The PipX protein, when not bound to its targets, has its signalling C-terminal helix in a flexed conformation

Alicia Forcada-Nadal,†,1 Martina Palomino-Schätzlein,‡,1 José L. Neira,#,§,1, Antonio Pineda-Lucena, †,&,* and Vicente Rubio†,¥,*

† Instituto de Biomedicina de Valencia, CSIC, Valencia, Spain; ‡ Centro de Investigación Príncipe Felipe, Valencia, Spain; # Instituto de Biología Molecular y Celular, Universidad Miguel Hernández, Elche (Alicante), Spain; § Instituto de Biocomputación y Física de Sistemas Complejos, Zaragoza, Spain; ¥ Group 739 of the CIBER de Enfermedades Raras (CIBERER-ISCIII), Valencia, Spain; & Instituto de Investigación Sanitaria La Fe, Hospital Universitario y Politécnico La Fe, Valencia, Spain.

1These three authors contributed equally to this work.

*Corresponding authors: Vicente Rubio, Instituto de Biomedicina de Valencia (IBV- CSIC); Phone: +34 96 3391772. Email: rubio@ibv.csic.es. Antonio Pineda-Lucena, Instituto de Investigación Sanitaria La Fe, Phone: +34 96 1246600. Email: pineda_ant@gva.es

Short title: Structure and mobility of isolated PipX

Keywords: dynamics, mobility, NMR, PipX, relaxation, structure.

Funding: This work was supported by the Spanish Ministerio de Economía y Competitividad (CTQ2013-4493-R to JLN; BFU2014-58229-P to VR) and by the Generalitat Valenciana (PrometeoII/2014/029 to VR).
**Abbreviations used:** 2-OG, 2-oxoglutarate; ASA, accessible solvent area; CSA, chemical shift anisotropy; D, translational diffusion coefficient; DOSY, diffusion ordered spectroscopy; GOGAT, glutamate synthase; GS, glutamine synthetase; MF, model-free; NtcA, nitrogen-responsive regulatory protein; RCI, random-coil index; RSDA, reduced spectral density approach; TLD, Tudor-like domain; TSP, sodium trimethylsilyl [2,2,3,3-$^2$H$_4$] propionate; $\tau_c$, global correlation time.
ABSTRACT

PipX, an 89-residue protein, acts as a co-activator of the global nitrogen regulator NtcA in cyanobacteria. NtcA-PipX interactions are regulated by 2-oxoglutarate (2-OG), an inverse indicator of the ammonia abundance, and by PII, a protein that binds to PipX at low 2-OG concentrations. The structure of PipX, when bound to NtcA or PII, consists of an N-terminal, five-stranded β-sheet (conforming a Tudor-like domain), and two long α-helices. These helices adopt either a flexed conformation, where they are in close contact and in an antiparallel mutual orientation, also packing against the β-sheet; or an open conformation (observed only in the PII-PipX complex) where the last α-helix moves apart from the rest of the protein. The aim of this work was to study the structure and dynamics of isolated PipX in solution by NMR. The backbone chemical shifts, the hydrogen-exchange and the NOE patterns indicated that the isolated, monomeric PipX structure was formed by an N-terminal five-stranded β-sheet and two C-terminal α-helices. Furthermore, the observed NOEs between the two helices, and of α-helix2 with β-strand2 suggested that PipX adopted a flexed conformation. The β-strands 1 and 5 were highly flexible, as shown by the lack of inter-strand backbone-backbone NOEs; in addition, the $^{15}$N-dynamics indicated that the C terminus of β-strand4 and the following β-turn (Phe42-Thr47), and the C-cap of α-helix1 (Arg70-Asn71) were particularly mobile. These two regions could act as hinges, allowing PipX to interact with its partners, including PlmA in the newly recognized PII-PipX-PlmA ternary complex.
PipX is an 89-residue protein which acts as a global regulator and a co-activator of the nitrogen regulator NtcA in cyanobacteria \(^1,^2\). Cyanobacteria, the ancestors of chloroplasts, not only played a crucial role in the appearance of oxygen in our atmosphere, but they presently have very important roles in solar energy fixation and biomass formation, feeding the trophic chain and being essential for ecological equilibrium of the biosphere \(^3,^4\). In these organisms, PipX shuttles between the carbon/nitrogen/energy-sensor/transducer/signalling protein P\(_\|\) and the global nitrogen gene expression regulator NtcA \(^2\), modulating cellular activities for adaptation to different levels of nitrogen abundance \(^5\) (Fig. 1 A). This abundance is inversely reflected in the amount of 2-oxoglutarate (2-OG), which is used by the glutamine synthetase (GS)-glutamate synthase (GOGAT) system when ammonia is incorporated into glutamate \(^6\). While 2-OG is an NtcA activator \(^7\), influencing the transcription of the NtcA regulon \(^8\), which includes many genes related to adaptation to ammonia scarcity \(^9\), the interaction between P\(_\|\) and PipX provides an additional regulatory mechanism. In fact, under nitrogen-limitation (Fig. 1 A), 2-OG binding to P\(_\|\) prevents the interaction of P\(_\|\) with PipX, leaving the latter capable of binding to NtcA and enhancing the activity of this transcriptional regulator \(^2\). Conversely, when 2-OG concentrations are low (a reflection of high ammonia levels) a single P\(_\|\) timer binds to three PipX molecules, excluding them from interaction with NtcA \(^10\) (Fig. 1 A). It is highly important for cyanobacteria to keep the PipX function strictly modulated, since its unregulated action is toxic \(^11-13\).

The NtcA-PipX complex \(^10\) is formed by a single NtcA dimer bound to two molecules of 2-OG and two of PipX (Fig. 1 A, right). Crystal structures show that the binding stabilizes the transcriptionally-active NtcA conformation. The interaction between the two proteins occurs through the N-terminal, highly bent \(\beta\)-sheet of PipX, which has the fold of a Tudor-like domain (TLD) (Fig. 1 A, B). This part of the molecule is also involved in interactions with P\(_\|\) \(^10\) (Fig. 1 A). The structure of the complex of the P\(_\|\) trimer with three PipX molecules \(^10,^14\) (Fig. 1 A, top left) shows that the
two C-terminal helices of PipX in this complex are accessible to the environment, rather than being sequestered by the PII trimer as it occurs in TLDs. In fact, in the PII-PipX complex formed by the Synechococcus elongatus proteins, the C-terminal α-helix2 was observed in two PipX monomers in extended conformation, where the helix is projected away from the complex (the so-called open-conformation), whereas it was found in the third monomer of the complex in the same conformation as in the PipX-NtcA complex (the so-called flexed conformation). In the latter conformation, both helices are packed against each other in an antiparallel arrangement (Fig. 1 B) and α-helix2 interacts with the β-sheet of the TLD. We have proposed that the mobile α-helix2 could provide an appropriate interacting regulatory element with other molecules, as shown in yeast three-hybrid assays. These assays suggest the formation of a ternary complex between PII-PipX and the putative transcriptional regulator PlmA (Fig. 1 A, left bottom), in which the crucial involvement of the C-terminal helix was shown by deletion and site-directed mutagenesis studies.

Although the crystal structures of NtcA-PipX and PII-PipX complexes have been described, the behaviour and structure of PipX in solution are still unknown. In this work, our goal was to describe the structure and dynamics of isolated PipX in solution by using NMR. The NOE pattern, the hydrogen-exchange behaviour and the chemical shifts of the backbone nuclei indicated that isolated PipX consisted of a five-stranded β-sheet at the N terminus of the protein, with two long α-helices on the C terminus, packed on one site of the β-sheet. Contacts between the two α-helices, as well as those of the last C-terminal one with the β-sheet proved that the structure of isolated PipX in solution corresponded to the flexed-conformation. The mobility of the whole protein within the picosecond-to-nanosecond time scale showed internal movements for some amide protons within the same range of the global correlation time ($\tau_c \sim 9$ ns) of PipX. Our analyses suggest that although isolated PipX was rigid in solution, there were several mobile residues, involved in binding to the PipX molecular partners, that could act as hinges, allowing the interaction of the protein with them.
FIGURE 1. PipX regulation and models of its known complexes. (A) The complexes formed by PipX with NtcA, P\textsubscript{II} and PlmA, depending on ammonia abundance and whose formation is inversely correlated with 2-OG abundance. The structures shown are those determined for the *Synechococcus elongatus* complexes of PipX with P\textsubscript{II} or NtcA-2OG (PDB files 2XG8 and 2XKO, respectively)\textsuperscript{10}. PipX is shown in pink cartoon, P\textsubscript{II} in yellow surface and NtcA in blue surface. The schematic model proposed\textsuperscript{15} for the P\textsubscript{II}-PipX-PlmA is shown to the bottom left, being at lower scale than the P\textsubscript{II}-PipX and NtcA-PipX models. (B) Top, PipX in *open conformation* as found in the P\textsubscript{II}-PipX complex. Secondary structure elements are indicated. Side-chains of Leu65 and Leu80 are labelled and highlighted in yellow. Bottom, PipX in the *flexed conformation* found in the P\textsubscript{II}-PipX complex (pink), showing also one T-loop of PII (sand-colored and labeled) and the \(\beta_4\)-\(\beta_5\) hairpin of one adjacent monomer of PipX (light pink). In addition to Leu65 and Leu80, side-chains of three dynamic residues (Leu39, Ile52 and Arg70; see Discussion) and Phe88 from the C-terminus are also shown, as well as Ser49 and Tyr46 of P\textsubscript{II} and Phe42 of the second PipX molecule, to highlight the interactions of the three molecules. In grey, a PipX molecule from the NtcA-PipX complex is superimposed. (C) Superposition of the \(\beta_4\)-\(\beta_5\) hairpins of the PipX molecules in the crystal structures of the *Synechococcus elongatus* PipX-NtcA and P\textsubscript{II}-PipX complexes. Grey and green, NtcA complex; other colors, P\textsubscript{II} complex. Figures were produced with Pymol\textsuperscript{71}. 
EXPERIMENTAL PROCEDURES

Materials. Deuterium oxide was obtained from Apollo Scientific (Stockport, UK), and sodium trimethylsilyl [2,2,3,3-2H4] propionate (TSP), 13C6H12O6 (13C-glucose) and 15NH4Cl were from Eurisotop (Saint-Aubin Cedex, France). Standard suppliers were used for all other chemicals. Water was deionized and purified on a Milli-Q (Millipore, Madrid, Spain) system.

Protein expression and purification. Synechococcus elongatus PCC7942 PipX (UniProtKB Q7X836) was expressed in E. coli Rosetta DE3 cells (Novagen, Barcelona, Spain) and was purified essentially as described 10 from plasmid pTrc99A 2 by using LB media. The 15N/13C-labelled PipX was produced in the same way except that the cell culture was carried out in M9 minimal medium containing 15NH4Cl (1 g/L) and 13C-glucose (2 g/L). Protein was found to be > 97% pure by SDS-PAGE (15% polyacrylamide) and Coomassie staining. The protein was dialyzed (3500 MW cutoff Slide-A-Lyzer Dialysis Cassette; Thermo Scientific) against 50 mM sodium phosphate buffer, 50 mM NaCl, at pH 6. Then, later on, it was concentrated to 2 mg/mL by centrifugal ultrafiltration devices (Amicon Ultra of 3 kDa cutoff; Millipore). Protein concentrations were determined according to Bradford 16 using a commercial reagent (Bio-Rad, Madrid, Spain) and bovine serum albumin as standard.

NMR spectroscopy. All NMR experiments were acquired at 290 K in a Bruker Avance-II 600 MHz spectrometer (Bruker GmbH, Germany), equipped with a triple resonance cryo-probe and z-pulse field gradients. The temperature of the probe was calibrated with methanol 17. All experiments were carried out in 50 mM sodium phosphate buffer, with 50 mM NaCl, at pH 6.0, and with a protein concentration of 200 μM, since higher concentrations led to precipitation. Calibration in the 1H dimension was carried out by using TSP, taking into account its pH-dependence 17; the calibration of the indirect dimensions was carried out as described 17.
(a) NMR assignment. Protein backbone assignment was done by using $^{15}$N/$^{13}$C-labelled samples. For this purpose, several multidimensional heteronuclear experiments were acquired: $^1$H-$^{15}$N HSQC, $^1$H-$^{13}$C HSQC, HNCA, HNCACB, CBCA(CO)NH, HCCH-TOCSY, $^1$H-$^{15}$N HSQC-NOESY, $^1$H-$^{15}$N HSQC-TOCSY, and 2D $^1$H-$^1$H NOESY and TOCSY $^{17,18}$. The four latter were used for side-chain assignment, and for delineating the PipX fold. Typically, 2 K data points were acquired in the $^1$H dimension, and 16 to 32 scans were acquired per experiment. Spectral width in the direct dimension was 7000 Hz and in the $^{15}$N was typically 2500 Hz and 6000 Hz for $^{13}$C. Spectra were processed by using Topspin 2.1 (Bruker GmbH, Germany) with a shifted square sine window function in all dimensions; the indirect dimensions were zero-filled two-times before apodization. Assignment was carried out with the help of Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). The TALOS+ webpage was used to predict dihedral angle restraints and the secondary structure of the protein. $^{19}$

(b) Translational diffusion NMR experiments (DOSY). Measurements were carried with an unlabelled sample. They were performed with the pulsed-gradient spin-echo sequence, varying the gradient strength ($G$) in sixteen lineal steps between 2 to 95 % of the total power of the probe gradient coil. The relationship between the translational self-diffusion coefficient, $D$, and the delays employed during acquisition $^{20}$ were used to determine $D$, by calculating the exponential factor in the plot of $I/I_0$ versus $G^2$ (Fig. S1) The methyl groups between 1.0 and 0.0 ppm were used for integration. Gradient strength was calibrated by using the value of $D$ for the residual proton water line in a sample containing 100 % D$_2$O in a 5-mm tube. Curve fitting was carried out by using Kaleidagraph (Abelbeck Software, Reading, PA, USA).

(c) Hydrogen-exchange experiments. Unlabelled protein was exchanged into D$_2$O using repeated centrifugal ultrafiltration at 278 K, and then the insoluble material was centrifuged away and discarded before sample transfer to an NMR tube, that was ice-cooled, until it was introduced into the magnet. A NOESY spectrum was acquired with 7000 Hz in both dimensions, 2 K in the
acquisition dimension and 128 experiments in F1; with 96 scans in each experiment: the $^1$H carrier was set on the residual H$_2$O signal. The spectrum was processed by using Topsin 2.1, zero-filled to 2048 $\times$ 512 in the F$_2$ and F$_1$ dimensions, respectively, and processed with a sine square window function in both dimensions.

(d) $^{15}$N relaxation experiments. The $^{15}$N-$T_1$ and $^{15}$N-$T_2$ relaxation and $^1$H-$^{15}$N {NOE} experiments were acquired by using enhanced sensitivity, gradient pulse sequences developed by Kay and co-workers $^{21}$. The $T_1$ was measured in an interleaved way with 11 inversion-recovery delays, varying from 5 to 1400 ms; $T_2$ was determined by collecting 9 time points ranging from 16 to 180 ms. For the $T_1$ and $T_2$ pulse sequences, the delay between transients was 6 s. At selected times, the $T_1$ and $T_2$ experiments were repeated to ensure measurement reproducibility. The $^1$H-$^{15}$N {NOE} was measured in duplicate experiments, recording interleaved spectra in the presence and in the absence of proton saturation (achieved by a series of 120º pulses separated by 5 ms delays), and acquiring the respective spectra with a saturation time of 6 s or, alternatively, incorporating a relaxation delay of 6 s. All spectra were recorded with 2048 $\times$ 256 complex matrices in the F$_2$ and F$_1$ dimensions, with 72 scans (NOE experiment) and 16 scans ($T_1$ and $T_2$ experiments) per F$_1$ experiment. Spectral widths of 1820 and 7000 Hz were used in F$_1$ and F$_2$, respectively; the $^{15}$N carrier was set at 115 ppm and that of $^1$H was set on the water signal in all the experiments. The spectra were zero-filled in the F$_1$ dimension four times in all the experiments, and processed by using a shifted square sine window function. The same window function was used through all the $T_1$ and $T_2$ experiments. Cross-peaks were measured as volumes with Sparky. The $T_1$ and $T_2$ values were determined by fitting the measured peak-heights with Kaleidagraph to the two-parameter function:

$$I(t) = I_0 \exp \left( -\frac{t}{T_{1,2}} \right)$$

where $I(t)$ is the peak volume after a delay $t$, and $I_0$ is the volume at zero time. Errors in the relaxation rates were derived from the fitting to Eq. (1).
The steady-state $^1$H-$^{15}$N \{NOE\} values were determined from the ratio of the peak intensities with and without proton saturation (i.e., NOE = $I_{\text{sat}}/I_{\text{nonsat}}$). The error of the NOE was determined on the basis of the measured background noise levels by using repeated experiments.

The transverse, $\eta_{xy}$ ($^1$H-$^{15}$N dipolar/$^{15}$N CSA), cross-correlation rate constants were measured with spectral widths of 1820 and 7000 Hz in F$_1$ and F$_2$, respectively, and the $^{15}$N carrier was set at 120 ppm in all the experiments. One set of cross-relaxation (which leads to $I_{\text{cross}}$) and auto-relaxation (which yields $I_{\text{auto}}$) experiments were acquired for each of the relaxation delays. Both sets of experiments consisted of 256 experiments in the F$_1$ dimension and 2 K data points in the F$_2$ dimension; 16 scans were acquired per F$_1$ experiment. The relaxation delays, $T$, were 13, 18, 22, 26 and 31 ms. Control experiments (acquiring the two experiments per relaxation time) were acquired with 24 scans and relaxation delays of 18, 22, 26 and 31 ms. The spectra were zero-filled in the F$_1$ dimension two times in all the experiments, and processed by using a shifted sine square window function for each relaxation time. The peak intensities were measured by using Topspin 2.1. Errors in the intensities were determined from the pairwise root-mean-squares deviations obtained by repeating experiments at two selected times. Control experiments were also carried out with a larger number of scans in each F$_1$ increment. The $\eta_{xy}$ was derived from the one-parameter fitting:

$$I_{\text{cross}}/I_{\text{auto}} = \tanh(\eta_{xy} T).$$

The data were fitted to such equation with Kaleidagraph. The values of the $\eta_{xy}$ were similar, within the experimental error, in both the current and control experiments.

The experimentally measured $R_2$ rates can be expressed by: $R_2 = R_2^0 + R_{\text{ex}}$, where the $R_{\text{ex}}$ is the transverse relaxation rate caused by conformational processes; and $R_2^0$ is the transverse relaxation rate, which does not contain any influence from those slow movements. Then, to quantify the microsecond-to-millisecond dynamics, it is necessary to obtain the $R_2^0$ independently. The $R_2^0$ is
obtained from the experimental value of $\eta_{xy}$, since $^{22,23}$:

$$R_2^0 = 1.2589 \cdot \eta_{xy} + 1.3[\text{NOE} - 1] \cdot R_1 \frac{\gamma_N}{\gamma_H},$$

where $\gamma_N$ and $\gamma_N$ are the gyromagnetic ratios of $^{15}$N and $^1$H, respectively.

(e) Rotational diffusion tensor. An initial estimation of the rotational correlation time, $\tau_c$, and the rotational diffusion tensor were obtained with TENSOR2 $^{24,25}$, from the subset of residues which fulfilled the following two criteria of: (a) having a $^1$H-$^{15}$N {NOE} $\geq$ 0.65; and (b) satisfying the inequality 

$$\frac{R_{2,j} - \langle R_2 \rangle}{\langle R_2 \rangle} - \frac{R_{1,j} - \langle R_1 \rangle}{\langle R_1 \rangle} < 1.5\sigma,$$

where $<R_j>$ and $<R_{j,i}>$ (with $j=1,2$ and $i$ being the residue number) are, respectively, the average rates and the individual $R_1 (=1/T_1)$ and $R_2 (=1/T_2)$ rates of the subset of remaining residues satisfying (a), and $\sigma$ is the standard deviation of:

$$\left| \frac{R_{2,j} - \langle R_2 \rangle}{\langle R_2 \rangle} - \frac{R_{1,j} - \langle R_1 \rangle}{\langle R_1 \rangle} \right|.$$ The determination of the rotational diffusion tensor was carried out by using the X-ray structure $^{10,14}$ of the flexed conformation (Protein DataBank file 2XKO, chain D). The rotational diffusion in the isotropic, axially symmetric or anisotropic schemes was explored by using 1000 Monte Carlo steps. Briefly, F-test analyses were performed to choose between the three diffusion models. A probability factor of 0.2, which indicates whether the probability of improvement in different fits when complexity increases is coincidental, was calculated.

(f) Analysis of relaxation data. The model-free (MF) approach $^{26-28}$, which assumes that the global diffusion of the protein (given by the global correlation time, $\tau_c$), and the internal motion of the N-H bond (given by $\tau_e$, the effective internal correlation time) are not coupled (i.e., correlated), was used to carry out the analysis in terms of the dynamic parameters related to the N-H bond vector motions. In this formalism, the square of the order parameter, $S^2$ (which measures the degree of restriction of fast internal motions), will approach to zero when the relaxation mechanism is dominated by the local internal motion, $\tau_e$, for a particular residue; conversely, when relaxation of a residue is dominated by the $\tau_c$ of the molecule, $S^2$ will approach unity. This formalism also accounts
for the effects of slow (microsecond-to-millisecond) conformational exchange (i.e., the $R_{ex}$), and incorporates two timescales of internal motions, with the global order parameter defined as:

$$S^2 = S_f^2 S_s^2,$$

where $S_f^2$ and $S_s^2$ are the order parameters for faster and slower motions, respectively. In general, the spectral density function at a particular frequency, $J(\omega)$, can be expressed, as:

$$J(\omega) = \frac{2}{5} \left( \frac{S_f^2 \tau_e}{1 + \omega^2 \tau_e^2} + \frac{(1 - S_f^2) \tau_e^2}{1 + \omega^2 \tau_e^2} \right),$$

where $\tau_e = \frac{\tau_c \tau_e}{\tau_c + \tau_e}$. If $S^2 = 1$ at $\omega = 0$ MHz, then the $J(0) = (2/5) \tau_c$, which sets a limit for the movements when the movement of an N-H bond is dominated by $\tau_c$.

Calculations of MF were carried out by using TENSOR2, with a Monte Carlo simulation of 1000 steps for those residues for which $R_1$, $R_2$ and $^1$H-$^{15}$N {NOE} measurements were available. The modelling of the internal dynamics of each N-H bond was carried out with the following five different models 24, 25, 29: (1) the $\tau_e$ of each amide proton was very fast and it was not relaxation-active; (2) the $\tau_e$ was relaxation-active (and thus, it must be calculated); (3) this model was identical to Model 1, except that the conformational exchange on a microsecond-to-millisecond timescale was taken into account (with the $R_{ex}$ parameter); (4) identical to Model 2, except for the inclusion of the $R_{ex}$ term; and (5) a model that included two kinds of internal motions: a very fast one and a very slow one (in the first four models the slow internal motions are assumed to be negligible, that is $S_f^2 = 1$).

TENSOR2 tests the five models against the three experimentally determined $R_1$, $R_2$ and $^1$H-$^{15}$N {NOE} for each residue. If the data can be reproduced satisfactorily from a model (within a 95% confidence limit, as determined from Monte-Carlo simulations) the program chooses that model and goes to the next residue. If for a particular residue the Models 2 or 3 do not lead to an improved fit good enough to satisfy the F-test (with a confidence of 0.2), then the previously failed, but marginal Model 1 is accepted.

Given the results in the fitting of the measured $R_1$, $R_2$ and $^1$H-$^{15}$N {NOE} with the MF (see Results), the reduced spectral density approach (RSDA) 30-35 was used to approximate the measured
$R_1$ and $R_2$ rates for each NH bond.; these rates can be expressed by the values of $J(\omega)$ at $\omega = 0$, $\omega_N$ and $0.87\omega_H$. The RSDA misses the physical meaning of the mobility of the N-H bond vectors, but it allows to obtain dynamic information on mobile biomolecules where it is not physically correct to separate the overall tumbling of the molecule from the internal bond motions. The MF approach and RSDA can be connected through the use of the generalized order parameter $(S^2)^{13}$:

$$S^2 = \frac{5}{2} (J(0) - J(\omega_N)) \left( 1 + \frac{\omega_N^2 \tau_c^2}{\omega_N^2 \tau_c^2} \right)$$ (2), and $\tau_c = \frac{1}{\omega_N} \left( \frac{(J(0) - J(\omega_N))}{J(\omega_N)} \right)^{1/2}$ (3).
RESULTS

Isolated PipX in solution adopted a monomeric flexed conformation

Recombinant $^{15}$N, $^{13}$C-labelled PipX was expressed in Rosetta DE3 *E. coli* in M9 minimal medium with yields of 1-5 mg/L culture. The 2D $^1$H-$^{15}$N HSQC spectrum of the construct was well-dispersed (Fig. 2), indicating that the protein was amenable to NMR structural studies. However, the sample was not very stable and during spectra acquisition the protein tended to precipitate, which did not allow working at higher concentrations than 200 µM. There was overlap of several amide protons (Fig. 2) and the methyl region of the 1D $^1$H-NMR spectrum did not show a large dispersion. Furthermore, due to the large number of aromatic residues (7 Phe residues and 8 Tyr out of 89 amino acids), there was also severe overlap in the region between 6.60 and 7.50 ppm, which made difficult to unambiguously assign some NOEs to particular aromatic protons. Keeping this in mind, 1,005 NOEs could be unambiguously assigned, but the number of long-range contacts (between residues $i$ and $j$, with $i-j > 3$) was only 119. In addition, hydrogen-exchange experiments only allowed the identification of sixteen hydrogen-bonded, or alternatively solvent-protected residues, namely, Leu27 to Thr30, Phe37, Phe38, Glu50, Val61 to Leu68, Glu76 (or Gln67) and Gln86. These results hampered the determination of a three-dimensional protein structure with a low target-function in CYANA 2.1. Therefore, we got insight into the structure of isolated PipX in solution based on the careful inspection of the observed NOEs and on the analyses of chemical shifts.

Backbone assignment was performed based on HNCA, HNCACB and CBCA(CO)NH spectra. Continuous sequential backbone-backbone NOEs (namely, the $\alpha$N($i$, $i+1$), $\beta$N($i$, $i+1$) and NN($i$, $i+1$)) facilitated the assignment for the majority of the backbone resonances, confirming the results of the 3D spectra. The two proline residues (Pro10 and Pro51) showed intense NOEs between their Hδ protons and the Hα of the preceding residue, as expected for a trans conformation. No evidence of minor cross-peaks corresponding to the cis conformation could be identified. The assignment has
been deposited in the Biological Magnetic Resonance Data Bank (BMRB) with accession code 27041.

FIGURE 2: 2D $^1$H-$^{15}$N HSQC spectrum of PipX. Residue numbers are indicated at the top. Cross-peaks with an asterisk correspond to unassigned residues belonging to the His-tag used for protein purification. The amide resonances of Ala2 and Ser3 were not unambiguously assigned. Experiments were acquired at pH 6.0 (sodium phosphate buffer, 50 mM), 50 mM NaCl in a 600 MHz Bruker magnet at 290 K.

A TALOS+ analysis of the chemical shifts suggested that the structure of PipX comprised five $\beta$-strands (Fig. S2): Leu7-Asn8, Leu14-Gln17, G1u26-Thr30, Leu36-Ala41 and Arg48-Glu50, and two $\alpha$-helices: Asn55-Ala72 and Leu74-Phe84. Most of the solvent-protected residues (Leu27 to Thr30, Phe37, Phe38, Glu50, Val61 to Leu68, Glu76 (or Gln67)) belonged to the predicted secondary structure regions. Analysis of the long-range NOEs (Fig. 3 A) showed that there were a large number of contacts among most of the TALOS+-predicted $\beta$-sheet residues. However, there were slight variations at the predicted termini of the strands with the experimental results, because of the absence of NOEs. Furthermore, residues in this region, encompassing the N terminus until
residue 51, could be easily assigned due to the strong sequential $\alpha_N(i, i+1)$ or even $\beta_N(i, i+1)$ contacts. Moreover, except for a restricted number of polypeptide regions, there was an almost complete absence of sequential $\text{NN}(i, i+1)$ contacts at this N-terminal region (Fig. 3 B). It is interesting to note the scarcity of inter-strand backbone-backbone NOEs of the $\beta$-strands 1 and 5 with the rest of the $\beta$-sheet scaffold (Fig. 3 A). In those strands, only side-chain-side-chain contacts (blue arrows in Fig. 3 A) with residues of the opposite $\beta$-strand were observed (Fig. 3 A); probably, this lack of NOEs was due to the flexibility of the two strands encasing the $\beta$-sheet. Altogether, we can conclude that the N-terminal region of PipX comprised a five-stranded antiparallel $\beta$-sheet (Fig. 3 A, see Discussion) in a close agreement with the prediction of TALOS+. The regions connecting the different strands (with the presence of $\text{NN}(i, i+1)$ contacts) would correspond to the $\beta$-turns or loops.

At the C-terminal region of PipX, encompassing residues Ile52 to Leu89, we could observe a large number of medium-strong sequential $\text{NN}(i, i+1)$ contacts, weak sequential $\alpha_N(i, i+1)$ NOEs and the presence of medium-range contacts, namely $\alpha\beta(i, i+3)$, $\alpha_N(i, i+3)$, $\text{NN}(i, i+3)$ and $\alpha_N(i, i+4)$ (Fig. 3 B). All these NOEs are characteristic of $\alpha$-helical structures. From the NOE pattern, the two helices spanning Ala57-Asp71 and Leu74-Phe88 were observed (see Discussion). In this case, we also observed a disagreement with the TALOS+ prediction, since the last helix was slightly shorter.

Based on these results, we can conclude that the overall secondary structure of isolated PipX basically matched that observed in the X-ray structures of the complexes with NtcA and $P_{ii}$ (Fig. 1 B, see Discussion). However, we wondered whether in solution the $\alpha$-helices adopted a flexed or an open conformation. Careful inspection of NOESY spectra, especially of the methyl-methyl region, suggested that the two helices were close to each other (Fig. 4). Interestingly enough, there was evidence of contacts between the methyl groups of Leu65 and Leu80, which have been
suggested to be key in fixing the flexed scaffolding of the helices \(^{15}\) (Fig. 1 B). Furthermore, the last helix (Leu74-Phe88) had NOE contacts with \(\beta\)-strand2 (Leu15-Ile18), indicating that it was packed against one side of the \(\beta\)-sheet (Fig. 4).

(A)
FIGURE 3: **Structure of isolated PipX in solution.** (A) Alignment of the five β-strands of isolated PipX. Long-distance NOEs observed in NOESY spectra between the connected protons are shown by continuous red double arrows. The blue arrows indicate NOEs observed between the side-chains of the connected residues. The dotted arrows indicate NOEs that could not be unambiguously detected due to overlapped cross-peaks or proximity to the residual water signal. Residues are numbered at the position of the Cα. (B) Summary of NMR data. NOEs are classified into strong, medium or weak, according to the thickness of the bar underneath the sequence. The corresponding sequential Hα NOEs with the following Hδ of the proline residue are indicated by an open bar in the row corresponding to the αN(i, i+1) NOEs. Dotted lines indicate those contacts that could not be unambiguously assigned due to overlapped signals or proximity to water resonance. The green rectangles at the bottom of the figure correspond to the hydrogen-exchange, solvent-protected amide protons. The asterisk in Ala57 indicates that this residue is repeated at the beginning of the next line of the sequence.

FIGURE 4: **NOE contacts between the two α-helices at the C terminus of PipX and with the β-strand2.** Long-distance NOEs observed in NOESY spectra are shown by continuous black double arrows. The red arrows indicate the hydrogen-bond scaffold of the two helices of PipX according to the X-ray structure (PDB files 2XG8, 2XK0)\(^{10}\). The β-strand2 (Leu15-Ile18) is highlighted in blue on the right side of the figure. Numbering of residues is indicated.
Finally, measurements of the $D$ of PipX yielded a value of $(8.1 \pm 0.1) \times 10^{-7}$ cm$^2$ s$^{-1}$. Assuming that the Stokes-Einstein equation can be used, and taking as reference the $D$ of dioxane measured under the same conditions ($\langle 5.4 \pm 0.3 \rangle \times 10^{-6}$ cm$^2$ s$^{-1}$, corresponding to a hydrodynamic radius of 2.12 Å$^{40}$), the hydrodynamic radius of PipX would be 13.6 Å. This value is within the range of $17 \pm 4$ Å predicted for a folded 89-residue-long protein$^{40}$. Furthermore, the experimentally determined $D$ is similar to the value of $8.76 \times 10^{-7}$ cm$^2$ s$^{-1}$ predicted by the program HYDROPRO$^{41}$ from the X-ray structure of the flexed conformation$^{10,14}$. Thus, we can conclude that PipX was a monomer in our conditions.

To sum up, isolated PipX in solution was a monomeric protein with an N-terminal, five-stranded $\beta$-sheet spanning 51 residues, which was flanked on one side by two long $\alpha$-helices, packed one against the other in an anti-parallel arrangement, and located at C-terminal region of the protein.

**Calculation of the global correlation time ($\tau_c$) of PipX**

Once we had assigned PipX and determined its overall fold, we decided to study its dynamics in the picosecond-to-nanosecond time regime. To characterize this dynamics it is important to determine first the global correlation time, $\tau_c$. To that end, we first estimated $\tau_c$, using two different theoretical approaches, and then we verified the obtained values with several experimental approaches, to validate which models fit better to the empiric data. The first theoretical approach was based on the correlation between $\tau_c$ and the accessible surface area (ASA)$^{42}$, which was estimated to be 5,927 Å$^2$ from the crystal structure of the flexed conformation$^{10,14}$, by using the VADAR webpage$^{43}$. From the two different figures of the correlation constant$^{42}$, $\tau_c$ values of $5.8 \pm 0.2$ and $6.6 \pm 0.3$ ns were estimated. The alternative theoretical approach by using HYDROPRO$^{41}$ and the X-ray structure of the flexed conformation$^{10,14}$ led to a prediction of $\tau_c = 9.1$ ns at 290 K. These theoretical values provided a valuable threshold for comparison with the experimental ones.
We then used four experimental approaches, based on NMR data, to estimate the $\tau_c$. In a first approach, $\tau_c$ was calculated from the $D$ and the expression\(^{44}\):

$$\tau_c = \frac{(2/9) \left( R_{\text{ref}}^h D_{\text{ref}}^h \right)^2}{D^3},$$

by using as $D_{\text{ref}}$ the measured translational coefficient for dioxane, and as $R_{\text{ref}}^h$ the hydrodynamic radius of this compound\(^ {40}\). These measurements yielded a $\tau_c = 5.5 \pm 0.3$ ns for PipX, which is smaller than that determined by using HYDROPRO, but not different from that obtained from ASA-based calculations.

Another experimental approach for determining $\tau_c$ was based on the $R_1$ and $R_2$ values, used in three different alternative ways. In a first method, $\tau_c$ was determined for each residue from Eq. (3) by using the RSDA. We observed a substantial variation among the values, within the range 8.0 to 12.5 ns; the mean value was 9.5 ns (with $\sigma = 0.8$ ns), close to the theoretical HYDROPRO estimate. The variation in the local correlation time among the residues reflected fast internal motions and slow segmental arrangement in addition to the overall rotational diffusion time; in fact, it must be expected that the individual correlation times should vary according to the angle between the corresponding NH bond vector and the longest axis of the protein\(^ {45}\).

In a second method, the values of the rotational diffusion tensor were obtained from the experimental measurements of the $R_1$ and $R_2$ by using TENSOR2\(^ {24,25}\), and then $\tau_c$ was estimated from $\tau_c = \left( 2D_{xx} + 2D_{yy} + 2D_{zz} \right)^{1/3}$, where the $D_{ii}$s are the corresponding diagonal components of such tensor. The statistical tests (F-tests with 10 and 5 % confidence) showed that the isotropically tumbling approach was simpler and far better than anisotropic tumbling ($\chi^2 = 6.1$ (for isotropic) versus 61.5 (at $p = 0.10$) and 68.3 (at $p = 0.05$) (for anisotropic)). The ratio of the principal components of the rotational diffusion tensor of PipX was 1:0.82:0.74, which is within the range expected for proteins tumbling isotropically in solution\(^ {21,24,25,46}\). The value of $\tau_c$ obtained by using
this approach was 9.2 ± 0.2 ns, which agrees with that determined from the RSDA and HYDROPRO.

A final strategy was based on the assumption that \( J'(\omega) \), at selected frequencies, can be expressed as a linear combination of the spectral frequency at \( \omega = 0, J(0) \). This assumption leads to a third degree equation in \( \tau \): 
\[
2\alpha \omega_0^2 \tau^3 + 5\beta \omega_0^2 \tau^2 + 2(\alpha - 1)\tau + 5\beta = 0 \quad 29-31, 
\]
where \( \alpha \) and \( \beta \) are the slope and y-axis intercept, respectively, of the linear fitting of \( J(\omega_N) \) versus \( J(0) \). The roots of this equation are the correlation times of the motions that contribute to the low frequency part of the spectral density functions (and among them, the \( \tau_c \)). We excluded from the \( J(\omega_N) \) and \( J(0) \) plot those residues for which conformational exchange made important contributions to \( J(0) \), and thus they are separated from the \( J(\omega_N) \) versus \( J(0) \) straight line defined by the other residues. The excluded residues were: Glu4, His9, Leu15, Ile18, Leu31, Leu36, Leu39, Arg45, Phe49, Ile52, Ala57, Arg58, Asp62, Leu65, Leu68, Arg70, Gln75, Phe88 and Leu89 (see Discussion). The slope, \( \alpha \), was \((2.4 \pm 0.6) \times 10^{-2} \) ns rad\(^{-1}\) and the y-intercept, \( \beta \), was \(0.18 \pm 0.02\) ns rad\(^{-1}\). The existence of this correlation between both spectral density functions indicates that there was a collective motion of the whole molecule due to the overall rotational tumbling \(^{31}\). Among the solutions of the third degree equation only 9.40 and 0.47 ns had physical meaning. The first of these values could be assigned to the overall tumbling of PipX, and the second one to an internal motion. The value of \( \tau_c \) obtained from this approach is in very good agreement with those obtained with the other two experimental methods that rely on \( R_1 \) and \( R_2 \) values and with the one theoretically obtained from HYDROPRO. This close agreement in \( \tau_c \) calculations supports isotropic tumbling of PipX, and indicates that the local mobility had little effect on the global tumbling of the molecule. We suggest that this mobility could affect the ASA and the translational diffusion movement of the chain in solution, by solvent-exposing polypeptide patches, and accounting for the lower \( \tau_c \) obtained from the ASA and the \( D \) calculations. Nevertheless, the average value of \( \tau_c \) of 9.3 ns at 290 K obtained in the other
estimations, which is the value used in the dynamic calculations below, is larger than expected for a protein of this size (with molecular weight around 10 kDa). Even so, larger values of $\tau_c$ have been attributed in other proteins to conformational exchange in the time scale of microseconds-to-milliseconds $^{47-49}$, anisotropy tumbling of the molecule $^{48}$ or coupling between the overall tumbling and local internal movements $^{50}$. It could be also thought that the large $\tau_c$ value was due to aggregation$^{17}$ in PipX; however, we can almost certainly rule out such possibility due to the DOSY measurements and the measured $D$ (see above).

The experimental picosecond-to-nanosecond mobility of PipX

Table 1 gives the mean values of the experimental relaxation data ($R_1$, $R_2$ and $^1$H-$^{15}$N [NOE]) for the $\beta$-sheet and $\alpha$-helical regions. Due to overlapped cross-peaks unambiguous data were not obtained for the following pairs of residues: Asn5/Asn71; Phe12/Arg69; Thr47/Lys85; Glu56/Asn79; Gln67/Glu76; and Asn78/Gln81. Furthermore, as proline residues do not contain amide protons, data for Pro10 and Pro51 could not be obtained.

In general, there were no large differences among the $R_1$ values for residues involved either in the structured regions or in the loops/turns between the elements of secondary structure (Fig. 5 A), with an average value $R_1$ for all residues of $1.44 \pm 0.11$ s$^{-1}$ (error is the standard deviation). Conversely, $R_2$ values showed larger variations, with an average value for all residues of $13.89 \pm 2.11$ s$^{-1}$ (mean value ± standard deviation). The residues exhibiting the largest $R_2$ values (residue/$R_2$ ± fitting error to Eq. (1) in s$^{-1}$: Ile18/17.3 ± 4.4, Leu31/17.7 ± 2.6, Leu36/17.2 ± 4.3, Leu39/18.6 ± 4.9, Ile52/18.9 ± 4.6 and Arg70/23.4 ± 7.5) (Fig. 5 B) were among those removed in the linear fitting of $J(\omega_N)$ versus $J(0)$ (see above), and all of them, except Leu31 and Ile52, were located at the structured elements of PipX (see Discussion). Their high $R_2$ rates indicate a conformational flexibility on slower time scales than the rest of the residues of the protein.
Table 1: Average relaxation data (± standard deviation) of residues in PipX.

<table>
<thead>
<tr>
<th>Structure a</th>
<th>$R_1$ (s$^{-1}$)</th>
<th>$R_2$ (s$^{-1}$)</th>
<th>NOE</th>
</tr>
</thead>
<tbody>
<tr>
<td>β–strand1 (6-8) (3)</td>
<td>1.35 ± 0.04</td>
<td>14.01 ± 1.61</td>
<td>0.64 ± 0.08</td>
</tr>
<tr>
<td>β–strand2 (15-18) (4)</td>
<td>1.52 ± 0.14</td>
<td>15.33 ± 1.64</td>
<td>0.85 ± 0.06</td>
</tr>
<tr>
<td>β–strand3 (27-30) (4)</td>
<td>1.49 ± 0.06</td>
<td>14.03 ± 1.40</td>
<td>0.80 ± 0.05</td>
</tr>
<tr>
<td>β–strand4 (36-39) (4)</td>
<td>1.38 ± 0.03</td>
<td>15.26 ± 3.17</td>
<td>0.77 ± 0.06</td>
</tr>
<tr>
<td>β–strand5 (47-49) (3)</td>
<td>1.39 ± 0.10</td>
<td>14.75 ± 0.73</td>
<td>0.73 ± 0.14</td>
</tr>
<tr>
<td>α–helix1 (57-71) (12)</td>
<td>1.44 ± 0.13</td>
<td>14.73 ± 3.01</td>
<td>0.81 ± 0.09</td>
</tr>
<tr>
<td>α–helix2 (74-88) (12)</td>
<td>1.46 ± 0.06</td>
<td>13.96 ± 1.66</td>
<td>0.78 ± 0.07</td>
</tr>
</tbody>
</table>

The figures in parenthesis indicate: (i) residues involved in the element of secondary structure; and, (ii) the number of residues taken into account to calculate the average value.

The $^{1}$H–$^{15}$N (NOE) values were larger for the elements of a defined secondary structure (except for the C-terminal region of β-strand4 and β-strand1) than for the loops or β-turns (Fig. 5 C). For most residues, the $^{1}$H–$^{15}$N NOEs were ≥ 0.65, indicating a rigid backbone, that is, the presence of low-amplitude motions on fast time scales for the N-H bonds. However, Ala2, Ser3 and Asn5 (N-terminus); Tyr6 (β-strand1); His9, Thr11, Tyr32 and Gln34 (at β-strands junctions); Phe42 and Asp43 (at the C terminus of β-strand4); and Ala44-Thr47 (in a β-turn connecting β-strands 4 and 5) had NOE values lower than 0.65 (see Discussion). Interestingly, β-strand1 had the lowest mean NOE value among all the elements of secondary structure (Table 1) suggesting that this strand had
low rigidity, thus explaining the scarcity of inter-strand backbone-backbone NOEs observed (Fig. 3A, see above).

FIGURE 5: Experimental relaxation data for PipX at 290 K and 600 MHz. (A) $R_1$; (B) $R_2$; (C) heteronuclear NOE; (D) $R_2$-$R_1$ plot. The arrows and blue rectangles at the top of the figure indicate the $\beta$-strands and $\alpha$-helical segments, respectively. Fitting errors to Eq. (1) for $R_1$ and $R_2$ are indicated. Errors in heteronuclear NOEs were determined from the measurements of the intensity of regions of the spectrum where no signal was observed. The errors in $R_2R_1$ have been obtained from propagation, assuming that the errors in $R_2$ and $R_1$ are independent; the horizontal line in panel D is the $<R_2R_1>$ after the exclusion of residues with NOE < 0.65.

In an isotropically tumbling molecule, without internal motions and with relaxation due only to dipole-dipole and chemical shift anisotropy (CSA) mechanisms, the maximum value for the $^1$H-$^{15}$N {NOE} value is 0.82 at 60.8 MHz. In PipX, a number of residues had experimental $^1$H-$^{15}$N {NOE} values that were larger than the theoretical maximum: Leu15, Tyr16, Ile18, Cys19, Ala29, Thr30,
Leu31, Phe38, Phe49, Ile52, Leu60, Asn63, Arg70, Tyr77, Leu80, Gln82 and Val83. Those large values cannot be associated to exchange phenomena (which can affect the intensity of the reference spectrum) between the amide proton and the solvent, as most of the above residues have an ASA for the amide proton smaller than 50 % in the X-ray structure (only Leu31, Arg70 and Gln82 had an ASA > 0.50). Furthermore, they cannot be due to short relaxation delays between the experiments, as the \(^1\)H-\(^15\)N NOE spectra were acquired with 6 s delays, long enough as judged from other dynamic studies of folded proteins. The most plausible explanation for that deviation is that the large values were due to experimental error. This deviation in the \(^1\)H-\(^15\)N \{NOE\} has also been observed in several model proteins for residues either involved in secondary structure or in disordered regions, and no satisfactory explanation has been provided.

To obtain an estimate of the apparent slower microsecond-to-millisecond motional processes experienced by some of the residues with large \(R_2\) rates (Ile18, Leu31, Leu36, Leu39, Ile52 and Arg70), we calculated the \(R_2R_1\) product since high \(R_2R_1\) values reflect the presence of conformational exchange, and low \(R_2R_1\) values monitor fast motions. We determined the \(<R_2R_1>\), the trimmed mean value after exclusion of residues with NOE < 0.65, which was 20.3 s\(^{-2}\). Residues with values larger than the trimmed mean by more than 1.5 \(\sigma\) were Leu39, Ile52 and Arg70 (Fig. 5 D), whereas those with values below this mean by more than 1.5 \(\sigma\), and thus affected by fast motions, were Leu14, Glu50, Asn63 and Leu89. Finally, we used the \(R_2R_1\) to estimate the averaged generalized order parameter, given by: \(S_{av}^2 = \sqrt{<R_2R_1> R_i R_{i}^{max}}\), where \(R_i R_{i}^{max}\) is the maximum value (that of Arg70). This equation yielded a \(S_{av}^2\) of 0.774.

Trying to get insight into the mobility of those residues with large \(R_2\) values, we measured the \(\eta_{xy}\). The values ranged from 4.9 ± 0.5 s\(^{-1}\) for Asn55 to 15 ± 3 s\(^{-1}\) for Leu15. As the \(\eta_{xy}\) does not rely on the \(R_{ex}\) contribution, we determined the transversal relaxation rate without any exchange contribution, i.e., the \(R_2^0\), and from that value, the conformational exchange term, \(R_{ex}\). The \(R_{ex}\) obtained with this approach was larger (> 4 s\(^{-1}\)) for residues: Ala33, Gln34 (in the loop between \(\beta-\)
strands 3 and 4); Leu39 (in β-strand4); Ile52 and Asn55 (in the loop connecting the β-sheet with the fist α-helix); Arg58 (at the N-cap of α-helix1); Arg70 (at the C-cap of α-helix1); and Gln75 and Phe88 (at both termini, respectively, of α-helix2). Therefore, the use of the η_{xy} approach to determine the exchange term further confirmed the \( R_2 R_1 \) results, and pinpointed to the presence of residues with a large mobility.

The dynamics of PipX as described by the MF formalism

The MF formalism was used to calculate the \( S^2 \) values of PipX and the possibility to detect slow microsecond-to-millisecond conformational processes as suggested by the analysis of \( R_2 R_1 \) and η_{xy}. In PipX, the dynamics of most residues could be fitted to Model 1, where the \( \tau_e \) of each amide proton is very fast and it is not relaxation-active. However, for a number of residues (Tyr6, Leu7, His9, Thr11, Leu15, Tyr16, Asp23, Lys25, Tyr32, Gln34, Phe42 to Thr47, Arg54, Ala57 and Leu89; see Discussion) the dynamics was fit to Model 2, which includes a local time, \( \tau_e \), that is dynamically active. In fact, all the calculated internal \( \tau_e \)s were in the range of ns, being the highest those of Leu15 and Tyr16 (9 ± 3 and 7 ± 4 ns, respectively) (Fig. 6 A), similar to the global \( \tau_c \) (9.2 ns). The other \( \tau_e \)s were in the range from 0.1 to 0.8 ns, which is within the time range of 0.4 ns obtained by the third-degree equation approach (see above). Thus, in all these residues the MF no longer applied, and this implies that the \( \tau_e \) and \( \tau_c \) times could not be separated.

It could be surprising that residues as Ile39, Ile52 and Arg70, with a large \( R_2 \) value (and \( R_{ex} \) as suggested by the η_{xy} analysis), were not fitted to Model 3 (with a \( R_{ex} \) term) (see Materials and Methods section). However, TENSOR2 works as follows^{24,25}. It tests the five models against the three experimentally determined \( R_1, R_2 \) and \( ^1H-^{15}N \) {NOE} for each residue. If the data can be reproduced satisfactorily from a model (within a 95 % confidence limit, as determined from Monte-Carlo simulations) the program chooses that model and goes to the next residue. If for a particular residue, the Models 2 or 3 do not lead to an improved fit good enough to satisfy the F-test (with an
confidence of 0.2), then the previously failed, but marginal Model 1 is kept. Thus, with the large fitting errors in $R_2$ for the above mentioned residues, TENSOR2 probably selected the marginal Model 1 as the best possible option, since Model 3 did not yield a statistically significant improvement in the fitting of the experimental $R_1$, $R_2$ and ¹H-¹⁵N {NOE}.

The flexible regions of PipX, that had the lowest values of $S^2$, were the C terminus, the $\beta$-turn and the loops connecting the secondary structure elements. On the other hand, the elements of defined secondary structure had the highest $S^2$ values, with average figures of 0.91±0.03, 0.91±0.06, 0.99±0.01, 0.93±0.02 and 0.94±0.03 for $\beta$-strands 1 to 5, respectively; and average values of 0.95±0.02 and 0.99±0.01 for $\alpha$-helices 1 and 2 (Fig. 6 B, blue line). These results suggest that the mobility of PipX was mainly governed by the $\tau_c$. It is interesting to note that the $S^2$ values of $\beta$-strand1 were similar to those of the other elements of secondary structure, even though the ¹H-¹⁵N {NOE} values were lower than those of the other strands (Table 1). However, Tyr6 and Leu7 (both in $\beta$-strand1) were fitted to Model 2, where their order parameters can be overestimated due to the additional fast movement⁴⁹.

We compared the values of $S^2$ from the MF formalism, with those predicted from the X-ray structure¹⁰,¹⁴ of the flexed conformation ⁵⁸ (Fig. 6 B, red line) and those from the RCI (random-coil index) method ⁵⁹ (which only uses the chemical shift values) (Fig. 6 C, red line); the theoretical values from the X-ray structure qualitatively mirrored those obtained by TENSOR2 (Fig. 6 B, black line), but they were consistently smaller. The same was observed when comparison was carried out with the order parameters only predicted from RCI method. Both predictions (X-ray and chemical shifts) and the results from TENSOR2 monitored a highly mobile region at the 40s residues (end of $\beta$-strand4 and the following $\beta$-turn); further, the theoretical prediction indicated that the region Arg70-Leu74 also showed a large mobility (this region comprises the C-cap of $\alpha$-helix1, see Discussion). The small $S^2$ values predicted from the crystal around Ser24-Lys25 could be explained because that region in the X-ray structure of PipX is poorly defined ¹⁰.
FIGURE 6: The MF and RSDA results for PipX. (A) $\tau_e$. (B) $S^2$, from the theoretical calculations (red) $^{58}$ and the one calculated by TENSOR2 (black line) $^{24,25}$; (C) $S^2$ calculated from $J(0)$ and $J(\omega_n)$ (black line) according to Eq. (2) and from the RCI method (red line), which only takes into account the values of the chemical shifts $^{59}$. Errors in $S^2$ determined from the spectral density functions are propagation errors; the others are from calculations in TENSOR2. The arrows and blue rectangles at the top of the figure indicate the $\beta$-strands and $\alpha$-helical segments, respectively.

The average mean value of $S^2$ derived from the MF was $0.941 \pm 0.008$, larger than that determined from the $R_2R_1$ approach $^{58}$ (see above). The difference could be explained by the
simplifications in the derivation of the $R_2R_1$ product, where it is assumed that: (i) no contribution from microsecond-to-millisecond conformational exchange intervenes; and, (ii) the variations in $J(\omega_N)$, due to the anisotropy, are small. Alternatively, for those residues fitting to Model 2, with a dynamic active $\tau_e$ (see above), the $S^2$ could be overestimated, since the order parameter is only sensitive to nanoseconds or faster motions. Furthermore, it is noteworthy that the MF did not predict residues having movements in the microsecond-to-millisecond timescale, as it would be expected for residues Leu39, Ile52, Arg70, due to their large $R_2R_1$ and the results of the $\eta_{xy}$ analysis; in fact, those three residues had $S^2$ values close to 1. As indicated above, this is probably due to the large errors associated with the $R_2$ measurements of those residues and the way how TENSOR2 chooses the Model.

Thus, taken together, the MF formalism suggests that PipX was a quite rigid molecule with a few localized residues experiencing an internal, dynamically active tumbling time.

The dynamics of PipX as described by the RSDA

As the MF results suggested that for some residues the $\tau_e$ and $\tau_c$ times could not be separated, we decided to use the reduced spectral density approach (RSDA) to obtain, if possible, more hints on the PipX protein dynamics. This approach provides insight into the motion of the NH bond vector by calculating the $J(\omega)$ values at three frequencies: 0, $\omega_N$ and 0.87 $\omega_H$, and not relying on any particular dynamic model to fit the experimental data. The $J(0.87 \omega_H)$ and $J(\omega_N)$ values are not affected by slow microseconds-to-millisecond motions, since they are calculated independently of the $R_2$ values. Thus, the high frequency $J(0.87 \omega_H)$ reports fast picosecond motions in flexible regions of proteins, with the largest and the smallest values being found, respectively, in highly mobile and rigid protein regions. In PipX, the high $J(0.87 \omega_H)$ values were found for the N and C termini, the loops connecting the strands, and the ends of $\beta$-strands 4 and 5 (Fig. 7 A), whereas the rest of the polypeptide did not exhibit a fast-motion behaviour, with similar $J(0.87 \omega_H)$ values. In the case of
$J(\omega N)$, its value usually decreases and increases with a raise of the internal flexibility, depending on whether the protein is small or large, respectively $^{31}$. This parameter was little informative in PipX, since its value was essentially constant for most of the residues (Fig. 7B). Finally, the value of $J(0)$ generally reports on motions in the nanosecond-to-millisecond range. For a protein moving isotropically with no internal motions, it can be shown that the $J(0)$ value should have a limit value of $2/5 \tau_c$ (see Materials and Methods)$^{31}$, with lower values indicating internal flexibility of the N-H bond, while larger $J(0)$ values pinpointing residues with slow conformational fluctuations. For a $\tau_c$ of 9.3 ns for PipX, the limit value for $J(0)$ would be 3.7 ns rad$^{-1}$ (Fig. 7 C). The residues with the smallest $J(0)$ were mainly at the C terminus and at regions between the elements of secondary structure, further supporting their high mobility (as suggested by the large values of $J(0.87\omega H)$).

Residues exhibiting large $J(0)$ values (> 1\sigma above the mean) were Leu31 and Leu36, in the $\beta$-strand3-$\beta$-strand4 connector; Leu39, in $\beta$-strand4; Ile52, at the beginning of the C-terminal region; and Arg70 at the C terminus of $\alpha$-helix1. None of these residues, except Ile52, exhibited $J(0.87\omega H)$ or $J(\omega N)$ spectral densities deviating from average values for the whole protein, suggesting that these large $J(0)$ values reflected conformational exchange motions on a microsecond-to-millisecond time scale, in agreement with $R_2R_1$ and the $\eta_{xy}$ analyses. These slow movements possibly reflected large segmental rearrangements rather than the mere breaking of single hydrogen-bonds. In the case of Ile52, the combination of large $J(0)$ and $J(\omega N)$ values showed the existence of complicated motional times at the high frequencies $^{50}$. 
FIGURE 7: The RSDA results for PipX. (A) $J(0.87 \omega_n)$; (B) $J(\omega_n)$; (C) $J(0)$, the line in the latter panel indicates the limit of $2/5 \tau_c$. The arrows and blue rectangles at the top of the figure indicate the $\beta$-strands and $\alpha$-helical segments, respectively. Errors in each spectral density function were obtained from error propagation.

Finally, the values of $S^2$ from RSDA (Eq. (2)) qualitatively resembled those obtained by using the MF (Fig. 6 C), especially in the region around 40s, where there was a decrease in the $S^2$ values, as well as at the C terminus. Furthermore, there were some residues with a nonphysical value of $S^2$ larger than 1, which was probably due to their large $J(0)$ values; similar findings have been observed in RSDA studies of highly mobile proteins $^{60,61}$. 
DISCUSSION

Our present results showed that when PipX was isolated in solution, that is, it is not bound to its target proteins, its two α-helices were in close proximity. This conclusion is based on the inter-helical NOEs found between the methyl groups of Leu65 and Leu80, located at α-helices 1 and 2, respectively (Fig. 1 B). These contacts were observed in the two PipX chains found in the crystal structure of the NtcA-PipX complex. They were also present in the PipX chain observed in the *flexed conformation* in the crystal structure of the PII-PipX complex of *Synechococcus elongatus*, whereas they were absent in the PipX molecules found in *open conformation* in this last complex (Fig. 1 B)\(^{10,14}\). We also obtained evidence that the C-terminal helix (residues Leu74-Phe88) was near β-strand2, closing on the β-sheet at that site, again showing structural features of the *flexed conformation* observed in the PipX complexes with NtcA and with P\(_{II}\). Therefore, all the NMR-derived evidence indicates that the structure in isolated PipX, in the absence of bound protein targets, is characterized by two C-terminal helices adopting a *flexed conformation* (Fig. 1 B, bottom).

Thus, the structure described in this work appears identical to the *flexed conformation* epitomized by PipX in the NtcA-PipX complex \(^{10}\). This finding could explain why isolated PipX does not interact with PlmA \(^{15}\), since such interaction appears to occur only *via* the *open conformation* of PipX, apparently directly involving the extended α-helix2 \(^{15}\) (Fig. 1 A, bottom left).

As the C-terminal helix is involved in a conformational change, it was reasonable to study the dynamics of PipX in a sufficiently wide time range. First, we shall discuss the possible relation between the fast (picosecond-to-millisecond) and slow (millisecond-to-second) dynamics. The comparison of the dynamics in the time scale provided by solvent-exchange (millisecond-to-second) with that provided by the picosecond-to-nanosecond showed that all residues with slow solvent-exchange (Fig. 3 B) had a high \(S^2\) parameter (Fig. 6 B, C); furthermore, all of them were fit to Model 1 in the Lipari-Szabo approach \(^{26,27}\). Exchange at those sites should take place during global
unfolding events, where residues seemed to have a restricted mobility in the picosecond-to-nanosecond time regime. Moreover, large $S^2$ values suggest increased enthalpic contribution to the thermodynamic stability of the folded state; and enthalpic contributions quench the movements in the picosecond-to-nanosecond time regime\textsuperscript{62}.

FIGURE 8. *Synechococcus elongatus* PipX sequence and secondary structure based on the crystallographic structures of the complexes with NtcA and P\textsubscript{II} and on the present data for the isolated protein. The sequence is in single-letter amino acid code, highlighting residues with special mobility. Residues in blue are those with their $J(0)$ and $J(\omega_N)$ functions were removed from calculation of $\tau_c$, according to the approach of Wagner and co-workers\textsuperscript{31,32}. Residues in red are those with low $^1$H-$^{15}$N NOE values. Residues in brown are those with both low $^1$H-$^{15}$N NOE values and whose $J(0)$ and $J(\omega_N)$ functions were removed from calculation of $\tau_c$. Residues in green are those whose $J(0)$ and $J(\omega_N)$ functions were removed from calculation of $\tau_c$ and they had high $R_2 R_1$ values\textsuperscript{57}. Underlined residues were those fitted to Model 2 in the MF formalism. Residue numbers are given under the sequence. The $\alpha$-helices are shown as cylinders and $\beta$-strands as arrows at the bottom, according to the crystals structures\textsuperscript{10}. PipX residues interacting with P\textsubscript{II} or with NtcA are labeled at the top with closed circles or inverted triangles, respectively. Upright triangles indicate residues believed, on the basis of site-directed mutagenesis, to be involved in PipX interactions with PlmA\textsuperscript{15}. The grey bar at the top right marks the deletion of $\alpha$-helix2 known to abolish interaction with PlmA\textsuperscript{15}.

Although the general mobility in the picosecond-to-nanosecond time scale of PipX was dominated by its overall tumbling, $\tau_{cs}$, there were flexible regions in the protein, with movements in the same range of $\tau_c$. In addition, residues Leu39, Ile52 and Arg70 appeared to have movements in the microsecond-to-millisecond time range, as suggested by the large values of $R_2$ (Fig. 5 B), $J(0)$ (Fig. 7 C), and the $\eta_{xy}$ approach to determine the $R_{ex}$ term. Flexibility in proteins is needed for proper
function, being associated with hinges, enzyme active-sites, recognition “hot-spots” 63-65, and allostERIC behaviour, often encompassing huge networks of evolutionarily conserved residue clusters that can span entire protein domains 66,67. In PipX, the mobile Leu39, at the end of \(\beta\)-strand4 (Fig. 8), is close to an edge of an extremely flexible region that comprises Phe42 to Thr47 (Fig. 1 C). In turn, Ile52, also a mobile residue, is at the junction of the \(\beta\)-sheet of the TLD with the two \(\alpha\)-helices (Fig. 8), being the sole residue in such polypeptide patch to show a large \(J(0)\). However, the residue with the apparent largest mobility \((J(0))\) was Arg70, which is located at the C-cap of \(\alpha\)-helix1 (where a small number of helical contacts were observed, Fig. 3 B), and showed the largest \(R_2R_1\) value (Fig. 5 D). Therefore, we suggest that the region around Arg70 could act as a picosecond-to-nanosecond “hinge” for the opening of the second helix, which thus can move apart from the rest of the protein (Fig. 1 B).

As the open conformation was only observed in the complex of PipX with \(P_{II}\) 10, it appears that \(P_{II}\) “opens” PipX. The other two mobile regions close to Leu39 and Ile52 do not seem to be involved in any “hinge” movement during PipX function, but these residues are close to amino acids whose dynamics was fit to Model 2 (Fig. 8). Furthermore, these two regions, together with the C-terminus of \(\alpha\)-helix2, cluster together in the crystal structures of the flexible conformation of PipX 10,14 (Fig. 1 B, bottom panel). Moreover, Pro51 and the two residues following Ile52 interact with the flexible T-loop of \(P_{II}\) in the PipX- \(P_{II}\) complex 10,14 (Fig. 1 B bottom), possibly providing a way by which the interaction with \(P_{II}\) could favour the release of the C-terminus of PipX, allowing the opening of PipX. In addition, the \(^{15}N\) dynamics indicated that the C terminus of \(\beta\)-strand4 and the following \(\beta\)-turn (Phe42-Thr47) were particularly flexible (Fig. 8). Interestingly, this region has different conformations in the complexes with NtcA and with \(P_{II}\) (Fig. 1 C). In the NtcA-PipX complex it is exposed and does not interact with NtcA 10. On the other hand, in the complex with \(P_{II}\), it interacts with the same \(\beta\)-turn of the adjacent molecule of the PipX trimer formed within this complex. This region also interacts with the flexible T-loop that modulates most of the \(P_{II}\) functions 5,10,11,14,68-70.
(Fig. 1 B, bottom). In fact, in the P_{II}-PipX complex the T-loop is sandwiched between this turn and the connection between the TLD and the C-terminal $\alpha$-helices that includes Ile52 $^{10,14}$ (Fig. 1 B, bottom). This endows this $\beta$-turn with regulatory potential, given the importance of the T-loop in the signalling of P_{II} functions, perhaps implicating it in changes that could favour the release of $\alpha$-helix2 into the open conformation. This release would result in extra regulatory features, as reflected in the formation of the novel ternary complex with both P_{II} and the putative transcriptional regulator PlmA $^{15}$.

The $\beta$-strand1 had very low $^1$H-$^{15}$N {NOE} values (Table 1) and exhibited few inter-strand backbone-backbone contacts (Fig. 3 A); in addition, two residues of this strand had an additional dynamic time (i.e., their dynamics were fitted to Model 2). This strand is involved in the interactions of PipX with P_{II} and NtcA in its corresponding complexes $^{10,14}$ (Fig. 8). In addition, in the PipX-P_{II} complex $^{10,14}$, this region is also involved in the connections between adjacent monomers in the PipX trimer formed when this protein is P_{II}-bound. Other PipX residues participating in binding to the target protein are Leu14 and Tyr16, both of which are close to amino acids with a large $R_2$ (Fig. 5 B) or where conformational exchange makes important contributions to $J(0)$; Leu31, which had a large $R_2$ rate and a large $J(0)$ value; and Tyr32, which is close to Leu31. The movements in these residues do not have to involve large amplitude collective motions (as those hypothesized to occur around Arg70), but they probably implicate more frequent small amplitude fast fluctuations which help to overcome the important energy barrier of a large conformational change. To sum up, it seems that all residues of PipX involved in binding to the biomolecular partners of PipX had a larger flexibility in different time scales (Fig. 8, the different coloured residues). In the end, it is important to pinpoint that this increase in flexibility exhibited at different time scales would result in an increase in the protein stability, since it favours the folded state by raising the entropy of the protein.

As already mentioned, the present findings provide the first structural insight of PipX in solution, free from other proteins like P_{II} and NtcA. Such insight has implications for understanding
the formation of the P_II-PipX-PlmA complex. Thus, the dynamic information provided by NMR, concerning the conformational arrangement of different parts of the protein chain, as well as mobility of individual residues, helps to understand the "opening" effect of P_II on PipX. This effect might be further enhanced by the concomitant presence of PlmA, which, by interacting with the open conformation of PipX in the P_II-PipX complex, would trap it. In any case, the present structural information might also be relevant to explain the formation of other biologically important complexes of PipX that conceivably could be formed in addition to those with PII, NtcA and PlmA, given the large PipX regulon revealed by high throughput transcriptomic studies.1

CONFLICT OF INTEREST

The authors declare they do not have any competing interest.

ACKNOWLEDGEMENTS

We thank the two reviewers for helpful suggestions and discussions. We thank Drs. Rodrigo Carbajo (Astra-Zeneca, Cambridge, UK), Sergio Gil-Caballero (CSIC, Seville), and Jorge Santoro and David Pantoja-Uceda (CSIC, Madrid) for helpful suggestions and discussions.

SUPPORTING INFORMATION AVAILABLE

A figure showing the results of the DOSY experiments is provided. The results of TALOS+ are shown in a diagram. This information is available free of charge at http://www.acs.pub.org.
REFERENCES


