synthetic NiV P<sub>XD</sub>.** **(A)** Coomassie blue staining of a 18% SDS-PAGE or native PAGE loaded with purified recombinant and synthetic NiV P<sub>XD</sub> (approximately 5  $\mu$ g each). **(B)** SEC-MALLS-RI analysis of the recombinant (P<sub>XD</sub><sup>R</sup>) and synthetic NiV P<sub>XD</sub> (P<sub>XD</sub><sup>S</sup>) samples. The left y axis represents the molecular mass and the right y axis represents the differential refractive index. The horizontal traces show the molecular masses calculated from light-scattering intensity at different angles and differential refractive indexes as a function of the elution volume. Samples were injected at either 1.2 mM (P<sub>XD</sub><sup>R</sup>) or 1.48 mM (P<sub>XD</sub><sup>S</sup>). **(C)** Far-UV CD spectra of recombinant and synthetic NiV P<sub>XD</sub> (each at 0.1 mg/mL) in 10 mM sodium phosphate pH 7. Spectra were recorded at 20°C. Data are representative of one out of two independent acquisitions.



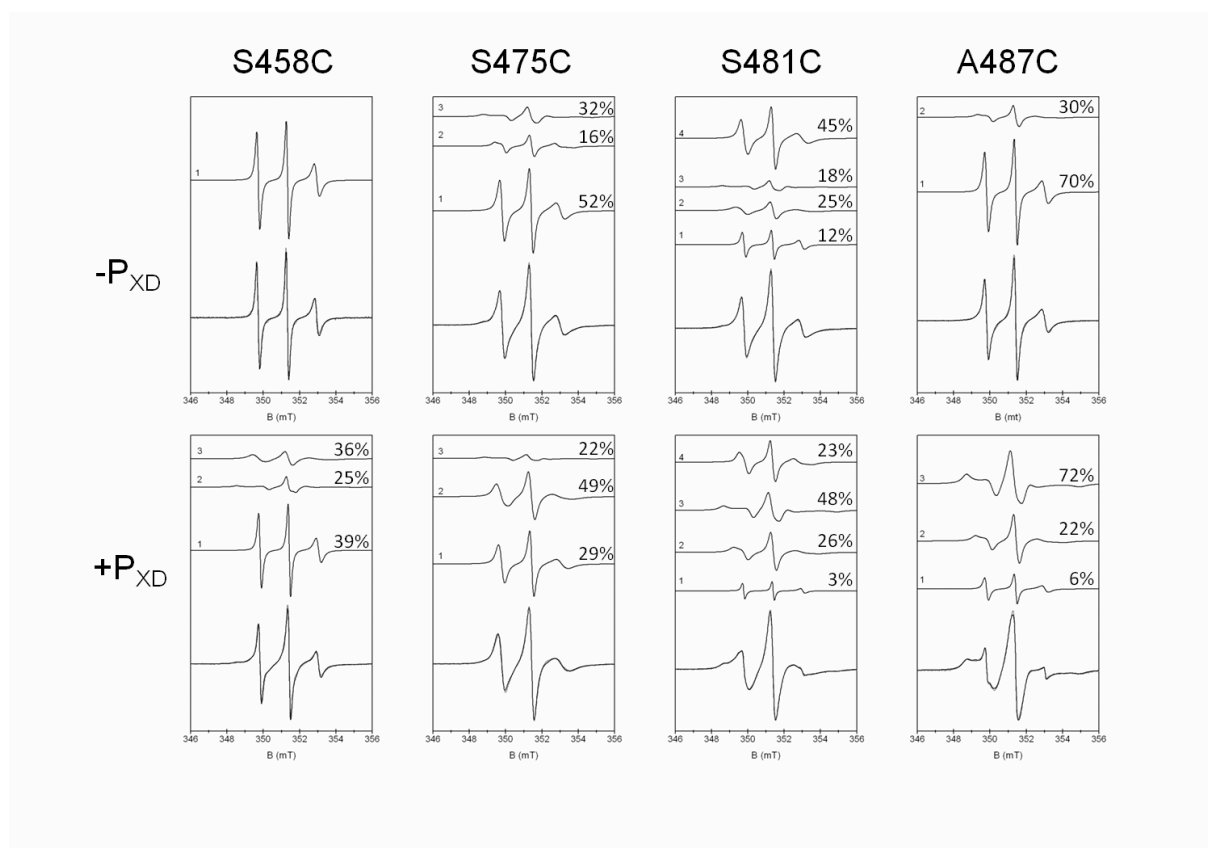
**Figure S2. Far-UV CD studies of parental and spin labeled cysteine  $N_{TAIL}$  variants.** CD spectra of parental and spin labeled HeV (**A**) and NiV (**B**)  $N_{TAIL}$  proteins at 0% and 20% TFE (v/v). Each spectrum is the mean of three independent acquisitions.



**Figure S3. Effect of TFE on spin label mobility.** Amplitude normalized room temperature EPR spectra of the spin labeled HeV **(A)** and NiV **(B)** N<sub>TAIL</sub> proteins at 0% or 20% TFE (v/v).



**Figure S4. Simulation of the RT EPR spectra of the MeV S491C, L496C and V517C spin labeled N<sub>TAIL</sub> variants in the presence of sucrose (30% v/v) and in the absence or presence of saturating amounts of P<sub>XD</sub>.** Simulation was performed using the EPRSIM-C software program (Strancar, Koklic et al. 2005). The individual simulated components are given at the top of each panel with the indication of their respective proportion. **At the bottom of each panel are shown the simulated spectra (black lines) resulting from the addition of each individual simulated component, superimposed on the experimental spectra (grey lines).**



**Figure S5. Simulation of the RT EPR spectra of the NiV S458C, S475C, S481C and A487C spin labeled N<sub>TAIL</sub> variants in the presence of sucrose (30% v/v) and in the absence or presence of saturating amounts of P<sub>XD</sub>.** Simulation was performed using the EPRSIM-C software program (Strancar, Koklic et al. 2005). The individual simulated components are given at the top of each panel with the indication of their respective proportion. **At the bottom of each panel are shown the simulated spectra (black lines) resulting from the addition of each individual simulated component, superimposed on the experimental spectra (grey lines).**