

# Intrinsic Dynamics of the Partly Unstructured PX Domain from the Sendai Virus RNA Polymerase Cofactor P

Klaartje Houben, Laurence Blanchard, Martin Blackledge, and Dominique Marion

Institut de Biologie Structurale Jean-Pierre Ebel, CNRS, CEA, UJF, UMR-5075, 38027 Grenoble cedex 1, France

**ABSTRACT** Despite their evident importance for function, dynamics of intrinsically unstructured proteins are poorly understood. Sendai virus phosphoprotein, cofactor of the RNA polymerase, contains a partly unstructured protein domain. The phosphoprotein X domain (PX) is responsible for binding the polymerase to the nucleocapsid assembling the viral RNA. For RNA synthesis, the interplay of the dynamics of the unstructured and structured PX subdomains is thought to drive progression of the RNA polymerase along the nucleocapsid. Here we present a detailed study of the dynamics of PX using hydrogen/deuterium exchange and different NMR relaxation measurements. In the unstructured subdomain, large amplitude fast motions were found to be fine-tuned by the presence of residues with short side chains. In the structured subdomain, where fast motions of both backbone and side chains are fairly restricted, the first helix undergoes slow conformational exchange corresponding to a local unfolding event. The other two helices, which represent the nucleocapsid binding site, were found to be more stable and to reorient with respect to each other, as probed by slow conformational exchange identified for residues on the third helix. The study illustrates the intrinsically differential dynamics of this partly unstructured protein and proposes the relation between these dynamics and its function.

## INTRODUCTION

Many proteins contain unstructured regions of considerable length, some of which are entirely lacking a persistent tertiary structure (1). These proteins, or protein regions, with no stable fold play important roles in numerous cellular processes and signaling events, where they make efficient use of both their high flexibility and extensive conformational freedom (2). More specifically, the unstructured character allows low affinity interactions, promotes rapid degradation of the protein, enlarges the interaction surface, enables binding to numerous structurally distinct partners, and facilitates posttranslational modifications by leaving all side chains readily accessible (3–5). In addition, unstructured protein regions can function as a string or “fishing line” between distinct protein domains. Despite their functional importance, intrinsically unstructured proteins have long remained beyond the reach of classical structural biology. These proteins can be most readily described as an ensemble of interchanging conformations, and detailed descriptions of the sampled conformational space as well as the dynamics of such unstructured ensembles can contribute to a better understanding of the important biological processes in which they are involved.

In this article we studied the dynamics of a partly unstructured protein from Sendai virus (SeV). This virus, also called murine parainfluenza virus (MPV1), is a prototype negative-strand RNA virus. SeV belongs to the *Paramyxoviridae* family, which also contains measles, mumps, and human parainfluenza viruses. In these viruses the RNA genome is assembled in a nucleocapsid, where it is protected by many copies of the nucleoprotein N and where each copy

of N is associated with exactly six nucleotides (6). The RNA polymerase for both replication and transcription of the viral RNA is a complex composed of the large protein L, which contains the polymerase activity and the phosphoprotein P, the tetrameric (7,8) cofactor, which places the polymerase on the N:RNA template (9) (Fig. 1). Positioning of the polymerase occurs through a highly dynamic interaction (10) of the C-terminal partly structured phosphoprotein X (PX) domain (11) of P (*dotted circle* in Fig. 1) and the unstructured tail of N ( $N_{\text{TAIL}}$ ) that becomes structured upon binding to PX. Interestingly, PX is an independent protein, since it is also expressed as a sole (monomeric) domain in infected cells (12). PX comprises a disordered N-terminal subdomain (474–515) (13) and a C-terminal subdomain (516–568) having a three-helix bundle structure (14). Whereas this bundle forms the  $N_{\text{TAIL}}$  interaction site (15), the unstructured domain is supposed to play a pivotal role in the dynamics of RNA synthesis: RNA synthesis occurs one nucleotide at a time, whereas displacement of P on the N:RNA template occurs in steps of subunits of N. Therefore, the P tetramer is proposed to “cart-wheel” across the N:RNAs, by on-and-off interactions of the PX domain with subsequent  $N_{\text{TAIL}}$  domains, whereas the polymerase transcribes or replicates the template RNA (10).

In this context we used NMR to study the dynamics of the PX domain. NMR has been shown to be a well-adapted technique for studying the dynamics of both structured and disordered protein states in solution at atomic resolution (16). Both  $^{15}\text{N}$  (17,18) and  $^2\text{H}$  (19,20) spin relaxation report on the fast fluctuation of the protein backbone and side chains, respectively. Slower motions, which are often related to conformational changes, can be probed by relaxation dispersion spectroscopy (21,22). In addition, experimental residual dipolar couplings (RDCs) can be reproduced remarkably well

Submitted March 13, 2007, and accepted for publication June 13, 2007.

Address reprint requests to Dominique Marion, E-mail: [dominique.marion@ibs.fr](mailto:dominique.marion@ibs.fr)

Editor: Arthur G. Palmer 3rd.

© 2007 by the Biophysical Society  
0006-3495/07/10/2830/15 \$2.00

doi: 10.1529/biophysj.107.108829

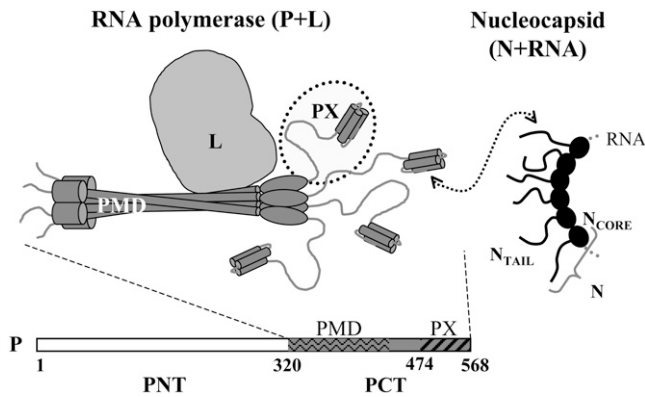


FIGURE 1 Schematic representation of the SeV replicative complex. The RNA polymerase is composed of phosphoprotein (P) and large protein L, whereas L (light gray) contains the polymerase activity, and the P tetramer (dark gray) is the cofactor that interacts with L and that is responsible for the interaction with the nucleoprotein N that protects the RNA. These functions reside in the phosphoprotein C-terminus (PCT). More specifically, the phosphoprotein multimerization domain (PMD) contains the interaction site for L, and the structured part of PX contains the interaction site for the  $N_{TAIL}$  domain of N. For L to transcribe or replicate the RNA, P has to progress along the nucleocapsid by transient interactions of its four PX domains with successive  $N_{TAIL}$  domains of the nucleocapsid.

using a statistical coil model to define the conformational space explored by the unstructured peptide chain (13,23). NMR constraints are also used to direct conformational sampling in either restrained molecular dynamics or Monte Carlo simulations (24,25). Analysis of NMR relaxation data to extract timescales and amplitudes of the underlying motions is not straightforward in the case of an unstructured protein, since the common decoupling of global and internal motions no longer has any relevance. Several methods have been proposed to circumvent this problem, for example, determination of an effective correlation time per residue (26, 27) or a distribution of correlation times (28). A more detailed interpretation of the relaxation rates of unstructured systems is provided by the iRED (isotropic reorientational eigenmode dynamics) method (29) developed by Brüschweiler and co-workers, which relies on a principal component analysis of the covariance matrix from a molecular dynamics trajectory, where the correlation times of the extracted eigenmodes are adjusted to optimally reproduce the experimental relaxation rates.

Here we used different NMR methods to describe the dynamics of this partly unstructured protein PX in detail. Backbone dynamics are probed by identifying infrequent events using hydrogen/deuterium (H/D) exchange experiments, ms- $\mu$ s motions measured by relaxation dispersion experiments, and fast ps-ns fluctuations extracted from  $^{15}\text{N}$  relaxation rates. In addition, fast side-chain dynamics were identified from  $^2\text{H}$  relaxation in  $\text{CH}_2\text{D}$  methyl groups, where an optimized pulse sequence was designed. The partly structured nature of this protein allows the dynamics of the two domains to be investigated simultaneously and to relate

them to their very distinct functions. Whereas the unstructured N-terminal subdomain acts as a flexible linker between the C-terminal subdomain and the P tetramer, the structured C-terminal subdomain forms a favorable interaction surface upon which the nucleocapsid  $N_{TAIL}$  domain can fold. There are not many examples of a detailed dynamic analysis of partly unstructured proteins, and this study highlights the importance of the intrinsic differential dynamics for the biological function of this class of proteins.

## MATERIALS AND METHODS

### Sample preparation

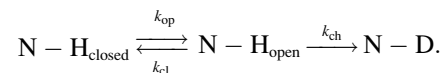
Isotope-labeled samples of the PX domain (474–568) of SeV phosphoprotein P, including an N-terminal His tag and a factor Xa cleavage site (MRGSHHHHHHIEGR), were prepared as described previously (30). A partially deuterated and uniformly  $^{15}\text{N}/^{13}\text{C}$ -labeled sample was prepared using 70%  $\text{D}_2\text{O}$  in the growth medium and *Escherichia coli* BL21(DE3) that were adapted to grow in 70%  $\text{D}_2\text{O}$ . From an electrospray mass spectrum the portion of deuteration was determined to be 60%. For the backbone dynamics both a  $\sim 1.3$  mM and  $\sim 0.2$  mM  $^{15}\text{N}$ -labeled sample were used. Deuterium relaxation experiments were performed on the triple-labeled (60%)  $^2\text{H}/^{15}\text{N}/^{13}\text{C}$  PX sample with an approximate concentration of 0.8 mM. Stereospecific assignments of the methyl groups were obtained from a 1 mM 10%  $^{13}\text{C}$ -labeled sample as described by Neri et al. (31). All samples were prepared in potassium phosphate buffer (50 mM) pH 6 with 500 mM NaCl. Integrity of the samples was verified by comparison of heteronuclear single quantum coherence (HSQC) spectra before and after sets of NMR experiments.

### NMR experiments

All spectra were recorded at 298 K on Varian Inova 600 MHz and/or 800 MHz spectrometers (Varian, Palo Alto, CA). The 800 MHz spectrometer was equipped with a cryogenically cooled probe. NMRPipe (32) was used to process the spectra, which were analyzed using NMRView (33). NMR assignments were previously reported (30) (BMRB-4999).

### Hydrogen/deuterium exchange

A freeze-dried 1 mM  $^{15}\text{N}$  PX sample was dissolved in  $\text{D}_2\text{O}$  immediately before acquisition of a series of  $^{15}\text{N}$ -SOFAS-HSQC spectra (34). The dead time between dissolving the protein and the first NMR experiment was 3 min. For each amide group the signal intensities in a  $^1\text{H}$ - $^{15}\text{N}$  correlation spectrum as a function of time were fitted to a single exponential decay to obtain the H/D exchange rates ( $k_{\text{ex}}$ ). Exchange rates were converted into protection factors ( $\log(k_{\text{ch}}/k_{\text{ex}})$ ) where  $k_{\text{ch}}$  are the intrinsic exchange rates for a specific residue in a certain sequence, as determined by Bai et al. (35,36):



In the same way equilibrium constants  $K_{\text{op}}$  ( $k_{\text{op}}/k_{\text{cl}}$ ) were calculated from

$$K_{\text{op}} = \frac{k_{\text{ex}}}{k_{\text{ch}}}, \quad (1)$$

which is true in the EX2 exchange regime ( $k_{\text{cl}} \gg k_{\text{ch}}$ ). From the average equilibrium constants for the three helices, average free energies were calculated according to

$$\Delta G = -RT \ln K. \quad (2)$$

### <sup>15</sup>N auto- and cross-relaxation measurements

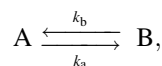
<sup>15</sup>N  $R_1$ ,  $R_2$  Carr-Purcell-Meiboom-Gill (CPMG) and heteronuclear nuclear Overhauser effect (hetNOE) relaxation experiments were recorded as described by Farrow et al. (37) using both the  $\sim 1.3$  and  $\sim 0.2$  mM <sup>15</sup>N-PX samples. Details on the used delays, scans etc. are summarized in the Supplementary Material. For the hetNOE, recorded only at 600 MHz, a 1.7 kHz WALTZ16 decoupling scheme centered at the amide proton frequencies was used to saturate the amide proton signals. During the  $R_1$  relaxation delays, cross correlated relaxation was suppressed by applying a 550  $\mu$ s (600 MHz) or 350  $\mu$ s (800 MHz) cosine-modulated 180° squared pulse every 5 ms with excitation maxima 2 kHz (600 MHz) or 2.8 kHz (800 MHz) from the carrier. The CPMG experiments were recorded with 180° <sup>15</sup>N pulses applied every 1.25 ms ( $\nu_{\text{CPMG}} = 400$  Hz).  $R_2$  relaxation rates at 800 MHz were also obtained from  $R_{1\rho}$  experiments using infrequent 180° <sup>1</sup>H pulses during the relaxation delay to suppress cross correlation pathways as proposed by Korzhnev et al. (38) and a 1.6 kHz on resonance <sup>15</sup>N spin-lock pulse. For both CPMG and  $R_{1\rho}$  experiments on the 800 MHz spectrometer, which is equipped with a cryogenically cooled probe, compensation pulses were applied during the recycle delay to keep the amount of radio frequency (RF) power equal for all experiments. Moreover, pulse lengths and powers were calibrated while sending the same amount of RF power to the probe as used during the relaxation experiments as a slight detuning of the probe was observed. Spectra at 600 (800) MHz were recorded with spectral widths of 6000 (8200) and 1500 (1600) Hz, offsets of 4.77 (4.77) and 116.9 (118.8) ppm in the <sup>1</sup>H and <sup>15</sup>N dimensions, respectively, and a recycle delay of 2 s.

Errors in the signal intensities were estimated from the duplicated points, with a minimal error of the standard deviation in the noise. Decay curves were fitted with the program Curvefit (<http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/software/curvefit.html>) using a two-parameter single exponential and a Monte Carlo simulation to estimate the errors in the fitted parameters. A model-free (39) analysis of the relaxation rates was performed with modelfree4.0 (<http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/software/modelfree.html>), using a value of 1.02 Å for  $r_{\text{NH}}$  and  $-170$  ppm for the amplitude of the <sup>15</sup>N CSA tensor. Error estimates were made from 300 Monte Carlo simulations.

Similarly <sup>15</sup>N  $R_1$  and  $R_2$  relaxation rates of the triple-labeled sample were measured to accurately determine the rotational correlation time  $\tau_c$  to be used for analysis of the <sup>2</sup>H relaxation analysis.

### <sup>15</sup>N relaxation dispersion measurements

<sup>15</sup>N relaxation dispersion measurements were performed with the  $\sim 1.3$  mM PX sample at both 600 and 800 MHz using a relaxation-compensated CPMG pulse sequence (40), where the CPMG period was implemented in a constant time (CT) manner (21). As discussed above, compensation pulses were applied during the recycle delay to keep the amount of RF power equal for all experiments on the 800 MHz spectrometer equipped with a cryogenically cooled probe. The CT period ( $T_{\text{CP}}$ ) was set to 40 ms and by changing the number of 180° (5 kHz (600 MHz)/4 kHz (800 MHz)) pulses during  $T_{\text{CP}}$ , spectra with 10 different CPMG frequencies ( $\nu_{\text{CPMG}}$ ) were recorded: 25 ( $\times 2$ ), 50, 75, 125, 200, 300, 400 ( $\times 2$ ), 600, 800, and 1000 Hz, where  $\nu_{\text{CPMG}} = 1/(4\tau_{\text{CP}})$  and  $2\tau_{\text{CP}}$  is the delay between two 180° <sup>15</sup>N pulses. Spectral widths were set to 6980 (8800)  $\times$  1430 (1900) Hz and number of complex points to 480 (512)  $\times$  120 (160) for <sup>1</sup>H and <sup>15</sup>N, respectively, at 600 (800) MHz. For the reference experiment, where  $T_{\text{CP}}$  was set to 0 ms, 16 (16) scans were used, whereas the CPMG experiments were recorded with 32 (24) scans. Using the signal intensity of the reference experiment ( $I_{\text{ref}}$ ), the effective transverse relaxation rates for each experiment could be determined using  $R_2^{\text{eff}}(\nu_{\text{CPMG}}) = (-1/T_{\text{CP}})\ln[I(\nu_{\text{CPMG}})/I_{\text{ref}}]$  (41). Since only residues in the structured part were found to have <sup>15</sup>N relaxation dispersion decays, it seemed reasonable to analyze the data assuming exchange between two states:



with an exchange rate  $k_{\text{ex}} = k_a + k_b$ , populations  $p_a = (k_b/k_{\text{ex}})$  and  $p_b = (k_a/k_{\text{ex}})$ , a chemical shift difference  $\Delta\omega = \omega_a - \omega_b$ , and transversal autorelaxation rates  $R_{2a}$  and  $R_{2b}$ . A simultaneous nonlinear least-squares fit of the data points at the two fields was performed using the following equation, where the autorelaxation rates were assumed to be the same in the two states, i.e.,  $R_{2a} = R_{2b} = R_2$ :

$$R_2^{\text{eff}} = R_2 + \frac{k_{\text{ex}}}{2} - \nu_{\text{CP}} \cosh^{-1}[D_+ \cosh(\eta_+) - D_- \cos(\eta_-)]$$

$$D_{\pm} = \frac{1}{2} \left[ \pm 1 + \frac{\psi + 2\delta\omega^2}{\sqrt{\psi^2 + \xi^2}} \right] \eta_{\pm} = \frac{\sqrt{2}}{4\nu_{\text{CP}}} \sqrt{\pm\psi + \sqrt{\psi^2 + \xi^2}}$$

$$\psi = (k_{\text{ex}}(2p_b - 1))^2 - \delta\omega^2 + 4p_b k_{\text{ex}}^2 (1 - p_b)$$

$$\xi = 2\delta\omega k_{\text{ex}}(2p_b - 1). \quad (3)$$

This equation is valid for intermediate to fast exchange. In some cases, however, the exchange contribution was very small ( $< 2$  Hz) and fitting of the data points failed. For these residues only an estimation of the  $R_{\text{ex}}$  value was determined from the difference between  $R_{2,\text{eff}}(\nu_{\text{CP}} = 25$  Hz) and  $R_{2,\text{eff}}(\nu_{\text{CP}} = 1000$  Hz). The residues for which fitting of the decay curves succeeded were grouped according to the magnitude of their exchange rate and their location in the protein structure, followed by a simultaneous fit using one unique exchange rate per group. Only when the  $\chi^2$  of the simultaneous fit was not larger than 1.2 times the sum of the  $\chi^2$  values of the individual fits was the simultaneous fit considered relevant. Errors in the fitted parameters were estimated from 100 Monte Carlo simulations.

### <sup>2</sup>H relaxation measurements

The five deuterium relaxation experiments were recorded both at 600 and 800 MHz using pulse programs provided by the group of L. E. Kay (19,42), where pulse scheme C (see Fig. 2 in Millet et al. (42)) was used for the  $R(D_+D_z + D_zD_+)$  experiment. In all cases, an ‘‘on the fly’’ correction for relaxation of the  $C_2H_2$  coherence was performed. The sequences to measure the three multiple quantum relaxation rates were slightly adapted to improve sensitivity, as discussed below. An overview of the acquired data sets at 600 and 800 MHz are given in Table S-1 of the Supplementary Material. Spectral widths and offsets at 600 (800) MHz were set to 8000 [9804]  $\times$  3000 [4000] Hz and 4.76 [4.79]  $\times$  19 [19.5] ppm for <sup>1</sup>H and <sup>13</sup>C, respectively. A <sup>2</sup>H 670 ( $\pm 45$ ) [835 ( $\pm 92$ )] Hz spin-lock pulse was used in the experiment to measure the  $R(D_+)$  relaxation rate. Calibration of the field strength was performed as proposed by Guenneugues et al. (43). <sup>2</sup>H hard pulses had a field strength of 1.8 [1.4] kHz. All deuterium pulses were given centered at 1.8 ppm.

Duplicated points were used to determine errors in the signal intensities using a minimal error equal to the standard deviation in the noise. Decay curves were fitted using Curvefit. Peaks that partly overlap were left out of the analysis. However, in the case of complete overlap and a proper fit of the relaxation decay, the same relaxation rate was assigned to the overlapping methyls (i.e., A-504 <sup>$\beta$</sup> /A-507 <sup>$\beta$</sup> , I-470 <sup>$\gamma^2$</sup> /I-481 <sup>$\gamma^2$</sup> /I-490 <sup>$\gamma^2$</sup> , I-481 <sup>$\delta^1$</sup> /I-490 <sup>$\delta^1$</sup> , and L-543 <sup>$\delta^1$</sup> /L-543 <sup>$\delta^2$</sup> ). Apart from the quadrupolar contribution, Millet et al. (42) showed that dipolar interactions between carbon and deuterium within the methyl group, as well as between proton and deuterium, contribute significantly to the double quantum relaxation rate. Corrections for these contributions can be made using the  $R(C_2H_2)$  rates. These were measured using 16 scans; 512  $\times$  84 complex points; 0.05, 5, 10, 15, 20, 30, 40, 60 ( $\times 2$ ), and 80 ms relaxation delays at 600 MHz and 16 scans; 512  $\times$  84 complex points; and 0.05 ( $\times 2$ ), 4.5, 9.5, 15, 21, 28, 36, 48, and 63 ms relaxation delays at 800 MHz. In addition, evolution under the <sup>1</sup>J<sub>CD</sub> coupling was shown to systematically reduce the measured double quantum relaxation rates but can be accounted for by a multiplication of the measured rates by a factor of 1.025. The <sup>2</sup>H double quantum relaxation rates used in the analysis in this article and listed in the Supplementary Material were all corrected in this way. As was shown by Skrynnikov et al. (44), the five deuterium relaxation rates are dominated by the deuterium quadrupolar interaction and thus solely

depend on the spectral density function at zero frequency (0) as well as the single ( $\omega_D$ ) and double ( $2\omega_D$ ) deuterium frequencies:

$$R(D_z) = \frac{3}{40} \left( \frac{e^2 q Q}{\hbar} \right)^2 [J(\omega_D) + 4J(2\omega_D)] \quad (4)$$

$$R(D_+) = \frac{1}{80} \left( \frac{e^2 q Q}{\hbar} \right)^2 [9J(0) + 15J(\omega_D) + 6J(2\omega_D)] \quad (5)$$

$$R(D_+^2) = R(D_x^2 - D_y^2) = \frac{3}{40} \left( \frac{e^2 q Q}{\hbar} \right)^2 [J(\omega_D) + 2J(2\omega_D)] \quad (6)$$

$$R(3D_z^2 - 2) = \frac{3}{40} \left( \frac{e^2 q Q}{\hbar} \right)^2 [3J(\omega_D)] \quad (7)$$

$$R(D_+ D_z + D_z D_+) = \frac{1}{80} \left( \frac{e^2 q Q}{\hbar} \right)^2 [9J(0) + 3J(\omega_D) + 6J(2\omega_D)], \quad (8)$$

where  $(e^2 q Q/\hbar)$  is the quadrupolar coupling constant, which was experimentally shown to have a rather uniform value of 167 kHz (45). Using these equations, five (0,  $\omega_D$ [600],  $\omega_D$ [800],  $2\omega_D$ [600], and  $2\omega_D$ [800]) spectral density values were extracted from the  $10^2$ H relaxation rates via a least-squares minimization algorithm in Octave. These spectral density values were concomitantly used in a model-free analysis, using the Lipari-Szabo (LS) (39) form for the spectral density function

$$J(\omega) = \alpha S_f^2 \frac{\tau_c}{1 + \omega^2 \tau_c^2} + (1 - \alpha S_f^2) \frac{\tau}{1 + \omega^2 \tau^2}, \quad (9)$$

with  $\tau_c$  the correlation time for global rotational diffusion,  $S_f^2$  the order parameter for fast internal motions with associated correlation time  $\tau_c$ ,  $\tau^{-1} = \tau_c^{-1} + \tau_e^{-1}$ ,  $\alpha = (3\cos^2\theta - 1)^2/4$ , and  $\theta$  the angle between the C-D bond and the methyl symmetry axis. For the unstructured part as well as for some methyl groups in the structured part, an effective correlation time, which takes into account both overall tumbling and local ns motions, was used. When statistically relevant an extended form of the spectral density function was used, which accounts for the presence of both fast and slow internal dynamics (46):

$$J(\omega) = \alpha S_f^2 S_s^2 \frac{\tau_c}{1 + \omega^2 \tau_c^2} + \alpha S_f^2 (1 - S_s^2) \frac{\tau_1}{1 + \omega^2 \tau_1^2} + S_s^2 (1 - \alpha S_f^2) \frac{\tau_2}{1 + \omega^2 \tau_2^2} + (1 - S_s^2) (1 - \alpha S_f^2) \frac{\tau_3}{1 + \omega^2 \tau_3^2}, \quad (10)$$

where  $\tau_1^{-1} = \tau_c^{-1} + \tau_s^{-1}$ ,  $\tau_2^{-1} = \tau_c^{-1} + \tau_e^{-1}$ ,  $\tau_3^{-1} = \tau_c^{-1} + \tau_s^{-1} + \tau_e^{-1}$ , and  $S_s^2$  is the order parameter for slow internal motions with associated correlation time  $\tau_s$ . Extraction of the model-free parameters as well as model selection using an F-test were performed in Octave. A grid search was used to define the initial values of the parameters to be fitted, followed by a nonlinear least-squares minimization algorithm. Error estimates were made from 500 Monte Carlo simulations. The values for the order parameters were constrained between zero and one to improve stability of the fit.

**Pulse sequence to measure  $R(D_+^2)$ ,  $R(3D_z^2 - 2)$ , and  $R(D_+ D_z + D_z D_+)$**

In the experiments to measure both  $R(D_z)$  and  $R(D_+)$  relaxation rates (19), the delay during which the magnetization evolves due to the one-bond

deuterium-carbon scalar coupling ( $^1J_{CD}$ ) has a value of  $1/(4^1J_{CD})$ . To measure the other three deuterium relaxation rates the coupling evolution time is adapted to  $1/(2^1J_{CD})$  according to Millet et al. (42) where different pulse schemes are used to select the coherence of interest. To refocus the  $^{13}C$ - $^{13}C$  coupling in a fully  $^{13}C$ -labeled protein, the total delay ( $2T_C$ ; see Supplementary Material Fig. S1) before and after the relaxation period is adjusted to a multiple of  $(1/^1J_{CC})$ , i.e.,  $\sim 29$  ms, whereas the time needed for the  $^1J_{CD}$  coupling to evolve is only  $1/2^1J_{CD}$ , i.e.,  $\sim 24$  ms. Therefore during the remaining 5 ms the  $^1J_{CD}$  needs to be refocused achieved by a  $180^\circ$   $^2H$  pulse in the sequence by Millet et al. (42). We instead applied a decoupling scheme during this period (see Supplementary Material Fig. S1), as in the  $R(D_z)$  and  $R(D_+)$  experiments (19).

## RESULTS

### Hydrogen/deuterium exchange

H/D exchange experiments of PX were performed at pH 6.0. Under these conditions exchange rates are in the EX2 regime, and protection factors ( $\log[1/K_{op}]$ ) can thus be calculated using Eq. 1. These factors, calculated from the H/D exchange rates ( $k_{ex}^{H/D}$ ) measured by dissolving freeze-dried  $^{15}N$  PX in  $D_2O$ , are shown in Fig. 2. All amide protons were exchanged within 2.5 h. Amide protons that exchanged within the dead time (3 min) of the experiment are indicated with negative bars. These amide protons were also exchanged within 200 s when mixing PX in  $H_2O$  with  $D_2O$  (see Supplementary Material Fig. S3). It could therefore be excluded that fast exchange resulted from slow protein refolding when dissolving the freeze-dried sample and validates the conclusion that these amide protons are not involved in stable hydrogen bonds.

All amide protons in the N-terminal part of PX exchanged within the dead time confirming the unstructured nature of this N-terminal subdomain. Several residues at the start of  $\alpha 1$  helix also exchange within the dead time of the experiment. The lack of a suitable hydrogen-bond partner explains this finding for residues 519–521. In contrast, the hydrogen bond between the amide proton of L-522 and the carbonyl of T-518, which should be present based on the three-dimensional structure of PX, is apparently opening rapidly. The other amides in helix  $\alpha 1$ , including the turn that follows comprising S-528 and S-529, are protected, reflecting the  $HN^i-O^{i-4}$  hydrogen bond pattern, as is also the case for helices  $\alpha 2$  and  $\alpha 3$ . However, the average protection factor of 2.0 for helix  $\alpha 1$  is significantly lower than the average values of 2.5 and 2.4 for  $\alpha 2$  and  $\alpha 3$ , respectively, indicating that this first helix has a lower stability. In the loop between helix  $\alpha 1$  and  $\alpha 2$ , both L-531 and S-532 are protected. The two amides are oriented toward the interior of the protein and although L-531 is hydrogen bonded with the carbonyl of S-529, S-532 is likely protected by the side-chain oxygens of E-535. Likewise, the amide of E-535 can be protected by the side chain of S-532. In the loop between helix  $\alpha 2$  and  $\alpha 3$ , only T-548 is protected, which is pointing toward the hydrophobic core and can form a hydrogen bond with the side chain of E-551.

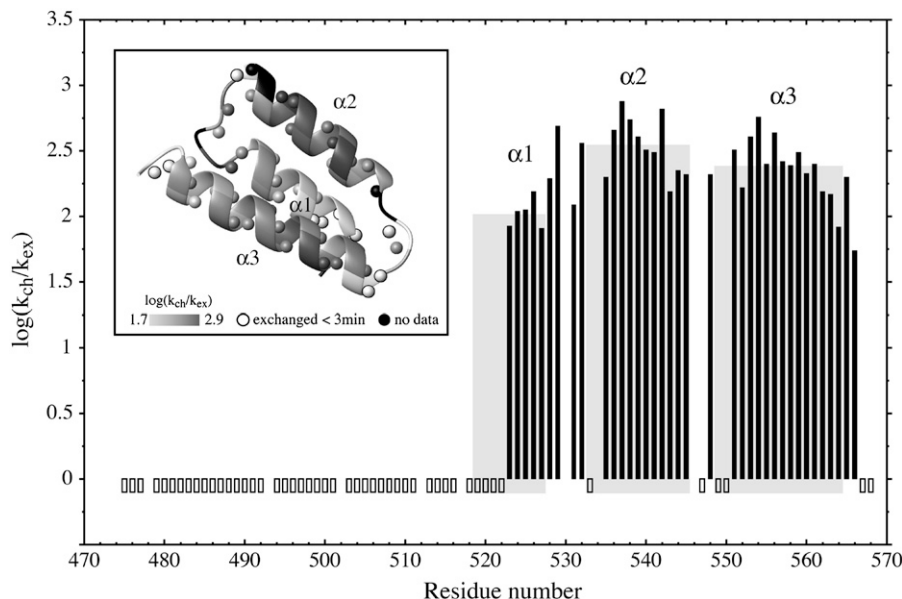


FIGURE 2 Protection factors of PX. Hydrogen/deuterium exchange rates were normalized using exchange rates of a random chain as proposed by Bai et al. (35,36) to obtain protection factors. These are represented by the black filled bars for the backbone amide protons. Open bars indicate exchange within the dead time (3 min) of the experiment. Gray squares indicate the three helices, where the height of the square indicates the average protection factor for the concerned helix. (Inset) Ribbon of the structured C-terminal subdomain of PX. Spheres represent backbone amide protons. Both the ribbon and the spheres are color coded from light gray (low protection factor) to dark gray (high protection factor). Black regions represent missing data due to overlap as well as one proline residue. White spheres with black rims are the amides that exchanged within the dead time of the experiment.

### Fast ps-ns backbone dynamics of PX

The  $^{15}\text{N}$  relaxation rates of PX at two fields and at two different protein concentrations are presented in Fig. 3. A dependence of the relaxation rates on the protein concentration is observed, with increased  $R_2$  and decreased  $R_1$  rates for the more concentrated sample. This indicates that PX, which was identified to be monomeric (14), starts to aggregate at higher protein concentrations.

As mentioned previously (14), low but mainly positive values of the hetNOE are found in the unstructured part of PX. We therefore postulated initially that there might be some residual structure causing the NOE values to be

positive (14). Here residual structure is considered distinct from the inherent properties of the amino acid chain and implies the presence of persistent local or global interactions. Recently we demonstrated that the backbone RDCs in the unstructured region of PX can very well be reproduced by an ensemble of random structures, using a database of residue-selective  $\phi/\psi$  propensities for coil regions (13). No persistent tertiary structure could thus be dissected from these data for this polypeptide chain.

Comparison with hetNOE values of other unstructured proteins or protein domains shows that although predominantly negative in some cases (phage  $\lambda\text{N}$  (47), tropoelastin

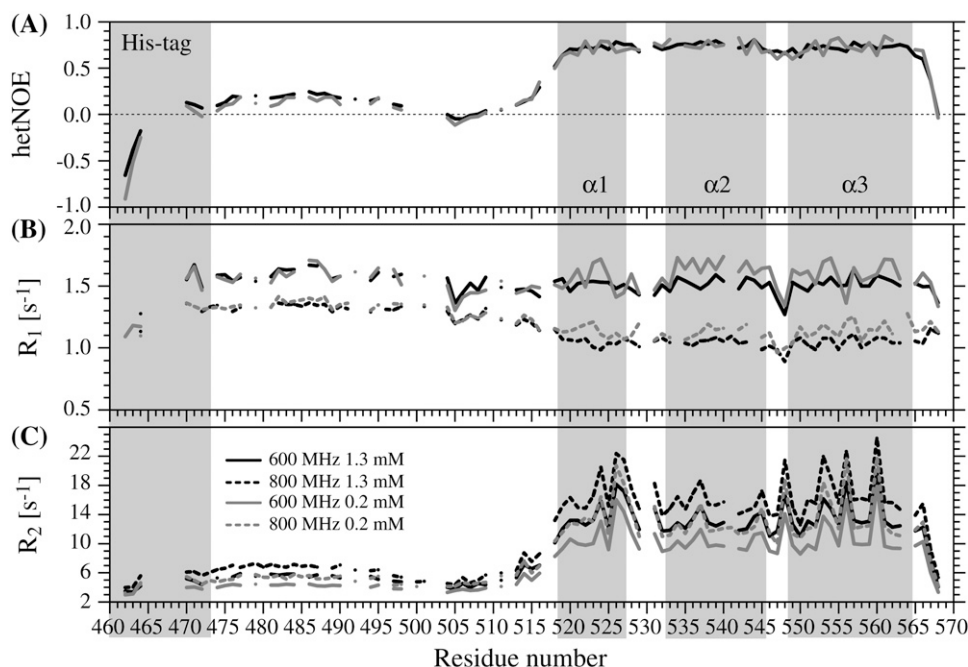


FIGURE 3  $^{15}\text{N}$  relaxation rates of PX at 298 K. Two sets of  $^{15}\text{N}$  hetNOE values (A), and four sets of  $R_1$  (B) and  $R_2$  (C) relaxation rates are presented that correspond to relaxation rates at a  $\sim 1.3$  mM (black) or  $\sim 0.2$  mM (gray) protein concentration and are recorded with a 600 MHz (solid) or 800 MHz (dashed) spectrometer. Error bars are omitted for clarity but can be found in the tables in the Supplementary Material. The three helices as well as the His tag are indicated by the gray areas.

D26 (48), ATF-2 transactivation domain (49)), positive values, even up to 0.4, are regularly found (fibronectin binding protein (50), pro-peptide of subtilisin (51), unfolded ubiquitin (52)). This indicates that the positive hetNOEs may, rather, be a result of an inherent property of the individual nature of the amino acids along the sequence. Indeed, a high percentage of large and charged side chains is present in the N-subdomain of PX, and the stretch (470–498) that has positive hetNOE values lacks both Ala and Gly residues, which evidently slightly restricts the amplitudes of fast motions. The region that contains a series of relatively small residues (A-504, S-505, N-506, A-507, S-508) does display very low and even negative NOE values. Similarly, lower  $R_1$  and  $R_2$  values are found for the residues in this region. Zooming in on the  $R_2$  relaxation rates in the unstructured part reveals that they follow a bell-shaped profile, with uniform values in the middle and lower values at both ends, as is expected for an unstructured polypeptide that has by nature a chain-like character (53). This bell-shaped profile tends to confirm the absence of residual structure, since the presence of local structure likely results in locally increased  $R_2$  rates (54).

Several residues in the structured part of PX have high transverse relaxation rates, which indicate the presence of slow conformational exchange ( $R_{ex}$ ). This is confirmed by the fact that lower transverse relaxation rates were found for these residues using a  $^{15}\text{N}$  spinlock field ( $R_{1\rho}$ ) instead of a CPMG sequence during the relaxation period (see Fig. 5, A and B). This contribution has been quantified using relaxation dispersion measurements, as discussed in more detail below.

Fitting of the relaxation data at a high protein concentration using the LS model-free approach led to artificially high order parameters ( $S_{ave}^2 = 0.93$ ) and long internal correlation times ( $>400$  ps) for some of the residues in the structured part. These features are clearly related to the observed aggregation (26), which is not taken into account by the LS model and is only implicitly parameterized in an increased overall correlation time (8.9 ns). The relaxation rates at the low protein concentration could, however, be analyzed with the model-free approach (Fig. 4). For the structured part, an isotropic diffusion model, using a  $\tau_c$  of 7.3 ns, was considered acceptable since  $D_{par}/D_{per}$  was found to be only 1.15.

The correlation time is slightly higher than what is expected for a 12.7 kDa globular protein. This might be due to the presence of aggregation, even at this low protein concentration. However, PX is not a globular protein, and another source of the higher correlation time might be the fact that in PX the two domains have rather similar mass, such that the overall diffusion of the structured part is dependent on the unstructured part. An extended conformation of the unstructured part would therefore result in completely different rotational diffusion of the protein than a conformation where the unstructured part is, for example, back-folded on the structured domain. Since the unstructured part is very flexible, these types of conformations are expected to interconvert rapidly. But since the timescale of this process is not a priori known, it is not necessarily true that the rotational diffusion of the structured part can be accurately described by one average correlation time. The order parameters ( $S^2$ ) in the well-structured area (excluding the C-terminus) have an

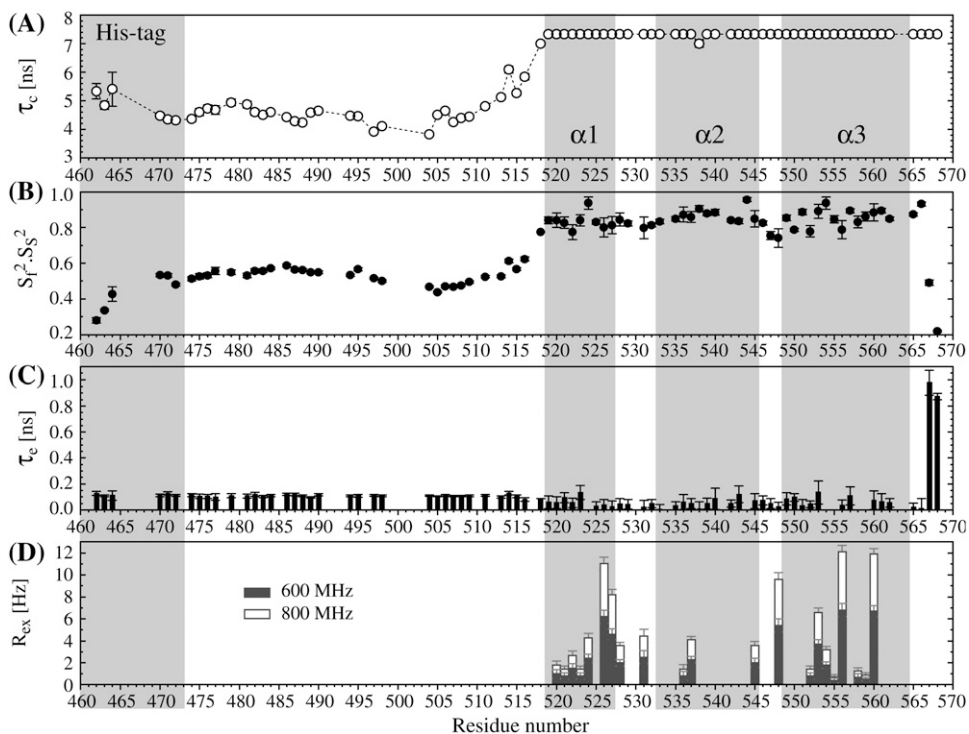


FIGURE 4 Model-free analysis of the PX  $^{15}\text{N}$  relaxation rates measured at both 600 and 800 MHz and a protein concentration of  $\sim 0.2$  mM. (A) Overall rotational correlation time  $\tau_c$ . Isotropic rotational diffusion with a correlation time of 7.3 ns was assumed for the structured part (519–568), whereas a local diffusion model was used for residues in the unstructured part. (B) Product of order parameters for fast ( $S_1^2$ ) and slow ( $S_2^2$ ) internal motions. An extended model including  $S_s^2$  was only used for T-567 ( $S_s^2 = 0.79$ ) and N-568 ( $S_s^2 = 0.68$ ). (C) Local correlation time ( $\tau_c$ ) for internal motion. (D)  $R_{ex}$  contribution to the  $R_2$  relaxation rates.

average value of  $0.85 \pm 0.05$ , indicating a rather rigid structure. Only some residues located in the loops between the helices as well as in the C-terminus have lower order parameters. In this subdomain a large number of residues were fitted using a model including a contribution of conformational exchange ( $R_{ex}$ ) to the transverse relaxation rate. These include all the residues for which the existence of conformational exchange was detected from the difference in CPMG and  $R_{1\rho}$ -derived transverse relaxation rates.

To fit the  $^{15}\text{N}$  relaxation rates in the unstructured part, three fitting parameters were used in the model-free analysis, i.e., an effective correlation time  $\tau_c^{\text{eff}}$ , an order parameter  $S^2$ , and a correlation time for fast local motions  $\tau_e$ . We are aware of the fact that by using this form of the spectral density function we oversimplify the potentially complex nature of dynamics in an unstructured polypeptide chain. In particular, the assumed absence of correlation between overall tumbling and the internal dynamics is not evident. However, the small number of observables and the sparse knowledge concerning the nature of dynamics in unstructured proteins restricts the use of more sophisticated analyses at this stage. Therefore it is important to keep in mind that the extracted fitting parameters give only a simplified view of the dynamics. The fitted values for the correlation time  $\tau_e$  for fast local motions are very similar over the whole of the unstructured chain ( $110 \pm 7$  ps).

The effective rotational correlation times as well as the order parameter, however, do slightly vary along the sequence. In particular, lower order parameters are found for residues 504–508. These are dictated by both the low hetNOE values and  $R_2$  relaxation rates, as discussed above. More specifically, this region can be presented as a flexible hinge between the structured and unstructured subdomains. Although the effective correlation time also varies in this region (3.8–4.5 ns), which is also the case for the N-terminal end, a rather uniform value of  $4.6 \pm 0.2$  ns is found for residues 474–

495. This again shows that the unstructured subdomain can be regarded as an independent module whose dynamics are not correlated with the dynamics of the structured part. The average order parameter for the unstructured subdomain is  $0.53 \pm 0.04$ , excluding the residues of the His tag. Order parameters in the same range were also found for other unstructured proteins (51,55–57) or protein domains (58).

### Slow $\mu\text{s}$ -ms backbone dynamics of PX

From both the difference between the transverse relaxation rates determined using a CPMG or  $R_{1\rho}$  sequence (Fig. 5, A and B) and the model-free analysis (Fig. 4 D), it is clear that several regions in the C-terminal structured subdomain of PX undergo slow  $\mu\text{s}$ -ms conformational exchange. Since the  $R_{ex}$  contribution is independent of the protein concentration (Fig. 5, A and B), it can be excluded that the exchange originates from aggregation. To further characterize this conformational exchange we recorded  $^{15}\text{N}$  relaxation dispersion measurements at both 600 and 800 MHz, where for sensitivity purposes the higher concentrated sample was used. Fitted relaxation dispersion curves of three residues in different parts of the structured subdomain of PX are presented in Fig. 6. Fig. 5 C shows an overview of all the residues for which conformational exchange was detected from the relaxation dispersion measurements. These residues closely correspond to the ones for which a significant difference between  $R_2(\text{CPMG})$  and  $R_2(R_{1\rho})$  was also found. For some residues (*black open bars* in Fig. 5 C) fitting of the data points failed due to a small, but significant  $R_{ex}$  value (as determined from the difference between  $R_2^{\text{eff}}$  (25 Hz) and  $R_2^{\text{eff}}$  (1000 Hz)). These residues are located in the middle and at the end of helix  $\alpha 2$ . The gray bars in Fig. 5 C represent residues for which curve fitting did succeed.

The first group of residues (group I in Fig. 5 C) is located at the end of helix  $\alpha 1$  (L-524, I-526, E-527) as well as in the

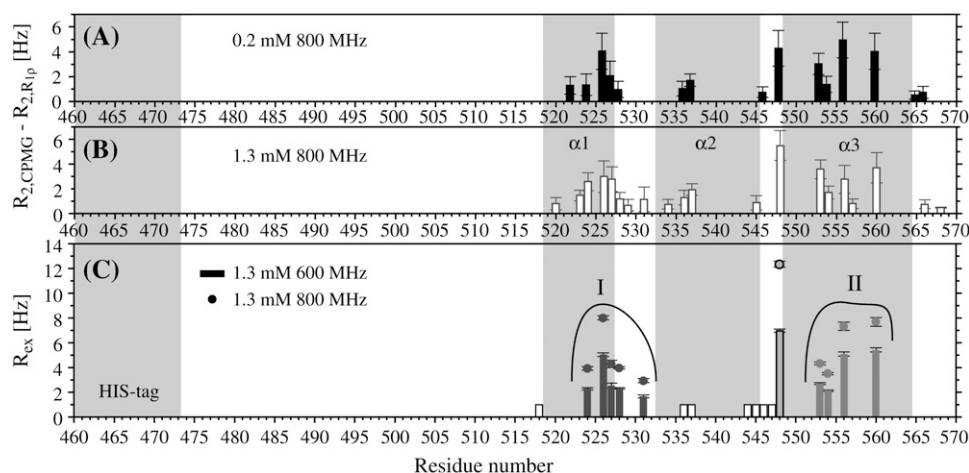


FIGURE 5 Contribution of conformational exchange ( $R_{ex}$ ) to the  $^{15}\text{N}$  transverse relaxation rates  $R_2$  of backbone amide groups in PX, determined from (A and B) the difference between  $R_2$  rates from CPMG and  $R_{1\rho}$  measurements at 800 MHz or from (C)  $^{15}\text{N}$  relaxation dispersion measurements at 600 (bars) and 800 (circles) MHz. Protein concentrations were  $\sim 0.2$  mM for A and  $\sim 1.3$  mM for B and C. In C, open bars represent residues for which dispersion curves were not fit due to low  $R_{ex}$  values. Gray bars/circles represent  $R_{ex}$  values extracted from fits of the relaxation dispersion curves. Curve fitting of residues in group I (L-524, I-526, E-527, S-528, and L-531) or II (K-553, A-554, M-556, and E-560) was performed simultaneously.

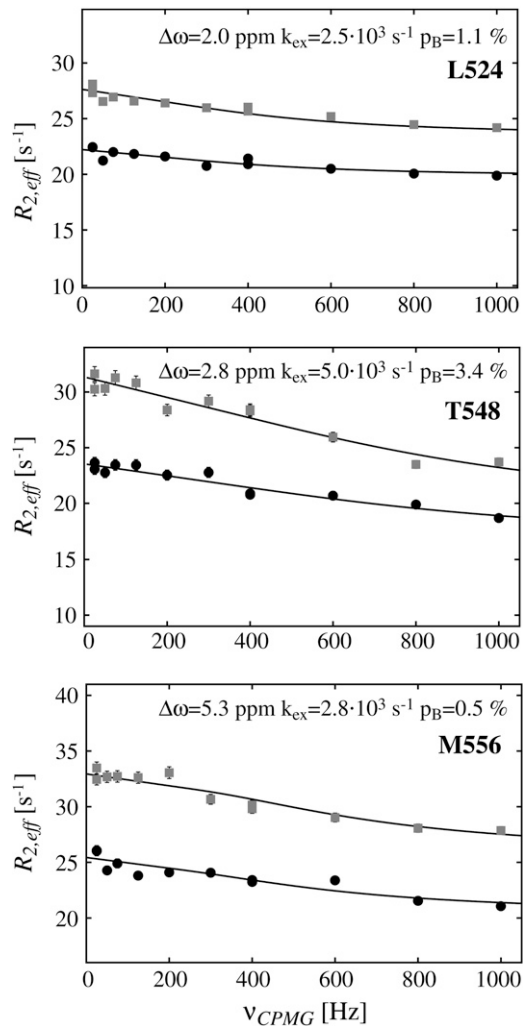


FIGURE 6  $^{15}\text{N}$  backbone relaxation dispersion curves of three residues in the structured part of PX at 600 (circles) and 800 (squares) MHz. Fitting parameters are indicated on top of the graphs, where for L-524 and M-556 values from a simultaneous fit of residues in group I and II (see Fig. 5), respectively, are indicated.

loop between  $\alpha 1$  and  $\alpha 2$  (S-528, L-531). Since exchange rates and populations after fitting the individual curves were in the same range, the dispersion curves of the five residues were fitted simultaneously, assuming that one process causes this exchange. The total  $\chi^2$  when fitting these five residues simultaneously increased only by a factor of 1.2 as compared to the sum of  $\chi^2$  of the individual fits. This resulted in an exchange rate of  $2538 \pm 142 \text{ s}^{-1}$  with a population of  $1.1\% \pm 0.1\%$  of the “excited” state ( $k_a = 30 \pm 2 \text{ s}^{-1}$ ,  $k_b = 2508 \pm 176 \text{ s}^{-1}$ ). This exchange very likely reflects an unfolding event of helix  $\alpha 1$ , which is also sensed by S-528 and L-531, whose amide protons are connected to this helix through their hydrogen bonds. The equilibrium constant of the exchange process can be extracted from the populations ( $K = p_a/p_b = 11.1 \times 10^{-3} \pm 1.0 \times 10^{-3}$ ) and thus the free energy  $\Delta G$ :  $11.1 \pm 0.2 \text{ kJ/mol}$ . This value corresponds

remarkably well to the free energies of these residues calculated from the H/D exchange rates using Eq. 2, ranging from 10.9 to 13.0 kJ/mol, confirming that the observed conformational exchange originates from the unfolding of helix  $\alpha 1$ . The absolute values of the chemical shift changes associated with the exchange process for the three residues at the end of helix  $\alpha 1$  correspond reasonably well with the chemical shift differences between the observed chemical shifts and random coil values (Table 1).

Indeed, for two other residues at the end of helix  $\alpha 1$  for which no exchange was observed (R-523, V-525), the chemical shift deviation from random coil is very low, possibly resulting in the presence of conformational exchange to be undetectable for these residues. In contrast, no exchange is observed for the residues in the beginning of the helix, whereas their chemical shifts do deviate from random coil. Presumably the exchange process in this part of  $\alpha 1$  is faster than in the second half of the helix and therefore invisible. From the H/D exchange experiments we had concluded that the amides in this part of helix  $\alpha 1$  are not involved in stable hydrogen bonds, as is the case for the residues that follow. The expected  $^1\text{H}^{\text{N}}\text{-}^1\text{H}^{\text{N}}(i,i+1)$  NOEs are relatively weak for this part of helix  $\alpha 1$  and long-range  $^1\text{H}\text{-}^1\text{H}$  NOEs were only detected for L-552 on the hydrophobic side of helix  $\alpha 1$  (14). In contrast, for the second part of helix  $\alpha 1$   $^1\text{H}^{\text{N}}\text{-}^1\text{H}^{\text{N}}(i,i+1)$  NOEs are strong, and for both V-525 and I-526 many long-range NOEs with residues on both helix  $\alpha 2$  (Y-539, V-540) and helix  $\alpha 3$  (V-559) were identified (14). Taken together these observations indicate that at the beginning of helix  $\alpha 1$  (519–522) fraying of the helix is a fast process due to the lack of stable hydrogen bonds and the absence of a strong hydrophobic network, which makes this process “invisible” for NMR relaxation measurements. The second part of the helix  $\alpha 1$  (523–527) does contain hydrogen bonds detectable from H/D exchange as well as a strong hydrophobic network with the core of the protein, making it energetically less favorable to unfold this second part of the helix.

In the turn between helix  $\alpha 2$  and  $\alpha 3$  a large contribution of conformational exchange was observed only for the backbone nitrogen of T-548, and smaller  $R_{\text{ex}}$  values were found

TABLE 1 Deviations from random coil of the  $^{15}\text{N}$  chemical shifts ( $\delta_{\text{N}}^{\text{obs}} - \delta_{\text{N}}^{\text{r.c.}}$ ) and absolute  $^{15}\text{N}$  chemical shift differences ( $\Delta\omega_{\text{N}}^{\text{fit}}$ ) related to the exchange process

|        | $\delta_{\text{N}}^{\text{obs}} - \delta_{\text{N}}^{\text{r.c.}}$ | $ \Delta\omega_{\text{N}}^{\text{fit}} ^{\dagger}$ |        | $\delta_{\text{N}}^{\text{obs}} - \delta_{\text{N}}^{\text{r.c.}}$ | $ \Delta\omega_{\text{N}}^{\text{fit}} ^{\dagger}$ |
|--------|--|--|--------|--|--|
| M-519* | -3.1 ppm   | –  | I-526* | -4.5 ppm   | 3.1 ppm  |
| H-520* | -2.0 ppm   | –  | E-527* | -1.8 ppm   | 2.1 ppm  |
| S-521* | -1.0 ppm   | –  | S-528  | -4.1 ppm   | 2.0 ppm  |
| L-522* | 1.3 ppm  | –  | S-529  | 2.8 ppm  | –  |
| R-523* | -0.1 ppm   | –  | P-530  | –  | –  |
| L-524* | -3.8 ppm   | 2.0 ppm  | L-531  | -4.3 ppm   | 1.7 ppm  |
| V-525* | -0.5 ppm   | –  |        |  |  |

\*Helix  $\alpha 1$ : 519–527.

$^{\dagger}$  Absolute chemical shift difference of exchange process.



for some of the neighboring residues. The amide proton of T-548 is most probably hydrogen bonded to side chain oxygen in E-551, but it is possible that in this rather flexible turn a local backbone conformational change resulting in a  $\phi/\psi$  flip of T-548 is allowed, which can cause the observed exchange.

In helix  $\alpha 3$  four residues (K-553, A-554, M-556, and E-560) were found to have a relaxation dispersion decay. Fitting the curves simultaneously increased the total  $\chi^2$  only by 1.14 with respect to fitting the individual curves, resulting in a  $k_{\text{ex}}$  of  $2857 \pm 163 \text{ s}^{-1}$  ( $k_a 15 \pm 1 \text{ s}^{-1}$ ,  $k_b 2842 \pm 212 \text{ s}^{-1}$ ). The underlying mechanism is unlikely to be the unfolding of helix  $\alpha 3$ , which is rather stable according to the H/D exchange data. The amides of residues K-553, M-556, and E-560 are all located on the same side of the helix and not far from the single aromatic residue Y-539 of the structured subdomain that is located on helix  $\alpha 2$ . Both M-556 and E-560, which are closest to the aromatic ring, have downfield-shifted  $^{15}\text{N}$  chemical shift values. Moreover, the absolute value of the chemical shift differences extracted from the fit of the relaxation dispersion profiles is larger for M-556 and E-560 (5.3 and 5.6 ppm, respectively) than for K-553 (3.5 ppm) and A-554 (3.0 ppm). It is therefore tempting to speculate that the observed exchange results from a change in orientation of the Y-539 ring with respect to helix  $\alpha 3$ . However, Y-539 is located in the hydrophobic core of the protein, and its four aromatic ring protons all have different chemical shifts, indicating the absence of fast flipping of this ring. Rather than a reorientation of the aromatic residue, it would thus be more likely that a reorientation of the two helices with respect to each other is the underlying event of the observed exchange.

## Methyl $^2\text{H}$ relaxation

*Pulse sequences to measure  $R(D_+^2)$ ,  $R(3D_z^2 - 2)$ , and  $R(D_+D_z + D_zD_+)$*

Application of a  $^2\text{H}$ -decoupling scheme in the sequences to measure  $R(D_+^2)$ ,  $R(3D_z^2 - 2)$ , or  $R(D_+D_z + D_zD_+)$  shortens the existence time of the antiphase term  $C_xH_zD_z$ , which relaxes faster than the in-phase term  $C_yH_z$ . Considering that the  $R(D_z)$  rates for the protein in this study are on the order of  $20 \text{ s}^{-1}$ , shortening the existence of the antiphase term in both  $^1\text{J}_{\text{CD}}$  evolution periods before and after the relaxation delay would roughly result in a signal increase of 20%. This corresponds well with the intensity gain of 20% found in practice (see Supplementary Material Fig. S1).

## $^2\text{H}$ relaxation rates

A selection of  $^2\text{H}$  relaxation decays at both fields as well as the extracted relaxation rates are presented in Fig. 7. Note that the axes of the relaxation rate graphics are discontinuous because only the methyl groups are shown. In general the relaxation decays are of good quality, with only rather noisy

double quantum signal decays. In addition, relaxation rates determined for L-522<sup>δ1</sup> have large error bars due to its low intensity broad signal in the  $^1\text{H}$ - $^{13}\text{C}$  correlation spectrum. The relaxation rates are scaled to show that the  $R(D_+D_z) > 5/3R(D_+) > 5/3R(3D_z^2 - 2) > 5/3R(D_z^2) > R(D_z)$  trend is fulfilled (59), with the exception of some double quantum rates that exceed the quadrupolar rate, which is clearly caused by the lower quality of the data for the double quantum relaxation rates.

In the unstructured part of PX the  $D_+D_z$  and  $D_+$  coherences clearly relax much slower than in the structured part. These relaxation rates are the only two that depend on the spectral density function at zero frequency ( $J(0)$ ), and the large difference in rates thus clearly reflects the difference in effective tumbling of the two subdomains. The entire set of 10 relaxation rates is a function of five spectral density values at different frequencies, i.e.,  $J(0)$ ,  $J(\omega_D^{600})$ ,  $J(\omega_D^{800})$ ,  $J(2\omega_D^{600})$ , and  $J(2\omega_D^{800})$ . These spectral density values were extracted from the relaxation rates. Back calculation of the relaxation rates from these spectral density values shows that the 10 relaxation rates are self-consistent (see Fig. S2 of the Supplementary Material). Only the double quantum relaxation rates at 600 MHz are not well reproduced by the extracted spectral density values. As mentioned above, these rates suffer from noisy data points. However, omitting this data set results in nearly identical spectral density values, showing the robustness of this large set of relaxation rates.

A model-free analysis of the spectral density values was performed. An overview of the extracted motional parameters for the side-chain methyl groups in PX is shown in Fig. 8. As in the model-free analysis for the  $^{15}\text{N}$  relaxation data, an effective residue-specific correlation time was used for residues in the unstructured part. Although this is an unrealistically simple model for side-chain motions in an unstructured protein chain, where a wide range of motions are likely superimposed, the extracted parameters are still indicative of relative amplitudes and timescales. For the structured part, the correlation time was fixed to 7.6 ns, as was determined from the  $^{15}\text{N}$  relaxation data on the same sample.

## Side-chain dynamics in the unstructured PX subdomain

In the unstructured part, the fitted residue-specific effective correlation times were found to be rather uniform, with an average value of  $3.2 \pm 0.1 \text{ ns}$  for the residues 481–507. This confirms the modular behavior of this protein, as was also found from the  $^{15}\text{N}$  relaxation data analysis. From the  $^{15}\text{N}$  relaxation data a ratio of 1.6 between the rotational correlation time of the structured subdomain and the average effective correlation time in the unstructured subdomain was found. The higher ratio of 2.4 found from  $^2\text{H}$  relaxation data is likely caused by the fact that the effective correlation time reports not only on the overall tumbling but also on internal ns motion, meaning that the side chains as compared to the

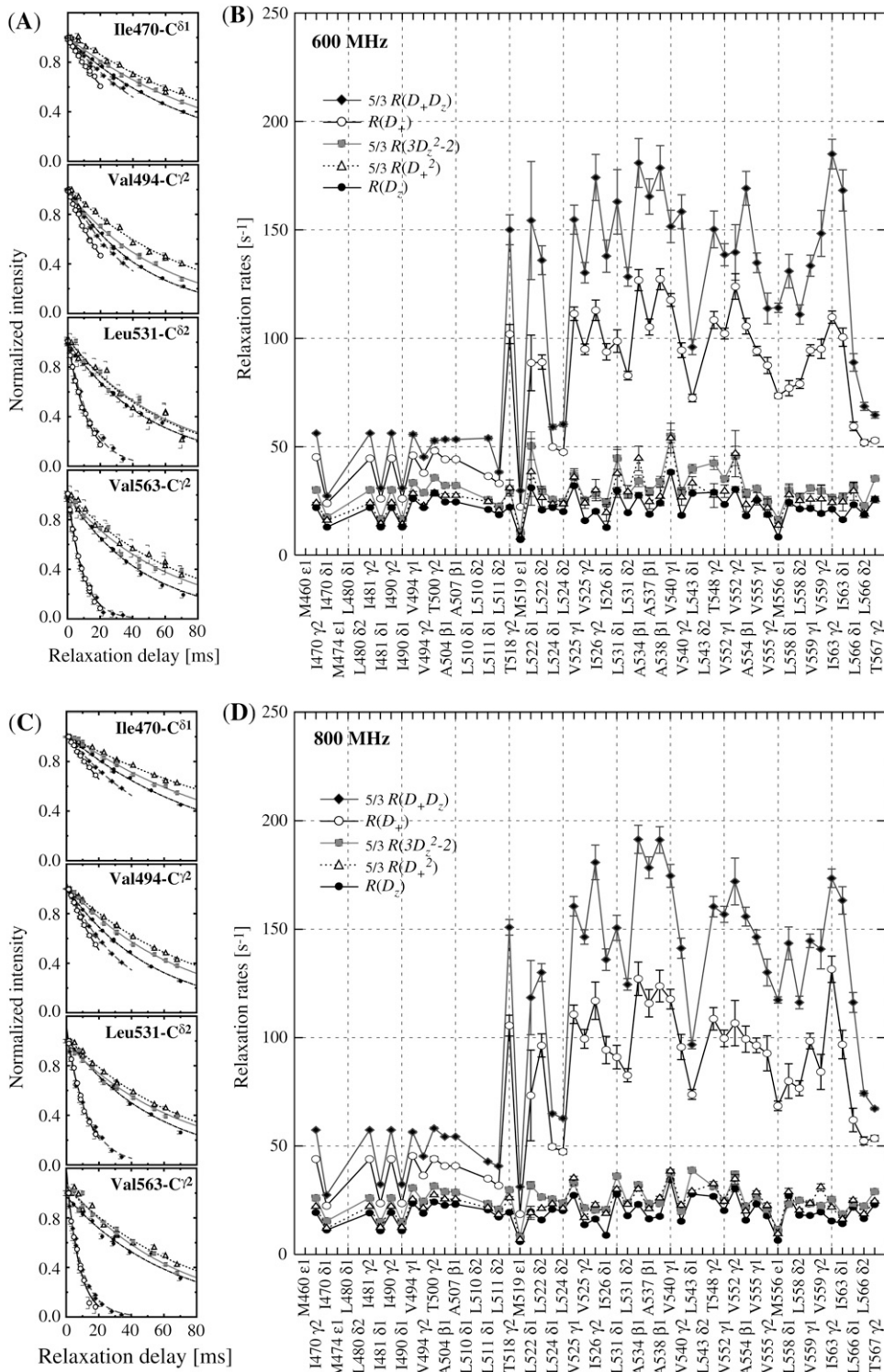


FIGURE 7 <sup>2</sup>H relaxation decay curves and relaxation rates of PX. The relaxation curves at both 600 (A) and 800 (C) MHz are from two residues in the unstructured N-terminal subdomain (I-470 and V-494) and two in the structured C-terminal subdomain. The fitted relaxation rates are shown at 600 (B) and 800 (D) MHz are scaled for comparison as indicated on the graph.

backbone experience additional slow ns motions. In the unstructured part the internal correlation times of fast methyl axis motion range from 20 to 60 ps and order parameters between 0.25 and 0.5. The variation in methyl axis order parameters is partly a result of the different lengths of the

methyl-bearing side chains. As is found for globular proteins, lower  $S_{axis}^2$  values are found for the  $\delta$  methyl groups of I and L in this unstructured chain than for the  $\gamma$  methyl groups of V, T, and I. However, two alanine residues (A-504 and A-507) have relatively low order parameters.

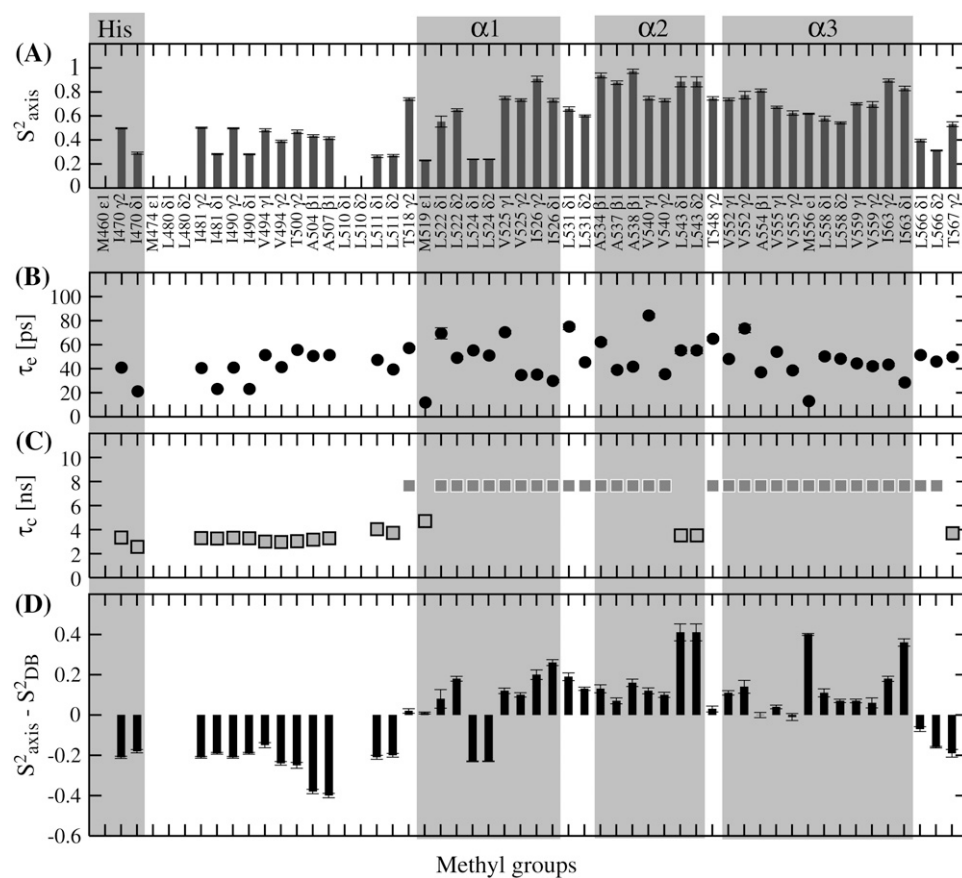


FIGURE 8 Methyl axis model-free parameters from  $^2\text{H}$  relaxation measurements of PX. Order parameters for fast internal motion  $S^2_{\text{axis}}$  of the methyl groups in PX are represented by the bars in A with the corresponding time-scales  $\tau_c$  in B. In C, the effective or global correlation times are represented by dark gray and light gray squares, respectively. The gray frames indicate the residues located in the His tag or the three  $\alpha$ -helices. (D) The order parameters of PX in A were normalized by subtracting a residue-specific average order parameter.

To account for the difference in side-chain length, Mittermaier et al. (60) have proposed to subtract a methyl-specific order parameter, as determined from a database analysis, and to use this normalized  $S^2$  that no longer depends on the length of the side chain. The normalized  $S^2_{\text{axis}}$  values of PX are shown in Fig. 8 D. Negative normalized order parameters indicate flexibility, whereas positive values indicate rigidity. The two alanines A-504 and A-507 both have high negative normalized order parameters. They are located in the stretch (504–508), which from the backbone  $^{15}\text{N}$  relaxation data was identified as a very flexible linker between the unstructured and structured domain; and this highly flexible character is clearly also sensed by the alanine  $\text{C}^\alpha\text{-C}^\beta$  axis. All other normalized order parameters in the unstructured part are also negative, meaning that the methyl groups are more flexible than on average in structured proteins. Remarkably, the order parameters of the two methyl groups in V-494 differ by 0.1. All the other valine order parameters in PX do not differ more than 0.05, and considering that the  $S^2_{\text{axis}}$  error bars of the V-494 methyls are on the order of 0.01, the difference of 0.1 is probably significant.

Differences between valine methyl order parameters of this order or larger have in some cases been observed in well-folded proteins (61–63). They likely originate from complex dynamics due to steric hindrance. In a flexible unstructured protein the existence of complex asymmetric dynamics is,

however, less intuitive. Choy et al. (20) measured methyl axis order parameters in  $\Delta 131\Delta$ , an unstructured fragment of staphylococcal nuclease. To circumvent problems of overlap in the  $^1\text{H}\text{-}^{13}\text{C}$  correlation spectrum, the  $^2\text{H}$  relaxation rates were measured using an experiment where the magnetization is transferred from the methyls to the backbone NH groups. In this way, an average order parameter for the methyls in V, as well as in I and L residues, was measured, and a possible difference between the methyl axis order parameters thus remained unnoticed. V-494 in PX is adjacent to P-493 and Y-495. Previously it was noted from RDC data that the backbone root mean-square deviation around prolines of the calculated ensemble of the unstructured subdomain of PX is considerably lower than at other places in the peptide chain (13). This was explained by the fact that conformational sampling is restricted due to the presence of the proline residue. Valine residues preferentially have a  $\chi_1$  angle of  $180^\circ$  with  $\text{C}^\gamma 1$  closer to its own carbonyl and  $\text{C}^\gamma 2$  closer to the amide group. It might be that the restricted conformational sampling, as a result of the proline that is attached to the amide of V-494, makes the environment of the two methyls unequal, which can even be accentuated by the presence of a large aromatic side chain on the other side. This asymmetry might be the origin of a complex motion with an effective director other than the  $\text{C}^\alpha\text{-C}^\beta$  axis, causing the difference in order parameters for the two methyl groups in V-494.

### Side-chain dynamics in the structured PX subdomain

The normalized methyl order parameters in the structured part are mainly positive, with the exception of L-524, which is protruding into the solvent, as well as the C-terminal residues L-566 and T-567 (Fig. 8 D). The internal correlation times are all between 30 and 75 ps, with faster motions for the two methyl groups in the long methionine side chains, for which internal timescales are around 12 ps (Fig. 8). Methyls that are located in the hydrophobic core all have order parameters higher than 0.6. The two methyl groups of L-543 resonate at the same frequency. For analysis of their relaxation rates a three-parameter fit was performed using an effective correlation time. It should be kept in mind that this effective correlation time takes into account the global  $\tau_c$  as well as both the amplitude and timescale of slow internal motions. Although ps timescale motions are rather restricted for the methyl groups of L-543 ( $S_{\text{axis}}^2 = 0.88$ ), the methyls undergo a slower ns timescale motion. This motion most probably is related to rotameric jumps, whereby the two methyl groups occupy the same rotameric conformations, which results in both methyls resonating at the same frequency.

In Fig. 9 the correlation between the backbone and corrected side-chain order parameters is shown. An average value of the methyl axis order parameters of V, L, and I methyl groups is used. This graph shows that although a correlation between backbone and side-chain motions is virtually absent in the structured subdomain, which is generally found for structured proteins (64), the correlation is more pronounced when including the methyls in the unstructured part. It should be noted, however, that the two lowest data points are from the two alanine residues in the flexible linker, whose side-chain order parameters report on the  $C^\alpha$ - $C^\beta$  axis, which can be considered a backbone vector.

Recently, we identified the interaction site on PX of the unstructured C-terminus ( $N_{\text{TAIL}}$ ) of the SeV nucleoprotein to

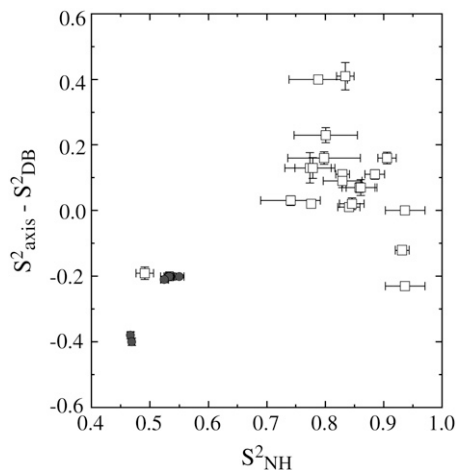


FIGURE 9 Comparison of normalized methyl order parameters and backbone order parameters in PX. White squares represent residues in the structured part, and gray circles residues in the unstructured part.

be located on the face created by helix  $\alpha 2$  and  $\alpha 3$  (65). All the methyl-bearing residues that are located on the  $N_{\text{TAIL}}$  interaction site (V-555, A-534, A-538, T-548, L-558) have zero or even positive normalized order parameters, creating a relative rigid interaction site. This is even the case for L-558, which is rather uncommon for such a long side chain that is solvent accessible.

## DISCUSSION

### Dynamics in unstructured proteins

This detailed analysis of the dynamics in a partly structured protein gives us the unique possibility to compare structured and unstructured protein motions within the same experimental system. Both  $^{15}\text{N}$  and  $^2\text{H}$  relaxation rates are in agreement with a protein that consists of two dynamically distinct parts. These two parts, one structured domain that does not display large-amplitude fast ps motions and one unstructured region that is very flexible on this timescale, are separated by a stretch of small residues. This stretch of residues displays even higher amplitude fast motions probably resulting from the fact that this part of the protein sequence lacks bulky residues and can therefore be considered as a flexible hinge that decorrelates the dynamics of the two domains. Describing PX as a two-domain protein when one of the domains has no structure may appear paradoxical since historically domains have been defined as having an independent persistent structure. For this kind of hybrid protein, however, differential dynamics are rather useful as a criterion for defining the different domains.

In the unstructured part of the protein the effective correlation times from both the backbone and side-chain relaxation analysis are rather uniform over a long stretch of residues. Only residues at the very N-terminal end, as well as residues close to the structured subdomain, have different effective correlation times. This may be interpreted in terms of a global motion with a certain persistence length (53), directed by the random coil polymer whose sequence in this region consists of several bulky amino and lacks both glycine and alanine residues. For the unstructured staphylococcal nuclease  $\Delta 131\Delta$  fragment, a large spread of effective correlation times from  $^{15}\text{N}$  and  $^2\text{H}$  relaxation data was found (20,57). The presence of a significant amount of tertiary structure was, however, identified for  $\Delta 131\Delta$  (66,67). This is, however, not the case for PX, where no residual tertiary structure has been detected in the unstructured part from both RDCs and small angle x-ray scattering data but can be represented as an ensemble of random conformers (13), which is in agreement with the relaxation data presented in this article. Comparison of the backbone and side-chain order parameters in the structured and unstructured subdomains of PX (Fig. 9) reveals that although no strong correlation between backbone and side-chain motions is found in the structured part, as was also shown for several other proteins (62), a weak correlation does

exist in the unstructured part. Such a correlation was also found for  $\Delta 131\Delta$  (20), and it seems that in contrast to structured proteins, the backbone in an unstructured polypeptide chain does, to a certain degree, direct the dynamics of its side chains.

### Distinct dynamics in the subdomains of PX in view of its function

PX is the C-terminal domain of SeV protein P, the cofactor of the RNA polymerase. The unstructured part of PX connects the C-terminal three-helix bundle of PX to the tetrameric coiled-coil domain that binds to L. Although L is the protein that possesses the RNA polymerase activity, the four PX C-terminal three-helix bundles bind the polymerase to the nucleocapsid through an interaction with the unstructured  $N_{TAIL}$  domain (Fig. 1). Whether the four PX domains interact simultaneously or alternatively with the nucleocapsid is not yet known, nor is how the polymerase subsequently proceeds along the nucleocapsid so that L can then synthesize the RNA nucleotide per nucleotide. The highly dynamic character of both the backbone and side chains of the unstructured subdomain of PX, as identified by this study, must, however, be crucial for this process, as this domain forms a flexible linker between the nucleocapsid binding site of the polymerase cofactor and the protein L binding site.

We show here that the unstructured part has the typical dynamic characteristics of a polypeptide that has by definition a chain-like character and that the stretch of small residues in between the two subdomains creates a very flexible hinge which breaks the elongated chain. Although the unstructured chain is important to reach for the unstructured  $N_{TAIL}$  domain, the hinge region allows reorientation of the helix bundle without any hindrance. Tuckis et al. (68) investigated the effect of engineered mutations in the PX domain within the full phosphoprotein on the polymerase activity in both transcription and replication *in vitro*. Interestingly, mutation of both N-506 and R-509 in the hinge region to alanine residues resulted in a complete loss of polymerase activity, whereas the polymerase was still able to bind nucleocapsids. Likely these mutations altered the dynamics in this region, which might in turn have resulted in a drastic change in the ability of the polymerase to efficiently progress along the nucleocapsid.

The small helix  $\alpha 1$  that follows the unstructured subdomain is not very stable, and the presence of a small percentage of an unfolded conformer was detected. In contrast,  $\alpha 2$  and  $\alpha 3$ , which form the  $N_{TAIL}$  binding site, are more stable helices. Slow conformational exchange was identified in the turn between these two helices as well as on the hydrophobic site of  $\alpha 3$ . The exchange observed for the residues on helix  $\alpha 3$  likely reflects a motion of helix  $\alpha 3$  with respect to helix  $\alpha 2$ . This motion can be crucial for an optimal insertion of the induced helix of  $N_{TAIL}$  to form a well-packed four-helical bundle. The binding site contains several methyl-containing

residues that, even though facing the solvent, display relatively restricted fast dynamics. This rather rigid interface forms a good mold for the  $N_{TAIL}$  to fold upon.

This study highlights the sophisticated design of proteins that optimally makes use of the wide range of dynamics that are possible in a polypeptide chain. Although in the unstructured subdomain both the flexibility and the elongated character of an unstructured chain are employed, the structured domain provides a rigid interaction surface that still allows small conformational changes for an optimal insertion.

### SUPPLEMENTARY MATERIAL

To view all of the supplemental files associated with this article, visit [www.biophysj.org](http://www.biophysj.org).

The authors thank Isabel Ayala for help with protein expression and purification, Adrien Favier for help with the setup of  $^2H$  relaxation experiments, Paul Schanda for help with SOFAST experiments, and both Carine van Heijenoort and Anthony Mittermaier for providing scripts to analyze the relaxation dispersion curves and  $^2H$  relaxation rates, respectively. Prof. R. W. Ruigrok is thanked for stimulating discussions and careful reading of the manuscript.

K.H. was funded by a fellowship of the Netherlands Organization for Scientific Research (NWO) and le Ministère Français délégué à la Recherche.

### REFERENCES

1. Uversky, V. N. 2002. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* 11:739–756.
2. Wright, P. E., and H. J. Dyson. 1999. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* 293:321–331.
3. Dyson, H. J., and P. E. Wright. 2005. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6:197–208.
4. Fink, A. L. 2005. Natively unfolded proteins. *Curr. Opin. Struct. Biol.* 15:35–41.
5. Dunker, A. K., C. J. Brown, J. D. Lawson, L. M. Iakoucheva, and Z. Obradovic. 2002. Intrinsic disorder and protein function. *Biochemistry.* 41:6573–6582.
6. Kolakofsky, D., T. Pelet, D. Garcin, S. Hausmann, J. Curran, and L. Roux. 1998. Paramyxovirus RNA synthesis and the requirement for hexamer genome length: the rule of six revisited. *J. Virol.* 72:891–899.
7. Tarbouriech, N., J. Curran, C. Ebel, R. W. Ruigrok, and W. P. Burmeister. 2000. On the domain structure and the polymerization state of the Sendai virus P protein. *Virology.* 266:99–109.
8. Tarbouriech, N., J. Curran, R. W. Ruigrok, and W. P. Burmeister. 2000. Tetrameric coiled coil domain of Sendai virus phosphoprotein. *Nat. Struct. Biol.* 7:777–781.
9. Horikami, S. M., and S. A. Moyer. 1995. Alternative amino acids at a single site in the Sendai virus L protein produce multiple defects in RNA synthesis *in vitro*. *Virology.* 211:577–582.
10. Kolakofsky, D., P. Le Mercier, F. Iseni, and D. Garcin. 2004. Viral DNA polymerase scanning and the gymnastics of Sendai virus RNA synthesis. *Virology.* 318:463–473.
11. Curran, J., T. Pelet, and D. Kolakofsky. 1994. An acidic activation-like domain of the Sendai virus P protein is required for RNA synthesis and encapsidation. *Virology.* 202:875–884.
12. Curran, J., and D. Kolakofsky. 1988. Scanning independent ribosomal initiation of the Sendai virus X protein. *EMBO J.* 7:2869–2874.

13. Bernado, P., L. Blanchard, P. Timmins, D. Marion, R. W. Ruigrok, and M. Blackledge. 2005. A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. *Proc. Natl. Acad. Sci. USA*. 102:17002–17007.
14. Blanchard, L., N. Tarbouriech, M. Blackledge, P. Timmins, W. P. Burmeister, R. W. Ruigrok, and D. Marion. 2004. Structure and dynamics of the nucleocapsid-binding domain of the Sendai virus phosphoprotein in solution. *Virology*. 319:201–211.
15. Curran, J., H. Homann, C. Buchholz, S. Rochat, W. Neubert, and D. Kolakofsky. 1993. The hypervariable C-terminal tail of the Sendai paramyxovirus nucleocapsid protein is required for template function but not for RNA encapsidation. *J. Virol.* 67:4358–4364.
16. Dyson, H. J., and P. E. Wright. 2004. Unfolded proteins and protein folding studied by NMR. *Chem. Rev.* 104:3607–3622.
17. Kay, L. E., D. A. Torchia, and A. Bax. 1989. Backbone dynamics of proteins as studied by  $^{15}\text{N}$  inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*. 28: 8972–8979.
18. Farrow, N. A., O. Zhang, J. D. Forman-Kay, and L. E. Kay. 1995. Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer. *Biochemistry*. 34:868–878.
19. Muhandiram, D. R., T. Yamazaki, B. D. Sykes, and L. E. Kay. 1995. Measurement of  $^2\text{H}$   $T_1$  and  $T_{1\rho}$  relaxation-times in uniformly  $^{13}\text{C}$ -labeled and fractionally  $^2\text{H}$ -labeled proteins in solution. *J. Am. Chem. Soc.* 117:11536–11544.
20. Choy, W. Y., D. Shortle, and L. E. Kay. 2003. Side chain dynamics in unfolded protein states: an NMR based  $^2\text{H}$  spin relaxation study of  $\Delta 131\Delta$ . *J. Am. Chem. Soc.* 125:1748–1758.
21. Tollinger, M., N. R. Skrynnikov, F. A. A. Mulder, J. D. Forman-Kay, and L. E. Kay. 2001. Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc.* 123:11341–11352.
22. Mulder, F. A. A., B. Hon, A. Mittermaier, F. W. Dahlquist, and L. E. Kay. 2002. Slow internal dynamics in proteins: application of NMR relaxation dispersion spectroscopy to methyl groups in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc.* 124:1443–1451.
23. Jha, A. K., A. Colubri, K. F. Freed, and T. R. Sosnick. 2005. Statistical coil model of the unfolded state: resolving the reconciliation problem. *Proc. Natl. Acad. Sci. USA*. 102:13099–13104.
24. Vendruscolo, M., E. Paci, M. Karplus, and C. M. Dobson. 2003. Structures and relative free energies of partially folded states of proteins. *Proc. Natl. Acad. Sci. USA*. 100:14817–14821.
25. Kristjansdottir, S., K. Lindorff-Larsen, W. Fieber, C. M. Dobson, M. Vendruscolo, and F. M. Poulsen. 2005. Formation of native and non-native interactions in ensembles of denatured ACBP molecules from paramagnetic relaxation enhancement studies. *J. Mol. Biol.* 347:1053–1062.
26. Schurr, J. M., H. P. Babcock, and B. S. Fujimoto. 1994. A test of the model-free formulas. Effects of anisotropic rotational diffusion and dimerization. *J. Magn. Reson. B*. 105:211–224.
27. Brutscher, B., R. Brüschweiler, and R. R. Ernst. 1997. Backbone dynamics and structural characterization of the partially folded A state of ubiquitin by  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  nuclear magnetic resonance spectroscopy. *Biochemistry*. 36:13043–13053.
28. Buevich, A. V., and J. Baum. 1999. Dynamics of unfolded proteins: incorporation of distributions of correlation times in the model free analysis of NMR relaxation data. *J. Am. Chem. Soc.* 121:8671–8672.
29. Prompers, J. J., and R. Brüschweiler. 2002. General framework for studying the dynamics of folded and nonfolded proteins by NMR relaxation spectroscopy and MD simulation. *J. Am. Chem. Soc.* 124: 4522–4534.
30. Marion, D., N. Tarbouriech, R. W. Ruigrok, W. P. Burmeister, and L. Blanchard. 2001. Assignment of the  $^1\text{H}$ ,  $^{15}\text{N}$  and  $^{13}\text{C}$  resonances of the nucleocapsid-binding domain of the Sendai virus phosphoprotein. *J. Biomol. NMR*. 21:75–76.
31. Neri, D., T. Szyperski, G. Otting, H. Senn, and K. Wuthrich. 1989. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional  $^{13}\text{C}$  labeling. *Biochemistry*. 28:7510–7516.
32. Delaglio, F., S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*. 6:277–293.
33. Johnson, B. A. 2004. Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.* 278:313–352.
34. Schanda, P., and B. Brutscher. 2005. Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J. Am. Chem. Soc.* 127:8014–8015.
35. Bai, Y., J. J. Englander, L. Mayne, J. S. Milne, and S. W. Englander. 1995. Thermodynamic parameters from hydrogen exchange measurements. *Methods Enzymol.* 259:344–356.
36. Bai, Y., J. S. Milne, L. Mayne, and S. W. Englander. 1993. Primary structure effects on peptide group hydrogen exchange. *Proteins*. 17: 75–86.
37. Farrow, N. A., R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, and L. E. Kay. 1994. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by  $^{15}\text{N}$  NMR relaxation. *Biochemistry*. 33:5984–6003.
38. Korzhnev, D. M., N. R. Skrynnikov, O. Millet, D. A. Torchia, and L. E. Kay. 2002. An NMR experiment for the accurate measurement of heteronuclear spin-lock relaxation rates. *J. Am. Chem. Soc.* 124:10743–10753.
39. Lipari, G., and A. Szabo. 1982. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of the experimental results. *J. Am. Chem. Soc.* 104:4559–4570.
40. Loria, J. P., M. Rance, and A. G. Palmer. 1999. A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *J. Am. Chem. Soc.* 121: 2331–2332.
41. Mulder, F. A. A., N. R. Skrynnikov, B. Hon, F. W. Dahlquist, and L. E. Kay. 2001. Measurement of slow (micro-s) time scale dynamics in protein side chains by  $^{15}\text{N}$  relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc.* 123:967–975.
42. Millet, O., D. R. Muhandiram, N. R. Skrynnikov, and L. E. Kay. 2002. Deuterium spin probes of side-chain dynamics in proteins. 1. Measurement of five relaxation rates per deuteron in  $^{13}\text{C}$ -labeled and fractionally  $^2\text{H}$ -enriched proteins in solution. *J. Am. Chem. Soc.* 124: 6439–6448.
43. Guenneugues, M., P. Berthault, and H. Desvaux. 1999. A method for determining B1 field inhomogeneity. Are the biases assumed in heteronuclear relaxation experiments usually underestimated? *J. Magn. Reson.* 136:118–126.
44. Skrynnikov, N. R., O. Millet, and L. E. Kay. 2002. Deuterium spin probes of side-chain dynamics in proteins. 2. Spectral density mapping and identification of nanosecond time-scale side-chain motions. *J. Am. Chem. Soc.* 124:6449–6460.
45. Mittermaier, A., and L. E. Kay. 1999. Measurement of methyl  $^2\text{H}$  quadrupolar couplings in oriented proteins. How uniform is the quadrupolar coupling constant? *J. Am. Chem. Soc.* 121:10608–10613.
46. Clore, G. M., A. Szabo, A. Bax, L. E. Kay, P. C. Driscoll, and A. M. Gronenborn. 1990. Deviations from the simple 2-parameter model-free approach to the interpretation of  $^{15}\text{N}$  nuclear magnetic-relaxation of proteins. *J. Am. Chem. Soc.* 112:4989–4991.
47. Prasch, S., S. Schwarz, A. Eisenmann, B. M. Wohrl, K. Schweimer, and P. Rosch. 2006. Interaction of the intrinsically unstructured phage lambda N protein with *Escherichia coli* NusA. *Biochemistry*. 45:4542–4549.
48. Mackay, J. P., L. D. Muiznieks, P. Toonkool, and A. S. Weiss. 2005. The hydrophobic domain 26 of human tropoelastin is unstructured in solution. *J. Struct. Biol.* 150:154–162.
49. Nagadoi, A., K. Nakazawa, H. Uda, K. Okuno, T. Maekawa, S. Ishii, and Y. Nishimura. 1999. Solution structure of the transactivation

- domain of ATF-2 comprising a zinc finger-like subdomain and a flexible subdomain. *J. Mol. Biol.* 287:593–607.
50. Penkett, C. J., C. Redfield, J. A. Jones, I. Dodd, J. Hubbard, R. A. Smith, L. J. Smith, and C. M. Dobson. 1998. Structural and dynamical characterization of a biologically active unfolded fibronectin-binding protein from *Staphylococcus aureus*. *Biochemistry*. 37:17054–17067.
  51. Buevich, A. V., U. P. Shinde, M. Inouye, and J. Baum. 2001. Backbone dynamics of the natively unfolded pro-peptide of subtilisin by heteronuclear NMR relaxation studies. *J. Biomol. NMR*. 20:233–249.
  52. Wirmer, J., W. Peti, and H. Schwalbe. 2006. Motional properties of unfolded ubiquitin: a model for a random coil protein. *J. Biomol. NMR*. 35:175–186.
  53. Schwalbe, H., K. M. Fiebig, M. Buck, J. A. Jones, S. B. Grimshaw, A. Spencer, S. J. Glaser, L. J. Smith, and C. M. Dobson. 1997. Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry*. 36:8977–8991.
  54. Klein-Seetharaman, J., M. Oikawa, S. B. Grimshaw, J. Wirmer, E. Duchardt, T. Ueda, T. Imoto, L. J. Smith, C. M. Dobson, and H. Schwalbe. 2002. Long-range interactions within a nonnative protein. *Science*. 295:1719–1722.
  55. Choy, W. Y., and L. E. Kay. 2003. Probing residual interactions in unfolded protein states using NMR spin relaxation techniques: an application to  $\Delta 131\Delta$ . *J. Am. Chem. Soc.* 125:11988–11992.
  56. Shojania, S., and J. D. O’Neil. 2006. HIV-1 Tat is a natively unfolded protein: the solution conformation and dynamics of reduced HIV-1 Tat (1–72) by NMR spectroscopy. *J. Biol. Chem.* 281:8347–8356.
  57. Alexandrescu, A. T., and D. Shortle. 1994. Backbone dynamics of a highly disordered 131 residue fragment of staphylococcal nuclease. *J. Mol. Biol.* 242:527–546.
  58. van Heijenoort, C., F. Penin, and E. Guittet. 1998. Dynamics of the DNA binding domain of the fructose repressor from the analysis of linear correlations between the  $^{15}\text{N}$ - $^1\text{H}$  bond spectral densities obtained by nuclear magnetic resonance spectroscopy. *Biochemistry*. 37:5060–5073.
  59. Jacobsen, J. P., H. K. Bildsøe, and K. Schaumburg. 1976. Application of density matrix formalism in NMR spectroscopy. II. The one-spin-1 case in anisotropic phase. *J. Magn. Reson.* 23:153–164.
  60. Mittermaier, A., A. R. Davidson, and L. E. Kay. 2003. Correlation between  $^2\text{H}$  NMR side-chain order parameters and sequence conservation in globular proteins. *J. Am. Chem. Soc.* 125:9004–9005.
  61. Millet, O., A. Mittermaier, D. Baker, and L. E. Kay. 2003. The effects of mutations on motions of side-chains in protein L studied by  $^2\text{H}$  NMR dynamics and scalar couplings. *J. Mol. Biol.* 329:551–563.
  62. Mittermaier, A., L. E. Kay, and J. D. Forman-Kay. 1999. Analysis of deuterium relaxation-derived methyl axis order parameters and correlation with local structure. *J. Biomol. NMR*. 13:181–185.
  63. Ishima, R., A. P. Petkova, J. M. Louis, and D. A. Torchia. 2001. Comparison of methyl rotation axis order parameters derived from model-free analyses of  $^2\text{H}$  and  $^{13}\text{C}$  longitudinal and transverse relaxation rates measured in the same protein sample. *J. Am. Chem. Soc.* 123:6164–6171.
  64. Igumenova, T. I., K. K. Frederick, and A. J. Wand. 2006. Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution. *Chem. Rev.* 106:1672–1699.
  65. Houben, K., D. Marion, N. Tarbouriech, R. W. Ruigrok, and L. Blanchard. 2007. Interaction of the C-terminal domains of Sendai virus N and P proteins: comparison of polymerase-nucleocapsid interaction within the paramyxovirus family. *J. Virol.* 81:6807–6816.
  66. Gillespie, J. R., and D. Shortle. 1997. Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J. Mol. Biol.* 268:170–184.
  67. Gillespie, J. R., and D. Shortle. 1997. Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol.* 268:158–169.
  68. Tuckis, J., S. Smallwood, J. A. Feller, and S. A. Moyer. 2002. The C-terminal 88 amino acids of the Sendai virus P protein have multiple functions separable by mutation. *J. Virol.* 76:68–77.