Conformational Analysis of the Partially Disordered Measles Virus N$_{TAIL}$-XD Complex by SDSL EPR Spectroscopy

Aleh Kavalenka,[†] Iztok Urbančič,[†] Valérie Belle,[‡] Sabrina Rouger,[§] Stéphanie Costanzo,[§] Sandra Kure,[†] André Fournel,[‡] Sonia Longhi,[§] Bruno Guigliarelli,[‡] and Janez Strancar[†]*

[†]Laboratory of Biophysics, Solid State Physics Department, Josef Stefan Institute, Ljubljana, Slovenia; [‡]Bioénergétique et Ingénierie des Protéines, UPR 9036 Centre National de la Recherche Scientifique et Universités Aix-Marseille I et II, Marseille, France; and [§]Architecture et Function des Macromolécules Biologiques, UMR 6098 Centre National de la Recherche Scientifique et Universités d'Aix-Marseille I et II, Marseille, France

ABSTRACT   To characterize the structure of dynamic protein systems, such as partly disordered protein complexes, we propose a novel approach that relies on a combination of site-directed spin-labeled electron paramagnetic resonance spectroscopy and modeling of local rotation conformational spaces. We applied this approach to the intrinsically disordered C-terminal domain of the measles virus nucleoprotein (N$_{TAIL}$) both free and in complex with the X domain (XD, aa 459–507) of the viral phosphoprotein. By comparing measured and modeled temperature-dependent restrictions of the side-chain conformational spaces of 12 SL cysteine-substituted N$_{TAIL}$ variants, we showed that the 490–500 region of N$_{TAIL}$ is prestructured in the absence of the partner, and were able to quantitatively estimate, for the first time to our knowledge, the extent of the α-helical sampling of the free form. In addition, we showed that the 505–525 region of N$_{TAIL}$ conserves a significant degree of freedom even in the bound form. The latter two findings provide a mechanistic explanation for the reported rather high affinity of the N$_{TAIL}$-XD binding reaction. Due to the nanosecond timescale of X-band EPR spectroscopy, we were also able to monitor the disordering in the 488–525 region of N$_{TAIL}$, in particular the unfolding of the α-helical region when the temperature was increased from 281 K to 310 K.

INTRODUCTION

Intrinsically disordered proteins (IDPs) are ubiquitary proteins that fulfill essential biological functions while being devoid of highly populated secondary and tertiary structure under physiological conditions and in the absence of a partner (1). For such proteins, whose function relies on a disordered state, a high-resolution static description is not adequate. Indeed, IDPs consist of dynamic ensembles of interconverting conformers that exert their biological function(s) by recognizing their binding partners through their disordered regions (1–12). Although some IDPs can carry out their function while remaining permanently disordered (e.g., entropic chains) (2), many of them undergo induced folding, i.e., a disorder-to-order transition upon binding to their physiological partners (6,11,13). Because of their inherent flexibility, IDPs generally fail to be crystallized in the absence of their partner(s). In those rare cases where crystallization of the free form is successful, it only leads to a single conformation snapshot that is not representative of the whole conformational ensemble (14).

Therefore, the dynamic description of such loosely packed proteins requires the use of experimental techniques endowed with a time window corresponding to the typical nanosecond-to-microsecond timescale of protein dynamics (defined as the lifetime of a given conformation of a backbone segment). Indeed, all of the dynamic events that take place over a timescale that is lower than the time window of the technique employed are averaged over their allowed states, and thus escape detection. Site-directed spin-labeled (SDSL) electron paramagnetic resonance (EPR) spectroscopy, which has a nanosecond time window (15), is an appropriate experimental approach. It relies on the covalent grafting of a nitroxide spin label at a selected protein site, followed by analysis of the EPR spectra of the SL protein under various conditions, including the presence of a partner protein.

Although the nanosecond EPR time window results in averaging of the local subnanosecond side-chain motions, this approach can be used to monitor backbone motions (16) or primarily the local rotational restrictions of an SL side chain (17); however, it fails to directly describe the local protein main-chain conformation. Therefore, molecular modeling, which is able to scan probable side-chain conformations over a few nanoseconds, must be used to infer structural conformations from SL motions. Notably, EPR sensitivity to the anisotropy of SL rotational motion, i.e., reorientation of the nitroxide group in space, can be used to accelerate modeling.

Here, we propose what we believe to be a novel approach for protein structure analysis that relies on a combination of SDSL EPR spectroscopy and modeling of local rotation conformational spaces. We applied this approach to a partly disordered model system, i.e., the intrinsically disordered C-terminal domain of the measles virus nucleoprotein (aa 401–525, N$_{TAIL}$) (18) both alone and in complex with the X domain of the viral phosphoprotein (aa 459–507, XD). The crystal structure of XD has been solved, and XD has been shown to trigger α-helical folding of N$_{TAIL}$ (19).

An α-helical molecular recognition element (α-MoRE) located within a conserved region of N_{TAIL} (Box2, aa 489–506) is required for this structural transition (20,21–24).

The structural data available to date were obtained from crystallographic studies of a chimeric construct consisting of XD and the 486–504 region of N_{TAIL} (25), as well as from a low-resolution, small-angle x-ray scattering (SAXS) model of the complex between XD and the entire N_{TAIL} domain (12) (see Fig. 1 A). This model revealed that most of N_{TAIL} (residues 401–488) remains disordered in the complex and does not establish contacts with XD, in contrast to the α-MoRE and the C-terminal region. Involvement of the C-terminal Box3 (aa 517–525) in binding to XD is further supported by surface plasmon resonance studies in which removal of this region resulted in a significant drop in the binding affinity (12).

We recently reported two studies in which we investigated the N_{TAIL}-XD interaction by SDSL EPR spectroscopy (26,27). In those studies, we targeted 14 sites scattered within N_{TAIL}, 12 of which are located within the proposed region of interaction with XD (aa 488–525; Fig. 1 B) (26,27). We showed that the mobility of the SLs grafted within the 488–502 region is reduced even in the absence of the partner, consistent with the existence of a transiently populated α-helix. The mobility of the SLs grafted within this region is dramatically reduced upon binding to XD, and the spin label grafted onto position 491 experiences a highly restricted mobility due to its location at the XD interface (see Fig. 1 A). In addition, we showed that XD triggers α-helical folding of the 488–502 region only, in contrast to the 505–522 region, which is only moderately affected by XD and does not gain α-helicity. Strikingly, for the latter region, the secondary-structure stabilizer trifluoroethanol (TFE) was shown to trigger reductions in spin-label mobility similar to those observed in the presence of XD, thus supporting a lack of direct contacts between Box3 and XD. The latter was also

recently confirmed by heteronuclear single quantum coherence experiments using ^{15}N-labeled XD and a Box3 peptide (28). Based on these data, we proposed a model in which the region downstream from the α-MoRE participates in the N_{TAIL}-XD binding reaction by 1), stabilizing the α-helical conformation of Box2 through the establishment of transient tertiary contacts with the latter; or 2), reducing the entropic penalty of the binding reaction thanks to its disordered nature even upon complex formation (27,28).

Although the N_{TAIL}-XD interaction has been the focus of numerous studies, the actual conformation of the region downstream from the α-MoRE in the complex is not known, and no structural description of the free form of N_{TAIL} is available. In particular, quantitative information on the extent of α-helical sampling of the α-MoRE in the absence of the partner is still lacking.

In this work, using a novel approach that allows structural information to be inferred by analysis of multiple SDSL EPR experimental data (29), we provide a structural description of the dynamic behavior of N_{TAIL} in both free and bound forms. By comparing the modeled and experimentally derived anisotropies of local rotational motions (17), as obtained with 12 SL N_{TAIL} variants (see Fig. 1 B), we were able to optimize backbone dihedral angles and infer probable nanosecond-stable N_{TAIL} secondary structures. Using this approach, we showed that the 490–500 region of N_{TAIL} is partly prestructured in the absence of the partner, and were able to quantitatively analyze for the first time, to our knowledge, the level of α-helical sampling within this region.

MATERIALS AND METHODS

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

Expression, purification, and spin labeling of cysteine-substituted N_{TAIL} variants were carried out as previously described (26,27). Expression and
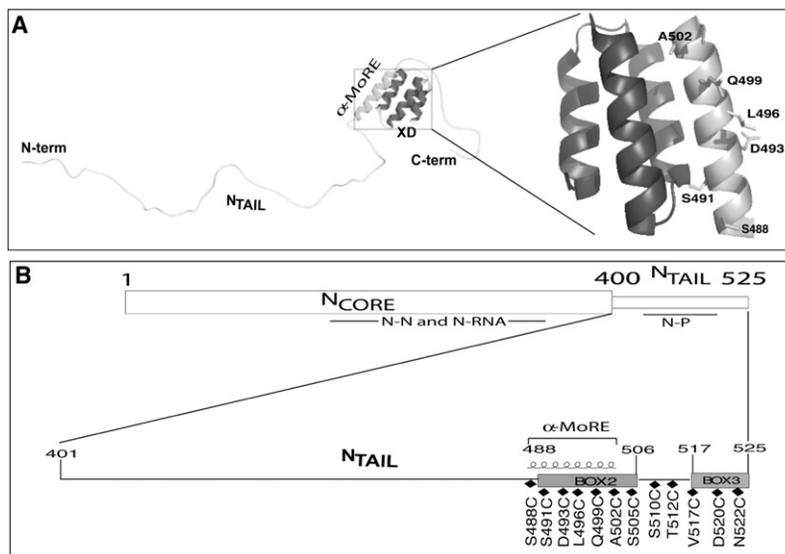
FIGURE 1 (A) Model of the N_{TAIL}-XD complex as derived from SAXS studies highlighting the role of the α-MoRE in the interaction with XD (12). The structure of the chimera between XD and the N_{TAIL} region encompassing residues 486–504 (PDB code: 1T60) is zoomed in, and the side chains of residues targeted for cysteine substitution and spin labeling are shown. The picture was drawn using PyMOL (DeLano Scientific, San Carlos, CA; http://www.pymol.org). (B) Organization of the N protein and schematic representation of positions targeted for cysteine substitution and spin labeling (diamonds) in the 12 SL N_{TAIL} variants selected for this study. The two regions of homology conserved in Morbillivirus members (Box2 and Box3) are also shown, as are the sites of interaction with N, with RNA and with P.

purification of XD were performed according to Johansson et al. (19). All samples were stored at 193 K. Protein concentrations were either calculated using OD_{280} measurements and the theoretical absorption coefficients ε (mg/mL.cm) at 280 nm, as obtained using the program ProtParam at the EXPASY server (http://www.expasy.ch/tools, Swiss Institute of Bioinformatics, Lausanne, Switzerland), or measured using the Biorad protein assay (Bio-Rad Laboratories, Hercules, CA).

EPR spectroscopy

EPR spectra were acquired using a liquid-N_2 cryostat and Elexsys E500 Bruker spectrometer (Bruker, Karlsruhe, Germany) at 9.3 GHz of 20 mW power. The magnetic field modulation frequency and amplitude were 100 kHz and 0.15 mT, respectively.

The concentration of the SL N_{TAIL} proteins was 100 μM. An XD molar excess was used to ensure 100% complex formation (26,27). All samples were recorded in the absence and presence of 30% v/v sucrose to decrease the contribution of protein rotation to the EPR spectral line shape.

Before the EPR spectra were recorded, each sample (20 μL) was preequilibrated in a water bath at 296 K for 10 min and then quickly transferred into a quartz capillary. Its EPR spectrum was measured at 296 K. Each sample was then cooled down to 279 K and equilibrated for 8 min before its EPR spectrum was measured at 279 K. Then the sample was heated up in two successive steps of 2 K each and equilibrated 3 min at each step before its EPR spectra were acquired at 281 K and 283 K. Afterward, the sample was heated up to 308 K and equilibrated for 8 min before EPR measurements were obtained at 308 K. Finally, the sample was heated up in two additional steps of 2 K each and equilibrated 3 min at each step before its EPR spectra were recorded at 310 K and 312 K. The temperature stability and accuracy were 0.3 K and 0.2 K, respectively. The number of acquisitions was adjusted to obtain a satisfactory signal/noise (S/N) ratio. Experimental and simulated EPR spectra are shown in Fig. S3, Fig. S4, and Fig. S5 of the Supporting Material.

Data analysis by motional pattern condensation

The experimental spectra were fitted within the fast restricted rotational motion approximation with two, three, or four spectral components depending on whether the S/N ratio was <100, 100–150, or >150, respectively (17). The spectral parameters $\{\vartheta, \varphi, \tau_c, W, p_A, prot\}$ of each component of the simulated spectra were simultaneously optimized with a 20-run hybrid evolutionary algorithm (dHEO) (30–32) for each of the 336 experimental spectra, followed by condensation with GHOST, an algorithm that filters and groups multiple solutions found during spectral optimization (31–33) (see Fig. S1). For each sample, all of the condensed solutions in terms of motional patterns were plotted as a function of the temperature, allowing identification of significant motional patterns. Small isolated patterns, as well as patterns that exhibited no similarity to patterns at nearby temperatures, were discarded. In addition, all patterns with a weight below 10% were also discarded from further analyses because they could not significantly contribute to the main results.

Finally, the local temperature sets, recorded in the ranges of 279 K–283 K and 308 K–312 K, were used to obtain average condensed solution pattern sets representative of the low (281 K) and high (310 K) temperatures. The experimental rotational restrictions, which restrict particular side-chain conformations, were then obtained from EPR spectra simulations in terms of free rotational space Ω_{exp} by analogy to Eq. 2 (see Fig. S1) and normalized rotational diffusion D_{exp} positional dependences (33) by Eq. 4 for D_{exp}.

Molecular modeling and structure optimization

Here, we propose a modeling procedure that allows global structural information to be inferred from local rotational restrictions data experimentally obtained by SDSL EPR in combination with motional pattern analysis. The overall procedure is schematically illustrated in Fig. S1. Two major approximations were made: 1), slower-than-nanosecond backbone motions

were approximated to be frozen in the nanosecond time window, whereas faster-than-nanosecond side-chain motions were averaged; and 2), EPR spectroscopy was taken to be sensitive only to the anisotropy of the local side-chain rotational motion, rather than to the exact atomic coordinates of the nitroxide. In this way, the SDSL experimental data were explained in terms of local rotational restrictions to guide a five-step optimization of nanosecond-stable protein structures, as described below.

Step 1: Calculation of protein secondary structure

Since the mass attached to the backbone measured from its closest terminal end is much larger than the mass of the attached side chain, the amplitude of rotational diffusion is expected to be much lower for the backbone. This approximation applies much better to positions that are far from the terminal ends than to those in the proximity of the termini. Considering the timescale of side-chain motions, this means that the side-chain rotational space is a priori not affected by the backbone rotational motion. Hence, the backbone motion can be approximated to be virtually frozen. Moreover, this approximation may be extended to backbone terminal ends if the free wobbling of the terminal ends is slowed down by the addition of osmolytes, such as sucrose, which increase the viscosity of the milieu. In an approximation of the nanosecond EPR timescale and a virtually static backbone, the backbone can be defined by a set of $2N_{aa}$-2 pairs of dihedral angles φ and ψ, one pair for each amino acid residue within a protein with N_{aa} amino acids. Possible dihedral angle pairs are restricted due to steric clashes of the backbone atoms by taking into account the minimal interatomic distances (van der Waals distances, contact distances) (34,35).

Because EPR spectroscopy is sensitive to the anisotropy of the conformational space of the SL, and not to the precise position of the atoms, atomic resolution of the modeling is not required. Consequently, the bond lengths and angles were fixed to the values reported in the literature (36,37). This approximation was used for both backbone and side-chain conformation calculations.

Step 2: Calculation of the side-chain rotational conformational spaces

For each amino acid side chain, including the SL cysteine side chain, the full conformational space was generated by discrete rotations around side-chain single bonds. The torsion potentials were modeled by a discrete set of equiprobable but not equidistant rotational states, such that their density increased proportionally to attain the one found in the rotamer library (38), adjusted to be applied at physiological temperatures. Self-overlapping conformations were discarded. However, the van der Waals distances for small groups of atoms (CH, CH_2, CH_3, NH, etc.) were decreased by 20% with respect to the values used in backbone modeling to simulate the nonuniform electron distribution in these groups and to allow more realistic contacts (34,39,40).

Step 3: Calculation of the restrictions to the rotational conformational space of the spin label

To restrict the rotational conformational space of a side chain, two main effects were considered: restrictions imposed by the backbone, and restrictions arising from neighboring amino acids. A restriction imposed on an arbitrary side-chain conformation is described by its decreased statistical weight p_i defined in Eq. 1 (see also below). Under the conditions of step 1, the backbone motion is approximated to be frozen. Hence, the statistical weight of a side-chain conformation, which overlaps with the backbone, is set to zero, i.e., such conformations are forbidden (the first factor in Eq. 1).

The statistical weight of an allowed conformation of an SL side chain is reduced if it shares physical space with the side chains of the neighboring amino acid residues. Since all side chains wobble rapidly over the nanosecond EPR time window, they occupy the same volume element only during a time period lower than the EPR timescale (i.e., subnanoseconds). Thus, if a conformation of an SL side chain overlaps with a certain number of conformations of the neighboring amino acid side chain(s), the statistical weight of that particular SL conformation is reduced by the ratio of the

number of overlapping conformations and the number of all possible conformations of the neighboring side chains that are allowed by the backbone overlap check (the second factor in Eq. 1). However, since the probability of an individual restricted conformation is calculated in the EPR time window, the probability of a transition between conformational states, which depends on the density of the local states, also has to be taken into account. This correction is approximated with a Gaussian spreading function that emulates the effective probability of getting into the ith state from a set of j (allowed) states that are in the local neighborhood. It penalizes lonely conformation states due to the finite time available to reach them. The r_i and r_j represent the appropriate nitroxide coordinate vectors, and σ is approximated to 10% of the size of rotational conformation space (0.14 nm):

$$p_i^* = \left\{ \begin{array}{l} 0, \text{ backbone overlap} \\ 1, \text{ no backbone overlap} \end{array} \right\}$$
$$\times \left[\prod_{k \in \text{ neighboring amino acids}} \left(1 - \frac{N_{\text{overlaps}}^k}{N_{\text{all}}^k} \right) \right], \, p_i = p_i^* \sum_j e^{\frac{-|r_i - r_j|^2}{2\sigma^2}}$$

(1)

Step 4: Characterization of the anisotropy of the rotational conformational space

For the purposes of comparison with the experimental values of the normalized free rotational space Ω_{exp}, we interpreted the results of the simulations in terms of the cone model parameters, i.e., the ϑ_0 and φ_0 angles, which respectively describe the amplitude and anisotropy of the SL rotational motion within a cone leading to the simulated Ω_{sim} (31–33):

$$\Omega_{\text{sim}} = \frac{\vartheta_0 \varphi_0}{(\pi/2)^2}$$

(2)

where ϑ_0 and φ_0 are calculated from average squared directional cosines, explicitly taking into account the probabilities of individual conformations (Eq. 1) as well as their orientations (for further details see the Supporting Material).

Step 5: Structural optimization

A stochastic optimization algorithm was used to tune the secondary structure of N_{TAIL} and the relative position of N_{TAIL} with respect to XD, such that the calculated local restrictions matched the features extracted from the experimental EPR data. The crystal structure of the chimeric construct consisting of XD and the 486–504 region of N_{TAIL} (PDB code: 1T6O) (25) was used as the starting model.

The parameters that were optimized are listed in Table S1. The optimization module was based on a stochastic algorithm of the Metropolis Monte Carlo family (41) with several elements of the Evolutionary Optimization (mutation operator, replacement operator, and elite) (42). Unlike conventional Evolutionary Optimization, each optimization run in this algorithm tunes a single structure. One run counts for 200 generations.

Note that although the structure of XD was kept fixed, the structure of N_{TAIL} was modified (via modification of the backbone dihedral angles), as were the position and orientation of N_{TAIL} relative to XD at each generation. For each new structure of the complex, the local restrictions were compared in terms of experimental and simulated free rotational space at each SL position along the protein sequence. The quality of the fit was evaluated as follows:

$$\chi^2 = \frac{1}{N} \sum_i \left(\frac{\Omega_{\text{exp},i} - \Omega_{\text{sim},i}}{\Delta\Omega_{\text{exp},i}} \right)^2$$

(3)

where N is the number of SL variants (sites), $\Omega_{\text{exp},i}$ and $\Omega_{\text{sim},i}$ respectively correspond to experimentally derived and simulated free rotational space values at the ith SL position, and $\Delta\Omega_{\text{exp},i}$ represents the inaccuracy of the

experimental free rotational space. The goodness-of-fit χ^2 served to guide the optimization routine and allow a structure to be accepted for the ensuing generation of structural evolution.

RESULTS

EPR experiments

Twelve SL N_{TAIL} variants, which were previously shown to possess an overall secondary structure content comparable to that of the wild-type (wt) form, and whose SL mobility is affected by XD (27), were selected for our study (see Fig. 1 B). Their EPR spectra were recorded at seven different temperatures (279 K, 281 K, 283 K, 296 K, 308 K, 310 K, and 312 K; see Materials and Methods) under four different conditions: in the presence or absence of XD, and with or without supplementing the buffer with 30% sucrose (see Fig. S3, Fig. S4, and Fig. S5), which was added to slow down the backbone motion. As previously noted (12,26), sucrose induced no secondary structure rearrangements within either wt or SL N_{TAIL} (see Fig. S7 and Fig. S8). Taking into account possible nonspecific labeling, and considering that the proteins probably possess more than one local conformation at a given site, the spectra were analyzed with a multicomponent model of asymmetrically restricted rotational motion of SLs (30,31,33,43). S/N ratios as high as ~200 allowed such an analysis, as well as optimization of the fitted spectra employing a multirun, multisolution hybrid evolutionary algorithm (31). The reduced χ^2 function (Eq. 3) was chosen to judge the goodness of fit, and for the final spectral fits it systematically turned out to be <3 noise amplitudes relative to the experimental spectra. The results from multiple spectra optimizations were condensed with the GHOST methodology (30–32) (see Fig. S6) and summarized in more condensed ''bubble'' plots (Fig. 2). In this way, the most significant and probable groups of motional patterns of spectral parameters, which represent the local restriction and mobility profiles along N_{TAIL}, were recognized for each temperature and each condition (altogether 336 motional pattern sets). These data were successively used as constraints for the structural optimization of the bound form of N_{TAIL} (see below).

Scanning motional restriction along N_{TAIL} positions

Since the differences between EPR spectra and hence between the detected motional pattern sets of the local temperature ranges (e.g., 279 K, 281 K, and 283 K) were not significant, we averaged out the detected motional patterns to derive the low (279 K, 281 K, and 283 K) and high (308 K, 310 K, and 312 K) temperatures, i.e., 281 K and 310 K, respectively. In this way, we were able to compare the local properties at two characteristic temperatures more accurately. In further analyses, two physical quantities were considered: 1), the free rotational space Ω_{exp}, a normalized product of the
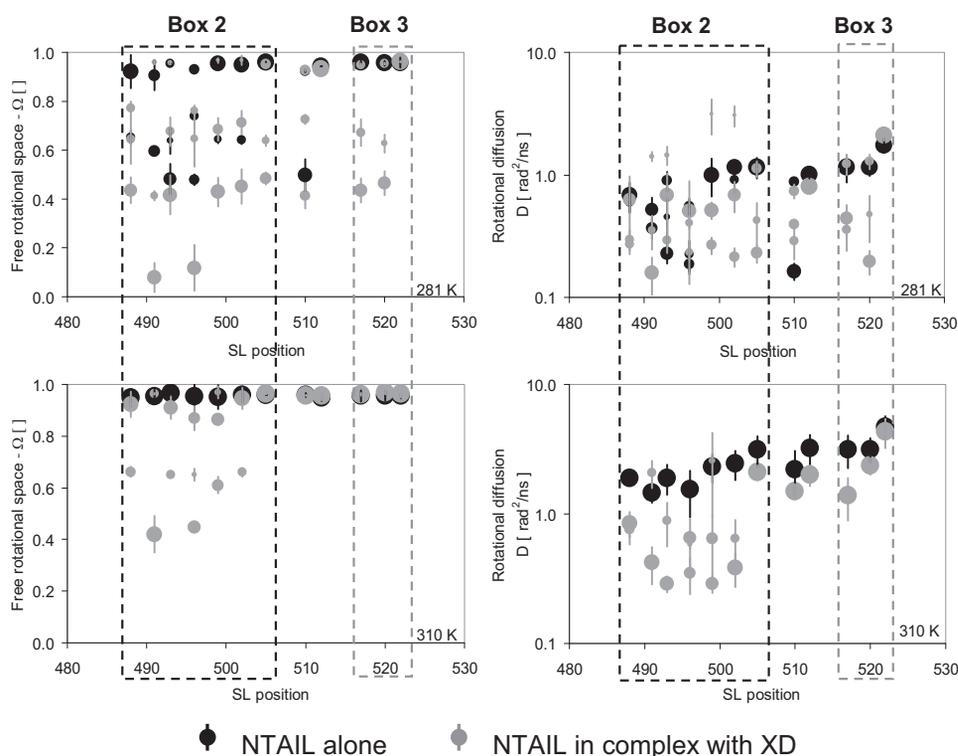
FIGURE 2 Experimental free rotational space Ω (*left panels*) and rotational diffusion D (*right panels*) obtained for the 12 SL N_{TAIL} protein variants either alone (*black circles*) or in complex with XD (*gray circles*) in the presence of 30% sucrose at 281 K (*top panels*) and at 310 K (*bottom panels*). Ω and D are plotted as a function of SL position. The area of the circles is proportional to the relative contribution of the motional patterns (spectral components) to the overall solution (total spectrum), and the vertical bars point the second moments of the solutions.

cone opening angle ϑ_0 and cone anisotropy angle φ_0 (Eq. 2); and 2), the normalized rotational diffusion D_{exp}, an angular mean-square displacement of the SL in the EPR time window, approximated with a quotient of a product of ϑ_0 and φ_0 angles over the effective rotational correlation time τ_c:

$$D_{exp} = \vartheta_0\varphi_0/4\tau_c \qquad (4)$$

The free rotational space Ω_{exp} measures the space angle left for local SL wobbling: high values of Ω (between 0.9 and 1) correspond to nearly unrestricted motional patterns of the SL, whereas very low values (<0.2) imply very high restrictions preventing side-chain conformational motions. Unlike the conventional correlation time τ_c, which is correlated to the free rotational space Ω (33), normalized diffusion D allows a more clear comparison of the local dynamics of various SL sites with possible different rotational restrictions. As such, normalized rotational diffusion is used to distinguish between the unstructured and structured segments of protein backbone.

In the absence of sucrose, the rotational motion was found to be unrestricted at high temperatures, irrespective of the presence or absence of XD (see spectral line shapes in Fig. S3, Fig. S4, and Fig. S5). Notably, the addition of 30% sucrose, which slows down the backbone global motion by a factor of 3, as judged by the decrease in the rotational diffusion D-values, brings significant restrictions (free rotational space < 0.8) at numerous N_{TAIL} positions at both low (Fig. 2, *top panels*) and high (Fig. 2, *bottom panels*) temperatures, and this effect is enhanced by the presence of XD (Fig. 2).

In the absence of XD, at high temperature there are no stable restrictions (which can be determined within the few-nanoseconds EPR time window) along all the N_{TAIL} positions (see Fig. 2 *bottom left*, *black circles*). At low temperature, some restrictions appear within the Box2 region (Fig. 2 *top left*, *black circles*). The normalized diffusion values obtained in the absence of XD reveal a very fast motion at high temperature (Fig. 2 *bottom right*, *black circles*), whereas the motion slows down by a factor of 2 in the Box3 region and by almost an order of magnitude for some Box2 positions at low temperature (Fig. 2 *top right*, *black circles*). Strikingly, the presence of XD stabilizes the local N_{TAIL} conformations, with this effect being particularly pronounced for Box2 positions. Indeed, local motional restrictions can already be detected in the presence of XD at high temperature (Fig. 2 *bottom left*, *gray circles*). The rotational diffusion slows down too, especially within Box2, where stable restrictions are detected (Fig. 2 *bottom right*, *gray circles*). At low temperature, even more restrictions are detected (Fig. 2, *top panels*, *gray circles*). The very strong restrictions at positions 491 and 496 indicate that the spin label is squeezed between two heavy objects, i.e., the main chains of N_{TAIL} and XD.

For many SL positions, several motional patterns are resolved with different restrictions (see the condensed motional patterns in Fig. 2). Since fast backbone rearrangements significantly increase the detected rotational conformational space (i.e., make it virtually unrestricted), the restricted local rotational space therefore represents stable protein conformation. On the contrary, an unrestricted one at a given amino acid position can indicate either a stable

conformation with open conformational space or a locally disordered protein state. To distinguish between these two cases, the restriction profile must be analyzed over a longer protein segment together with the weights of motional patterns. The weights are expected to be position-independent for a particular protein conformation. Indeed, our results (Fig. 2) can be interpreted in this way, revealing one disordered conformation with a weight of ~40% and two ordered conformations with weights of 50% and 10%. Since the latter is close to the sensitivity limit (~10%) of the method, we focused on the major structured protein conformation. However, such a simplistic interpretation does not exclude less-populated conformation(s) with mixed restricted/unrestricted motional patterns.

Conformational space modeling and protein structure optimization

To describe the N_{TAIL}-XD complex at the structural level, we performed structural modeling and searched for the best-fit structures in terms of their agreement with the experimental EPR data at 281 K and 310 K (see Materials and Methods).

First, we calculated the free rotational spaces for the six N_{TAIL} positions (S488, S491, D493, L496, Q499, and A502) occurring within the N_{TAIL} segment (aa 486–504) bound to XD in the crystal structure of the chimera (PDB code: 1T6O) (25). As shown in Fig. 3 A, positions S491, Q499, and A502 in the chimera structure are completely restricted, forbidding any alternate conformation of an SL side chain, whereas the other three positions appear to be less restricted. When we compared these restrictions with those obtained from the experimental EPR data at low temperature and in the presence of XD (Fig. 2 top left, gray circles), we found a reasonable agreement for positions 488, 491, and 493, but significant discrepancies for positions 496, 499, and 502 (Fig. 3 A and Fig. 2 top left, gray circles).

Second, in an attempt to understand what kind of structural features could explain this discrepancy, we optimized the structure of N_{TAIL} in the N_{TAIL}-XD complex by fitting the modeled restrictions (29) to the most probable
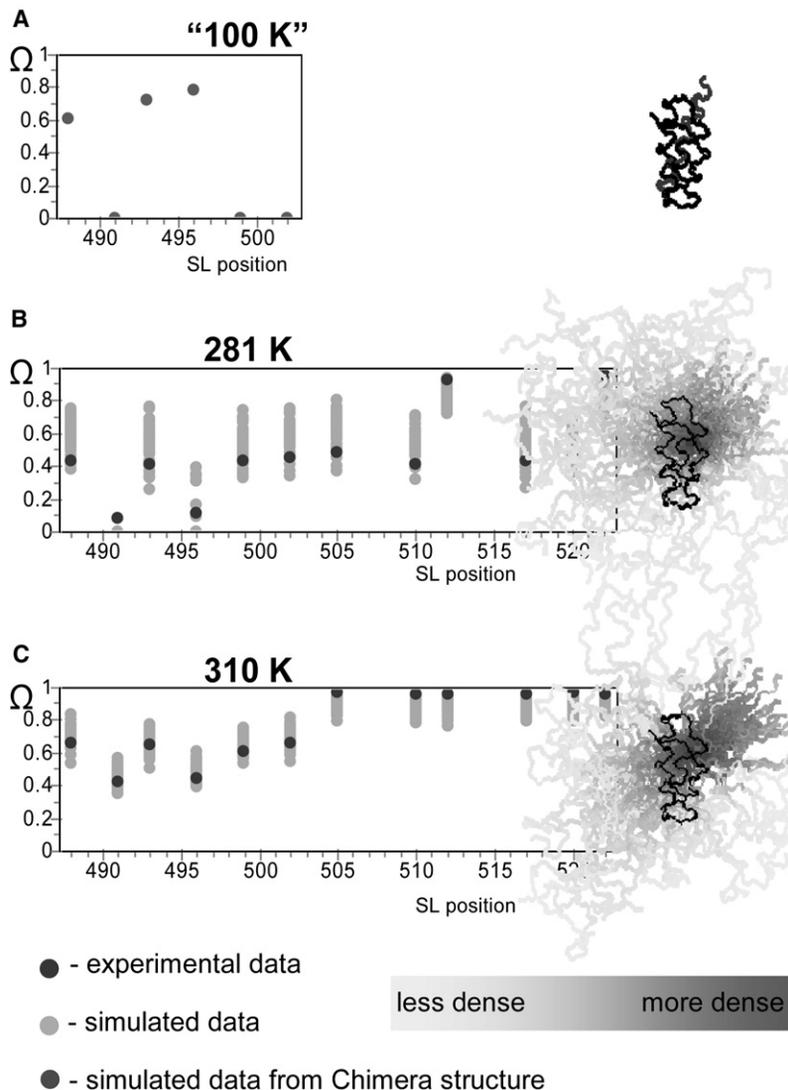
FIGURE 3 Results of the optimization of the structure of the 488–525 region of N_{TAIL} in complex with XD as obtained by fitting the calculated local restrictions to the restrictions profiles obtained from SDSL EPR. (A) Local restrictions calculated for the fixed structure of the chimera of N_{TAIL} (aa 486–504) and XD. Panels B and C show the population of the 50 best-fit structures of the N_{TAIL}-XD complex as obtained after structural modeling at 281 K (B) and 310 K (C). Residues 488–525 of N_{TAIL} are depicted with a color gradient ranging from light to dark gray with decreasing structural density, and XD is shown in black.

restrictions detected by SDSL EPR. To obtain a starting model of the 488–525 region of N_{TAIL}, we used the available structural data for the 488–504 region of N_{TAIL} and modeled the missing region by initially imposing on it an α-helical conformation (backbone dihedral angles $\varphi_{initial} = -57°$ and $\psi_{initial} = -47°$). This initial model of the 488–525 structure of N_{TAIL} was then optimized against the experimental restriction profiles extracted from the EPR data at low (281 K) and high (310 K) temperatures. The best 50 fits resulting from 1000 optimization runs, together with the corresponding structural models of the N_{TAIL}-XD complex at both 281 K and 310 K, are shown in Fig. 3, B and C. The orientation of the complex was defined by locating the three helical chains of XD from the crystal structure of the chimera (*black*) in the front and drawing the N_{TAIL} structures in the background. The obtained population of structures of the N_{TAIL}-XD complex at 281 K and 310 K shows that whereas the Box2 region adopts a similar conformation in the 50 structures, the C-terminal region of N_{TAIL}, including Box3, samples a much wider conformational space. To unveil possible XD- and temperature-induced changes in the secondary structure of N_{TAIL}, the 100 resulting best-fit structures were analyzed in terms of their α-helicity. For that purpose, the backbone dihedral angles of the best-fit structures were collected for each amino acid position in Ramachandran plots (35), and the percentage of α-helical angle pairs was computed for each amino acid position (Fig. 4). Helical segments are clearly detectable within the Box2 region of

both the free (Fig. 4 A) and the bound (Fig. 4 B) form of N_{TAIL} at 281 K, as well as for the bound form at 310 K (Fig. 4 C). Because of the lack of stable restrictions, no helical segments could be detected for the free form of N_{TAIL} at 310 K.

DISCUSSION

Two features of SL EPR spectroscopy—namely, its nanosecond time window and sensitivity to rotational motion anisotropy—are ideally suited to monitor backbone motions that occur in the nanoseconds timescale, as well as to detect restrictions in side-chain rotational conformational spaces that can be further translated into local structural constraints. Moreover, if the backbone experiences a structural transition from a folded conformation to a random coil, the timescale of the backbone dynamics suddenly speeds up and begins to overlap with the timescale of the side chains. In such cases, EPR is also perfectly suited to detect possible destabilizations of the backbone structure, as illustrated in this study, where these events were monitored for the bound form of N_{TAIL}. Detection of transient backbone conformations can be further enhanced by the use of sucrose, which slows down the backbone motion. Basically, the addition of sucrose and the lowering of the temperature both slow down the protein backbone motion. However, a temperature decrease also affects the side-chain motion, as judged by the rotational diffusion.
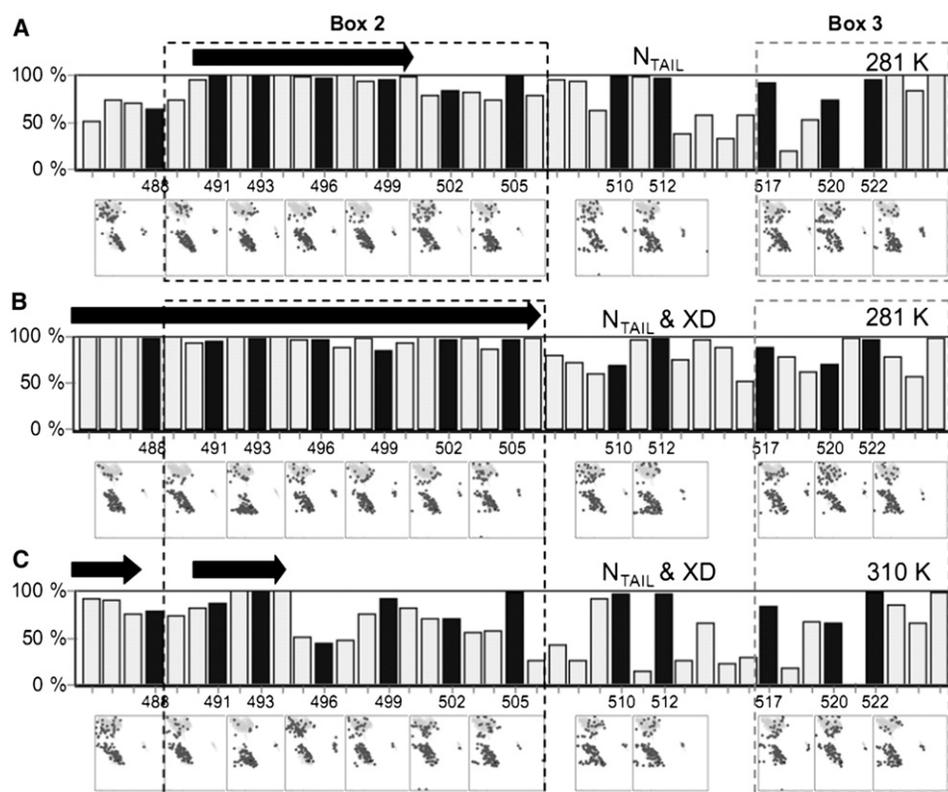
FIGURE 4 Analysis of the secondary structures of the 100 best-fit structures of N_{TAIL} alone at 281 K (*A*), and of N_{TAIL} in complex with XD at 281 K (*B*) and 310 K (*C*). The height of the bars represents the amount of the helical conformations among the 100 best-fit structures. Black and gray bars correspond to SL and unlabeled positions, respectively. The Ramachandran plots shown below each SL position represent the corresponding distribution of the backbone dihedral angles φ and ψ of the 100 best-fit structures. The arrows correspond to regions with an α-helical conformation, where a chain segment is considered to adopt an α-helical conformation if it includes at least four amino acids with at least 80% of the dihedral angles φ and ψ distribution in the near α-helical region of the Ramachandran plot.

In the presence of sucrose, the lifetime of backbone conformations therefore increases from a few nanoseconds to up to 10 or even a few tens of nanoseconds, thus making SDSL EPR an ideally suited method for detecting structures sampled only over a few nanoseconds.

In this study, in the absence of the partner and at high temperature (310 K), all spin labels, irrespective of their positions, experienced no restrictions (Fig. 2 *bottom left*, *black circles*), whereas significant restrictions appeared within Box2 at 281 K (Fig. 2 *top left*, *black circles*). If the number of degrees of freedom (9° of freedom per bond due to spatial displacements, rotations, and vibrations) in this eight-residue-long region is approximated to 1000, the thermal energy difference would correspond to the breaking of ~5–6 hydrogen bonds, a value close to that required to break 7–8 hydrogen bonds stabilizing the α-helix within Box2. These results argue for the existence of a preformed α-helix within Box2, which is stable at 281 K for at least a few nanoseconds even in the absence of the partner, as already suggested by previous EPR studies (27).

In the bound form, the structure of N_{TAIL} becomes much more stable, as judged by the restrictions in the SL conformational space (Fig. 2, *left panels*, *gray circles*). Under these conditions, a few sites appear to be almost fully restricted, consistent with a direct overlap of the spin label with the conformational space and backbone of XD, and hence with a strong interaction in the neighborhood.

In addition to a pronounced XD-induced stabilization of the α-helix within Box2, which was also appreciable by means of secondary structure analyses (see Fig. 4 B), a weaker stabilization effect within Box3 was also detected. Indeed, stable motional restrictions were detected at 281 K (Fig. 2 *top left*, *gray circles*), although the XD impact was not sufficiently pronounced to stabilize this region at 310 K (Fig. 2 *bottom left*, *gray circles*). Since the temperature drop itself was shown not to induce a stabilization of this region (cf. Fig. 2 *bottom left* with Fig. 2 *top left*, *black circles*), these results imply that XD plays a role in stabilizing Box3. In light of the reported lack of a stable interaction between Box3 and XD (27,28), the motional restriction of Box3 brought by XD likely arises from a gain of rigidity resulting from α-helical folding of the neighboring Box2 region and/or by the establishment of transient, weak contacts with other N_{TAIL} or XD regions.

To obtain further insights into the actual conformation of N_{TAIL} within the complex, we performed structural modeling. It would have been feasible and straightforward to use molecular-dynamics simulations if we had not included experimental data fitting in the structure determination. Since it takes several weeks to perform a few 10-ns simulations on a small cluster, it would have been excessively time-consuming to solve the inverse spectral problem by searching for the spectral parameters that provide a good fit between the simulated and experimental spectra. On the other hand, to explain restrictions determined by SDSL EPR, the full-time trajecto-

ries of all the atoms in the system are not required—only some of the time averages are needed. Since EPR is not sensitive to atom coordinates, but rather to the anisotropy of the SL rotational space, atomic precision is not required either. Therefore, the problem can be solved by an approach based on side-chain conformational-space modeling (29), which can, however, be applied only near physiological temperatures. Under such conditions, the entropic contribution to the free energy becomes more important and side chains are not stacked in some of the energetically favorable conformers (rotamers). At this point, therefore, the modeling approximation exactly meets the experimental conditions, enabling protein structure characterization based on the detection and analysis of the rotational conformational space restrictions under physiological conditions.

Using an optimization method that combines two stochastic optimization methods—namely, Monte Carlo individual optimization and hybrid evolutionary population-based optimization—we were able to tune the dihedral angles of the N_{TAIL} protein backbone and its relative orientation to XD to meet the constraints obtained from SDSL EPR data at different temperatures. By fitting the modeled restrictions of the local conformational spaces to the measured ones, we derived a family of structures that are equally well in agreement with the experimental data (Fig. 3, B and C). The most striking result is the reorienting of the N_{TAIL} Box2 α-helix relative to the XD helices as a function of the temperature increase. As shown in Fig. 3 B, the ''281 K'' structure is much less strictly positioned than the starting model (i.e., the crystal structure of the chimera, which encompasses only the 486–504 region of N_{TAIL} and was determined at 100 K), and, on average, the N_{TAIL} helix tilts away from the three XD helices (Fig. 3 B, *black structure*). At 310 K (Fig. 3 C), this phenomenon is even more pronounced because the tilt is more marked and clearly defined. By exploring the overlaps between the side-chain conformational spaces of N_{TAIL} and XD, we found that the tilt is actually a result of minimization of the rotational-spaces overlap, i.e., maximization of the entropy. More disordered side chains of different protein molecules increase the steric repulsions between the two protein molecules. Therefore, the family of structures modeled at high temperature represents the tendency of the flexible N_{TAIL} protein to minimize the side-chain contacts with XD (Fig. 3). It is therefore obvious that at physiological temperatures, such as 310 K, structure predictions based on free-energy minimization should also involve an entropic component in addition to the energy component.

Based on these findings, we propose that the region downstream from Box2 conserves a high conformational freedom in the complex, despite the XD-induced gain of rigidity within Box2. In agreement with previous SAXS data (see Fig. 2 in Bourhis et al. (12)), we propose that the C-terminus of N_{TAIL} dynamically samples the conformational space in the proximity of XD without establishing stable contact(s)

with the latter. Because of its intrinsic flexibility, this region can contribute to the interaction with XD by reducing the entropic penalty of the binding reaction, as previously postulated based on the analysis of EPR spectra in the presence of TFE and XD (27), and as discussed by Tompa and Fuxreiter (44).

This study also shows that the α-MoRE is preconfigured in the absence of the partner, as deduced from previous EPR studies (27). The possibility that such a preconfigured α-MoRE could artifactually arise from the introduction of the spin label was checked and ruled out by previous circular dichroism studies that showed that the circular dichroism spectra of the SL N_{TAIL} variants can be superimposed on that of the wt form (27). These studies also pointed out the presence of a residual secondary structure within wt N_{TAIL} consistent with a premolten, globule-like conformation that was ascribed mainly to the folding potential of Box2 (12,18). Here, we quantitatively estimated, for the first time to our knowledge, the extent of α-helical sampling of Box2 on a nanosecond timescale, and showed that the 490–500 region of N_{TAIL} adopts an α-helical conformation in almost 100% of N_{TAIL} conformers in solution (Fig. 4 *A*) at 281 K. The presence of this prestructured α-helix, together with the significant flexibility of the regions downstream and upstream from Box2, likely favors the binding reaction by reducing the entropic penalty. The limited entropic cost of the binding reaction would account for the rather high affinity ($K_D = 100$ nM) that was experimentally determined for the N_{TAIL}-XD binding reaction (12).

By increasing the temperature from 281 K to 310 K, we were also able to show that the structure of Box2 becomes less helical (Fig. 4, *B* and *C*), with only a small part of Box2 persisting in an α-helical conformation despite the presence of XD. Indeed, at 310 K, only the 485–494 region, which includes three SL positions (488, 491, and 493), was shown to conserve an α-helical conformation. We can speculate that this region includes the main anchoring site responsible for binding to the XD interface, and that it may serve as the starting points in the folding process induced by XD.

As a final methodological note, we should point out that the quality of structure determination based on EPR data obviously depends on the number of SL sites. Since the SL rotational space is twice as big as the average rotational space of amino acid side chains, its restrictions arise not only from the first amino acid neighbor (forward and backward), but also from several amino acids in the vicinity that are not restricted to the same backbone. Therefore, grafting spin labels at positions separated by as few as two or three residues should ensure accuracy while keeping to a minimum the number of sites targeted for SDSL. Indeed, by using 12 SL variants in this study, we were able to derive temperature-dependent backbone conformations for a highly dynamic protein region of 40 residues in length, with subnanometer accuracy. This corresponds to the accuracy of relative positions of the backbone atoms in a local protein

segment, and should not be confused with the distribution of different structures, which in the case of IDPs can be rather large. We can speculate that in the absence of a strong interaction site, such as that within Box2, which provides strong constraints for the structural modeling, additional SL variants probably would have been required to accurately resolve the N_{TAIL} orientation relative to XD.

CONCLUSIONS

Although low-temperature (energy-minimized) conformational states are widely used in high-resolution protein structure determination, they are conceptually ineffective for characterizing a quickly rearranging IDP under physiological conditions. To solve this problem, we developed an alternative structural methodology that allows fast motional conformational events to be taken into account. By taking advantage of the nanosecond time window of EPR spectroscopy and computing local conformational space restrictions, we were able to achieve a structural description of a partly disordered protein complex consisting of a family of short-lived conformations, and to monitor temperature-induced structural changes. Beyond paving the way toward the structural characterization of other IDPs, this work extends our previous data on the N_{TAIL}-XD interaction and thus contributes to the field of protein-protein interactions and, in particular, to the rapidly growing field of IDPs.

SUPPORTING MATERIAL

REFERENCES

1. Dyson, H. J., and P. E. Wright. 2005. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol*. 6:197–208.

2. Dunker, A. K., J. D. Lawson, …, Z. Obradovic. 2001. Intrinsically disordered protein. *J. Mol. Graph. Model*. 19:26–59.

3. Tompa, P. 2002. Intrinsically unstructured proteins. *Trends Biochem. Sci*. 27:527–533.

4. Uversky, V. N. 2002. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci*. 11:739–756.

5. Fink, A. L. 2005. Natively unfolded proteins. *Curr. Opin. Struct. Biol*. 15:35–41.

6. Uversky, V. N., C. J. Oldfield, and A. K. Dunker. 2005. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J. Mol. Recognit.* 18:343–384.

7. Receveur-Bréchot, V., J. M. Bourhis, …, S. Longhi. 2006. Assessing protein disorder and induced folding. *Proteins.* 62:24–45.

8. Ferron, F., S. Longhi, …, D. Karlin. 2006. A practical overview of protein disorder prediction methods. *Proteins.* 65:1–14.

9. Bourhis, J. M., B. Canard, and S. Longhi. 2007. Predicting protein disorder and induced folding: from theoretical principles to practical applications. *Curr. Protein Pept. Sci.* 8:135–149.

10. Dunker, A. K., I. Silman, …, J. L. Sussman. 2008. Function and structure of IDPs. *Curr. Opin. Struct. Biol.* 18:756–764.

11. Wright, P. E., and H. J. Dyson. 2009. Linking folding and binding. *Curr. Opin. Struct. Biol.* 19:31–38.

12. Bourhis, J. M., V. Receveur-Bréchot, …, S. Longhi. 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.

13. Dyson, H. J., and P. E. Wright. 2002. Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* 12:54–60.

14. Timsit, Y., F. Allemand, …, M. Springer. 2006. Coexistence of two protein folding states in the crystal structure of ribosomal protein L20. *EMBO Rep.* 7:1013–1018.

15. Hoff, A. J. 1989. Advanced EPR: Applications in Biology and Biochemistry. Elsevier, Amsterdam/New York.

16. Columbus, L., and W. L. Hubbell. 2004. Mapping backbone dynamics in solution with site-directed spin labeling: GCN4-58 bZip free and bound to DNA. *Biochemistry.* 43:7273–7287.

17. Stopar, D., R. B. Spruijt, and M. A. Hemminga. 2006. Anchoring mechanisms of membrane-associated M13 major coat protein. *Chem. Phys. Lipids.* 141:83–93.

18. Longhi, S., V. Receveur-Bréchot, …, B. Canard. 2003. The C-terminal domain of the measles virus nucleoprotein is intrinsically disordered and folds upon binding to the C-terminal moiety of the phosphoprotein. *J. Biol. Chem.* 278:18638–18648.

19. Johansson, K., J. M. Bourhis, …, S. Longhi. 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.

20. Bourhis, J., K. Johansson, …, S. Longhi. 2004. The C-terminal domain of measles virus nucleoprotein belongs to the class of IDPs that fold upon binding to their physiological partner. *Virus Res.* 99:157–167.

21. Bourhis, J. M., B. Canard, and S. Longhi. 2005. Désordre structural au sein du complexe réplicatif du virus de la rougeole: implications fonctionnelles. *Virologie.* 9:367–383.

22. Bourhis, J. M., B. Canard, and S. Longhi. 2006. Structural disorder within the replicative complex of measles virus: functional implications. *Virology.* 344:94–110.

23. Bourhis, J. M., and S. Longhi. 2007. Measles virus nucleoprotein: structural organization and functional role of the intrinsically disordered C-terminal domain. *In* Measles Virus Nucleoprotein. S. Longhi, editor. Nova Publishers, Hauppage, NY. 1–35.

24. Longhi, S. 2009. Nucleocapsid structure and function. *Curr. Top. Microbiol. Immunol.* 329:103–128.

25. Kingston, R. L., D. J. Hamel, …, B. W. Matthews. 2004. Structural basis for the attachment of a paramyxoviral polymerase to its template. *Proc. Natl. Acad. Sci. USA.* 101:8301–8306.

26. Morin, B., J. M. Bourhis, …, S. Longhi. 2006. Assessing induced folding of an *IDP* by SDSL EPR spectroscopy. *J. Phys. Chem. B.* 110:20596–20608.

27. Belle, V., S. Rouger, …, S. Longhi. 2008. Mapping α-helical induced folding within the intrinsically disordered C-terminal domain of the measles virus nucleoprotein by SDSL EPR spectroscopy. *Proteins.* 73:973–988.

28. Bernard, C., S. Gely, …, H. Darbon. 2009. Interaction between the C-terminal domains of N and P proteins of measles virus investigated by NMR. *FEBS Lett.* 583:1084–1089.

29. Štrancar, J., A. Kavalenka, …, M. A. Hemminga. 2009. Analysis of side chain rotational restrictions of membrane-embedded proteins by spin-label ESR spectroscopy. *J. Magn. Reson.* 197:245–248.

30. Filipic, B., and J. Strancar. 2001. Tuning EPR spectral parameters with a genetic algorithm. *Appl. Soft Comput.* 1:83–90.

31. Strancar, J., T. Koklic, …, M. A. Hemminga. 2005. Spin label EPR-based characterization of biosystem complexity. *J. Chem. Inf. Model.* 45:394–406.

32. Kavalenka, A. A., B. Filipic, …, J. Strancar. 2005. Speeding up a genetic algorithm for EPR-based spin label characterization of biosystem complexity. *J. Chem. Inf. Model.* 45:1628–1635.

33. Stopar, D., J. Strancar, …, M. A. Hemminga. 2006. Motional restrictions of membrane proteins: a site-directed spin labeling study. *Biophys. J.* 91:3341–3348.

34. Ho, B. K., A. Thomas, and R. Brasseur. 2003. Revisiting the Ramachandran plot: hard-sphere repulsion, electrostatics, and H-bonding in the α-helix. *Protein Sci.* 12:2508–2522.

35. Ramachandran, G. N., C. Ramakrishnan, and V. Sasisekharan. 1963. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* 7:95–99.

36. Engh, R., and R. Huber. 1991. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A.* 47:392–400.

37. Word, J. M., S. C. Lovell, …, D. C. Richardson. 1999. Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. *J. Mol. Biol.* 285:1711–1733.

38. Lovell, S. C., J. M. Word, …, D. C. Richardson. 2000. The penultimate rotamer library. *Proteins.* 40:389–408.

39. Tombolato, F., A. Ferrarini, and J. H. Freed. 2006. Dynamics of the nitroxide side chain in spin-labeled proteins. *J. Phys. Chem. B.* 110:26248–26259.

40. Grigoryan, G., A. Ochoa, and A. E. Keating. 2007. Computing van der Waals energies in the context of the rotamer approximation. *Proteins.* 68:863–878.

41. Kirkpatrick, S., C. D. Gelatt, Jr., and M. P. Vecchi. 1983. Optimization by simulated annealing. *Science.* 220:671–680.

42. Fogel, D. B., T. Bäck, and Z. Michalewicz. 2000. Evolutionary computation. Institute of Physics Publishing, Bristol/Philadelphia.

43. Stopar, D., J. Strancar, …, M. A. Hemminga. 2005. Exploring the local conformational space of a membrane protein by site-directed spin labeling. *J. Chem. Inf. Model.* 45:1621–1627.

44. Tompa, P., and M. Fuxreiter. 2008. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* 33:2–8.