



Structural and functional characterization of the PDZ domain of the human phosphatase PTPN3 and its interaction with the human papillomavirus E6 oncoprotein

Mariano Genera, Damien Samson, Bertrand Raynal, Ahmed Haouz, Bruno Baron, Catherine Simenel, Raphaël Guérois, Nicolas Wolff, Célia Caillet-Saguy

► To cite this version:

Mariano Genera, Damien Samson, Bertrand Raynal, Ahmed Haouz, Bruno Baron, et al.. Structural and functional characterization of the PDZ domain of the human phosphatase PTPN3 and its interaction with the human papillomavirus E6 oncoprotein. Scientific Reports, 2019, 9, pp.7438. 10.1038/s41598-019-43932-x . cea-02146733

HAL Id: cea-02146733

<https://cea.hal.science/cea-02146733>

Submitted on 4 Jun 2019

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

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



Distributed under a Creative Commons Attribution 4.0 International License

SCIENTIFIC REPORTS

OPEN

Structural and functional characterization of the PDZ domain of the human phosphatase PTPN3 and its interaction with the human papillomavirus E6 oncoprotein

Mariano Genera^{1,2}, Damien Samson³, Bertrand Raynal⁴, Ahmed Haouz⁵, Bruno Baron⁴, Catherine Simenel³, Raphael Guerois⁶, Nicolas Wolff¹ & Célia Caillet-Saguy¹

The human protein tyrosine phosphatase non-receptor type 3 (PTPN3) is a PDZ (PSD-95/Dlg/ZO-1) domain-containing phosphatase with a tumor-suppressive or a tumor-promoting role in many cancers. Interestingly, the high-risk genital human papillomavirus (HPV) types 16 and 18 target the PDZ domain of PTPN3. The presence of a PDZ binding motif (PBM) on E6 confers interaction with a number of different cellular PDZ domain-containing proteins and is a marker of high oncogenic potential. Here, we report the molecular basis of interaction between the PDZ domain of PTPN3 and the PBM of the HPV E6 protein. We combined biophysical, NMR and X-ray experiments to investigate the structural and functional properties of the PDZ domain of PTPN3. We showed that the C-terminal sequences from viral proteins encompassing a PBM interact with PTPN3-PDZ with similar affinities to the endogenous PTPN3 ligand MAP kinase p38 γ . PBM binding stabilizes the PDZ domain of PTPN3. We solved the X-ray structure of the PDZ domain of PTPN3 in complex with the PBM of the HPV E6 protein. The crystal structure and the NMR chemical shift mapping of the PTPN3-PDZ/peptide complex allowed us to pinpoint the main structural determinants of recognition of the C-terminal sequence of the E6 protein and the long-range perturbations induced upon PBM binding.

Protein tyrosine phosphatases (PTPs) play critical roles in cell signaling pathways. They are known to regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation. Together with kinases, they control the balance of phosphorylated species, enabling specific and varied signaling responses.

The human protein tyrosine phosphatase non-receptor type 3 (PTPN3) and the protein tyrosine phosphatase non-receptor type 4 (PTPN4) compose the NT5 type of non-transmembrane PTPs. These large modular proteins feature a N-terminal FERM (Band 4.1, Ezrin, Radixin, and Moesin) domain that is responsible for the localization of these enzymes to submembranous sites in the cell, a PDZ (PSD-95/Dlg/ZO-1) domain that allows specific binding to C-terminal recognition sequences in target proteins, and a C-terminal catalytic tyrosine phosphatase domain (Fig. 1). Three different isoforms of PTPN3 are known to be produced by alternative splicing, one being the full-length protein (UniProt P26045-1), whereas the FERM domain is partially or totally missing in the other two (UniProt P26045-2 and P26045-3). PTPN3 is localized in the cytoplasm and enriched at the plasma membrane, although in absence of the FERM domain it is exclusively cytoplasmic¹.

PTPN3 has been linked to many cancers either with a tumor-suppressive or a tumor-promoting role². Interestingly, PTPN3 is targeted by oncoviruses such as the high-risk genital human papillomaviruses (HPV)

¹Récepteurs-Canaux, Institut Pasteur, UMR 3571, CNRS, F-75724, Paris, France. ²Sorbonne Université, Complexité du Vivant, F-75005, Paris, France. ³RMN des biomolécules, Institut Pasteur, UMR 3528, CNRS, F-75724, Paris, France.

⁴Plate-forme de Biophysique Moléculaire, Institut Pasteur, UMR 3528, CNRS, F-75724, Paris, France. ⁵Plate-forme de Cristallographie, Institut Pasteur UMR 3528, CNRS, F-75724, Paris, France. ⁶Institut de Biologie Intégrative de la Cellule (I2BC), CEA, CNRS, Université Paris-Sud, Université Paris-Saclay, 91190, Gif-sur-Yvette, Cedex, France. Correspondence and requests for materials should be addressed to C.C.-S. (email: celia.caillet-saguy@pasteur.fr)

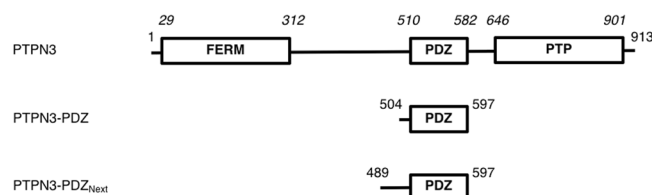


Figure 1. Schematic representation of the PTPN3 constructs. Numbers on both extremities indicate the boundary residues for each construct; Numbers in italic above the schematic construct of full-length PTPN3 correspond to the boundary residues of each protein domain from Uniprot P26045.

Peptides	Sequences	K _D (μM)	T _m (°C)
p38γ PBM	SWARVSKETPL	26 ± 25	45
HBVc PBM	RRRRSQSRESQC	29 ± 24	47
HPV16E6 PBM	RSSRTTRRETQL	53 ± 31	52
HPV18E6 PBM	RQERLQRRETQV	37 ± 20	49

Table 1. Peptide sequences, affinities for PTPN3-PDZ and melting temperature of complexes. The errors are the standard deviations of all the K_D values derived from 8 to 14 curves fitted for each titration.

types 16 and 18³ and the hepatitis B virus (HBV)⁴. HPV types 16 (HPV16) and 18 (HPV18) are responsible of 70% of cervical cancers and precancerous cervical lesions⁵.

The E6 proteins of HPV16 and HPV18 possess a conserved C-terminal PDZ binding motif (PBM), which mediates interaction with some cellular PDZ domain-containing proteins. The ability of E6 to bind to PDZ domain-containing proteins correlates with its oncogenic potential⁶. The PDZ domain of PTPN3 (PTPN3-PDZ) has been identified as a target of the E6 PBM of HPV16 and 18. This PDZ-PBM interaction results in the ubiquitin-mediated proteasomal degradation of PTPN3. Therefore, the endogenous levels of PTPN3 are particularly low in HPV-positive cervical carcinoma cell lines³.

To date, the structure of PTPN3-PDZ is not known and the PDZ-mediated interactions between PTPN3 and viral partners have not been described. Indeed, the few biophysical and structural studies on PTPN3 have focused on the PTP domain in complex with phospho-peptide substrates derived either from the mitogen-activated protein kinase (MAPK) p38γ (also known as MAPK12)⁷ or the Epidermal growth factor receptor substrate 15 (Eps15)⁸. We have previously shown that the PDZ domain of the closely related PTPN4 inhibits its phosphatase activity^{9,10}. Similarly, the inhibitory role of the PDZ domain of PTPN3 on the catalytic activity has been reported using a phospho-p38γ peptide as a substrate for kinetics experiments⁷.

In this study, we report the structural and functional analyses of the PDZ domain of PTPN3. We investigated the molecular mechanism of interaction between the PDZ domain of PTPN3 and the PBM of cellular and viral partners (Table 1), the MAP kinase p38γ, the HBV core protein and the HPV E6 protein. We employed an integrative approach based on biophysical, nuclear magnetic resonance (NMR) and X-ray crystallography experiments. We showed that the C-terminal PBM sequences of the HPV E6 protein interact with PTPN3-PDZ with similar affinities to those of the endogenous PTPN3 ligand p38γ and the HBV Core protein⁴. By solving the X-ray structure and mapping the NMR chemical shift changes of PTPN3-PDZ in complex with the PBM of HPV16 E6, we identified the molecular basis of recognition of the C-terminal sequence of E6 protein and the induced perturbations that spread from the binding groove.

Results

Quality control of the recombinant PTPN3-PDZ domain. Presently, there is no structural information on PTPN3-PDZ other than that provided by homology modelling⁷. To bridge this gap, we performed different structural and biophysical studies on the recombinant PTPN3-PDZ construct (Fig. 1) expressed in *Escherichia coli*.

To estimate the secondary structure content of the PDZ construct, we performed circular dichroism (CD) measurements in the far-UV (195–240 nm). The PTPN3-PDZ spectrum (Fig. 2A) appeared to be of mixed α-helix and β-sheet content, which upon deconvolution indicated 16% α-helix and 32% β-sheet. The secondary structure content extracted from the X-ray structure of the PDZ domain of its close homologue PTPN4 (PTPN4-PDZ) (PDB id 3NFK)¹¹ is 17% α-helix and 31% β-sheet. Thus, the PTPN3-PDZ construct presents a secondary structure content similar to that of PTPN4-PDZ.

We assessed the tertiary folding of PTPN3-PDZ by NMR by performing ¹⁵N heteronuclear single quantum correlation (HSQC) experiment. The high dispersion of resonances in the HSQC spectrum of PTPN3-PDZ (Fig. 2B) is indicative of a mainly well-folded construct in solution.

Stability and oligomeric state of PTPN3-PDZ. To evaluate the stability of PTPN3-PDZ, we performed differential scanning calorimetry (DSC) (Fig. 2C). A single endothermic peak was observed in the DSC thermogram, with a thermal transition midpoint (T_m) of 41 °C. As a comparison, we performed DSC on PTPN4-PDZ. We obtained an endothermic peak with a T_m of 47 °C (Fig. 2C). The data show that PTPN3-PDZ is less stable in

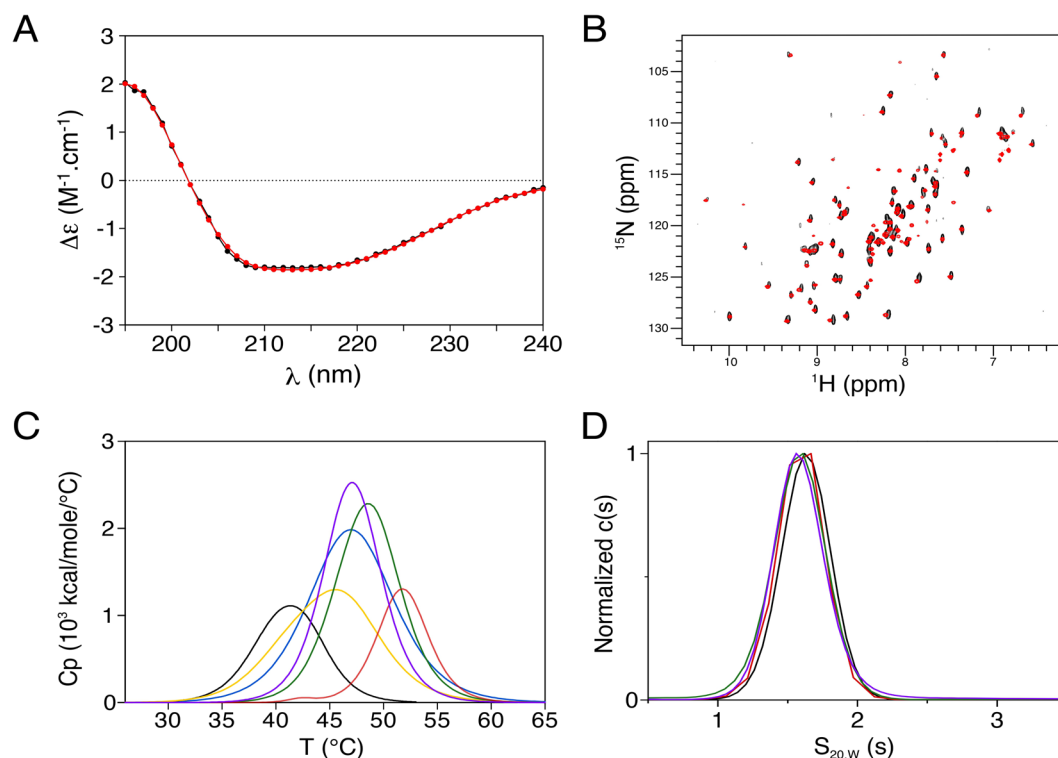


Figure 2. Stability, oligomerization and structural characterization of PTPN3-PDZ. **(A)** Far-UV CD spectrum (195–240 nm) of PTPN3-PDZ. The experimental spectrum of PTPN3-PDZ and the best fit obtained with the CONTINLL algorithm are represented as black and red lines, respectively. **(B)** Superimposed ^1H , ^{15}N HSQC spectra of PTPN3-PDZ (black) and PTPN3-PDZ_{Next} (red). **(C)** DSC thermograms of PTPN3-PDZ unbound and complexed to PBM peptides. Free PTPN3-PDZ curve is shown as a black line, and PTPN3-PDZ complexed to p38 γ PBM, HBVc PBM and HPV16E6 PBM and HPV18E6 PBM curves are shown as yellow, purple, green and red lines, respectively. The DSC thermogram of free PTPN4-PDZ is colored in blue. **(D)** Sedimentation coefficient distributions of PTPN3-PDZ_{Next} unbound and complexed to viral PBM peptides. The black, red, green and purple lines correspond to PTPN3-PDZ_{Next} unbound and complexed to HPV18E6PBM, HPV16E6 PBM and HBVc PBM, respectively.

PTPN3-PDZ						
	10 μM	25 μM	70 μM	70 μM + HBVc PBM		
MW theoretical (kDa)	10.5	10.5	10.5	12.0		
MW estimated (kDa)	12.4	11.3–25.9 ^a	nc	13.5		
Sed coef ($S_{0,w,20}$) (S)	1.5	1.5–2.6 ^a	nc	1.6		
PTPN3-PDZ _{Next}						
	17.5 μM	35 μM	70 μM	70 μM + HPV16E6 PBM	70 μM + HPV18E6 PBM	70 μM + HBVc PBM
MW theoretical (kDa)	12.7	12.7	12.7	14.1	14.1	14.2
MW estimated (kDa)	12.7	12.5	15	13.7	14.7	15.6
Sed coef ($S_{0,w,20}$) (S)	1.6	1.7	1.6	1.6	1.6	1.6

Table 2. Hydrodynamic parameters of PTPN3-PDZ derived from the analysis of analytical ultracentrifugation.

^aTwo peaks are present; nc: not calculated because of a continuum of species.

solution than PTPN4-PDZ. This thermal instability of the PTPN3-PDZ construct could be responsible for its aggregation previously detected by NMR. Indeed, we observed a significant loss of peak intensities in the HSQC spectrum of PTPN3-PDZ related to aggregation after 10 h at 20 °C (data not shown).

To estimate the oligomeric state of the PTPN3-PDZ construct, analytical ultracentrifugation (AUC) experiments were performed (Table 2). Three different concentrations of the protein were tested, 10, 25 and 70 μM . Only one single peak was detected at 10 μM with a sedimentation coefficient of 1.5S in agreement with a monomeric form of 10.5 kDa. At 25 μM , an additional peak was detected with a sedimentation coefficient of 2.6S corresponding to a dimeric form. At a concentration of 70 μM , higher oligomeric species could be detected. Thus, the AUC data of the PTPN3-PDZ construct reveal a concentration-dependent equilibrium between monomer, dimer and higher oligomeric states with only monomers detected at a concentration of 10 μM .

PBM-mediated stabilization of the PTPN3-PDZ. To check whether the PBM binding to the PDZ domain could affect the stability of PTPN3-PDZ, we performed DSC and AUC on PTPN3-PDZ in complex with the p38 γ PBM and the viral PBMs of HBV Core and HPV 16 and 18 E6 proteins, namely HBVc PBM, HPV16E6 PBM and HPV18E6 PBM respectively (Table 1). These PBMs have been shown to interact with the PTPN3-PDZ^{3,4,7}. A single endothermic peak was observed in the DSC thermograms of all complexes (Fig. 2C) with a T_m of 45 °C, 47 °C, 52 °C and 49 °C for PTPN3-PDZ complexed to p38 γ PBM, HBVc PBM, HPV16E6 PBM and HPV18E6 PBM, respectively (Table 1). For all the PBMs tested, an increase of 4 to 11 °C in the T_m was observed, showing that the PBM binding onto PTPN3-PDZ stabilizes the domain, whether the PBM is of cellular or viral origin.

In addition, AUC experiments of the PTPN3-PDZ construct at 70 μ M complexed to HBVc PBM show only one species with a sedimentation coefficient of 1.6S corresponding to a one-to-one complex with the peptide, without higher oligomeric species (Table 2). Thus, the PBM binding onto PTPN3-PDZ prevents oligomerization at high concentration.

Affinities of the viral and the cellular PBMs for PTPN3-PDZ. The affinities of HPV16E6 PBM, HPV18E6 PBM, p38 γ PBM and HBVc PBM for PTPN3-PDZ were measured by NMR titration. We followed ¹H, ¹⁵N chemical shift perturbations of PTPN3-PDZ signals in the ¹H-¹⁵N HSQC spectra as a function of increasing concentrations of the PBM peptide to determine the dissociation constant (K_D) (Fig. S1). The p38 γ PBM peptide binds to PTPN3-PDZ with a K_D value of 26 μ M (Table 1). This affinity is 16-fold lower than the one previously reported for PTPN4-PDZ (K_D of 1.6 μ M)¹⁰. We obtained K_D values of 29 μ M, 53 μ M and 37 μ M for PTPN3-PDZ with HBVc PBM, HPV16E6 PBM and HPV18E6 PBM, respectively (Table 1). The K_D values are all in the same tenth-of-micromolar range for the viral PBMs and are close to the one of the cellular partner p38 γ . The measured affinities fall in the standard 0.1–100 μ M range¹² for PDZ-PBM interactions. To rule out the possibility of an effect of self-association of PTPN3-PDZ on the K_D values, we compared the K_D values with the HPV16E6 and HPV18E6 PBM peptides fitted from the NMR data with the ones estimated from the binding intensities (BIs) obtained for the HPV16E6 and HPV18E6 PBM peptides measured from a quantitative screening assay against the human PDZome library, in which each PDZ domain is at a concentration of 4 μ M and in excess of PBM peptides¹². The BIs are directly related to the K_d. For PTPN3-PDZ, the estimated K_D values are 46 and 96 μ M for the HPV16E6 and HPV18E6 PBMs respectively, in good agreement with our experimental results.

Stabilization of PTPN3-PDZ for structural studies. The instability and tendency to oligomerize of PTPN3-PDZ made this construct unsuitable for crystallogenes. To gain in stability and protein expression yield, we extended the N-terminal extremity of PTPN3-PDZ by 15-residues with its wild-type upstream sequence (PTPN3-PDZ_{Next} in Fig. 1). These extensions have been reported to provide structural stability to some PDZ containing proteins^{13,14}. The PTPN3-PDZ production yield was increased two-fold after extension from 0.9 mg to 1.8 mg of purified protein per liter of culture. The stability of PTPN3-PDZ_{Next} was then evaluated by NMR HSQC spectra recorded during a week at 20 °C. Spectra show a similar pattern of resonances than in the HSQC spectrum of PTPN3-PDZ without any signs of aggregation, such as a loss of peak intensity (Fig. 2B). The additional peaks of the N-terminal extension residues fall in the spectral region characteristic of unfolded protein and no chemical shift differences between PTPN3-PDZ_{Next} and the shorter form are detected. Thus, the N-terminal extension is not structured in the PTPN3-PDZ_{Next} and does not interact with the PDZ domain.

In addition, one single peak was detected by AUC for PTPN3-PDZ_{Next} at concentrations between 17.5 μ M and 70 μ M with a sedimentation coefficient of 1.6–1.7S, in agreement with a monomeric form of 12.7 kDa (Table 2) and devoid of higher oligomeric species. All these data indicate that the extension of PTPN3-PDZ resulted in a stable and folded PDZ domain in solution.

AUC experiments were also performed on the PTPN3-PDZ_{Next} construct at 70 μ M complexed to HBVc PBM, HPV16E6 PBM and HPV18E6 PBM (Table 2 and Fig. 2D). As expected, only one species with a sedimentation coefficient of 1.6S is detected in all cases, corresponding to a one-to-one complex.

Crystal structure of the PDZ domain of PTPN3 in complex with the viral HPV16E6 PBM. To investigate the molecular determinants of the interaction of PTPN3 with the HPV16 E6 protein, we solved the crystal structure of the complex formed by PTPN3-PDZ_{Next} and HPV16E6 PBM by molecular replacement at 2.19 Å resolution (Table 3). The structure factors and coordinates have been deposited in the Protein data Bank under accession code 6HKS.

PTPN3-PDZ_{Next} adopts a typical PDZ fold, with a β -sandwich comprising five β strands and two α helices (Fig. 3A). The PBM ligand binds in a hydrophobic cleft formed by the β 2-strand, the α 2-helix and the “GLGF” loop (Figs 3A,B). HPV16E6 PBM binds to the PDZ domain as an anti-parallel extension of the β 2-strand domain in a conventional mode. There is no electron density corresponding to the N-terminal extension of PTPN3-PDZ_{Next} in the crystal, which indicates a disordered region, in agreement with the NMR results that show an unfolded extension.

We also determined a 3D structural model with the CS-ROSETTA approach¹⁵ using the ¹⁵N, ¹³C α , ¹³C β NMR backbone and ¹³C β resonances of PTPN3-PDZ_{Next} complexed to HPV16E6 PBM (BMRB accession number 27645). The CS-Rosetta modeling using solution NMR data revealed a very similar conformation to the crystal structure, with a low root mean square deviation (rmsd) of 1.11 Å for the backbone atoms between the crystal structure and the averaged solution model originated from the 10 lowest energy models (Fig. 4C). Thus, 3D models of PDZ-PTPN3 determined by CS-Rosetta reveal a conformation in solution that matches the crystal structure of the PDZ domain of the complex PTPN3-PDZ_{Next} and HPV16E6 PBM (Fig. 3B). The AUC sedimentation velocity was also back-calculated using the crystal structure¹⁶. We obtained a calculated sedimentation coefficient value

	PTPN3-PDZ _{Next} /HPV16E6 PBM
Data collection	
Space group	P 1 2 1 1
Unit cell (a, b, c) (Å)	46.62, 77.43, 130.03
α, β, γ (°)	90.00, 90.14, 90.00
Resolution (Å)	46.62–2.19 (2.23–2.19)
No. of reflections (total/unique)	324034/47024
Redundancy	6.9 (6.5)
Completeness (%)	99.1 (95.7)
$I/\sigma(I)$	8.17 (1.11)
R_{meas}	0.14 (1.71)
CC(1/2)	99.6 (72.6)
Refinement	
Resolution (Å)	43.34–2.19
No. reflections	46869
$R_{\text{work}}/R_{\text{free}}$	0.195/0.245
No. of protein atoms/ligand atoms	4402/350
No. of solvent/hetero-atoms	235/9
Rmsd bond lengths (Å)	0.008
Rmsd bond angles (°)	0.949
Wilson B-factors	47.6
Ramachandran plot (favored/disallowed)*	97.4/0.4
PDB code	6HKS

Table 3. Data collection and refinement statistics. Values in parenthesis correspond to the highest resolution shell. $R_{\text{meas}} = \sum h(n/n-1)^{1/2} \sum_i |I_i(h) - \langle I(h) \rangle| / \sum h \sum_i I_i(h)$, where $I_i(h)$ and $\langle I(h) \rangle$ are the i th and the mean measurement, respectively, of the intensity of reflection h . $R_{\text{work}} = \sum h ||F_{\text{obs}}(h)| - |F_{\text{calc}}(h)|| / \sum h |F_{\text{obs}}(h)|$, where $F_{\text{obs}}(h)$ and $F_{\text{calc}}(h)$ are the observed and calculated structure factors, respectively. No I/σ cutoff was applied. R_{free} is the R value obtained for a test set of reflections consisting of a randomly selected 5% subset of the data set excluded from refinement. *Categories were defined by MolProbity.

of 1.59S, which is in good agreement with the experimental measurement of 1.6S. Thus, the overall structure of the domain in the crystal is consistent with its arrangement observed in solution.

Then, we compared the PDZ domains of PTPN3 and PTPN4, which share 71% of sequence identity (Fig. 4A). Both structures of PTPN3-PDZ and PTPN4-PDZ are highly similar, with a very low rmsd value of 0.34 Å for the backbone atoms (PTPN4 PDB ID 5EYZ)¹⁰ (Fig. 4B).

Structural insights on the recognition of the PBM of HPV16 E6 protein. The PDZ domain of PTPN3 possesses the interaction network specific to class I PDZ domains and recognizes the consensus sequence S/T-X- Φ -COOH, where X is any residue, and Φ is a hydrophobic residue¹⁷. A clear electron density map was seen for only the last seven C-terminal residues of HPV16E6 PBM (-TRRETQL_{COOH}) that are inserted into the binding groove (Fig. 3A,B). Similarly, only the last 5 or 6 C-terminal residues of PBM peptides are visible in the crystal PTPN4-PDZ structures in complex with PBM peptides (PDB IDs 3NFL, 3NFK, 5EZO, 5EYZ in references^{10,11}).

The interactions of the PBM residues at positions 0 and -2 with PTPN3-PDZ are quite similar to the bonding patterns already observed in the complex between PDZ-PTPN4 and the viral peptide Cyto13-att (-GETRL_{COOH}) derived from the attenuated rabies virus, or the optimized pro-apoptotic 13-amino acids peptide Cyto8-RETEV (-RETEV_{COOH}) (PDB IDs 3NFK and 5EYZ respectively)^{10,11}. Indeed, the C-terminal carboxylate of the leucine (L_0) of the HPV16E6 PBM forms three H-bonds with the amide nitrogens of F521, G522 and F523 of the “GLGF motif” on PTPN3-PDZ (Fig. 3B) as observed for F528, G529 and F530 of PTPN4-PDZ with the L_0 of the Cyto13-att (-GETRL_{COOH}) PBM¹¹. The hydroxyl group of threonine at position -2 forms a hydrogen bond with the N ϵ 2 of the conserved histidine H572 from the α 2-helix of PTPN3-PDZ (Fig. 3B). Electron acceptors such as serine and threonine are therefore preferred at this position for the class I PDZ domains.

At position -1 of HPV16E6 PBM, the glutamine (Q_{-1}) side chain well defined and forms a H-bond with a water molecule that is also H-bonded to the N δ 2 of N524 of PTPN3-PDZ (Fig. 3B), whereas it is exposed to the solvent. Interestingly, in the complex of PTPN4-PDZ with the PBM of the glutamate receptor subunit GluN2A (PDB ID 3NFL), which presents a D in position -1¹¹, the D₋₁ forms a H-bond with the N531 side-chain amine group of PTPN4-PDZ, the equivalent of N524 of PTPN3-PDZ.

A glutamate at position -3 (E_{-3}) is conserved in all viral PBMs targeting PTPN3 and is also present in the one of p38 γ . The side-chain carboxyl of E_{-3} forms a bifurcated H-bond with the hydroxyl of the conserved S538 (S545 for PTPN4) and N524 (N531 for PTPN4) amine group (Figs 3B and 4A). E_{-3} is also stabilized by hydrophobic contacts involving its C β -C γ carbon chain and the long aliphatic side chain of K526 as observed in PTPN4 (K533) bound to Cyto8retev (PDB 5EYZ).

Finally, the guanidinium nitrogens of arginine at position -4 (R_{-4}) form ionic bonds with the carboxylate oxygens of D573 (D580 in PTPN4)(Fig. 3B).

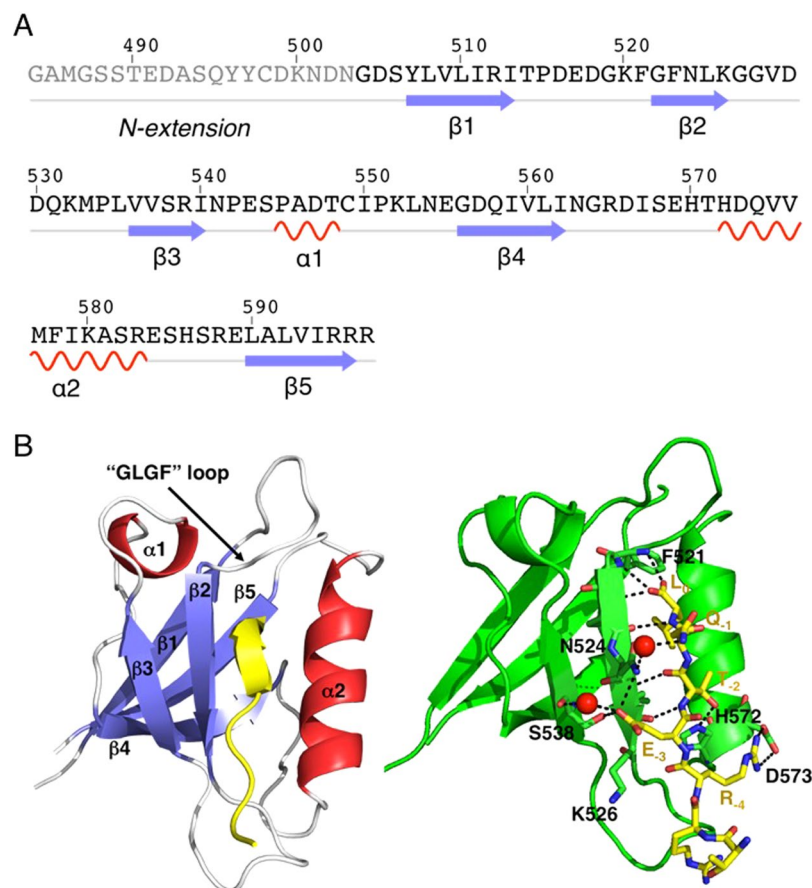


Figure 3. X-ray structure of PTPN3-PDZ_{Next} bound to HPV16E6 PBM. **(A)** Sequence and secondary-structure elements in PTPN3-PDZ_{Next}. No electron density was observed for the residues shown in gray. **(B)** Left panel: Overall structure of PTPN3-PDZ_{Next} bound to HPV16E6 PBM displayed as a ribbon. The structure is colored according to secondary structure elements (β -strands in blue, helices in red, and loops in white). The peptide (in yellow) enters the PDZ-binding pocket delimited by the β 2 strand, the α 2 helix, and the “GLGF” loop (⁵¹⁹GKFGF⁵²³). Right panel: Detailed view of the PTPN3-PDZ bound to HPV16E6 PBM. Important residues are shown as sticks in CPK colors, molecules of water are represented as red spheres, and intermolecular H-bonds and polar contacts are shown as black dashed lines.

To determine short- and long-range perturbations in PTPN3-PDZ upon PBM binding, we analyzed by solution NMR the chemical shifts of the free PTPN3-PDZ in solution and its complex with peptide HPV16E6 PBM, comparing their ¹H, ¹⁵N HSQC spectra. We assigned 97% of the HN resonances, 99% of the C α resonances, 90% of the C β resonances and 89% of the CO resonances (89 non-Proline residues over 94) of PTPN3-PDZ complexed to HPV16E6 PBM (BMRB accession number 27645). Chemical shift changes ($\Delta\delta$) in PTPN3-PDZ spectra upon complex formation provide insights at atomic level on the residues involved in the interaction with the HPV16E6 PBM and on potential distal effects. Two types of signals were useful for the analysis of the ¹H, ¹⁵N chemical shift mapping (Fig. 5A): (1) signals experiencing significant chemical shift changes ($\Delta\delta > 0.15$ ppm); (2) signals that undoubtedly disappear from their original well-resolved position in the spectrum upon complex formation and are therefore severely affected when binding to the PBM (either due to large chemical shift changes or to severe line broadening effects caused by exchange). Residues corresponding to type 1 are shown in blue in Fig. 5A,B, while seven residues of type 2 are colored in red. Nine residues, whose behavior could not be safely defined mainly because they fall in crowded spectral regions, are colored in gray.

As expected, we identified differences in backbone NH chemical shifts between free PTPN3-PDZ and when bound to HPV16E6 PBM. Residues G522, G527 and H572, are among the most affected and are located in the PBM binding site. I559 and L590 also experience a strong effect. Indeed, they make hydrophobic contacts of 5.4 Å and 4 Å with F523 and F521 of the “GLGF motif”, respectively. E589 (in red), the direct neighbor of L590, displays perturbations both in chemical shift or intensity. Some residues of α 2-helix in the binding site are affected: D573 and F578 (in blue) and V575 (in red). We find also residues such as F521 and F523 of the “GLGF” motif affected in their NMR resonances. D530 in the β 2- β 3 loop forms a H-bond with H572, which interacts directly with the PBM peptide (position T₋₂).

Interestingly, G556 (in red) is part of a path of perturbation starting from G527 (in red) and G528 (in blue) at the end of β 2-strand in the binding groove, transmitted to the close V536 (in red) at the beginning of β 3-strand, affected

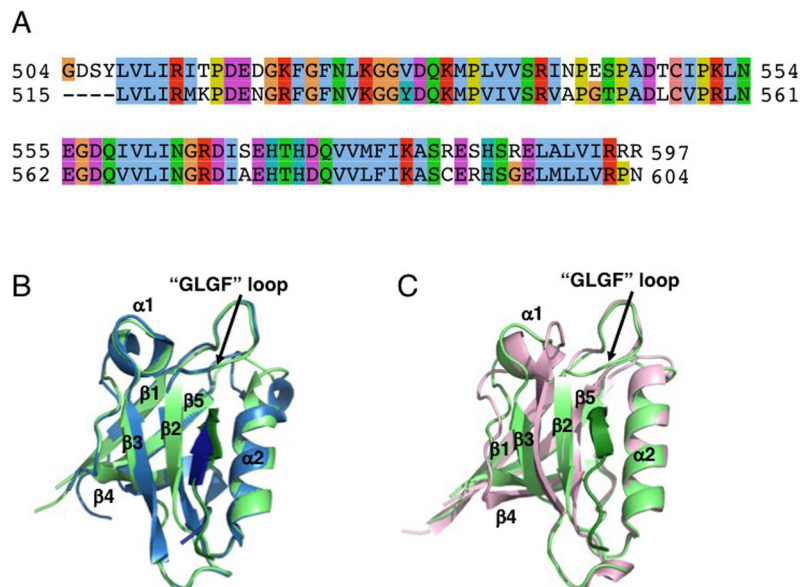


Figure 4. Structural comparison of the X-ray structure of PTPN3-PDZ_{Next} bound to HPV16E6 PBM with PTPN4-PDZ complexed to Cyto8retev, and with the CS-ROSETTA model of PTPN3-PDZ. (A) Clustalw2 alignment of PTPN3-PDZ (top) and PTPN4-PDZ (bottom). (B) Superposition of the structures of PTPN3-PDZ_{Next} bound to HPV16E6 PBM (light green/green) with PTPN4-PDZ bound to Cyto8retev (PDB 5EYZ) (light blue/blue). (C) Superposition of PTPN3-PDZ_{Next} bound to HPV16E6 PBM (light green/green) with the CS-ROSETTA model of PTPN3-PDZ (pink).

in signal intensity. The perturbation of V536 seems to be transferred to the close G556 on the β 4-strand. I559 (in red) on β 4-strand is also affected, possibly experiencing the perturbation of the F523 located 5 Å away of the “GLGF motif”. Finally, the path of perturbation finishes with R595 at the C-terminus of the PDZ domain (in blue). Thus, long-range perturbations seem to be induced upon PBM binding and transmitted through an interconnected network of residues throughout PTPN3-PDZ. This network involves the β 2, β 3 and β 4 strands and the C-terminus.

Discussion

PTPN3 is a PDZ domain-containing phosphatase that has been demonstrated to function as a tumor suppressor or, conversely, as an oncoprotein in a context-dependent manner. It integrates a number of signaling pathways via interactions mediated by its PDZ domain^{18–20}. The PDZ domain of PTPN3 is also a cellular target of oncoviruses. Until now, no structural data on this domain were available. We investigated the structural and functional properties of PTPN3-PDZ and characterized its interaction with the PBMs of one cellular partner, the MAP kinase p38 γ , and of its viral partners, the HPV E6 protein and the HBV core protein. We defined in particular the molecular basis of the PDZ-PBM interaction between HPV E6 and PTPN3.

Structural properties of the PDZ domain of PTPN3. The crystal structure of PTPN3-PDZ_{Next} complexed to HPV16E6 PBM reveals a typical PDZ fold with five β -strands and two α -helices (Fig. 3A,B). It was previously reported that some PDZ domains can be stabilized by extensions of either the N-, C-, or both termini of a PDZ²¹. These extended regions were required for correct folding and ligand binding. For example, the PDZ1 of the membrane-associated guanylate kinase inverted MAGI-1 is unstable without N- and C-terminal extensions of 14 and 26 residues, respectively, even though these sequences are unstructured in both free and ligand-bound states²². Our minimal PTPN3-PDZ construct, deduced from sequence alignments with PDZ domains of known structure (Fig. 1), started at residue 504, while the first N-terminal secondary structure element observed in the crystal structure of PTPN3-PDZ_{Next}, the β 1-strand, started at residue 507 and not at 510 as expected from the sequence alignments. We showed that this PTPN3-PDZ construct is folded but unstable. We were able to increase its stability and production yield, while decreasing its tendency to auto-associate, by the addition of an unfolded 15-residue N-terminal extension to the minimal delimitation of PTPN3-PDZ. 3D structural models of PDZ-PTPN3 in complex with HPV16E6 PBM determined by CS-Rosetta using backbone and ¹³C β NMR chemical shift assignments revealed a conformation in solution congruent with the X-ray structure of the PDZ domain.

Our AUC data showed that PTPN3-PDZ auto-association is prevented by PBM binding, as already reported for other PDZ domains²³. The molecular mechanism by which PBM binding interferes with PTPN3 PDZ auto-association remains unknown. Our NMR results on long-range perturbations upon PBM binding are consistent with a network involving the β 2, β 3 and β 4 strands and the C-terminus residues throughout PTPN3-PDZ. This connected pathway communicates the binding event to regions that are distal to the binding cleft. These findings are compatible with previous data on PDZ domains suggesting that energetic pathways within PDZ domains may support allostery^{24–27}. A regulation controlled by the equilibrium between PBM binding and dimerization is attractive since over 30% of PDZ domains are known to form dimers in solution²⁸.

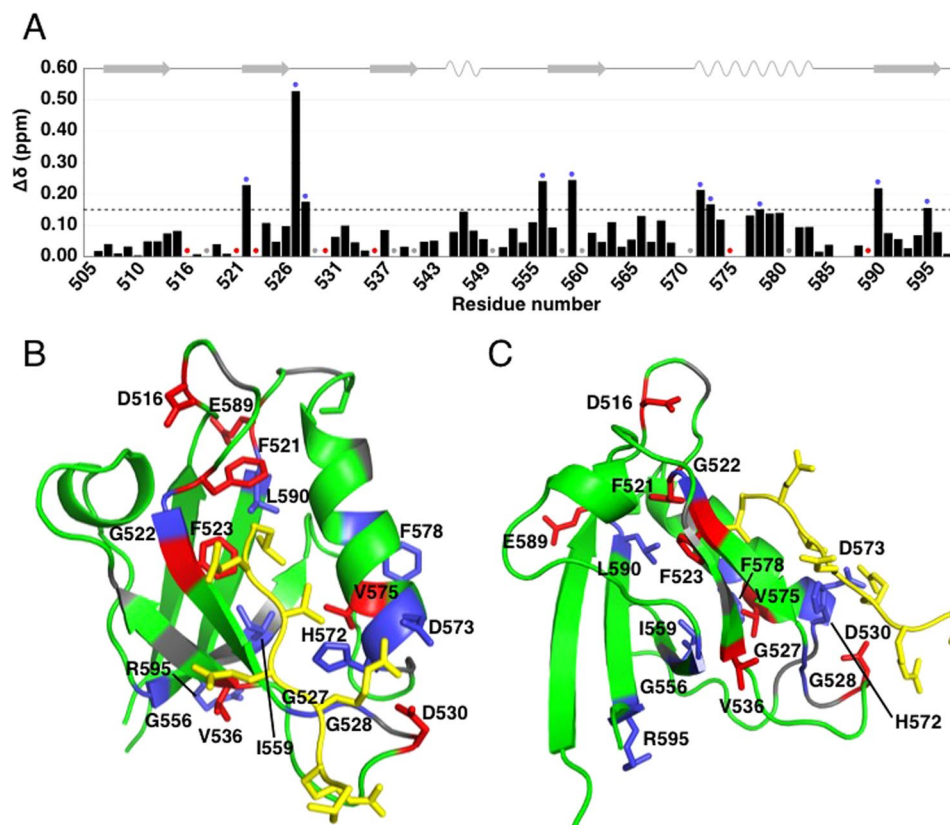


Figure 5. $\Delta\delta$ between free PTPN3-PDZ and PTPN3-PDZ bound to HPV16E6 PBM. **(A)** $\Delta\delta$ values (^1H , ^{15}N) computed as $\Delta\delta = [(\Delta\delta\text{H})^2 + (\Delta\delta\text{N} \cdot 0.15)^2]^{1/2}$. The dashed black line represents the threshold used to determine the perturbed region with $\Delta\delta > 0.15$ ppm; blue circles mark residues perturbed with $\Delta\delta > 0.15$ ppm; red circles mark residues with no resonance in free PTPN3-PDZ; grey circles mark residues whose behavior could not be safely defined, mainly because they fall in crowded spectral regions. **(B)** Two views of the structure of PTPN3-PDZ bound to HPV16E6 PBM (PDB ID 6HKS) with different colors depending on $\Delta\delta$ values. Blue: $\Delta\delta > 0.15$ ppm; red: resonances absent from the free PTPN3-PDZ spectrum; grey: residues whose resonances fall in crowded spectral regions. HPV16E6 PBM peptide is represented in yellow. Side chains of affected residues are represented as sticks.

The PDZ domain of PTPN3 as a target of viruses. We showed that the viral PBMs interact with PTPN3-PDZ with similar affinities to that of the endogenous PTPN3 ligand p38 γ . Despite variations in the PBM sequences, all the affinities fall in the 25–55 μM range (Table 1), in agreement with affinities previously reported for HPV E6 PBM PDZ binders¹². This result is consistent with the hypothesis of a competition between cellular PBM-containing partners and PBMs on viral proteins to bind PTPN3 and hijack signalling pathways in infected cells. E6 binding to PTPN3 should be favoured in HPV infected cells since E6 is highly expressed in HPV-derived cancer cells. Only the C-termini of E6 oncoproteins in high-risk HPV strains contain PBMs, and this is a marker of high oncogenic potential. The presence of a PBM on E6 confers the capacity to interact with various cellular PDZ domain-containing proteins including PTPN3²⁹. Furthermore, this interaction causes degradation of most of its targets through the ubiquitin–proteasome pathway³⁰. Therefore, PDZ-protein interactions with E6 PBM not only disrupt protein–protein interactions but also promote degradation of the PDZ-containing protein. Our results raise the question of how specificity is achieved in this interaction, and whether or not PTPN3 PDZ domain is able to discriminate its binding partners.

Molecular basis and specificities of the interaction with the PDZ domain of PTPN3. The crystal structure and the NMR chemical shift mapping of PDZ-PTPN3 in complex with the PBM of the E6 oncogenic protein of HPV16 highlight the main structural determinants of recognition of the C-terminal sequence of the E6 protein. HPV16E6 PBM binding to PTPN3-PDZ is consistent with the binding mode observed for canonical class I PDZ domains, where L_0 and T_{-2} are involved in key interactions with the PDZ domain. Indeed, L_0 interacts with the conserved GLGF motif, which is crucial for the hydrogen bond coordination of the terminal carboxylate group during the PDZ domain–PBM ligand interaction. Moreover, the T_{-2} forms a hydrogen bond with a conserved Histidine or Arginine in the PDZ domain, signature of class I PDZ domains^{14,31}. The histidine, H572 in PTPN3, is very well conserved in the PTPN3 and PTPN4 orthologs (Fig. 6), and was also identified as a preferred



Figure 6. Sequence alignment of PTPN3 and PTPN4 from different species.

residue for E6 binding by the alignment of 209 PDZ domains ranked according to their highest E6 binding intensity, as determined by a high-throughput assay¹².

At position −1 of HPV16E6 PBM, a glutamine forms a H-bond with a water molecule that is also H-bonded to the N524 of the PDZ domain of PTPN3, whereas the N531 of PTPN4 forms a H-bond with the D in position −1 in the PBM of the cellular partner GluN2A-16¹¹. This asparagine is very well conserved in the orthologs of PTPN3 and PTPN4, while a serine or threonine is more commonly found at this position in the PDZome.

N524 and S538 (S545 for PTPN4) form also H-bonds with the E in position −3 of the HPV16E6 PBM (Figs 3B and 4A). S538 and N524 are strictly conserved in PTPN3 and PTPN4 orthologues, with only 2 and 1 exceptions respectively (Fig. 6); although in the PDZome, S or T residues are often located at this position 524, and in position 538, we often find S or T but also K, H, Q and R. E_{−3} is also stabilized by K526 in PTPN3. This K is conserved in PTPN3 and PTPN4 orthologues but only a few K are present at this position in the human PDZome and I, V, A, S and M are more often found. We previously reported that one single mutation (Q to E) at position −3 of the PBM of the envelope G protein of rabies virus (RABV) switches the fate of the infected cell from survival to apoptosis³². Indeed, the Q to E change increases the number and change the pattern of cellular partners for the PBM in the infected cells, enabling the attenuated RABV strain G protein to interact with PTPN4 inducing cell death. PTPN4-PDZ is able to discriminate between these peptides, displaying a substantially higher affinity for E rather than Q in position −3. Altogether, these observations and the fact that a glutamic acid is often found in position −3 in all PBMs interacting with PTPN3 and PTPN4^{3,4,10,11,19,20,32,33} suggest that interactions with this E in position −3 are a critical determinant of specificity of PDZ/PBM interaction in the NT5 group comprising PTPN3 and PTPN4 phosphatases.

Finally, the arginine in position −4 is linked by ionic bonds to D573 in PTPN3-PDZ (D580 in PTPN4-PDZ) complexed to HPV16E6 PBM. This interaction is most likely maintained in the complexes with viral HPV18E6 PBM, HBVC PBMs, and even with the p38γ PBM, which presents a lysine instead of an arginine. D573 is well conserved in the orthologous PTPN3 and PTPN4 with only 3 exceptions, yet it is not the most frequently found amino acid at this position in the human PDZome. It might also be a specificity determinant common to the PDZ of PTPN3 and PTPN4 phosphatases. These two closely-related non-receptor PTPs that compose the NT5 sub-family emerged in holozoa³⁴ and duplicated in vertebrates with 54% of sequence identity, which rises to 71% for their PDZ domains. However, the expression pattern, substrates and interacting partners of PTPN3 and PTPN4 have a limited overlap.

Methods

Production and purification of recombinant proteins and PBM ligands. PTPN3-PDZ and PTPN3-PDZ_{Next} (Fig. 1) are encoded as an N-terminal glutathione S-transferase (GST) tagged protein in a pDEST15 expression plasmid and a pGST//2 (derived from pGEX-4T-1; Amersham) expression plasmid respectively. A TEV cleavage site is introduced between the N-terminal tag and the protein sequence. The vectors were used to transform *E. coli* BL21 Star (DE3) star cells (Invitrogen, Carlsbad, CA, USA). Uniformly ¹⁵N-labeled, ¹³C, ¹⁵N-labeled and unlabeled PTPN3-PDZ and PTPN3-PDZ_{Next} constructs were expressed and purified as previously described⁹.

Briefly, harvested cells were resuspended in buffer A (50 mM Tris/HCl, pH 7.5, 150 mM NaCl), 2 mM β-mercaptoethanol and protease inhibitor cocktail (Roche), and then disrupted in a French press. The clarified supernatants were loaded onto a GST column (GSTrap HP, GE Healthcare) and washed with the same buffer. The GST tag was cleaved by overnight incubation at 4 °C by TEV protease (1% mol/mol) directly injected into the column. The eluted fractions containing the protein were pooled and loaded onto a size exclusion column (HiLoad Superdex 75 pg; GE) equilibrated with buffer A with 0.5 mM Tris(2-carboxyethyl)phosphine (TCEP). For crystallogenesis of PTPN3-PDZ_{Next}, the same protocol was followed, replacing the Tris/HCl in buffer A by 20 mM HEPES pH 8 on the size exclusion chromatography step. Purified proteins were concentrated using centrifugal filter devices (Vivaspin, Sartorius). Protein concentration was estimated from its absorbance at 280 nm.

The peptides, p38γ PBM, HBVc PBM, HPV16E6 PBM and HPV18E6 PBM, were synthesized in solid phase using Fmoc strategy (Proteogenix) and resuspended in H₂O.

CD experiments. All CD measurements were acquired with an Aviv 215 spectropolarimeter. Far-UV (195–240 nm) spectra were recorded at 25 °C on 8.4 μM PTPN3-PDZ samples in a cylindrical cell with a 0.2-mm path-length. Ellipticity was measured every 1 nm. The final spectrum of the protein sample was obtained by averaging three successive scans and subtracting the baseline spectrum of the buffer recorded under the same conditions. The CONTIN program was used for quantitative decomposition of the far-UV CD spectrum³⁵.

NMR experiments. The NMR samples for the PTPN3-PDZ and PTPN3-PDZ_{Next} constructs were prepared in buffer A with 0.5 mM TCEP and D₂O (5–10% vol:vol). All NMR experiments were performed on a 600-MHz Varian NMR System spectrometer equipped with a triple resonance ¹H/¹³C/¹⁵N cryoprobe.

The NMR titration experiments to measure PTPN3-PDZ·PBM peptide affinities and the NMR experiments for backbone assignment of PTPN3-PDZ in complex with HPV16E6 PBM were performed with the PTPN3-PDZ construct at 15 °C. Briefly, the unlabeled peptides (stock solutions ranging from 2.8 to 5.7 mM) at pH 7.5 were added stepwise in a sample initially containing 240–260 μL of ¹⁵N-labeled PTPN3-PDZ at a concentration of 95 or 149 μM. A series of ¹H, ¹⁵N HSQC spectra was recorded for the different titration points. Chemical shift changes were calculated using the free PTPN3-PDZ signals as a reference. Chemical shift differences in the cross-peaks by titration were calculated using the relationship: $\Delta\delta = ((\Delta\delta_{\text{HN}})^2 + (0.15 \times \Delta\delta_{\text{N}})^2)^{0.5}$ where $\Delta\delta_{\text{HN}}$ and $\Delta\delta_{\text{N}}$ are changes in the ¹H and ¹⁵N chemical shifts in ppm, respectively. The weighting factor of 0.15 was used to adjust the relative magnitudes of the amide nitrogen chemical shift range and the amide proton chemical shift range. Signals broaden in the moderate fast-exchange regime observed with PTPN3-PDZ and the PBM peptides, increasing the experimental errors on the chemical shift measurements used for the fitting of the K_D. The following of chemical shift changes during titrations and the fitting of curves were performed with the CcpNmr Analysis software³⁶. A pool of 8 to 14 peaks with the best fit for each titration were kept to deduce the K_D, and the errors are the standard deviations of all the K_D values fitted from the curves.

The sequence specific ¹H^N, ¹⁵N, ¹³C^α, ¹³C^β and ¹³CO resonance assignments were performed using TROSY-based versions of the following experiments: 2D ¹H, ¹⁵N HSQC and 3D HNCO, HNCA, HN(CO)CA, CBCA(CO)NH and HNCACB. The spectra were processed with NMRPipe³⁷ and subsequently analysed in CcpNmr Analysis software³⁶. The chemical shifts have been deposited in the BioMagResBank (BMRB) under accession number 27645.

A series of ¹H, ¹⁵N HSQC spectra was recorded over a week at 20 °C for a sample of PTPN3-PDZ_{Next} at 170 μM to follow the stability of the construct.

DSC experiments. DSC experiments were carried out on a Microcal VP-capillary DSC (Malvern). 18, 21, 94 and 28 μM of PTPN4-PDZ, free PTPN3-PDZ, free PTPN3-PDZ_{Next} and PTPN3-PDZ complexed with peptides respectively were scanned from 10 °C to 90–100 °C with a scan rate of 200 °C/h. Large excess of peptides were used: about 600 μM for p38γ PBM, about 300 μM for HBVc PBM and HPV16E6 PBM and about 200 μM for HPV18E6 PBM. The buffer A was repeatedly scanned to ensure a stable buffer baseline. The buffer baseline was subtracted from the protein thermogram. Subsequent to normalization of the data by the protein concentration, a non-linear curve fitting algorithm was employed to obtain the T_m of the transition. All experiments were repeated three times.

AUC experiments. Sedimentation velocity experiments were carried out at 20 °C using a Beckman Coulter XL-I centrifuge equipped with a AN60-Ti rotor. Various protein and protein-peptide complex concentrations (protein concentrations ranging from 12.5 to 70 μM) samples were centrifuged for 17 h at 42000 rpm. Data were analyzed with SEDFIT 15.1 using a continuous size distribution c(S) model. The partial specific volume, the viscosity and the density of the samples were calculated with SEDNTERP. The processed data were used to obtain values of sedimentation coefficients at null concentration in our experimental conditions (S₀) and to get the standard sedimentation coefficients in water at 20 °C (S_{0,w,20}) (Table 2).

Crystallisation, data collection, and structure determination. The HPV16E6 PBM peptide used for co-crystallization was added in excess to form >95% of the complex with the protein. The PDZ domain-peptide complex for crystallization was generated by mixing PTPN3-PDZ_{Next} and the peptide at a ratio of 1:2. Initial screening of crystallization conditions was carried out by the vapor diffusion method using a MosquitoTM nanoliter-dispensing system (TTP Labtech). Sitting drops were set up using 400 nL of a 1:1 mixture of each sample protein and crystallization solutions (672 different commercially available conditions) equilibrated against a 150 µL reservoir in multiwell plates (Greiner Bio-One). The crystallization plates were stored at 4 °C in a RockImager1000® (Formulatrix) automated imaging system to monitor crystal growth. The best crystals were obtained by mixing 200 nL of PTPN3-PDZ_{Next}- HPV16E6 PBM complex solution (concentration of the PDZ domain at 4.8 mg/mL) in 20 mM HEPES pH 8, 150 mM NaCl, 0.5 mM TCEP mixed with 200 nL of reservoir solution containing 20% w/v PEG 3350, 0.2 mM KI at pH 7. Crystals were then flash-cooled in liquid nitrogen using Paratone-paraffin 50%(V/V)/50%(V/V) oil as the cryoprotectant.

X-ray diffraction data were collected at a wavelength of 0.979 Å on the beamline PROXIMA-2A at Synchrotron SOLEIL (St. Aubin, France). The data were processed with XDS³⁸ and Xdsme³⁹, and other programs from the CCP4 suite⁴⁰. The structures were solved by molecular replacement with PHASER⁴¹ using the search atomic model of PTPN4-PDZ (PDB ID 5EZ0). The locations of the bound peptides were determined from a $F_o - F_c$ difference electron density maps. Models were rebuilt using COOT⁴², and refinement was done with phenix.refine of the PHENIX suite⁴³. The overall assessment of model quality was performed using MolProbity⁴⁴. The crystal parameters, data collection statistics, and final refinement statistics are shown in Table 3. All structural figures were generated with the PyMOL Molecular Graphics System, Version 1.7 (Schrödinger).

Sequence alignment. Sequence of PTPN3 (accession number NP_002820.3) was used as query on the InterEvolAlign server⁴⁵ to retrieve one single homolog per species assessed as probable ortholog through a reciprocal blast search procedure against the non-redundant database. Retrieved full-length sequences were re-aligned using MAFFT⁴⁶ and displayed using Jalview⁴⁷.

Data Availability

All data generated or analysed during this study are included in this published article.

References

- Gjörloff-Wingren, A. *et al.* Subcellular localization of intracellular protein tyrosine phosphatases in T cells. *Eur. J. Immunol.* **30**, 2412–2421 (2000).
- Hendriks, W. J. A. J. & Böhmer, F.-D. *Non-transmembrane PTPs in Cancer in Protein Tyrosine Phosphatases in Cancer* (ed. Neel B. G. & Tonks N. K.). (Springer, 2016).
- Töpfer, S., Müller-Schiffmann, A., Matentzoglou, K., Scheffner, M. & Steger, G. Protein tyrosine phosphatase H1 is a target of the E6 oncoprotein of high-risk genital human papillomaviruses. *J. Gen. Virol.* **88**, 2956–2965 (2007).
- Hsu, E.-C. *et al.* Suppression of hepatitis B viral gene expression by protein-tyrosine phosphatase PTPN3. *J. Biomed. Sci.* **14**, 731–744 (2007).
- Human papillomavirus (HPV) and cervical cancer. Available at: [http://www.who.int/en/news-room/fact-sheets/detail/human-papillomavirus-\(hpv\)-and-cervical-cancer](http://www.who.int/en/news-room/fact-sheets/detail/human-papillomavirus-(hpv)-and-cervical-cancer). (Accessed: 17th September 2018).
- Thomas, M., Myers, M. P., Massimi, P., Guarnaccia, C. & Banks, L. Analysis of Multiple HPV E6 PDZ Interactions Defines Type-Specific PDZ Fingerprints That Predict Oncogenic Potential. *PLoS Pathog.* **12**, e1005766 (2016).
- Chen, K.-E. *et al.* Reciprocal allosteric regulation of p38γ and PTPN3 involves a PDZ domain-modulated complex formation. *Sci. Signal.* **7**, ra98 (2014).
- Chen, K.-E. *et al.* Substrate specificity and plasticity of FERM-containing protein tyrosine phosphatases. *Struct. Lond. Engl.* **1993**(23), 653–664 (2015).
- Maisonneuve, P. *et al.* Regulation of the catalytic activity of the human phosphatase PTPN4 by its PDZ domain. *FEBS J.* **281**, 4852–4865 (2014).
- Maisonneuve, P. *et al.* Molecular Basis of The Interaction of the Human Protein Tyrosine Phosphatase Non-receptor Type 4 (PTPN4) with the Mitogen-Activated Protein Kinase p38γ. *J. Biol. Chem.*, <https://doi.org/10.1074/jbc.M115.707208> (2016).
- Babault, N. *et al.* Peptides targeting the PDZ domain of PTPN4 are efficient inducers of glioblastoma cell death. *Struct. Lond. Engl.* **1993**(19), 1518–1524 (2011).
- Vincentelli, R. *et al.* Quantifying domain-ligand affinities and specificities by high-throughput holdup assay. *Nat. Methods* **12**, 787–793 (2015).
- Christopherson, K. S., Hillier, B. J., Lim, W. A. & Bredt, D. S. PSD-95 assembles a ternary complex with the N-methyl-D-aspartic acid receptor and a bivalent neuronal NO synthase PDZ domain. *J. Biol. Chem.* **274**, 27467–27473 (1999).
- Manjunath, G. P., Ramanujam, P. L. & Galande, S. Structure function relations in PDZ-domain-containing proteins: Implications for protein networks in cellular signalling. *J. Biosci.* **43**, 155–171 (2018).
- Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
- Ortega, A., Amorós, D. & García de la Torre, J. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* **101**, 892–898 (2011).
- Lee, H.-J. & Zheng, J. J. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun. Signal.* **8**, 8 (2010).
- Li, M.-Y. *et al.* Protein tyrosine phosphatase PTPN3 inhibits lung cancer cell proliferation and migration by promoting EGFR endocytic degradation. *Oncogene* **34**, 3791–3803 (2015).
- Hou, S.-W. *et al.* PTPH1 dephosphorylates and cooperates with p38γ MAPK to increase ras oncogenesis through PDZ-mediated interaction. *Cancer Res.* **70**, 2901–2910 (2010).
- Zheng, Y., Schlondorff, J. & Blobel, C. P. Evidence for regulation of the tumor necrosis factor α-converter (TACE) by protein-tyrosine phosphatase PTPH1. *J. Biol. Chem.* **277**, 42463–42470 (2002).
- Wang, C. K., Pan, L., Chen, J. & Zhang, M. Extensions of PDZ domains as important structural and functional elements. *Protein Cell* **1**, 737–751 (2010).
- Charbonnier, S. *et al.* The structural and dynamic response of MAGI-1 PDZ1 with noncanonical domain boundaries to the binding of human papillomavirus E6. *J. Mol. Biol.* **406**, 745–763 (2011).

23. Terrien, E. *et al.* Interference with the PTEN-MAST2 interaction by a viral protein leads to cellular relocalization of PTEN. *Sci. Signal.* **5**, ra58 (2012).
24. Delhommel, F. *et al.* Deciphering the unconventional peptide binding to the PDZ domain of MAST2. *Biochem. J.* **469**, 159–168 (2015).
25. Fuentes, E. J., Der, C. J. & Lee, A. L. Ligand-dependent Dynamics and Intramolecular Signaling in a PDZ Domain. *J. Mol. Biol.* **335**, 1105–1115 (2004).
26. Hultqvist, G. *et al.* Energetic Pathway Sampling in a Protein Interaction Domain. *Structure* **21**, 1193–1202 (2013).
27. van den Berk, L. C. J. *et al.* An Allosteric Intramolecular PDZ–PDZ Interaction Modulates PTP-BL PDZ2 Binding Specificity†. *Biochemistry* **46**, 13629–13637 (2007).
28. Chang, B. H. *et al.* A Systematic Family-wide Investigation Reveals that ~30% of Mammalian PDZ Domains Engage in PDZ-PDZ Interactions. *Chem. Biol.* **18**, 1143–1152 (2011).
29. Ganti, K. *et al.* The Human Papillomavirus E6 PDZ Binding Motif: From Life Cycle to Malignancy. *Viruses* **7**, 3530–3551 (2015).
30. Javier, R. T. & Rice, A. P. Emerging theme: cellular PDZ proteins as common targets of pathogenic viruses. *J. Virol.* **85**, 11544–11556 (2011).
31. Liu, X. & Fuentes, E. J. Emerging Themes in PDZ Domain Signaling: Structure, Function, and Inhibition. In *International Review of Cell and Molecular Biology*, <https://doi.org/10.1016/bs.ircmb.2018.05.013> (Elsevier, 2018).
32. Préhaud, C. *et al.* Attenuation of rabies virulence: takeover by the cytoplasmic domain of its envelope protein. *Sci. Signal.* **3**, ra5 (2010).
33. Hironaka, K., Umemori, H., Tezuka, T., Mishina, M. & Yamamoto, T. The protein-tyrosine phosphatase PTPMEG interacts with glutamate receptor delta 2 and epsilon subunits. *J. Biol. Chem.* **275**, 16167–16173 (2000).
34. Chen, M. J., Dixon, J. E. & Manning, G. Genomics and evolution of protein phosphatases. *Sci. Signal* **10**, eaag1796 (2017).
35. Sreerama, N. & Woody, R. W. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260 (2000).
36. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
37. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
38. Kabsch, W. X. D. S. & Acta Crystallogr., D. Biol. Crystallogr. **66**, 125–132 (2010).
39. Legrandp, Aishima, J. & CV-GPhL. Legrandp/Xdsme: Version Working With Xds Version: June 1, 2017 And Previous., <https://doi.org/10.5281/zenodo.837885> (2017).
40. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
41. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 32–41 (2007).
42. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
43. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
44. Lovell, S. C. *et al.* Structure validation by C α geometry: phi, psi and C β deviation. *Proteins* **50**, 437–450 (2003).
45. Faure, G., Andreani, J. & Guerois, R. InterEvol database: exploring the structure and evolution of protein complex interfaces. *Nucleic Acids Res.* **40**, D847–856 (2012).
46. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
47. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinforma. Oxf. Engl.* **25**, 1189–1191 (2009).

Acknowledgements

Mariano Genera was supported by the Pasteur - Paris University (PPU) International PhD Program. The authors are grateful to the staff of the Crystallography Platform and the Molecular Biophysics Platform at Institut Pasteur for robot-driven crystallization screening and biophysics expertise respectively. We thank Sébastien Brûlé and Ariel Mechaly for helpful discussions. We acknowledge the Synchrotron SOLEIL (St Aubin, France) staff for assistance and advice during data collection on PROXIMA-2A beamline.

Author Contributions

M.G., N.W. and C.C.S. planned the experiments. M.G., D.S., B.R., A.H., B.B., C.S., R.G. and C.C.S. performed the experiments. M.G., D.S., B.R., B.B., R.G. and C.C.S. analysed the data. M.G., N.W. and C.C.S. wrote the paper.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-43932-x>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019