

Article

Structural Analysis of the 42 kDa Parvulin of *Trypanosoma brucei*

Edisa Rehic¹, Dana Hoenig¹, Bianca E. Kamba¹, Anna Goehring¹, Eckhard Hofmann² ,
Raphael Gasper^{2,3} , Anja Matena¹ and Peter Bayer^{1,*} 

¹ University Duisburg-Essen, Research Group Structural and Medicinal Biochemistry, Centre for Medical Biotechnology (ZMB), University of Duisburg-Essen, 45117 Essen, Germany; edisa.rehic@uni-due.de (E.R.); dana.hoenig@uni-due.de (D.H.); bianca.kamba@uni-due.de (B.E.K.); Anna.goehring@uni-due.de (A.G.); anja.matena@uni-due.de (A.M.)

² Protein Crystallography, Faculty of Biology and Biotechnology, Ruhr University Bochum, 44801 Bochum, Germany; eckhard.hofmann@bph.rub.de

³ Max Planck Institute for Molecular Physiology, 44227 Dortmund, Germany; raphael.gasper@mpi-dortmund.mpg.de

* Correspondence: peter.bayer@uni-due.de; Tel.: +49-0201-183-4677

Received: 19 February 2019; Accepted: 28 February 2019; Published: 7 March 2019



Abstract: *Trypanosoma brucei* is a unicellular eukaryotic parasite, which causes the African sleeping sickness in humans. The recently discovered trypanosomal protein Parvulin 42 (*TbPar42*) plays a key role in parasite cell proliferation. Homologues of this two-domain protein are exclusively found in protozoa species. *TbPar42* exhibits an N-terminal forkhead associated (FHA)-domain and a peptidyl-prolyl-*cis/trans*-isomerase (PPIase) domain, both connected by a linker. Using NMR and X-ray analysis as well as activity assays, we report on the structures of the single domains of *TbPar42*, discuss their intra-molecular interplay, and give some initial hints as to potential cellular functions of the protein.

Keywords: *Trypanosoma brucei*; Parvulin; *cis/trans* Isomerase; PPIase; FHA

1. Introduction

Trypanosomes are unicellular eukaryotic parasites belonging to the class Kinetoplastea. Infection with trypanosomes causes various diseases in humans, including Chagas disease (*Trypanosoma cruzi*) and the sleeping sickness disease (*Trypanosoma brucei*). *Trypanosoma brucei* is transmitted to its human host by the bite of a tsetse fly carrying this parasite in its salivary gland. After transmission, the parasite can enter the cerebral fluid through the blood vessel walls, causing the life-threatening effects of sleeping sickness. In the blood stream, *T. brucei* cells proliferate and are taken up again by the bite of a tsetse fly, where the parasite cells (procyclic state) in the intestine transform their energy metabolism from anaerobic glycolytic to aerobic processes and migrate to the salivary gland. Once the parasite entered the central nervous system of its human host, therapy against the outbreak of sleeping sickness is challenging.

Recently, the nuclear protein *TbPar42* (molecular mass of 42 kDa) has been identified, of which RNAi-induced knockdown in procyclic cells of *T. brucei* inhibits proliferation of the parasite, suggesting a key role of the protein in cell growth [1,2]. Orthologous proteins are found in *T. cruzi* and *Leishmania* [2], as well as in *Chlamydomonas* [3]. Protein sequence alignment identified *TbPar42* as a two-domain protein, exhibiting an N-terminal flanked forkhead association (FHA) domain and a putative peptidyl-prolyl-*cis/trans*-isomerase (PPIase) domain [4] connected by a linker region. FHA domains are generally known as phosphopeptide recognition modules [5], although the *TbPar42* FHA

domain has not yet been proven to have affinity to phosphorylated substrates. FHA domains play a critical role in cellular events such as DNA damage, DNA replication, or cell cycle progression [5–7]. Peptidyl-prolyl *cis/trans* isomerases are enzymes that accelerate the adjustment of the equilibrium of Xaa-Pro (Xaa being any amino acid) moieties within substrates. Sequence alignments showed that the *TbPar42* PPIase domain belongs to the family of parvulin PPIases [8], of which the human orthologues *hPin1* and *hPar14/17* are structurally and functionally characterized [9]. In contrast, the three dimensional structure and the function of the FHA and PPIase domains of *TbPar42* are unknown. Although the PPIase domain shows a relatively high sequence alignment to *hPin1* (37.5% identity; 56.7% homology; see Supplementary Figure S3), Goh and coworkers have demonstrated that the PPIase domain lacks *cis/trans* isomerase activity against phosphorylated substrate peptides [1]. So far, no further functional roles of *TbPar42* have been suggested.

We used nuclear magnetic resonance spectroscopy (NMR) and X-ray analysis to evaluate the structures and dynamics of the isolated domains of *TbPar42* (1–383) and discuss their intra-molecular interplay. Moreover, we demonstrated the recognition of phosphorylated peptides by conserved motifs within the FHA domain, and tested the catalytic activity of the PPIase domain against a broad variety of substrates. Similarity searches using DALI [10] suggest structural homology of the FHA domain of *TbPar42* to FHA domains of nuclear RNA-binding proteins such as Dawdle/SNIP1, suggesting a cellular role of *TbPar42* in RNA binding/processing events.

2. Materials and Methods

2.1. Expression of *TbPar42* Constructs

The DNA sequence of *TbPar42* from *T. brucei* (SwissProt database no. Q57XM6) was purchased with *Escherichia coli* adapted codon usage in a PEX-K vector from Eurofins Genomics (Ebersberg, Germany). The constructs *TbPar42*-full length (1–383), N-terminal extended FHA domain (NterFHA) (1–177), linker region (172–266), and PPIase domain (264–383) were PCR-amplified using DreamTaq DNA polymerase (Thermo Fisher Scientific, Darmstadt, Germany), the oligonucleotides listed in Table S5, and the synthesized *TbPar42* DNA as a template. The constructs were Apa1/Xho1 (Fast Digest from Fermentas, Germany) cloned into a modified pET-41 expression vector as described before [11], with an N-terminal GST-His₆ fusion and a PreScission protease cleavage site. Before transformation into *E. coli* BL21(DE3)T1r (Sigma-Aldrich, Darmstadt, Germany) the constructs were verified by Sanger sequencing.

For production of unlabeled protein, 50 mL LB overnight culture (50 µg/mL kanamycin) was harvested, resuspended in 1 L LB medium, and grown up to OD₆₀₀ = 0.8 at 37 °C, 160 rpm. For the expression of isotopically-labeled ¹⁵N-or ¹³C-¹⁵N protein, 500 mL LB medium with kanamycin was inoculated with 2% [*v/v*] of the overnight culture and grown to OD₆₀₀ = 0.8. Subsequently, cells were pelleted, resuspended in 2 L M9 minimal medium supplemented with 1 g/L [¹⁵N] ammonium chloride and/or 4 g/L [¹³C] glucose, and grown to OD₆₀₀ = 0.8. After induction of protein expression by addition of 0.2 mM IPTG, cells were incubated overnight at 30 °C (*TbPar42*-fl), at 25 °C (NterFHA and linker), and at 37 °C (PPIase), followed by centrifugation (6000 rpm, 20 min, 4 °C). The triple labeled ²H-¹³C-¹⁵N-*TbPar42*-fl protein was produced in M9 minimal medium where water was substituted with 99.8% D₂O. The protein expression took 40 h at 30 °C after induction. Pelleted cells were resuspended in PBS buffer (NterFHA) or 50 mM BisTris, 150 mM NaCl pH 6.8 buffer supplemented with 1 mM PMSF for the other constructs, lysed with lysozyme (1 mg/mL) and disrupted by sonication. Upon ultracentrifugation (35,000 rpm, 60 min, 4 °C), the protein was purified by GSH affinity chromatography. PreScission protease digestion was used to cleave the GST fusion prior to size exclusion chromatography in 50 mM potassium phosphate, 150 mM potassium chloride, pH 6.3. For NMR studies, the protein was dialyzed against 50 mM potassium phosphate buffer pH 6.3, 90%/10% (*v/v*) H₂O/D₂O or 100% D₂O, without additional salt.

2.2. Nuclear Magnetic Resonance (NMR) Spectroscopy

All spectra were recorded at 27 °C on a 700 MHz Ultrashield NMR spectrometer (Bruker, Rheinstetten, Germany) equipped with a triple ($^1\text{H}/^{13}\text{C}/^{15}\text{N}$) cryoprobe. The NMR samples contained 200–600 μM of unlabeled or labeled protein, dissolved in potassium phosphate buffer. To the full length *TbPar42* sample, as well as to the linker construct, 2 mM dithiothreitol was added to the NMR sample. For the calibration of proton resonances 2,2-dimethyl-2-silapentane 5-sulfonate was used as an internal standard. ^{15}N and ^{13}C resonances were calibrated according to the IUPAC procedure [12,13]. Spectra needed for the assignment were recorded using standard Bruker library pulse sequences (except for the ^1H - ^{15}N -HSQC-nuclear overhauser enhancement spectroscopy (NOESY)). Spectra were processed with the software Topspin 3.0 (Bruker) and analyzed with CcpNmr analysis [14]. For the assignment of the backbone and aliphatic side chain ^1H , ^{13}C , and ^{15}N resonances, a set of 2D and 3D spectra (^1H - ^{15}N -HSQC, ^1H - ^{13}C -HSQC, ^1H - ^{15}N -TOCSY-HSQC, HNCACB, CBCACONH, HNCA, HN(CO)CA, HNCO, HN(CA)CO) were recorded and analyzed. Aromatic side chain ^1H and ^{13}C chemical shifts were determined using two-dimensional ^1H - ^{13}C -HSQC, ^1H - ^1H -correlation spectroscopy (COSY), TOCSY and NOESY spectra (in H_2O und D_2O), and 3D ^1H - ^{13}C -TOCSY-HSQC and ^1H - ^{13}C -NOESY-HSQC spectra. Interproton distance constraints were obtained from the 2D ^1H -homonuclear NOESY spectrum and the 3D ^1H - ^{15}N -NOESY-HSQC. Hydrogen bond donors were derived from ^1H - ^{15}N -HSQC spectra by proton/deuteron exchange experiments, after the protein was lyophilized and dissolved in 100% D_2O . Acceptors were initially determined from homologue structures and, if applicable, were corrected during the calculation procedure. Hydrogen bonds were set as distance restraints of 1.7–2.6 for H-O and 2.7–3.5 for N-O. Torsion angle restraints were calculated with TALOS using sequence and chemical shift data as input [15]. Structure calculation from a random template structure was performed with the CYANA 2.1 software package including automated NOE assignment, simulated annealing, and torsion angle dynamic algorithm [16]. An ensemble of 60 structures was calculated, and the 10 structures with the lowest CYANA target function value [17] were selected. The secondary structure elements of the *TbPar42*-domains were defined on the basis of NOE connectivities [18], calculated distances, and H-bonding, as well as using chemical shift indices [19–21]. Peptides for NMR studies were purchased from CASLO (Lyngby, Denmark).

2.3. NMR Titration Experiments

To 440 μM ^{15}N -labeled NterFHA or PPIase in 50 mM potassium phosphate buffer, pH 6.26, 90%/10% (*v/v*) H_2O / D_2O , the peptide Suc-Ala-pThr-Pro-Ala-NH₂, dissolved in the same buffer, was added stepwise to a final concentration of 13 mM. After each step a ^1H - ^{15}N -HSQC or SOFAST spectrum was recorded on a 700 MHz Ultrashield NMR spectrometer at 27 °C. The spectra were analyzed with the CcpNmr Analysis software, and the chemical shift perturbations were calculated with following equation [22]:

$$\Delta\delta_{total} = \sqrt{(\Delta\delta_H)^2 + (0.154 \cdot \Delta\delta_N)^2}$$

Residues with chemical shifts ≥ 0.015 ppm (PPIase) or ≥ 0.04 ppm (NterFHA) were represented on the related protein structure using PyMOL [23]. For titrations of the peptide Ala-Glu-Ala-pThr-Glu-Ala-Xaa (Xaa representing Asp, Ile, Val, Ser, or Gln) to the NterFHA domain, the protein concentration was 250 μM and the peptides were titrated to a final concentration of 8 mM.

2.4. EXSY Experiment

EXSY spectra were recorded for a high concentrated peptide probe (Suc-Ala-pThr-Pro-Ala-NH₂) of 2.4 mM in 25 mM potassium phosphate buffer, pH 6.26, 100% D_2O at 27 °C on a 700 MHz Ultrashield NMR spectrometer, in the presence and absence of 50 μM *hPin*-PPIase or *TbPar42*-PPIase. The pulse

sequence used for the spectra recording was from the standard Bruker library, noesygp19 with a mixing time of 400 ms, a relaxation delay of 1 s, and 16 scans.

2.5. Relaxation Measurements

The longitudinal (R_1) and transverse relaxation (R_2) rates were estimated by signal intensity measurements from a series of ^1H - ^{15}N -HSQC spectra with varying relaxation delays. Ten different delay times were used to determine the R_1 rate: 20, 100, 140, 180, 200, 300, 500, 750, 1000, and 1400 ms, whereas nine evolution times 10, 30, 50, 70, 90, 110, 130, 170, and 190 were used to estimate R_2 . R_1 and R_2 correspond to the reciprocal relaxation times T_1 and T_2 , which were calculated from intensity measurements for each residue according to Reference [24]. The ^1H - ^{15}N -NOE for each amino acid was determined from the ratio of signal intensities of a saturated and an unsaturated spectrum. The two spectra were extracted out of a simultaneously recorded spectrum, recorded with a modified hsqcnoef3gpsi pulse sequence. All measurements were recorded at 27 °C on a 700 MHz Ultrashield NMR spectrometer.

2.6. Sample Preparation for *TbPar42*-PPIase Crystallization Experiments

For crystallization, *TbPar42*-PPIase (264–383) was purified by size exclusion chromatography using a Superdex_{TM} 75 16/60 pg (GE Healthcare, Freiburg, Germany) column equilibrated with 50 mM Tris/HCl, pH 7.0, 150 mM NaCl, 1 mM DTT. The fractions containing pure *TbPar42*-PPIase were pooled and concentrated to 15–20 mg/mL. Crystallization conditions were screened by the hanging drop vapor diffusion method, mixing 2 μL reservoir solution with 2 μL protein solution incubated at 20 °C. Crystals used for structure determination grew at 20 °C in drop, with 23% PEG 4000, 100 mM sodium acetate as the reservoir. Prior to freezing in liquid nitrogen, crystals were soaked in 50% PEG 4000 for cryoprotection.

2.7. Data Collection and Structure Determination

Data sets were collected at 100 K on beamline P13 at PETRAIII of the Deutsches Elektronen Synchrotron (Hamburg, Germany) using a Pilatus 6M detector. The data were processed using the XDS suite and converted into appropriate formats by XDSCONV [25,26]. The *TbPPIase* structure was solved by molecular replacement, using human Pin1 as a search model (PDB: 1Pin, residues 50–163) in MOLREP [27]. Refinement was done with REFMAC [28] and PHENIX [29]. For manual model building, the molecular graphics program COOT was employed [30]. The final model was deposited in the PDB with the accession code 6GMP.

2.8. PPIase Activity

A protease coupled isomerase assay [31,32] was used to measure the activity of the PPIase for peptides with a Suc-Ala-Xaa-Pro-Phe-pNA scaffold (pNA-paranitroaniline; Xaa-represents pSer, pThr, Arg, Lys, Glu, Asp, Phe, Gln, or Ser), synthesized by CASLO or ChinaPeptides. The peptides were dissolved in 0.5 M LiCl in 2,2,2-trifluoroethanol to a concentration of 15 mM, followed by overnight incubation. The protease α -chymotrypsin (35 μM) was pre-incubated with 200 nM or 10 μM *TbPar42*-PPIase (as positive control 20 nM *hPin1*-PPIase was used) for 5 min in PBS buffer, pH 6.8 at 10 °C. For the mutants, 2 μM and 10 μM protein were used. Directly after the addition of 75 μM peptide, the assay was followed by monitoring the absorbance of the cleaved pNA at 390 nm. The reaction rate constant of the protease and the cis-isomer content of the peptide were determined from the uncatalyzed reaction for each substrate. The observed curves were analyzed with GraphPadPrism 5.04, using a bi-exponential reaction fit.

3. Results

3.1. NMR Measurements of Full-Length *TbPar42* (1–383)

Due to the high molecular weight of *TbPar42* (1–383), structural investigations of the protein by NMR spectroscopy were only possible using ^2H - ^{13}C - ^{15}N isotopically labeled protein harvested from *E. coli* grown in deuterium oxide. Although nearly 96% of all expected amide resonances were present in the ^1H - ^{15}N -total correlation spectroscopy (TROSY) spectra (Figure 1A), the low protein yield upon recombinant production hampered a complete structural analysis. We therefore decided to dissect the full-length protein in separated domains, which were more easily amenable for NMR spectroscopy. These domains were primarily defined using sequence alignment and similarity searches (BLAST). After several rounds of trials and optimization (expression values, NMR spectra appearance), three feasible domain constructs of *TbPar42* (the N-terminal extended FHA domain (1–177; NterFHA), the putative linker region (172–266; Loop_FHA_PPIase), and the putative catalytic domain (264–383; *TbPar42*-PPIase) were separately cloned, produced, and further analyzed (Figure 1B).

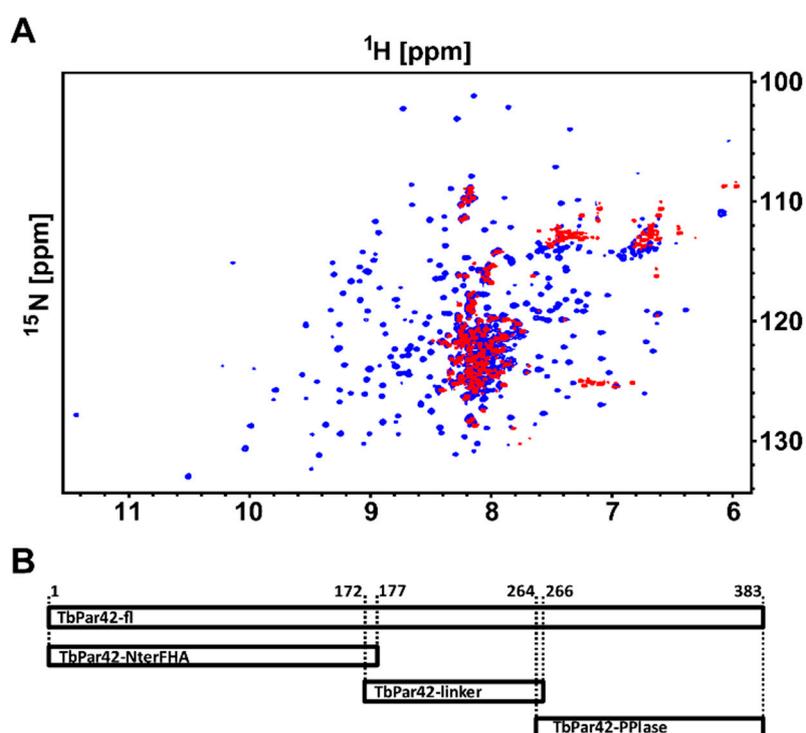


Figure 1. Nuclear magnetic resonance (NMR) spectroscopy and construct design of *TbPar42*. (A) Superposition of a ^1H - ^{15}N -total correlation spectroscopy (TROSY) of ^2H - ^{13}C - ^{15}N -*TbPar42* (blue) and a ^1H - ^{15}N -hetero single quantum coherence (HSQC) of ^{15}N -*TbPar42* (red). (B) Design of expression constructs (horizontal bars) for full length protein (*TbPar42*-fl), N-terminus with forkhead association (FHA) domain (*TbPar42*-NterFHA) and PPIase domain (*TbPar42*-PPIase). The corresponding residue numbers are given at the top of Figure 1B.

3.2. NMR Structure Calculation of *TbPar42*-PPIase

60 structures of the *TbPar42*-PPIase domain were calculated using 3023 distance NOEs, 206 torsion angle restraints, and 46 hydrogen bonds. A statistical analysis of the NMR ensemble (PDB-ID: 2N87; BMRB-ID: 19904) is shown in Table S1. Ten structures with the lowest energy conformations are presented in Figure 2A.

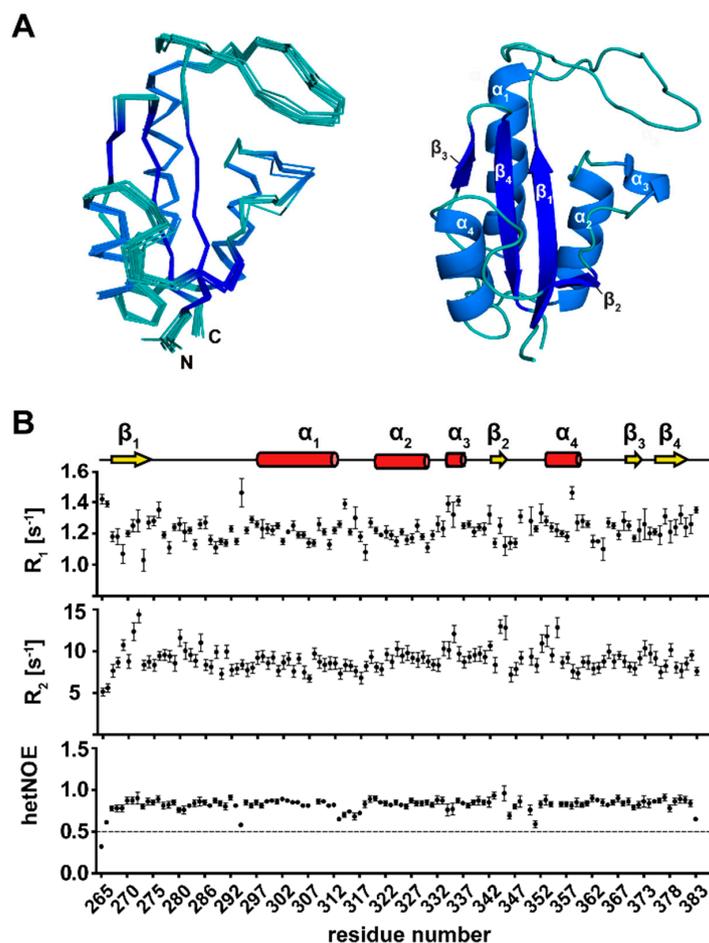


Figure 2. Structural and dynamical data of the PPIase domain of *TbPar42*. (A) Ensemble of the 10 lowest energy structures (left) and ribbon representation of *TbPar42*-PPIase (right). The four α -helices and four β -sheets of the parvulin fold are indicated. (B) Dynamics of the PPIase domain. The secondary structure elements are indicated at the top of the plot. Below, the longitudinal (R_1) and transversal (R_2) relaxation rates, as well as the HetNOE of the amino group of residues, within the domain are plotted.

3.3. Secondary Structure Elements of *TbPar42*-PPIase

The PPIase is buildup of four α -helices and four β -strands (Figure 2A,B). The first N-terminal β -strand comprises amino acids from Arg267 to Val274, and is followed by a long loop, which extends from Lys275 to Arg296. The adjacent α -helix 1 (Ser297 to His312) is separated through a small stretch of residues from α -helix 2 (Leu320–Phe330). Within the following thirteen residues, α -helix 3, which comprises only four residues (Gly334–Lys337), was identified by chemical shift data. Due to a lack of a typical helix-NOE pattern and fast exchanging amide protons in this helix, it is likely to be considered as α -helical turn. The second β -strand from Gly343 to Val345 is connected to α -helix 4 (Gly353 to Phe359) by a short loop. β -strand 3 is formed from Val370 to Thr372, and is followed by a short loop that turns into the fourth β -strand (Gly375–Glu381). Analysis of ^{15}N -relaxation data reflecting the motion of the backbone N-H vector (Figure 2B) indicated that all secondary structure elements are rigid, and even the extended PPIase loop regions (β_1 - α_1 and β_2 - α_4) lack significant internal mobility.

3.4. The Tertiary Structure of *TbPar42*-PPIase Comprises the Typical Parvulin Fold

The PPIase adopts a globular fold, which is characteristic for parvulin type proteins [33]. The central β -sheet, composed of four antiparallel β -strands (β_3 - β_4 - β_1 - β_2), is surrounded by four α -helices (Figure 2A). The dipole axes of these helices align parallel with the C–N orientation vector of the β -sheet. According to the DALI program, the solution structure of the PPIase domain of human

parvulin *hPin1* (PDB-ID: 1NMW), an isomerase acting on phosphorylated substrates, has the highest similarity to the *TbPar42*-PPIase (Z-score of 13.9; C α r.m.s.d: 1.58 Å).

A comparison of both structures (Figure 3A,B) indicates that well-conserved residues of Pin1-like parvulins, e.g., the histidine motif (His59 and His157 in *hPin1*; His271 and His377 in *TbPar42*), as well as a characteristic cysteine within the catalytic cleft (Cys113 in *hPin1*; Cys333 in *TbPar42*), are located at the same spatial positions within both folds. The *hPin1* Ser154 (Figure 3A), described as a part of the catalytic tetrad [34] is replaced by Leu374 in *TbPar42* (Figure 3B). In the catalytically active but non-phosphorylation specific parvulin *hPar14*, a phenylalanine (Phe120) can be found at this position (Figure 3C). Thr372, the counterpart of which in *Par14* was associated with the catalytic network [35], is also conserved in the *TbPar42*-PPIase structure. Residues Leu122, Met130, and Phe134 of *Pin1*, which are located at the concave side of the β -sheet core, constitute a hydrophobic binding pocket for the substrate's proline ring moiety. At the same spatial positions, Leu342 and Phe354 can be found in the *TbPar42*-PPIase structure, whereas the methionine is exchanged by a tyrosine (Tyr350), Figure 3D. Phosphorylation-specific parvulins exhibit an extended α 1- β 1 loop, with a basic amino acid cluster important for the binding of the phosphate group of substrates. A homologous loop is also present in the structure of *TbPar42*-PPIase (Figure 3E).

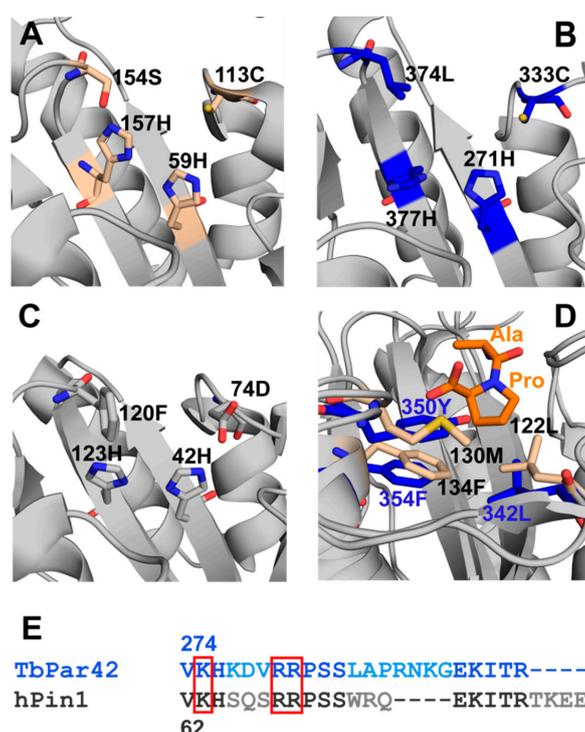


Figure 3. Structural insights into the PPIase domain of *TbPar42*. Comparison of the catalytic tetrad (prominent residues are represented as sticks) of (A) *hPin1*, PDB:1NMW, (B) *TbPar42*, PDB:2N87, and (C) *hPar14*, PDB:3UI4. (D) Overlay of the solution structure of *TbPar42*-PPIase (PDB:2N87) and the crystal structure of *hPin1* (PDB:1PIN), with its bound Ala-Pro dipeptide in orange. Hydrophobic residues of the proline-binding pocket are represented as sticks in blue for *TbPar42*, and in beige for *hPin1*. (E) Sequence alignment of the β 1- α 1 loop between *TbPar42* (blue) and *hPin1* (black). The phosphate group binding residues of *hPin1* are framed (red). A full alignment of both PPIase domains is found in Figure S3.

3.5. *TbPar42*-PPIase Lacks Catalytic PPIase Activity

Although the PPIase domain of *TbPar42* adopts a parvulin fold and reflects structural elements of *Pin1*-type proteins [1], it is catalytically inactive.

The protein failed to compensate the loss of ESS1 function in yeast (as *hPin1* does), and also lacked isomerase activity against a pThr containing substrate, which was used to verify the activities of *TbPin1* and *AtPin1* [1]. We confirmed this lack of activity of *TbPar42*-PPIase by exchange spectroscopy (EXSY) experiments using an alternative substrate (Figure 4A,B), and by isomerase activity studies using a protease coupled assay (Figure 4C). In the latter experiment, a considerable ensemble of substrates containing pSer or pThr residues preceding proline were used to track activity. In addition, peptides containing negatively charged (Glu, Asp) positively charged (Arg), or uncharged (Ser, Gln, Phe) residues at the Xaa-Pro motif were assayed to test a putative sequence dependency of *TbPar42*'s substrate activity for non-phosphorylated motifs (Figure 4C(right)). However, no *cis/trans*-activity could be detected against any of these substrates. From all the current data, we may conclude that the PPIase domain of *TbPar42* seems to have no enzymatic function within the protein. Although *TbPar42* seems to be catalytically inactive, titration experiments with a Suc-Ala-pThr-Pro-Ala-NH₂ peptide showed chemical shift perturbations in the 2D ¹H-¹⁵N hetero-single-quantum-coherence (HSQC) spectra of the PPIase domain, Figure 4D. Thus, the substrate was capable of binding to the enzyme. Residues affected by peptide addition were in near proximity to the proposed proline binding site (Thr265, Gly334, Val345, Thr349, Glu352, Gly353, Phe359, Glu373), but K_D values were in the millimolar range (~10 mM).

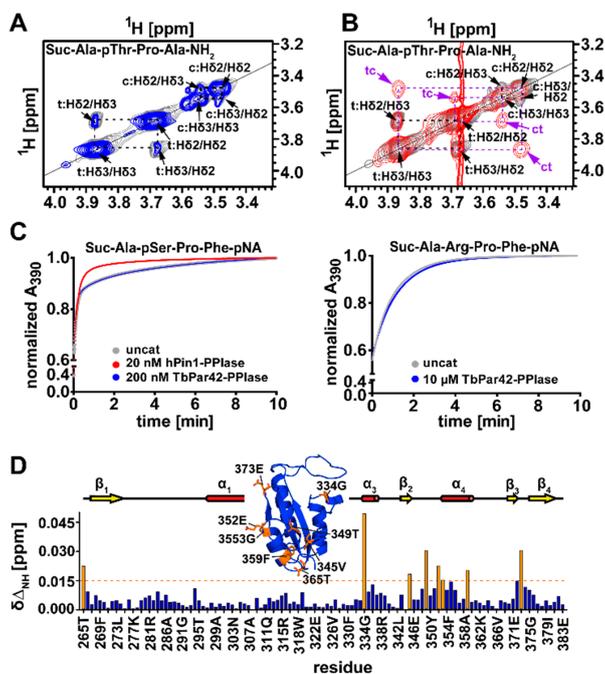


Figure 4. Catalytic activity and substrate binding of the PPIase domain. Exchange spectroscopy (EXSY) spectra of the proline-H δ protons of Suc-Ala-pThr-Pro-Ala-NH₂ peptide. (A) in the presence of *TbPar42*-PPIase; (B) in the presence of *hPin1*. The *cis* and *trans* conformers are indicated by c and t, and the resulting exchange peaks generated exclusively in B during isomerization are represented in purple by ct and tc. (C) Absorption curves monitoring the time dependent cleavage of pNA-containing model peptides (left: Suc-Ala-pSer-Pro-Phe-pNA (*hPin1*) and right: Suc-Ala-Arg-Pro-Phe-pNA (*TbPar42*-PPIase)) in a protease-coupled isomerase assay in the presence of *TbPar42*-PPIase (blue) or *hPin1* (red). The gray curve belongs to the uncatalyzed reaction without enzyme. A steep initial slope describes a fast isomerization. (D) Diagram of NH chemical shift changes ($\Delta\delta_{NH}$) observed in an ¹⁵N-HSQC spectrum of *TbPar42*-PPIase versus amino acid sequence after addition of 13 mM Suc-Ala-pThr-Pro-Ala-NH₂ peptide. Residues with a chemical shift larger than 0.015 ppm (orange dash line) are colored orange and mapped to the structure of the *TbPar42*-PPIase domain. The secondary structure of *TbPar42*-PPIase with helix represented in red, β -strand in yellow, and coil as line is shown on top of the diagram.

3.6. What Is the Structural Origin of the Lack of Catalytic Activity of *TbPar42*-PPIase?

The structural fold and the spatial organization of amino acids in the putative catalytic center of *TbPar42*-PPIase resemble the typical active side of a Pin1-type parvulin isomerase. However, despite the fact that binding of the model peptide Suc-Ala-pThr-Pro-Ala-NH₂ seems to occur at the catalytic cleft close to the proline binding pocket (Figure 4D), no isomerase activity could be detected under the tested conditions (Figure 4A,C). Certainly, the dissociation constant of the peptide was only in the millimolar range, but this is no sufficient argument to explain the lack of activity, as comparable weak binding affinities have been pointed out for active parvulins and their substrates [36,37]. Thus, insufficient binding of substrates may not be the origin of catalytic inactivity.

More reasonably, the lack of activity may originate from structural alterations within the catalytic cleft or within the phosphoryl group binding loop when compared to active orthologous or homologous PPIases [38,39]. Goh et al. mentioned that the conservation of residues in the β 2- α 4 region of *TbPar42*, which is part of the substrate-binding pocket, is low when compared to homologous regions of other parvulins. According to our structural model of *TbPar42*-PPIase, there are two significant alterations of residues in the catalytic core region of the protein when compared to *hPin1*. Position 350 of the proline binding pocket is taken by a tyrosine rather than a methionine, and position 374 of the catalytic site is captured by a leucine (serine in *hPin1*). To examine the role of these residues on the activity of the PPIase, we mutated these positions in *TbPar42* (*TbPar42*-PPIase_{Tyr350Met}, *TbPar42*-PPIase_{Leu374Ser}) and *hPin1* (*hPin1*_{Met130Tyr}, *hPin1*_{Ser154Leu}), and tested the catalytic efficacy of the mutants in a protein coupled assay (Figure 5). As *hPin1*_{Met130Tyr} was still active (Figure 5C), this mutation is not expected to be responsible for the lack of activity in *TbPar42*-PPIase. This was in accordance with a still inactive *TbPar42*-PPIase_{Tyr350Met} mutant (Figure 5A). As expected [35], the mutant *hPin1*_{Ser154Leu} showed no catalytic activity (Figure 5D). However, no gain-of-function for the *TbPar42*-PPIase_{Leu374Ser} mutant was observed.

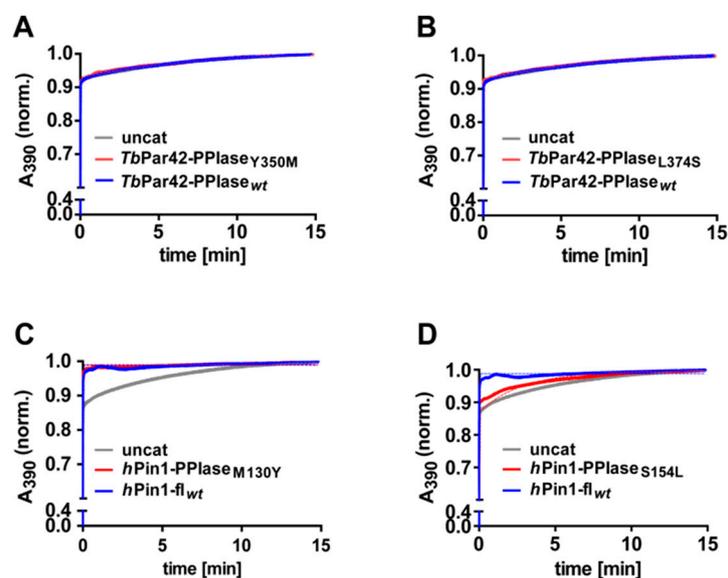


Figure 5. PPIase assays using 75 μ M of the peptide Suc-Ala-pThr-Pro-Phe-pNA in the presence and absence of different PPIase constructs. Absorption curves of uncatalyzed reactions are highlighted in gray color. Catalyzed reactions are shown (A) in presence of 10 μ M *TbPar42*-PPIase_{Y350M} (red) and 10 μ M *TbPar42*-PPIase_{wt} (blue); (B) in presence of 10 μ M *TbPar42*-PPIase_{L374S} (red) and 10 μ M *TbPar42*-PPIase_{wt} (blue); (C) in presence of 10 μ M *hPin1*-PPIase_{M130Y} (red) and 2 μ M full length (fl) *hPin1*-fl_{wt} (blue); (D) in presence of 10 μ M *hPin1*-PPIase_{S154L} (red) and 2 μ M full length (fl) *hPin1*-fl_{wt} (blue).

Beside residues in the catalytic cleft and proline binding pocket, three basic amino acids (basic triad) residing in the $\beta 1$ - $\alpha 1$ -loop in Pin-type parvulins are functionally important in *hPin1*. By binding to the phosphoryl group of the substrate, these residues serve as anchor points for the bond rotation occurring in the isomerization step (reviewed in Reference [9]). Thus, not only absence of this basic triad, but changes in their arrangement and its relation to the spatial organization of residues in the catalytic center or in the proline-binding pocket may abolish catalytic activity. In *TbPar42*, the triad of these putative phosphoryl-binding residues Lys275, Arg280 and Arg281 is conserved with respect to sequence position. Despite this fact, and in contrast to *TbPin1* [40] and other Pin1-type parvulins, no significant chemical shift perturbations were detected in a HSQC spectrum for resonances of the basic triad of the $\beta 1$ - $\alpha 1$ phosphoryl group binding loop of *TbPar42*-PPIase (Figure 4) upon addition of a phosphorylated peptide. This indicates the absence of an interaction of the triad with the phosphate group of the peptide. We wondered if structural differences in the arrangement of this loop with relation to the catalytic active site residues would prevent binding of the phosphoryl group. Goh and coworkers suggested that an insertion in the $\beta 1$ - $\alpha 1$ region of the protein may affect catalytic activity. To uncover discrepancies in the spatial arrangement, we compared the overall fold of *TbPar42*-PPIase to those of the known phosphate-binding relatives *TbPin* and *hPin1*.

As we wanted to focus on the structural orientation of the side chains of the basic triad and only few NMR constraints (NOEs) exist for the side chain atoms of these residues, we complemented our structural data by crystalizing the *TbPar42*-PPIase and resolving its structure using X-ray analysis (PDB-code: 6GMP; Table S2) (see also Figures S4 and S5). After superimposition of the catalytic core residues (Figure 6), the spatial orientation of the proteins does not differ with respect to position and orientation. This excludes that the arrangement of phosphoryl group binding residues is responsible for the absence of activity within *TbPar42*. Thus, the structural reason for the absence of activity remains unknown. To our knowledge, no parvulins from higher eukaryotic organisms have been discovered so far, of which the PPIase domains lack catalytic activity. However, the bacterial PPIase domains of the periplasmic chaperones PpID [41], as well as one PPIase domain of SurA (Par1) [42], are also lacking catalytic activity. Similar to the PPIase domain of *TbPar42*, the isomerase activity of PpID “could not be generated by substitutions at the peptide binding site” [41]. In case of SurA Par1, the domain was found to cooperate with the second domain Par2 in substrate folding reactions.

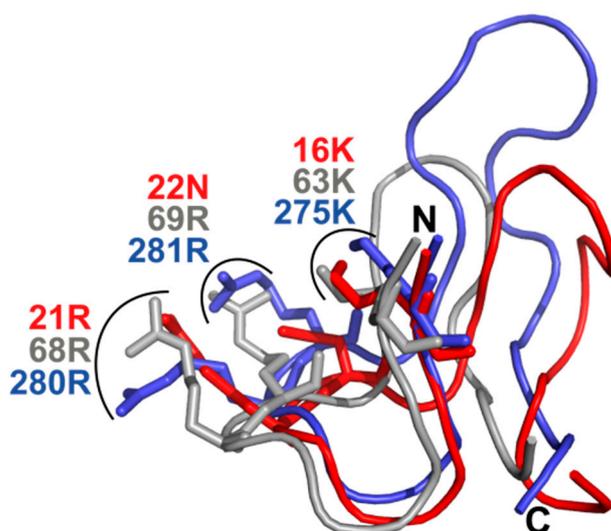


Figure 6. Spatial orientation of the $\beta 1$ - $\alpha 1$ loop regions when superimposing the catalytic core residues of *TbPar42* (H271, C333, L374, H377), *TbPin1* (H11, C68, S109, H112) and *hPin1* (H59, C113, S154, H157); *TbPar42* (PDB:6GMP) (blue), *hPin1* (PDB:1NMW) (gray), and *TbPin1* (PDB:2LJ4) (red).

3.7. NMR Structure Calculation of the N-Terminal Extended FHA Domain

The solution structure of the NterFHA-domain of *TbPar42* was calculated using 2667 distance NOEs, 262 torsion angle restraints, and 42 hydrogen bonds (Table S3). An overlay of the 10 final lowest energy structures is shown in Figure 7A. The ensemble is deposited in the Protein Data Bank (PDB code: 2N84), while chemical assignments and shifts were transferred to the BMRB database with the accession number 25834. Although all 10 structures of the N-terminal extended FHA domain (1–177) fit the experimental restraints, the r.m.s.d value for the backbone atoms of the ensemble was 12.5 Å. It dropped down to 0.33 Å when the flexible terminal residues (1–33) (Figure 7B) and the adjacent three helices α_1 (Glu33 to Asn37), α_2 (Ile42 to Val45), and α_3 (Pro84–Tyr57) were excluded from calculation, indicating an excellent structural resolution of the FHA-domain (residues 58–173). Including α_3 in the calculations increased the r.m.s.d value to 0.8 Å for the backbone atoms of the ensemble, while the r.m.s.d value increased to 1.62 Å when additional α_1 and α_2 were included (Table S3).

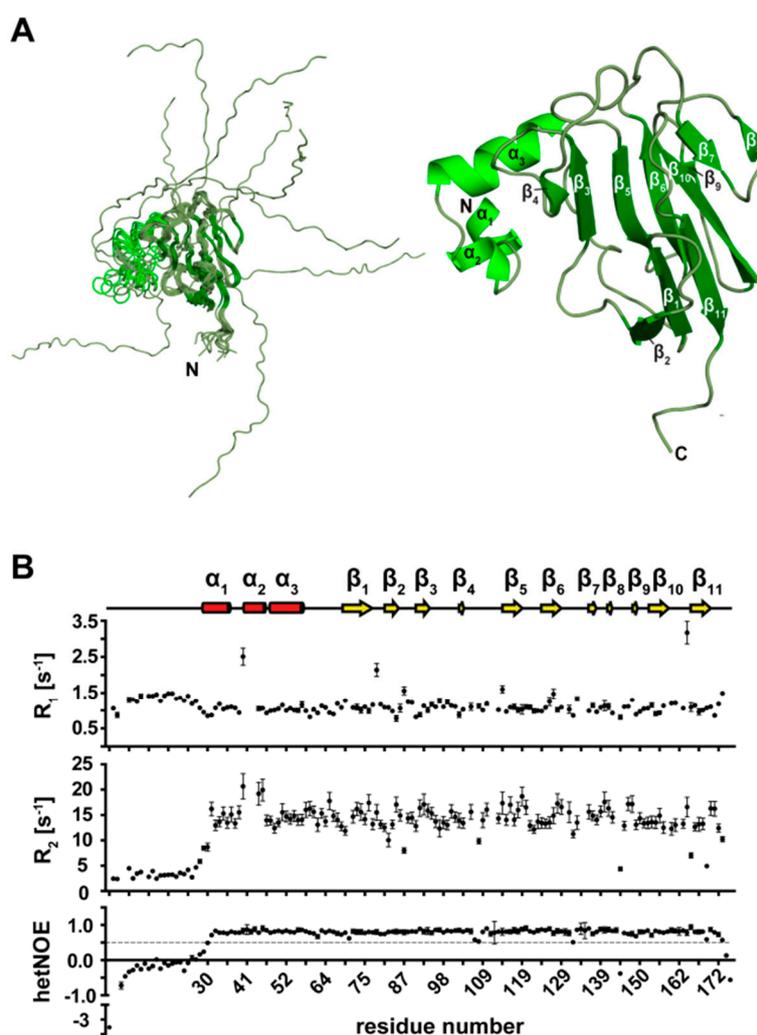


Figure 7. Structural and dynamic data of the FHA domain of *TbPar42*. (A) Ensemble of the 10 lowest energy structures of the N-terminal extended FHA domain (1–177) (left) and ribbon representation of *TbPar42*-FHA domain without flexible terminal residues 1–33 (right). (B) Dynamics of the NterFHA construct. The secondary structure elements are indicated at the top of the plot. Below, the longitudinal (R_1) and transversal (R_2) relaxation rates, as well as the HetNOE of the amino group of residues within the domain, are plotted.

3.8. Secondary Structure Elements of the N-Terminal Extended FHA Domain (1–177)

The FHA domain of *TbPar42* was analyzed together with the N-terminus of the protein. The low number of long-range NOEs per residue, as well as low R2 amide relaxation values and negative hetNOE data of the first 30 residues within the N-terminus, point towards a random coil conformation of this segment (Figure 7B). Based on distances and typical NOE connectivities, three well-characterized helices $\alpha 1$ to $\alpha 3$ (Glu33–Asn37, Ile42–Val45, Pro48–Tyr57) adjacent to the unstructured region were detected. Via a long loop (Phe58–Ala70), helix $\alpha 3$ is connected to an array of 11 sequentially arranged β -strands. The first $\beta 1$ strand extends from Cys71 to Arg77. The two following strands, $\beta 1$ and $\beta 2$, are both four residues long (Leu82 to Gly86 and Phe92–Gly96), whereas $\beta 4$ covers just two residues (Tyr103 to Val104). $\beta 5$ starts at Ala115 and extends to His120. $\beta 6$ (Cys125–Asp130) is connected to $\beta 5$ via a turn (Gly121–Arg124). The chemical shift index indicates a helix between strands $\beta 6$ and $\beta 7$ (Val137–Leu139), but, due to the fast exchanging amide protons and the missing helical NOE connections, the stretch (Leu131–Gly136) corresponds rather to a turn region. Hydrogen bonds to the other strands helped to identify β -strands $\beta 8$ (Asn142–Arg143) and $\beta 9$ (Pro149–Leu15). The last two strands extend from Gly155 to Phe160 ($\beta 10$) and from Val166 to Leu171 ($\beta 11$). Residues (Gly172–Ser177) following the last β -strand exhibit random coil character.

3.9. Tertiary Structure of *TbPar42-NterFHA*_{1–177}

Like other known FHA structures, the FHA of *TbPar42* adopts the typical β -sandwich fold, which is formed by two twisted β -sheets consisting of five ($\beta 2$ – $\beta 1$ – $\beta 11$ – $\beta 10$ – $\beta 9$) and six ($\beta 4$ – $\beta 3$ – $\beta 5$ – $\beta 6$ – $\beta 7$ – $\beta 8$) β -strands, respectively, Figure 3B. In both β -sheets, the adjoining β -strands are oriented antiparallel to each other, with the exception of $\beta 4$, which proceeds parallel to the neighboring $\beta 3$. Due to ^{15}N relaxation data, the loops connecting β -strands 4/5 and 6/7 exhibit a higher flexibility compared to the other loops or turns within the FHA structure. Except for the 30 N-terminal residues and a few residues at the very C-terminal end, the structure has an overall rigid character. In contrast to the compact β -stranded fold, there are only few long-range NOE restraints emerging from the helical part of the protein (Glu33 to Tyr57). Therein, Gln54 of $\alpha 3$ exhibits more than 40 NOE connectivities (e.g., to Phe93 of $\beta 3$ and to residues Val100, Cys101 and Asp102 within the $\beta 3$ – $\beta 4$ loop). According to its molecular interactions, $\alpha 3$ is tightly attached to the β -stranded FHA core, in contrast to $\alpha 1$ and $\alpha 2$, where such NOE contacts are absent.

3.10. *TbPar42-NterFHA*_{1–177} binds pThr-Pro Moieties within Peptides

As FHA domains are known as phospho-threonine recognition and binding modules [5], we examined the affinity of the ^{15}N labeled NterFHA to the phosphorylated model peptide Suc-Ala-pThr-Pro-Ala-NH₂ by performing NMR titration experiments. Within the recorded 2D ^1H - ^{15}N HSQC spectrum, the presence of peptide affected the chemical shifts of resonances from residues of several loop-regions of the FHA core, Figure 8A. Perturbed chemical shifts were observed for residues in the $\beta 6$ – $\beta 7$ loop (Ser133–Gly136), for isoleucine 144 in the $\beta 8$ – $\beta 9$ loop, serine 112 and surrounding residues in the $\beta 4$ – $\beta 5$ loop (Glu106, His107, Ile110–His114), for arginine 97 and serine 98 in the $\beta 3$ – $\beta 4$ loop, and for serine 163 in the $\beta 10$ – $\beta 11$ loop, Figure 3D. Structural and sequential alignments with other FHA domains indicate that perturbation mainly affects those resonances which are either essential for direct binding to phosphopeptide ligands, or for the maintenance of the binding surface (Figure 8B). Arginine 97 from the $\beta 3$ – $\beta 4$ loop as well as serine 112 from the $\beta 4$ – $\beta 5$ loop are expected to form a hydrogen bond network with the phosphate group of the ligand's pThr moiety. The conserved glycine 96 and histidine 114 at the end of $\beta 3$ and adjacent to strand 5 are most likely crucial for the structural stability of the binding site [43]. A significant chemical shift perturbation could also be observed for Asn135 and surrounding residues in the $\beta 6$ – $\beta 7$ loop. In some FHA domains, like in Kanadaplin [44], this position is substituted by a histidine residue. The asparagine/histidine is important for interactions to residues flanking the pThr in the ligand [45]. Although the *TbPar42*-FHA

binds the above mentioned phosphorylated model peptide, binding affinity is in the lower millimolar range (~4 mM).

Tight binding of FHA domains to phosphorylated motifs greatly depends on the ligand's residues flanking the phosphorylated threonine, particularly on the third position (pT+3) following pThr [43]. According to this so called “pT+3” rule, FHA domains can be divided in pTxxD or pTxxI/V (x representing any amino acid) binding modules. To test if *TbPar42*-FHA has a preference for a certain residue at this position, we generated synthetic peptides and tested their binding affinity towards the FHA domain by performing NMR titration experiments. As the proposed peptide-binding surface of the FHA is positively charged (Figure 8C, we used the sequence Ala-Glu-Ala-pThr-Glu-Ala-Xaa as the peptide scaffold and generated derivatives by altering position Xaa (Asp, Ile, Val, Ser, or Gln). Addition of these peptides to the protein perturbs NMR resonances of residues located in the same regions as observed on binding of Suc-Ala-pThr-Pro-Ala-NH₂, but, in addition, induces chemical shift changes within resonances of amino acids in the β10-β11 loop region. However, no significant improvement in binding affinity compared to the Suc-Ala-pThr-Pro-Ala-NH₂ peptide could be measured (Table S4).

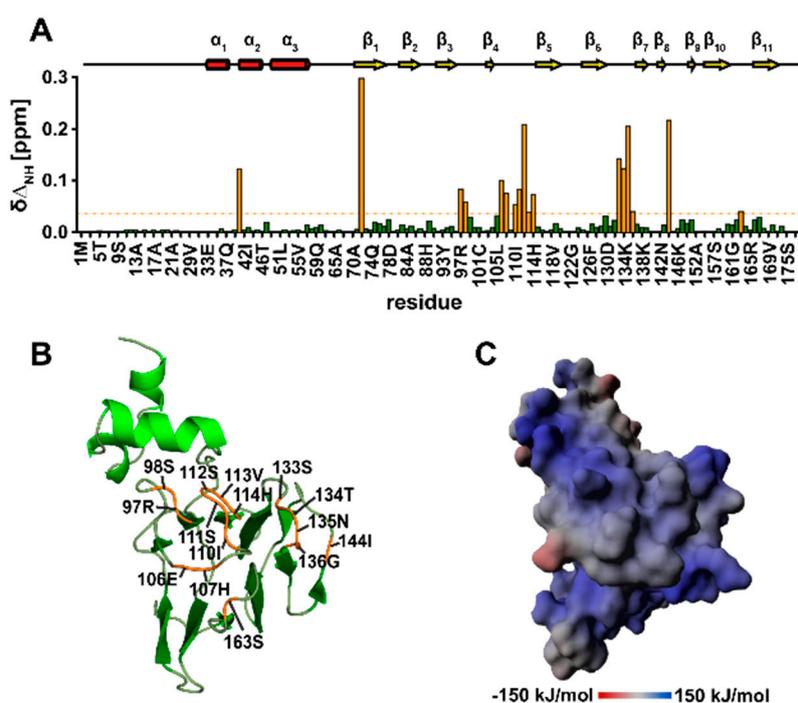


Figure 8. The phosphorylated peptide Suc-Ala-pThr-Pro-Ala-NH₂ is recognized by conserved phosphate binding loops of *TbPar42*-FHA. (A) (top) Secondary structure elements of NterFHA. (bottom) Diagram showing NH chemical shift changes ($\Delta\delta_{\text{NH}}$) versus the amino acid sequence of *TbPar42*_NterFHA₁₋₁₇₇ domain after addition of 13 mM peptide. Chemical shift perturbations observed for His41 and His72 may reflect tiny changes in pKs values, rather than indicating strong effects upon peptide addition [35]. Residues with a chemical shift larger than 0.04 ppm (orange dash line) are colored orange. (B) Residues of *TbPar42*-FHA₃₃₋₁₇₇ showing chemical shift changes larger than 0.04 ppm in the presence of phosphorylated peptide are indicated and are orange colored. (C) Electrostatic surface of the NterFHA₃₃₋₁₇₇ generated with YASARA using the Particle Mesh Ewald method. The intensity of the electrostatic potential is gradually colored from dark red (negative) over grey (neutral) to dark blue (positive), representing energy levels from -150 to $+150$ kJ/mol. The orientation of the structure is as in (B).

3.11. The Isolated Linker Region between the NterFHA and PPIase Domain Is Flexible and Unfolded

As we were able to completely assign the amide resonances in the ¹H-¹⁵N-HSQC spectra of the NterFHA₁₋₁₇₇ (Figure 9B) and PPIase₂₆₄₋₃₈₃ (Figure 9A) domains, we expressed and purified the

^{13}C - ^{15}N -labeled linker region (residues 172–266) connecting the two domains. The signal dispersion in the ^1H - ^{15}N -HSQC spectrum, as well as the chemical shift index analysis of the linker, indicate that the region is mostly random coil (Figure 9C). The backbone amide groups of the N-terminal stretch Pro174–Lys178 of the linker are absent in the ^1H - ^{15}N -HSQC spectrum, resulting in 96.6% assignment completeness. NMR shift data indicate a long α -helix extending from Arg179 to Val197 (Figure S1). Upon calculation of a Rosetta model (Figure S2) [46], five putative transient helices are predicted within this linker regions.

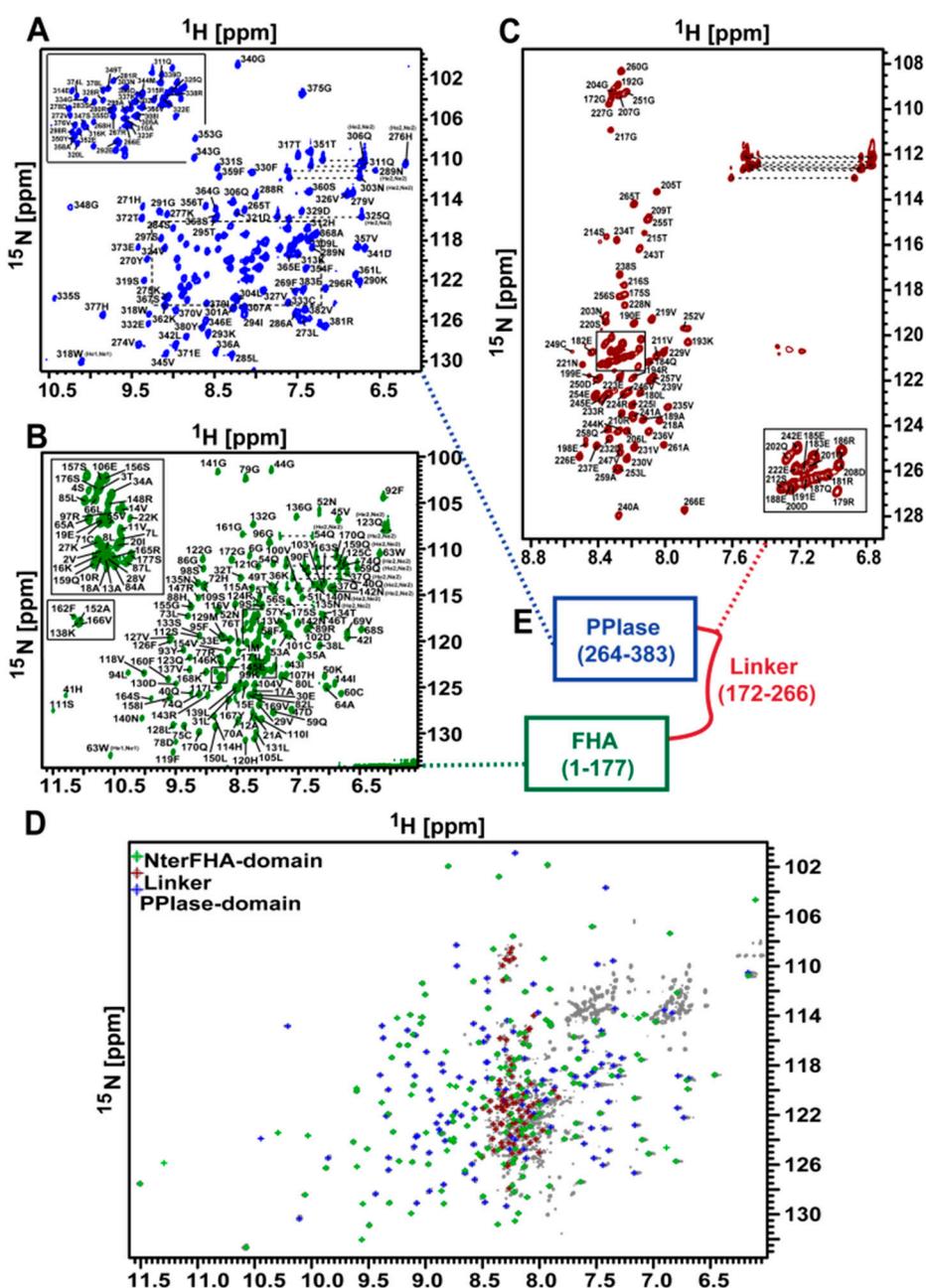


Figure 9. The domains of *TbPar42* do not interact with each other (A) Fully assigned ^1H - ^{15}N -HSQC spectrum of the PPIase domain₂₆₄₋₃₈₃. (B) Fully assigned ^1H - ^{15}N -HSQC spectrum of the N-terminal extended FHA domain₁₋₁₇₇. (C) 96.6% assigned ^1H - ^{15}N -HSQC spectrum of the linker₁₇₂₋₂₆₆ region. (D) Overlay of the *TbPar42* full-length TROSY spectrum with the HSQC spectra of the separately analyzed domains and the linker of *TbPar42*. (E) Scheme of domain organization.

3.12. There Is No Interaction between the Domains of *TbPar42*

Due to the high molecular weight of the protein, we analyzed the domains of *TbPar42* as isolated constructs. However, the structures and activities of these isolated domains may be modulated by intra-domain interactions within the protein. To investigate such putative modulations, the 2D-TROSY spectrum of triple labeled ^2H - ^{13}C - ^{15}N -*TbPar42* protein was superimposed with the assigned ^1H - ^{15}N -HSQC spectra of the isolated constructs recorded under identical conditions, Figure 9D. Interaction of domains should lead to significant chemical shift changes in the full-length protein when compared to its isolated domains. However, beside a few terminal residues of the isolated constructs, 353 signals from all three HSQC spectra perfectly matched to the TROSY spectrum, indicating that there are no interactions of the individual domains nor conformational rearrangements within the linker region under the tested buffer conditions.

4. Discussion

4.1. What is the Cellular Role of the FHA Domain of *TbPar42*?

*TbPar42*₁₋₁₇₇ resembles the structural fold of a typical FHA domain. Multi-domain proteins where a FHA domain precedes a parvulin PPIase domain are unique to unicellular organisms (e.g., *Trypanosoma* and *Leishmania* [2], as well as in *Chlamydomonas* [3] and *Dyctiostelium* [2]). The FHA domain of *TbPar42* seems to be involved in the binding of phosphorylated target proteins, as could be proven by NMR titration experiments with model ligand peptides. The dissociation constants of the tested peptides, however, were only in the millimolar range, demonstrating most likely that the optimal binding sequence for this protein domain remains to be elucidated. Nevertheless, binding occurred in regions found to interact with target proteins in homologous and related FHA structures (Figure 8A,B). As the *TbPar42* protein is unique to parasitic protozoa organisms and no interaction partners have been published so far, we can only speculate about its functional role by relying on similar and orthologous domains and proteins in other organisms.

In order to find orthologous proteins for the FHA domain of *TbPar42*, we ran a DALI-search that depicted the highest Z-score (13 and 12.2), as well as the highest sequence identity (37%), to the FHA domains of the proteins Kanadapin (human) and DAWDLE (DDL; *Arabidopsis thaliana*). Both representatives are nuclear-localized proteins. While almost nothing is known about the function of Kanadapin, DDL appears to act in multiple developmental processes such as growth of root and shoot, as well as in floral morphogenesis and fertility [47]. Dawdle was reported to stabilize the interaction between the protein Dicer-like 1 (DCL1) and a pre-miRNA [48], and therefore is involved in the biogenesis of small RNAs and in gene silencing. Its human orthologue, Smad nuclear interacting protein 1 (SNIP1), has been attributed a similar function [48]. Interestingly, the human *hPar14* parvulin protein was also found to be involved in ribosome biogenesis and RNA processing [49], and binds to double-stranded nuclear acids too. In addition to the cellular functions of DDL in *Arabidopsis*, the human SNIP1 constitutes an inhibitor of TGF- β and NF- κ B signaling pathways by competing with the TGF- β canonical signaling protein Smad4 and the NF- κ B transcription factor p65/RelA for binding to the transcriptional coactivator p300 [50,51]. Noticeably, human parvulins also act in these signaling pathways. *hPin1* protein has been demonstrated to be involved in a variety of these signaling events, e.g., of the canonical (Smad signaling) and non-canonical (Ras/ERK and PI3K/Akt) pathways of TGF- β (as reviewed in References [52,53]) and of NF κ B [54]. In addition, *hPar14* was found to act in insulin activated PI3K-Akt signaling (reviewed in Reference [9]). The fact that human parvulins are involved in cellular events in which orthologous proteins of the parvulin *TbPar42* are also involved may lead towards a role of *TbPar42* in TGF- β and NF κ B signaling, as well as RNA biogenesis and stability. These activities may be important for *Trypanosoma* to control host defense mechanisms and immunity.

4.2. *TbPar42* May Interact as a Protein Recruitment Platform

The domains of *TbPar42* are lined up as a string of pearls. They seem to act independently of each other, or they may interact only after binding to a yet unknown binding partner. Following this hypothesis, *TbPar42* would function as a specific dynamic recruitment platform, a scaffold protein that allows weakly associating proteins to be engaged in higher order assemblies [55]. Such assemblies allow a cell to gain spatiotemporal control over protein activity. Many cellular events, such as signal transduction cascades or gene activation and its control processes, involve such transient buildup of higher-order protein complexes. SNIP1 and the related eukaryotic parvulin proteins such as *hPar14* and *hPar17* have already been demonstrated to be involved in such events and assemblies [49,56–58]. An alternative explanation for the absence of domain interaction involves post-translational modification (PTM) of the protein, especially of its linker region. A Netphos-3.1 data search [59] predicted a plethora of phosphorylation sites within this region (Figure 10). In the N-terminus, only one such site (Thr32 by CKII) was found.

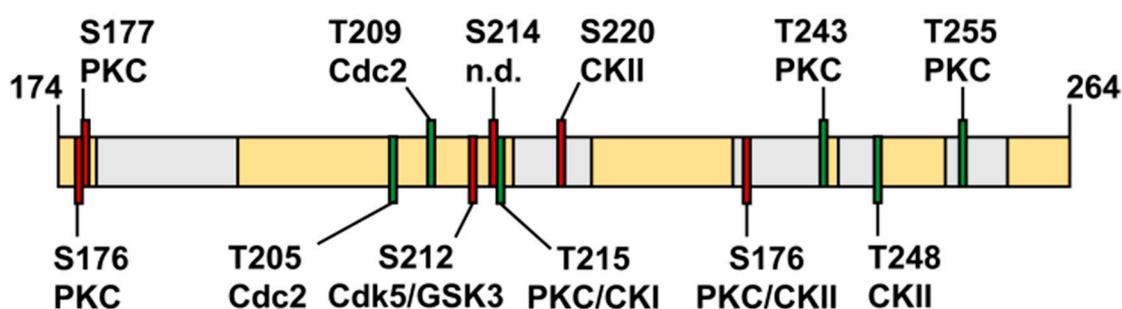


Figure 10. Phosphorylation sites predicted by Netphos-3b1. Schematic drawing of the linker region (residues 174 to 264) in yellow. Putative helices are represented by gray horizontal bars. Predicted phosphorylation sites are depicted as red (serine) and green (threonine) vertical bars. The respective phosphorylated residues and their putative modifying kinases are marked (PKC = protein kinase C; CKI/CKII = casein kinase I/II; Cdc/Cdk = cyclin-dependent kinase; GSK3 = Glycogensynthase kinase 3).

Most of the detected phosphorylation sites in the linker are at the beginning or within the helical regions predicted in our Rosetta model (Figure S2). Post-translational modifications often act in stabilizing transient helices, or in preventing the formation of such secondary structure elements. The amount and sequential arrangement of PTMs within a mainly unstructured large linker region (according to our NMR data) characterize this part of the protein as intrinsically disordered region (IDR) [60,61]. Such regions play a general regulatory role in signaling and controlling pathways [62–65].

Noticeably, all (but *cdc2*) kinases, which, according to the Netphos prediction, execute the post-translational modifications of the *TbPar42* linker region (Figure 10), are involved in Wnt-signaling [66,67], and are active in cytoskeleton organization, mitotic regulation, neuronal patterning, and cell fate decision. Many of these events have been demonstrated to be influenced by parvulins Pin1 and Par14/17 [9,68,69] in human cells. The IDR character of *TbPar42* and the involvement of relatives from other organisms in protein assemblies seems to be supportive for the recruitment platform hypothesis of this protein.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2218-273X/9/3/93/s1>, Table S1: Structural statistics for the deposited ensemble of 10 lowest energy PPIase structures, Table S2: Structural statistics for the deposited PPIase structure solved by X-ray analysis, Table S3: Structural statistics for the deposited ensemble of 10 lowest energy NterFHA structures, Table S4: Mean K_D values after titration of Ala-Glu-Ala-pThr-Ala-Glu-Xaa-peptides and Suc-Ala-pThr-Pro-Ala-NH₂ to NterFHA of *TbPar42*, Table S5: Oligonucleotides used as forward and reverse primers for PCR-amplification of the constructs. Figure S1: NMR shift data of the linker region of *TbPar42* (A) and secondary structure predictions, Figure S2: CS Rosetta model of the *TbPar42* linker region, Figure S3: Sequence alignment of the PPIase domains of *TbPar42* and *hPin1*, Figure

S4: NMR structure (PDB-ID: 2N87) of the PPIase domain of *TbPar42*, Figure S5: Comparison of X-ray structure (PDB-ID: 6GMP) and NMR structure (PDB-ID:2N87) of the PPIase domain of *TbPar42*.

Author Contributions: Conceptualization, E.R., D.H. and P.B.; methodology, E.R., D.H., B.E.K., E.H., R.G. and A.M.; validation, E.R., D.H., A.G., E.H., R.G. and P.B.; formal analysis, E.R. and B.E.K.; investigation, E.R., D.H., A.M. and P.B.; resources, P.B.; writing—original draft preparation, E.R., D.H. and P.B.; writing—review and editing, E.R., D.H., B.E.K., A.G., E.H., R.G., A.M. and P.B.; visualization, E.R., D.H. and A.M.; supervision, P.B.; project administration, P.B.

Funding: This research was funded by the Deutsche Forschungsgemeinschaft (DFG) (GRK 1431/2 and 1431/3).

Acknowledgments: We thank Tina Gerdes, Alma Rute and Peter Binz for excellent technical support. We thank the beamline staff of P13 at DESY as well as Astrid Port (University of Düsseldorf) for support of sample preparation and data collection.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Goh, J.Y.; Lai, C.-Y.; Tan, L.C.; Yang, D.; He, C.Y.; Liou, Y.-C. Functional characterization of two novel parvulins in *Trypanosoma brucei*. *FEBS Lett.* **2010**, *584*, 2901–2908. [[CrossRef](#)] [[PubMed](#)]
2. Erben, E.D.; Valguarnera, E.; Nardelli, S.; Chung, J.; Daum, S.; Potenza, M.; Schenkman, S.; Téllez-Iñón, M.T. Identification of an atypical peptidyl-prolyl cis/trans isomerase from trypanosomatids. *Biochim. Biophys. Acta* **2010**, *1803*, 1028–1037. [[CrossRef](#)] [[PubMed](#)]
3. Vallon, O. Chlamydomonas Immunophilins and Parvulins: Survey and Critical Assessment of Gene Models. *Eukaryot. Cell* **2005**, *4*, 230–241. [[CrossRef](#)] [[PubMed](#)]
4. Andreotti, A.H. Native state proline isomerization: An intrinsic molecular switch. *Biochemistry* **2003**, *42*, 9515–9524. [[CrossRef](#)] [[PubMed](#)]
5. Hammet, A.; Pike, B.L.; McNeese, C.J.; Conlan, L.A.; Tennis, N.; Heierhorst, J. FHA domains as phospho-threonine binding modules in cell signaling. *IUBMB Life* **2003**, *55*, 23–27. [[CrossRef](#)] [[PubMed](#)]
6. Coquelle, N.; Glover, J.N.M. FHA domain pThr binding specificity: it's all about me. *Structure* **2010**, *18*, 1549–1550. [[CrossRef](#)] [[PubMed](#)]
7. Mahajan, A.; Yuan, C.; Lee, H.; Chen, E.S.-W.; Wu, P.-Y.; Tsai, M.-D. Structure and function of the phosphothreonine-specific FHA domain. *Sci. Signal.* **2008**, *1*, re12. [[CrossRef](#)] [[PubMed](#)]
8. Scholz, C.; Rahfeld, J.; Fischer, G.; Schmid, F.X. Catalysis of protein folding by parvulin. *J. Mol. Biol.* **1997**, *273*, 752–762. [[CrossRef](#)] [[PubMed](#)]
9. Matena, A.; Rehic, E.; Hönig, D.; Kamba, B.; Bayer, P. Structure and function of the human parvulins Pin1 and Par14/17. *Biol. Chem.* **2018**, *399*, 101–125. [[CrossRef](#)] [[PubMed](#)]
10. Holm, L.; Laakso, L.M. Dali server update. *Nucleic Acids Res.* **2016**, *44*, W351–W355. [[CrossRef](#)] [[PubMed](#)]
11. Grum, D.; van den Boom, J.; Neumann, D.; Matena, A.; Link, N.M.; Mueller, J.W. A heterodimer of human 3'-phospho-adenosine-5'-phosphosulphate (PAPS) synthases is a new sulphate activating complex. *Biochem. Biophys. Res. Commun.* **2010**, *395*, 420–425. [[CrossRef](#)] [[PubMed](#)]
12. Wishart, D.; Bigam, C.; Yao, J.; Abildgaard, F.; Dyson, H.J.; Oldfield, E.; Markley, J.; Sykes, B. ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **1995**, *6*, 135–140. [[CrossRef](#)] [[PubMed](#)]
13. Markley, J.L.; Bax, A.; Arata, Y.; Hilbers, C.W.; Kaptein, R.; Sykes, B.D.; Wright, P.E.; Wüthrich, K. Recommendations for the presentation of NMR structures of proteins and nucleic acids. IUPAC-IUBMB-IUPAB Inter-Union Task Group on the Standardization of Data Bases of Protein and Nucleic Acid Structures Determined by NMR Spectroscopy. *J. Biomol. NMR* **1998**, *12*, 1–23. [[CrossRef](#)] [[PubMed](#)]
14. López-Méndez, B.; Güntert, P. Automated protein structure determination from NMR spectra. *J. Am. Chem. Soc.* **2006**, *128*, 13112–13122. [[CrossRef](#)] [[PubMed](#)]
15. Cornilescu, G.; Delaglio, F.; Bax, A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **1999**, *13*, 289–302. [[CrossRef](#)] [[PubMed](#)]
16. Herrmann, T.; Güntert, P.; Wüthrich, K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **2002**, *319*, 209–227. [[CrossRef](#)]
17. Güntert, P. Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* **2004**, *278*, 353–378. [[CrossRef](#)] [[PubMed](#)]

18. Wagner, G.; Neuhaus, D.; Wörgötter, E.; Vašák, M.; Kägi, J.H.R.; Wüthrich, K. Nuclear magnetic resonance identification of “half-turn” and 310-helix secondary structure in rabbit liver metallothionein-2. *J. Mol. Biol.* **1986**, *187*, 131–135. [[CrossRef](#)]
19. Wishart, D.S.; Sykes, B.D.; Richards, F.M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **1991**, *222*, 311–333. [[CrossRef](#)]
20. Wishart, D.S.; Sykes, B.D.; Richards, F.M. The chemical shift index: A fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **1992**, *31*, 1647–1651. [[CrossRef](#)] [[PubMed](#)]
21. Wishart, D.S.; Sykes, B.D. The ¹³C chemical-shift index: A simple method for the identification of protein secondary structure using ¹³C chemical-shift data. *J. Biomol. NMR* **1994**, *4*, 171–180. [[CrossRef](#)] [[PubMed](#)]
22. Ayed, A.; Mulder, F.A.; Yi, G.S.; Lu, Y.; Kay, L.E.; Arrowsmith, C.H. Latent and active p53 are identical in conformation. *Nat. Struct. Biol.* **2001**, *8*, 756–760. [[CrossRef](#)] [[PubMed](#)]
23. Delano, W.L. The PyMOL Molecular Graphics System. 2002. Available online: <http://www.pymol.org> (accessed on February 2019).
24. Kay, L.E.; Nicholson, L.K.; Delaglio, F.; Bax, A.; Torchia, D.A. Pulse sequences for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear T1 and T2 values in proteins. *J. Magn. Reson.* **1992**, *97*, 359–375. [[CrossRef](#)]
25. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D* **2010**, *66*, 133–144. [[CrossRef](#)] [[PubMed](#)]
26. Kabsch, W. XDS. *Acta Crystallogr. Sect. D* **2010**, *66*, 125–132. [[CrossRef](#)] [[PubMed](#)]
27. Vagin, A.; Teplyakov, A. MOLREP: An Automated Program for Molecular Replacement. *J. Appl. Cryst.* **1997**, *30*, 1022–1025. [[CrossRef](#)]
28. Murshudov, G.N.; Vagin, A.A.; Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D* **1997**, *53*, 240–255. [[CrossRef](#)] [[PubMed](#)]
29. Adams, P.D.; Afonine, P.V.; Bunkóczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.-W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D* **2010**, *66*, 213–221. [[CrossRef](#)] [[PubMed](#)]
30. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D* **2004**, *60*, 2126–2132. [[CrossRef](#)] [[PubMed](#)]
31. Fischer, G.; Bang, H.; Mech, C. Nachweis einer Enzymkatalyse für die cis-trans-Isomerisierung der Peptidbindung in prolinhaltigen Peptiden. *Biomed. Biochim. Acta* **1984**, *43*, 1101–1111. [[PubMed](#)]
32. Kofron, J.L.; Kuzmic, P.; Kishore, V.; Colón-Bonilla, E.; Rich, D.H. Determination of kinetic constants for peptidyl prolyl cis-trans isomerases by an improved spectrophotometric assay. *Biochemistry* **1991**, *30*, 6127–6134. [[CrossRef](#)] [[PubMed](#)]
33. Fanghänel, J.; Fischer, G. Insights into the catalytic mechanism of peptidyl prolyl cis/trans isomerases. *Front. Biosci.* **2004**, *9*, 3453–3478. [[CrossRef](#)] [[PubMed](#)]
34. Ranganathan, R.; Lu, K.P.; Hunter, T.; Noel, J.P. Structural and Functional Analysis of the Mitotic Rotamase Pin1 Suggests Substrate Recognition Is Phosphorylation Dependent. *Cell* **1997**, *89*, 875–886. [[CrossRef](#)]
35. Mueller, J.W.; Link, N.M.; Matena, A.; Hoppstock, L.; Rüppel, A.; Bayer, P.; Blankenfeldt, W. Crystallographic proof for an extended hydrogen-bonding network in small prolyl isomerases. *J. Am. Chem. Soc.* **2011**, *133*, 20096–20099. [[CrossRef](#)] [[PubMed](#)]
36. Bayer, E.; Goettch, S.; Mueller, J.W.; Griewel, B.; Guiberman, E.; Mayr, L.M.; Bayer, P. Structural analysis of the mitotic regulator hPin1 in solution: Insights into domain architecture and substrate binding. *J. Biol. Chem.* **2003**, *278*, 26183–26193. [[CrossRef](#)] [[PubMed](#)]
37. Heikkinen, O.; Seppala, R.; Tossavainen, H.; Heikkinen, S.; Koskela, H.; Permi, P.; Kilpeläinen, I. Solution structure of the parvulin-type PPIase domain of *Staphylococcus aureus* PrsA—Implications for the catalytic mechanism of parvulins. *BMC Struct. Biol.* **2009**, *9*, 17. [[CrossRef](#)] [[PubMed](#)]
38. Daum, S.; Erdmann, F.; Fischer, G.; Féaux de Lacroix, B.; Hessamian-Alinejad, A.; Houben, S.; Frank, W.; Braun, M. Aryl indanyl ketones: Efficient inhibitors of the human peptidyl prolyl cis/trans isomerase Pin1. *Angew. Chem.* **2006**, *45*, 7454–7458. [[CrossRef](#)] [[PubMed](#)]
39. Behrsin, C.D.; Bailey, M.L.; Bateman, K.S.; Hamilton, K.S.; Wahl, L.M.; Brandl, C.J.; Shilton, B.H.; Litchfield, D.W. Functionally important residues in the peptidyl-prolyl isomerase Pin1 revealed by unigenic evolution. *J. Mol. Biol.* **2007**, *365*, 1143–1162. [[CrossRef](#)] [[PubMed](#)]

40. Sun, L.; Wu, X.; Peng, Y.; Goh, J.Y.; Liou, Y.-C.; Lin, D.; Zhao, Y. Solution structural analysis of the single-domain parvulin TbPin1. *PLoS ONE* **2012**, *7*, e43017. [[CrossRef](#)] [[PubMed](#)]
41. Weinger, U.; Jakob, R.P.; Kovermann, M.; Balbach, J.; Schmid, F.X. The prolyl isomerase domain of PpiD from *Escherichia coli* shows a parvulin fold but is devoid of catalytic activity. *Protein Sci.* **2010**, *19*, 6–18. [[CrossRef](#)] [[PubMed](#)]
42. Behrens, S.; Maier, R.; de Cock, H.; Schmid, F.X.; Gross, C.A. The SurA periplasmic PPIase lacking its parvulin domains functions in vivo and has chaperone activity. *EMBO J.* **2001**, *20*, 285–294. [[CrossRef](#)] [[PubMed](#)]
43. Durocher, D.; Taylor, I.A.; Sarbassova, D.; Haire, L.F.; Westcott, S.L.; Jackson, S.P.; Smerdon, S.J.; Yaffe, M.B. The Molecular Basis of FHA Domain:Phosphopeptide Binding Specificity and Implications for Phospho-Dependent Signaling Mechanisms. *Mol. Cell* **2000**, *6*, 1169–1182. [[CrossRef](#)]
44. Xu, Q.; Deller, M.C.; Nielsen, T.K.; Grant, J.C.; Lesley, S.A.; Elsliger, M.-A.; Deacon, A.M.; Wilson, I.A. Structural insights into the recognition of phosphopeptide by the FHA domain of kanadaplin. *PLoS ONE* **2014**, *9*, e107309. [[CrossRef](#)] [[PubMed](#)]
45. Pennell, S.; Westcott, S.; Ortiz-Lombardía, M.; Patel, D.; Li, J.; Nott, T.J.; Mohammed, D.; Buxton, R.S.; Yaffe, M.B.; Verma, C.; et al. Structural and functional analysis of phosphothreonine-dependent FHA domain interactions. *Structure* **2010**, *18*, 1587–1595. [[CrossRef](#)] [[PubMed](#)]
46. Lange, O.F.; Rossi, P.; Sgourakis, N.G.; Song, Y.; Lee, H.-W.; Aramini, J.M.; Ertekin, A.; Xiao, R.; Acton, T.B.; Montelione, G.T.; et al. Determination of solution structures of proteins up to 40 kDa using CS-Rosetta with sparse NMR data from deuterated samples. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 10873–10878. [[CrossRef](#)] [[PubMed](#)]
47. Morris, E.R.; Chevalier, D.; Walker, J.C. DAWDLE, a forkhead-associated domain gene, regulates multiple aspects of plant development. *Plant Physiol.* **2006**, *141*, 932–941. [[CrossRef](#)] [[PubMed](#)]
48. Yu, B.; Bi, L.; Zheng, B.; Ji, L.; Chevalier, D.; Agarwal, M.; Ramachandran, V.; Li, W.; Lagrange, T.; Walker, J.C.; et al. The FHA domain proteins DAWDLE in *Arabidopsis* and SNIP1 in humans act in small RNA biogenesis. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10073–10078. [[CrossRef](#)] [[PubMed](#)]
49. Fujiyama-Nakamura, S.; Yoshikawa, H.; Homma, K.; Hayano, T.; Tsujimura-Takahashi, T.; Izumikawa, K.; Ishikawa, H.; Miyazawa, N.; Yanagida, M.; Miura, Y.; et al. Parvulin (Par14), a peptidyl-prolyl cis-trans isomerase, is a novel rRNA processing factor that evolved in the metazoan lineage. *Mol. Cell. Proteom.* **2009**, *8*, 1552–1565. [[CrossRef](#)] [[PubMed](#)]
50. Kim, R.H.; Flanders, K.C.; Birkey Reffey, S.; Anderson, L.A.; Duckett, C.S.; Perkins, N.D.; Roberts, A.B. SNIP1 inhibits NF-kappa B signaling by competing for its binding to the C/H1 domain of CBP/p300 transcriptional co-activators. *J. Biol. Chem.* **2001**, *276*, 46297–46304. [[CrossRef](#)] [[PubMed](#)]
51. Kim, R.H.; Wang, D.; Tsang, M.; Martin, J.; Huff, C.; de Caestecker, M.P.; Parks, W.T.; Meng, X.; Lechleider, R.J.; Wang, T.; et al. A novel Smad nuclear interacting protein, SNIP1, suppresses p300-dependent TGF- β signal transduction. *Genes Dev.* **2000**, *14*, 1605–1616. [[PubMed](#)]
52. Zhou, X.Z.; Lu, K.P. The isomerase PIN1 controls numerous cancer-driving pathways and is a unique drug target. *Nat. Rev. Cancer* **2016**, *16*, 463–478. [[CrossRef](#)] [[PubMed](#)]
53. Nakatsu, Y.; Matsunaga, Y.; Yamamotoya, T.; Ueda, K.; Inoue, Y.; Mori, K.; Sakoda, H.; Fujishiro, M.; Ono, H.; Kushiya, A.; et al. Physiological and Pathogenic Roles of Prolyl Isomerase Pin1 in Metabolic Regulations via Multiple Signal Transduction Pathway Modulations. *Int. J. Mol. Sci.* **2016**, *17*. [[CrossRef](#)] [[PubMed](#)]
54. Ryo, A.; Suizu, F.; Yoshida, Y.; Perrem, K.; Liou, Y.-C.; Wulf, G.; Rottapel, R.; Yamaoka, S.; Lu, K.P. Regulation of NF-kappaB signaling by Pin1-dependent prolyl isomerization and ubiquitin-mediated proteolysis of p65/RelA. *Mol. Cell* **2003**, *12*, 1413–1426. [[CrossRef](#)]
55. Wijnands, S.P.W.; Engelen, W.; Lafleur, R.P.M.; Meijer, E.W.; Merks, M. Controlling protein activity by dynamic recruitment on a supramolecular polymer platform. *Nat. Commun.* **2018**, *9*, 65. [[CrossRef](#)] [[PubMed](#)]
56. Thiele, A.; Krentzlin, K.; Erdmann, F.; Rauh, D.; Hause, G.; Zerweck, J.; Kilka, S.; Pösel, S.; Fischer, G.; Schutkowski, M.; et al. Parvulin 17 promotes microtubule assembly by its peptidyl-prolyl cis/trans isomerase activity. *J. Mol. Biol.* **2011**, *411*, 896–909. [[CrossRef](#)] [[PubMed](#)]
57. Zhang, J.; Nakatsu, Y.; Shinjo, T.; Guo, Y.; Sakoda, H.; Yamamotoya, T.; Otani, Y.; Okubo, H.; Kushiya, A.; Fujishiro, M.; et al. Par14 protein associates with insulin receptor substrate 1 (IRS-1), thereby enhancing insulin-induced IRS-1 phosphorylation and metabolic actions. *J. Biol. Chem.* **2013**, *288*, 20692–20701. [[CrossRef](#)] [[PubMed](#)]

58. Chen, L.-L.; Lin, H.-P.; Zhou, W.-J.; He, C.-X.; Zhang, Z.-Y.; Cheng, Z.-L.; Song, J.-B.; Liu, P.; Chen, X.-Y.; Xia, Y.-K.; et al. SNIP1 Recruits TET2 to Regulate c-MYC Target Genes and Cellular DNA Damage Response. *Cell Rep.* **2018**, *25*, 1485–1500. [[CrossRef](#)] [[PubMed](#)]
59. Blom, N.; Gammeltoft, S.; Brunak, S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* **1999**, *294*, 1351–1362. [[CrossRef](#)] [[PubMed](#)]
60. Dunker, A.K.; Obradovic, Z.; Romero, P.; Garner, E.C.; Brown, C.J. Intrinsic protein disorder in complete genomes. *Genome Inform.* **2000**, *11*, 161–171.
61. Romero, P.; Obradovic, Z.; Li, X.; Garner, E.C.; Brown, C.J.; Dunker, A.K. Sequence complexity of disordered protein. *Proteins* **2001**, *42*, 38–48. [[CrossRef](#)]
62. Dyson, H.J.; Wright, P.E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197–208. [[CrossRef](#)] [[PubMed](#)]
63. Tompa, P. Intrinsically disordered proteins: A 10-year recap. *Trends Biochem. Sci.* **2012**, *37*, 509–516. [[CrossRef](#)] [[PubMed](#)]
64. Uversky, V.N. Disorder in the lifetime of a protein. *Intrinsically Disord. Proteins* **2013**, *1*, e26782. [[CrossRef](#)] [[PubMed](#)]
65. Trusch, F.; Matena, A.; Vuk, M.; Koerver, L.; Knævelsrud, H.; Freemont, P.S.; Meyer, H.; Bayer, P. The N-terminal Region of the Ubiquitin Regulatory X (UBX) Domain-containing Protein 1 (UBXD1) Modulates Interdomain Communication within the Valosin-containing Protein p97. *J. Biol. Chem.* **2015**, *290*, 29414–29427. [[CrossRef](#)] [[PubMed](#)]
66. Komiya, Y.; Habas, R. Wnt signal transduction pathways. *Organogenesis* **2008**, *4*, 68–75. [[CrossRef](#)] [[PubMed](#)]
67. Marinou, K.; Christodoulides, C.; Antoniadis, C.; Koutsilieris, M. Wnt signaling in cardiovascular physiology. *Trends Endocrinol. Metab.* **2012**, *23*, 628–636. [[CrossRef](#)] [[PubMed](#)]
68. Ryo, A.; Liou, Y.-C.; Lu, K.P.; Wulf, G. Prolyl isomerase Pin1: A catalyst for oncogenesis and a potential therapeutic target in cancer. *J. Cell Sci.* **2003**, *116*, 773–783. [[CrossRef](#)] [[PubMed](#)]
69. Burgardt, N.I.; Schmidt, A.; Manns, A.; Schutkowski, A.; Jahreis, G.; Lin, Y.-J.; Schulze, B.; Masch, A.; Lücke, C.; Weiwad, M. Parvulin 17-catalyzed Tubulin Polymerization Is Regulated by Calmodulin in a Calcium-dependent Manner. *J. Biol. Chem.* **2015**, *290*, 16708–16722. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).