

Mapping and Characterization of the Interaction Interface between Two Polypyrimidine-Tract Binding Proteins and a Nova-Type Protein of *Solanum tuberosum*

Shweta Shah¹, Nathaniel M. Butler², David J. Hannapel², A. Gururaj Rao^{1*}

¹ Roy J. Carver Department of Biochemistry Biophysics and Molecular Biology, Iowa State University, Ames, Iowa, United States of America, ² Plant Biology Major, Iowa State University, Ames, Iowa, United States of America

Abstract

Polypyrimidine tract-binding (PTB) proteins are RNA-binding proteins that generally contain four RNA recognition motifs (RRMs). In potato, six cDNAs encoding full-length PTB proteins have been identified. In the present study Nova1-like protein, designated StNova1, was identified as a potential interacting partner of the StPTB proteins via yeast two-hybrid screening. Nova protein is a RNA-binding protein that contains three K-homology (KH) domains. In humans, these proteins are involved in regulation of neuronal RNA metabolism but the role of Nova-like proteins in plants is poorly understood. We have validated this interaction and mapped the protein binding region on StNova1 and StPTB1 and -6 using a novel domain interaction phage display (DIPP) technique. The interaction between the two RNA-binding proteins StPTB1/6 and StNova1 is mediated through linker regions that are distinctly separated from the RRM. Furthermore, using a random 21-mer phage-peptide library, we have identified a number of peptides with the consensus sequence motif [S/G][V/I][L/V]G that recognize the StPTB proteins. One over-represented peptide that recognizes StPTB6 contains the GVLGPWP sequence that is similar to the GIGGRYP sequence in the glycine-rich linker region between the KH2 and KH3 domains of StNova1. We show, through site-specific mutations, the importance of glycine and proline residues in StNova1-StPTB interactions.

Citation: Shah S, Butler NM, Hannapel DJ, Rao AG (2013) Mapping and Characterization of the Interaction Interface between Two Polypyrimidine-Tract Binding Proteins and a Nova-Type Protein of *Solanum tuberosum*. PLoS ONE 8(5): e64783. doi:10.1371/journal.pone.0064783

Editor: Beata G. Vertessy, Institute of Enzymology of the Hungarian Academy of Science, Hungary

Received: February 7, 2013; **Accepted:** April 18, 2013; **Published:** May 24, 2013

Copyright: © 2013 Shah et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported by the National Science Foundation Plant Genome Research Program (Grant no. DBI-0820659) and National Research Initiative (Grant no. 2008-02806) from the USDA National Institute of Food and Agriculture. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: gururao@iastate.edu

Introduction

Polypyrimidine tract-binding (PTB) proteins are RNA-binding proteins that have been well characterized and studied in vertebrates. They typically contain four RNA recognition motifs (RRM) that are separated by three linker regions and all the RRM domains have been shown to be involved in RNA binding. The RRM domains interact with sequence in UTRs containing groups of four cytosine/uracil motifs at least four nucleotides in length [1,2]. The PTBs participate in multiple regulatory functions in mRNA metabolism including polyadenylation, 3' end formation, translation from internal ribosomal entry sites, RNA localization and stability, and alternative splicing (AS) [2,3,4]. In some instances, the function of PTB proteins is mediated via primary interactions with other regulatory proteins. Thus, PTB proteins have been reported to interact with proteins such as Raver, Nova-1, Nova-2, and MRG15 proteins [5,6,7,8,9,10]. Raver1 is a protein that promotes AS of α -tropomyosin and consists of three RRM domains [11]. The Raver1 protein has an extended C-terminal region with four conserved PTB interaction motifs containing the consensus sequence [S/G][I/L]LGxxP [8,11]. The Nova (Neuron Oncological Ventral Antigen) proteins are neuron specific RNA binding proteins containing three RNA-binding domains, referred to as K-homology or KH domains. The first and second domains are in tandem arrangement, followed by a large spacer region and

the third domain at the C-terminal of the protein. These proteins are also involved in controlling AS [5]. The Nova1-like protein of *Arabidopsis thaliana*, designated BTR1 (binding to ToMV RNA 1L), binds specifically to terminal regions of genomic RNA of the tomato mosaic virus (ToMV) [12]. These terminal regions contain important regulatory elements for translation and RNA replication. In planta analysis suggests that by binding to the viral RNA, BTR1 regulates replication and cell-to-cell movement of ToMV [12]. The Nova-1 family of genes is highly conserved within the plant kingdom with orthologs in numerous diverse species, including *Glycine max*, *Hordeum vulgare*, *Medicago truncatula*, *Oryza sativa*, *Populus trichocarpa*, *Ricinus communis*, *Solanum lycopersicum*, *Vitis vinifera*, and *Zea mays*. Despite this conservation, however, with the exception of BTR1, very little is known about the function of these plant orthologs.

PTB proteins and other RNA-binding proteins involved in diverse aspects of RNA metabolism also occur in plants [13,14]. In *Arabidopsis thaliana* there are three PTB-homologs (*APT*B1–3). *At*PTB1 and -2 contain three RRM, whereas *At*PTB3 contains four. Using a transcriptome-wide analysis, AS activity was observed for *APT*B1 and -2 but no activity was detected for the distantly related *APT*B3 [15]. Importantly, their localization to distinct cellular compartments (nucleus, cytosol and processing bodies) attest to their multifunctional role in developmental processes [16]. PTB-related proteins in *Arabidopsis* have also been

implicated in pollen germination possibly through their function in processes of RNA metabolism [17]. In pumpkin, RBP50, a phloem-mobile PTB protein, similar to *AtPTB3*, functions as a non-cell-autonomous RNA-binding protein, is highly enriched in the phloem sap and forms the core of a large ribonucleoprotein complex containing at least sixteen proteins and six RNAs [13].

Recently in potato (*Solanum tuberosum*), six cDNAs that encode full-length PTB proteins have been identified [18]. Of these, two are *AtPTB3*-types containing four RRM s (*StPTB1* and *StPTB6*) and four are 3-RRM types (*StPTB7*, TC201749, B1920231 and TC218925). It has been demonstrated that the *StPTB1* and *-6* genes have distinct expression patterns and respond to both developmental and environmental cues [18]. To elucidate the function of these two RBP50-like proteins in potato and identify putative protein partners that may facilitate their binding to phloem RNAs, we have focused our attention on protein-protein interactions mediated by *StPTB1* and *-6*. In the present work, using *StPTB1* protein as bait in yeast two-hybrid experiments, we have identified a Nova-1 like protein (*StNova1*) and characterized the interaction between the two PTB proteins *StPTB1* and *-6*, and *StNova1* using multiple techniques. We also describe the use of a modified phage display technique, termed **Domain Interaction Phage Panning (DIPP)**, to identify the binding site interface between the two proteins and provide molecular insights for further mutagenesis experiments.

Materials and Methods

Yeast Two-hybrid Screening

The yeast two-hybrid screening was performed as described previously [19]. Briefly, the full length *StPTB1* was cloned into the pBridge vector (CLONTECH) and used as bait to screen $\sim 10^6$ transformants from a leaf cDNA library of potato (*Solanum tuberosum* cv Désirée) in pAD-GAL4-2.1 (Stratagene, La Jolla, CA). Positive interactions were confirmed by co-transforming yeast strain AH109 with each purified pAD plasmid and pBridge: *StPTB1* and plating on *-Leu/-Trp* (transformation control).

Plasmid Construction

The *StPTB1* and *StPTB6* were cloned into pE-SUMO vector (LifeSensors) using the primers as mentioned in Table S1. The forward primer incorporated BsaI site and the reverse primer incorporated BsaI and XbaI sites. The PCR products were ligated into pE-SUMO vector that adds N-terminus SUMO and 6xH tag to PTB proteins. The *StNova1* was cloned into modified pGEX4T-1-TEV vector for ligation independent cloning using the primer mentioned in Table S1. For cloning *StNova1* domain (D5), region between 571–780 was amplified by PCR using *StNova1* cDNA as a template. The forward primer incorporated SfiI site and reverse primer has NotI site (see Table S1). The PCR fragments were digested with SfiI and NotI and ligated into similarly digested pGEX4T-1. Sub-regions (designated as S1, S2 and S3) of domain D5 were also similarly cloned into pGEX4T-1 vector.

Protein Expression and Purification

Proteins were expressed overnight in Rosetta2 (DE3) pLysS (Novagen) cells at 20°C with 0.5 mM IPTG induction. For purification of Glutathione S-transferase (GST) fused protein, cell pellets were suspended in buffer A (1×Tris-buffer saline pH 7.4, 0.1% Triton X-100, 1 mM DTT and 1 mM PMSF) and lysed by sonication on ice. The lysate was centrifuged at 13,000 rpm for 20 min and clarified supernatant was loaded onto glutathione agarose (Pierce) resins that were already equilibrated with buffer

A. The unbound protein was washed with buffer B (50 mM Tris-HCl pH-8 and 1 mM DTT) and elution was done with buffer B containing 10 mM glutathione. For 6xH tagged proteins, the cell pellets were lysed into buffer C (50 mM Tris-HCl pH 8.0, 40 mM imidazole, 0.1% triton X-100, 1 mM DTT and 1 mM PMSF) and the protein was purified using Ni-NTA superflow resins. After washing the unbound proteins with buffer B containing 40 mM imidazole, the protein was eluted with Buffer B containing 150 mM imidazole. Protein concentrations were determined by the Bradford method [20].

Pull Down Assays

The MBP-*StNova1* protein was used as bait to pull out SUMO-*StPTB* proteins. Purified MBP-*StNova1* (20 µg) was bound on amylose resin (20 µl) and washed three times with wash/binding buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1 mM DTT, 0.1% NP-40). The resin-bound MBP-*StNova1* protein was subsequently added to a solution containing SUMO-*StPTB* protein (30 µg) in binding buffer in a final volume of 500 µl, followed by incubation with rotation for 1 h at room temperature. As a negative control MBP alone was incubated with SUMO-*StPTB* and SUMO proteins. After washing four times with wash/binding buffer to remove unbound protein, the resin was boiled at 95°C for 5 min in SDS-PAGE sample buffer, proteins resolved by 12% SDS-PAGE and analyzed by coomassie staining or by western blots using anti-SUMO primary antibody (Rockland Immunochemicals, 1:2000) and polyclonal alkaline phosphatase (AP) conjugated anti-rabbit secondary antibody (Sigma, 1:20,000). Colorimetric detection of secondary antibody was performed using the AP substrate kit (Bio-Rad).

Construction of Phage-displayed Overlapping Domains of *StNova1*

Six overlapping domains of *StNova1* (D1–D6) comprising amino acid residues 1–76 (D1), 39–129 (D2), 62–129(D3), 110–205(D4), 190–260(D5), 246–326(D6) were generated by PCR using *StNova1* cDNA as a template and forward and reverse primers to incorporate SfiI and NotI sites (Table S1). The PCR fragments were then digested with SfiI and NotI, ligated into a similarly digested pCANTAB 5E phagemid vector and subsequently transformed into *E. Coli* XLI-Blue cells (Stratagene). Single colonies from each plate were inoculated into 5 ml 2YT/carb/Amp media, grown to OD at 600 $\lambda \sim 0.2$ –0.3 at 37°C and then infected with helper phage-VCSM13 (Stratagene). After 1 hour the cell culture was transferred to 25 ml 2YT/Kan media and further incubated overnight at 37°C. After removing cell debris by centrifugation, phage particles were precipitated from the supernatant using 7.5 ml of 20% PEG solution containing NaCl. The precipitate was re-suspended in 1 ml of PBS and phage concentration determined by measuring absorbance at 268 λ (OD₂₆₈ = 1.0 for a solution containing 5×10^{12} phage per ml). A “domain library” was prepared by mixing equal concentrations of individual phage-domains.

Domain Interaction Phage Panning (DIPP)

SUMO fused PTB1, PTB6 and SUMO (control) proteins were immobilized in the wells of a nunc maxisorp ELISA plate by aliquoting 100 µl of each protein at a concentration of 10 µg/ml in 50 mM NaHCO₃, pH 9.6, at room temperature with gentle rotation for 2 h. Wells were then blocked with PBS containing 0.2% BSA for 1 h followed by 3 washings with PBS containing 0.05% Tween 20 (PBST) and then incubated with 100 µl of domain phage library for 3 h at room temperature with gentle

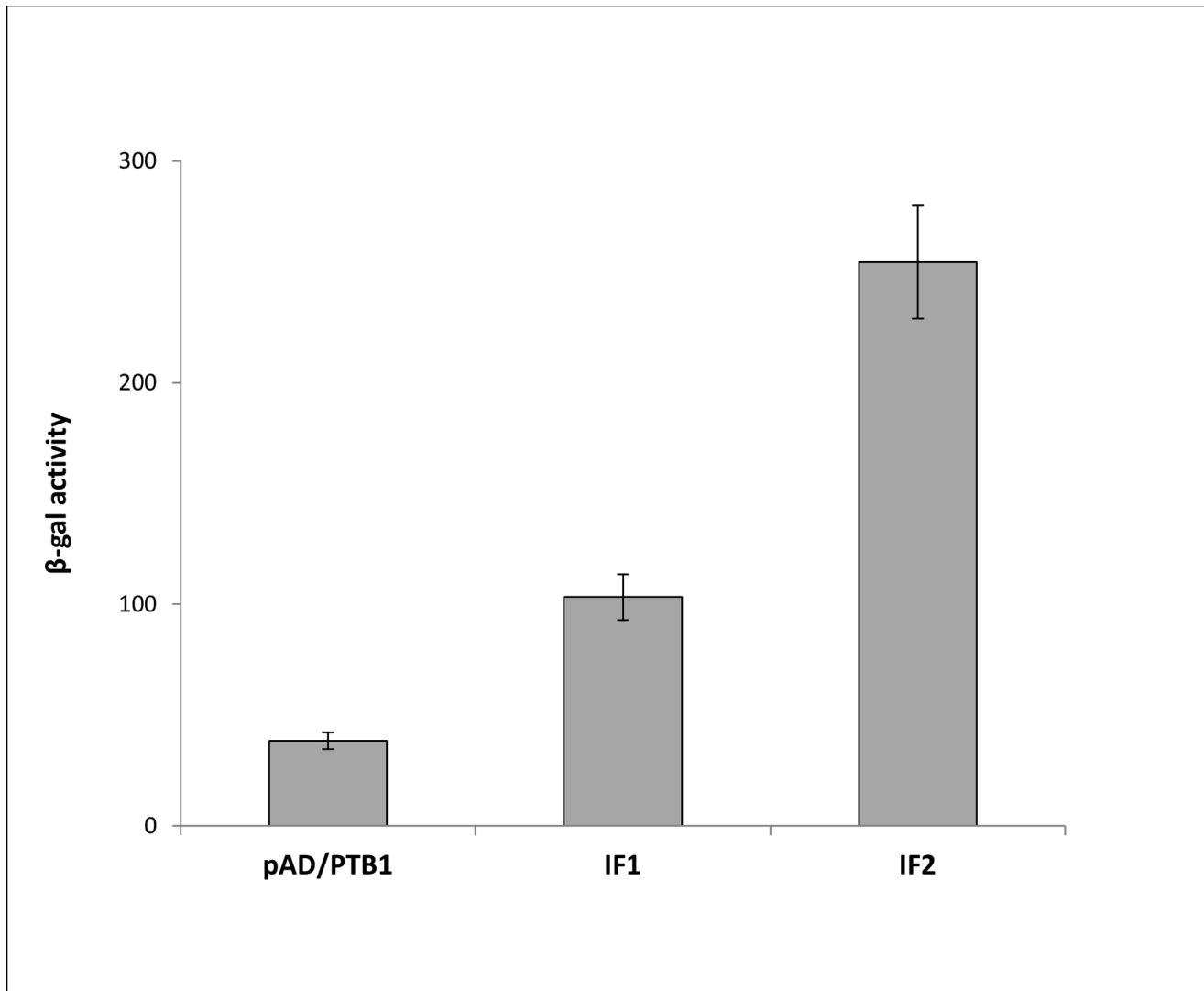


Figure 1. Yeast 2-hybrid interaction of *S*PTB1 with IF1 and IF2. Induction of β -galactosidase activity is used to assess the degree of interaction. Interaction of *S*PTB1 expressed in pBridge with an empty pAD vector is shown as a negative control. Both IF1 and IF2 show statistically significant differences in β -galactosidase activity compared to the PTB1+ pAD negative control. doi:10.1371/journal.pone.0064783.g001

rotation. After removing unbound phage by washing 5 times with PBST, bound phage was eluted by incubating with 500 μ l of 0.1 M HCl for 5 min at room temperature with shaking. The eluted phage was immediately neutralized by the addition of 1/3rd phage volume of 1 M Tris-HCl buffer pH 8.0, followed by infection of XL1-Blue cells (grown to <0.6 OD) with the phage. Infected XL1-Blue cells were incubated at 37°C for 20 min followed by addition of the helper phage (VCSM13) and again incubating at 37°C for 30 min. The phage infected XL1-Blue cells were transferred into a conical flask containing 50 ml of 2YT media containing 10 μ g/ml of tetracycline and 100 μ g/ml ampicillin that was further incubated at 37°C overnight with shaking at 210 rpm. Phage preparation was carried out as described earlier. This process yielded the first round of enriched phage. The entire process was repeated for one more cycle and after the 2nd round, phage infected XL1-Blue cells were grown on 2YT ampicillin plates. Plasmids were prepared from 20 randomly selected colonies and their DNA sequenced.

Phage ELISA for Binding Specificity

For phage ELISAs, 100 μ l of GST fused proteins and GST (control) (10 μ g/ml in 50 mM NaHCO₃, pH 9.6) were immobilized in the wells of an ELISA plate at room temperature with gentle rotation for 2 h. Plate was then washed two times with PBS followed by blocking with PBST containing 0.2% BSA for 1 h. Subsequently, after three washings with PBST, wells were incubated with the 100 μ l of appropriate domain phage diluted in PBST containing 0.2% BSA for 2 h at room temperature with gentle shaking. The plate was again washed three times with PBST followed by incubation with anti-M13 HRP conjugated antibody for 1 h. After washing four times with PBST, bound phage in each well was detected by incubating with 50 μ l of substrate solution (0.01% hydrogen peroxide +0.8 mg/ml o-Phenylenediamine dihydrochloride) for ~10 min. Reactions were terminated by the addition of 50 μ l of 3 M HCl and measuring the absorbance of the developed yellow color at 490 nm. Data were generated by subtracting the absorbance of the control (SUMO or GST) wells from the experimental wells.

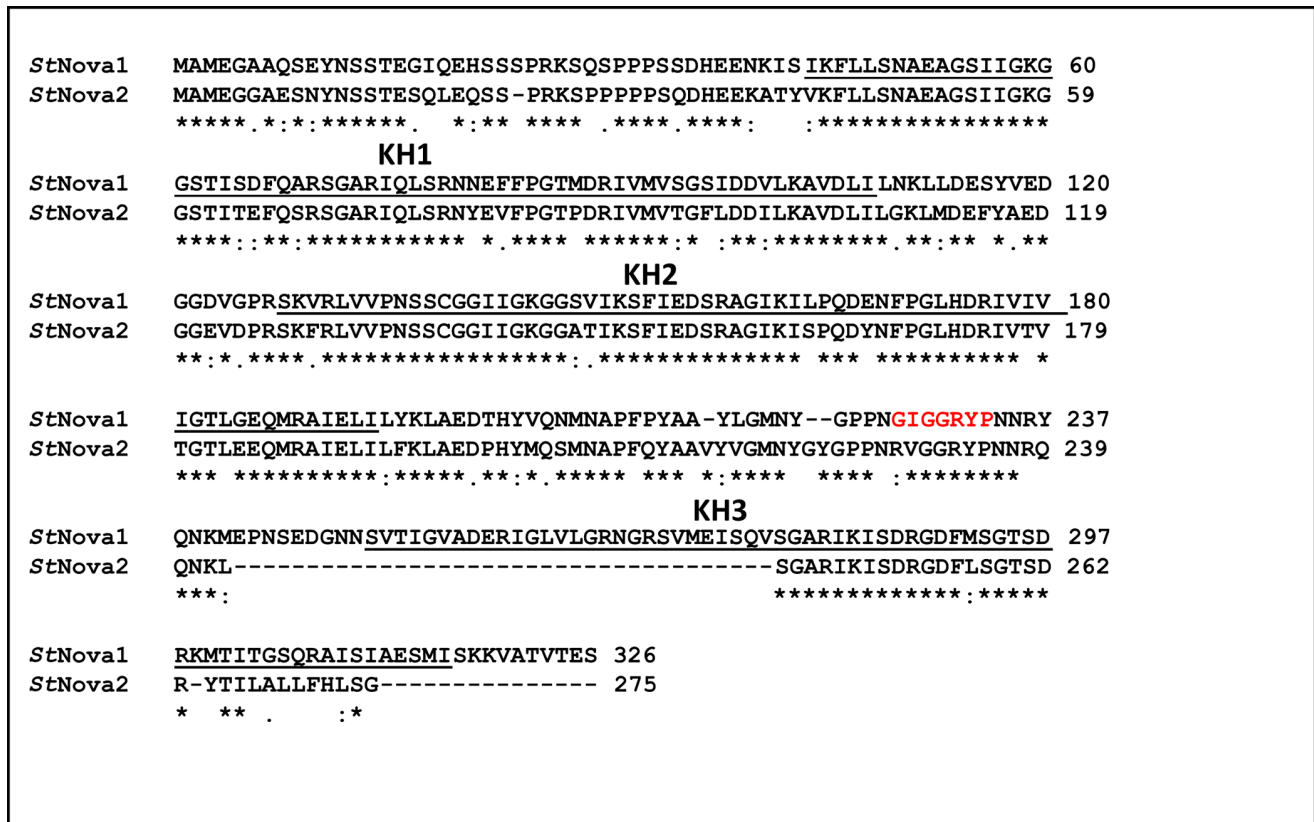


Figure 2. Amino acid sequence alignment of two isoforms of *StNova* proteins. The KH domains are underlined. Note the missing portion of the third KH domain in *StNova2*. The GxxGxxP motif present on *StNova1* is highlighted in red. doi:10.1371/journal.pone.0064783.g002

Protein Overlay Assays

For protein overlay assays, 2 µg of GST fused recombinant proteins were separated on 12% SDS-PAGE and blotted on to PVDF membrane. Blots were then blocked with 5% milk protein in Tris-buffered saline +0.1% Tween 20 (TBST) for 1 h. Blots were then washed three times with TBST (5 min) and incubated with the SUMO-fused protein or SUMO protein (for control) at a concentration of 25 nM for 1 h at room temperature. The binding of SUMO-fused proteins and SUMO proteins were detected by using anti-SUMO primary antibody as described in the section on pull down assays.

Results

Identification of PTB Interacting Proteins

SPTB1 and *SPTB6* are similar to CmRBP50 (>84% sequence identity among the proteins), the core protein in a phloem-mobile protein-RNA complex containing 16 proteins and six RNAs [13]. Therefore, it is reasonable to infer that the potato PTB proteins function similarly through protein-protein and protein-RNA interactions. To identify potential proteins interacting with *SPTB* proteins we performed a yeast two-hybrid screening against leaf proteins using *SPTB1* as bait. Based on the induction of β-galactosidase activity (Fig. 1), two interacting clones, designated as IF1 and IF2, were identified. Sequence analysis revealed that IF2 exhibited a sequence match (22.8% identity/~38% similarity) to the mammalian RNA-binding protein, Nova, containing three KH domains [5] and is therefore referred to as *StNova1*. The second clone, IF1, was found to be a putative phosphonucleolar

protein containing three RRM domains but is not discussed further in this paper. We cloned the full-length 35.2 kDa *StNova1* from a potato cDNA library and also identified an isoform, *StNova2*, which is missing a portion of the third KH domain (Fig. 2).

Interaction between *StPTB* Protein and *StNova1* Protein Analyzed by Pull-down Assays

In a parallel approach we assessed the interaction between *SPTB* and *StNova1* proteins by pull-down assays. For this purpose, *SPTB* proteins were expressed and purified as 6His-SUMO fusion and *StNova1* as MBP fusion. Both the SUMO-PTB proteins were pulled down by MBP-*StNova1* (Fig. 3A, lanes 1 and 2). The specificity of this interaction was confirmed by the absence of interaction of MBP-*StNova1* with SUMO protein alone (Fig. 3A, lane 3). In contrast, the MBP protein by itself was unable to pull-down the SUMO-PTB proteins (Fig. 3A, lanes 4 and 5). The presence of SUMO-PTB in lanes 1, 2, 6 and 7 was also confirmed with western blots using polyclonal, anti-SUMO antibody (Fig. 3B).

Phage Panning of *StPTB* Proteins

The interaction between mammalian PTB and Raver1 protein is well studied and it has been shown that PTB recognize a consensus binding motif [S/G][I/L]LGxxP present on the C-terminus of Raver1 [8,21] as well as Raver2 protein [7]. In order to ascertain if the *SPTB* proteins also recognize a specific binding motif, we used a random 21-mer phage peptide library to screen against both *SPTB1* and *SPTB6* as described in Materials and Methods. The phage display technique is widely used to identify peptide ligands, probe structure-function relationships in proteins

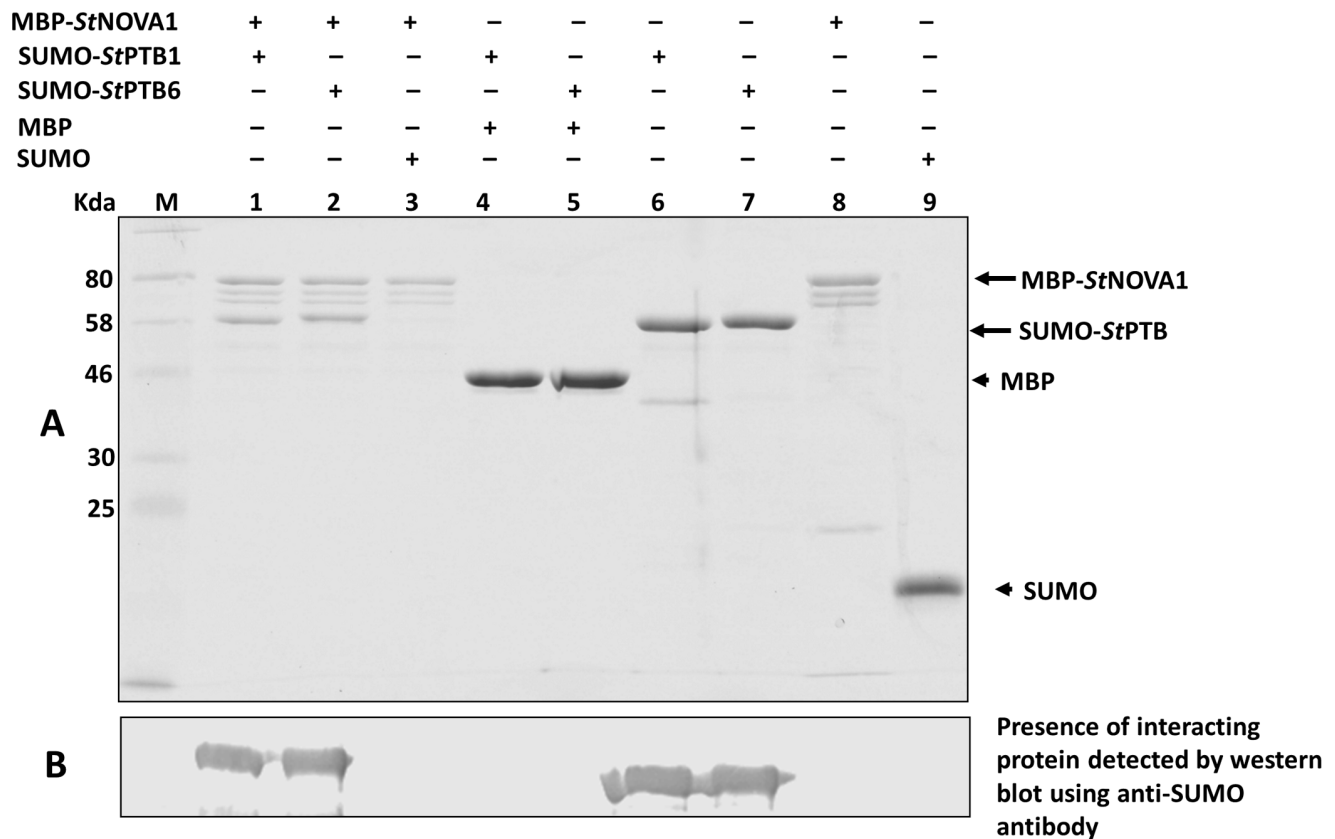


Figure 3. Interaction between *StPTB* and *StNova1* proteins demonstrated by pull-down assays. (A) Coomassie stained SDS PAGE showing interaction between MBP-*StNova1* protein with SUMO-*StPTB1* (lane 1) and SUMO-*StPTB6* (lane 2) while the controls MBP-*StNova1* and SUMO (lane 3) and MBP and *StPTB* proteins (lane 4 and 5) do not show any interaction. Lanes 6–9 shows the purified *StPTB1*, *StPTB6*, MBP-*StNova1* and SUMO proteins respectively used for interaction. (B) Western blot using anti-SUMO antibody demonstrating the presence of interacting protein in lanes 1, 2, 6 and 7.

doi:10.1371/journal.pone.0064783.g003

and map protein-protein interaction surfaces [22]. Typically, phage-display screening yields peptides that contain linear sequence/motifs that have the potential to interact with the target molecule. Although the probability of finding the exact sequence of a binding partner is rather low, consensus analysis facilitates the identification of natural candidates from the database. Thus, after four rounds of panning, phages were obtained that were specifically enriched for binding with *StPTB1* and *StPTB6*. One hundred clones were randomly selected, phagemids purified and checked for binding specificity against the respective target proteins in a phage-ELISA. Approximately 40 clones reacted positive for each protein and the derived peptide sequences are shown in Fig. 4. A single dominant peptide sequence, VNVQRYRMD**GVLGPWP**GYNLE (35 representations) recognized *StPTB6* (Fig. 4A, top line) and the dominant sequence recognizing *StPTB1* (11 representations) had the sequence LYN**SMPSILG**VWRPSTSRFPD (Fig. 4B, top line). From an analysis of all peptides recognizing both *StPTB1* and -6 (Fig. 4), a consensus binding peptide motif with the sequence [S/G][V/I][L/V]G was identified by WebLogo analysis [23]. Interestingly, in the case of *StPTB6*, the recognition peptide contained not only the consensus sequence but also harbored the sequence ‘GVLGxxP’ that is similar to the mammalian PTB binding motif [S/G][I/L]LGxxP present on Raver1 protein. However, a search of the plant database did not reveal any Raver-like protein. Intriguingly, an amino acid sequence scan of *StNova1* revealed the

sequence GIGGRYP located in the linker region between KH2 and KH3 domains that is similar to the GVLGPWP peptide identified in panning against *StPTB6*. It is worth noting that human Nova protein does not contain the [S/G][I/L]LGxxP like motif and, from the crystal structure of the Nova-RNA complex [24], it is inferred that the proline-alanine-glycine rich linker sequence between the KH2 and KH3 domains represents the protein-protein interaction surface.

Determination of *StPTB* and *StNova1* Protein Interaction Interface via DIPP

To further delineate the specific sites of interaction between *StPTB1/StPTB6* and *StNova1*, we used a modified novel phage display technique termed DIPP. As described in Materials and Methods, *StNova1* was divided into six different overlapping domains D1–D6 (Fig. 5), cloned into phagemid vector and homogeneous phage particles prepared for each domain. SUMO-*StPTB1* and SUMO-*StPTB6* (and SUMO control) were immobilized on an ELISA plate and panned against a “domain phage library” prepared by mixing equal concentrations of individual phage subdomains. Additionally, phage-ELISA was performed with the individual domains. In the experiment with individual domains it is evident that only D5 (residues 190–260 in the linker region between KH2 and KH3) is capable of binding to both the PTB proteins (Fig. 6, top and middle panels). Further confirmation of the specificity of this interaction was obtained by protein overlay

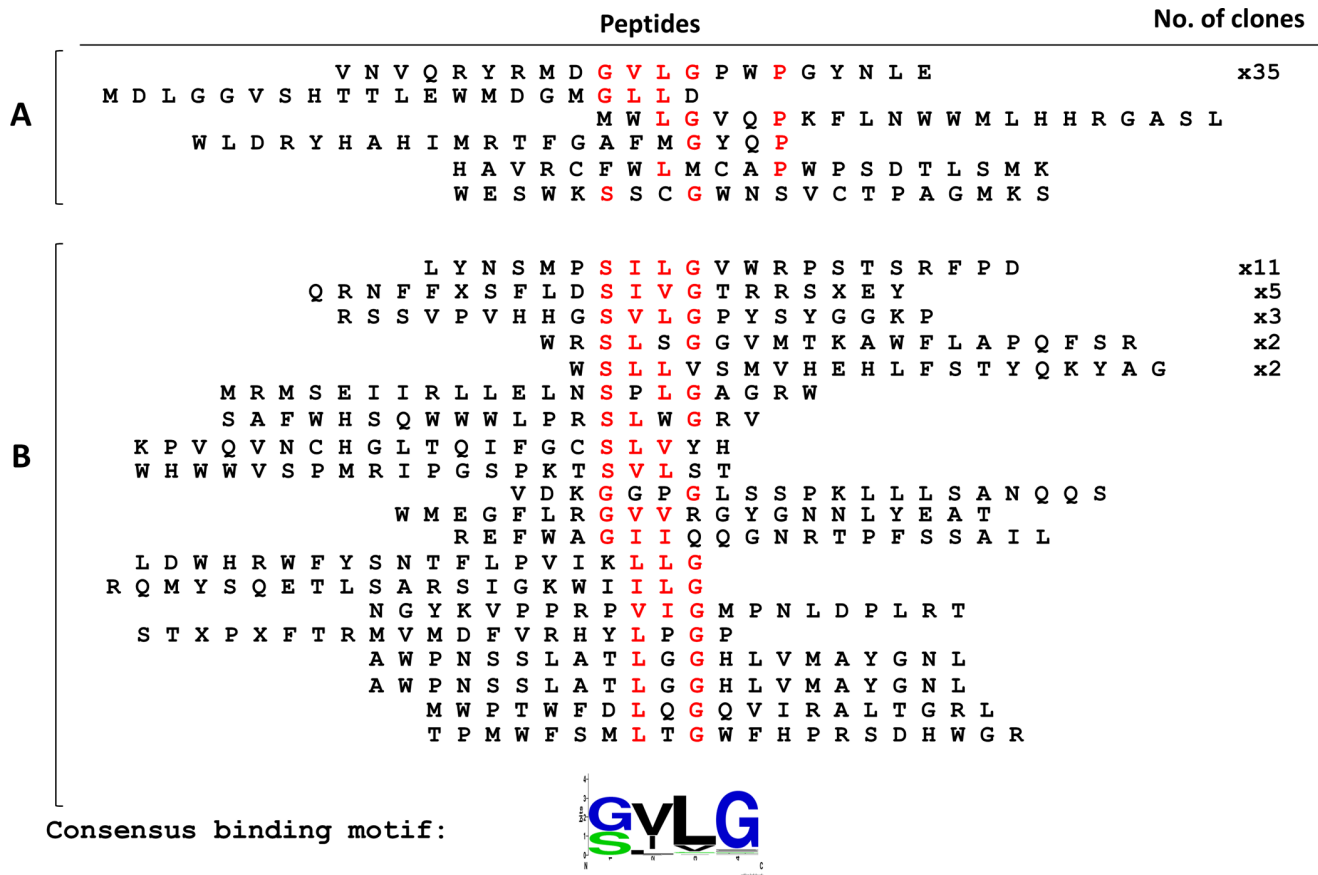


Figure 4. Phage-peptide library panning of StPTB proteins. (A) phage-peptides specifically enriched for binding with StPTB6 protein after four rounds of panning (B) phage-peptides specifically enriched for binding with StPTB1 protein after four rounds of panning. A consensus binding motif 'S/G[V/I][L/V]G' was identified for StPTB protein from these peptide sequence using Weblogo [23]. doi:10.1371/journal.pone.0064783.g004

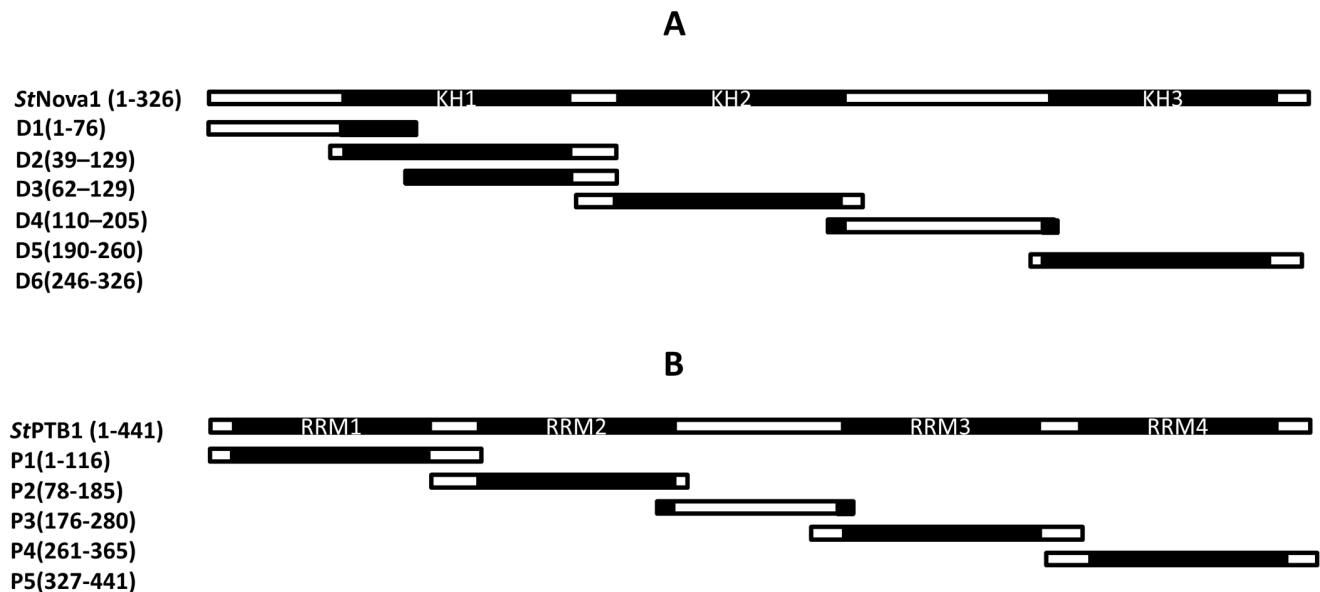


Figure 5. Schematic diagram of StNova1 (A) and StPTB1 proteins (B) showing the arrangement of their respective KH domains and RRM and their overlapping regions used to construct domain phage library. doi:10.1371/journal.pone.0064783.g005

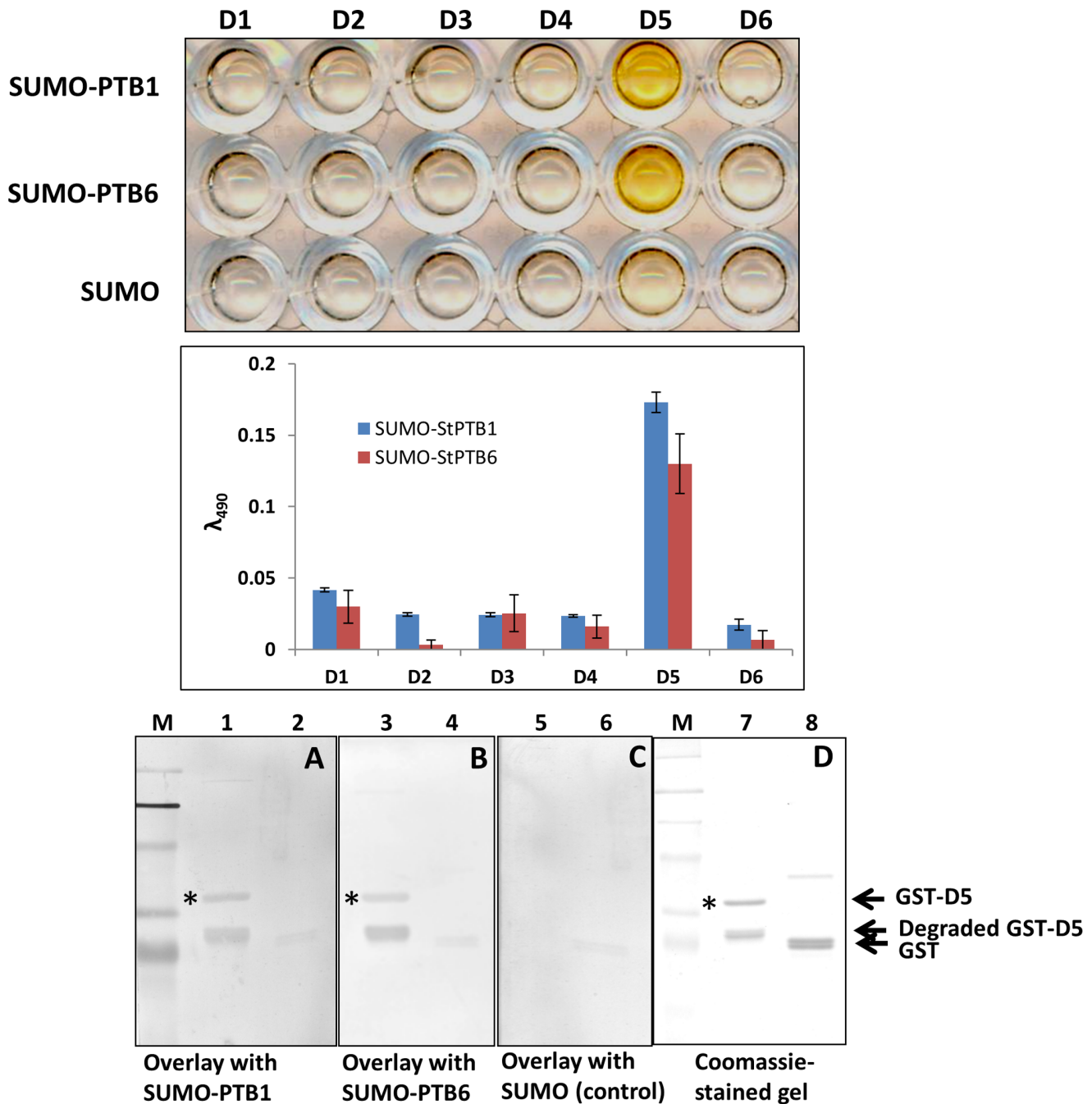


Figure 6. Mapping protein-protein interaction site on *StNova1* protein. (Top panel) Phage ELISA using *StNova1* single domain phage clones (D1 to D6) against SUMO-*StPTB1*, SUMO-*StPTB6* and SUMO control. (Middle panel) Bar diagram showing quantification of phage binding. (Bottom panel) **A.** SUMO-*StPTB1* protein overlay demonstrating its binding with GST-D5 region (lane 1, indicated with *), GST alone as a control shows no binding (lane 2) **B.** SUMO-*StPTB6* protein overlay demonstrating its binding with GST-D5 region (lane 3, indicated with *), GST alone as a control shows no binding (lane 4) **C.** SUMO protein (as a control) overlay does not bind to either GST-D5 (lane 5) or GST alone (lane 6) **D.** Coomassie stained SDS PAGE showing protein inputs, GST-D5 (lane 7, indicated with *) and GST (lane 8) used for blotting PVDF membrane for protein overlay experiments.
doi:10.1371/journal.pone.0064783.g006

assays. Towards this end, the D5 region of *StNova1* was expressed as GST fusion and purified GST-D5 and GST (control) were blotted onto a PVDF membrane. The membrane was then overlaid with SUMO-PTB1/PTB6 protein and SUMO protein (control) followed by detection with anti-SUMO antibody as described in Materials and Methods. *StPTB1* and *StPTB6* were able to specifically bind with D5 region of *StNova1* (Fig. 6, bottom

panel A and B). In a reciprocal experiment, five individual phage domains (P1–P5) and a domain library were prepared for the *StPTB1* protein (Fig. 5B) and experiments performed against immobilized GST-*StNova1*. With the individual subdomains, P3 (residues 176–280 in the linker region between RRM2 and RRM3) displayed the strongest specific binding (Fig. 7 top and middle panel). In a protein overlay assay, GST-D5 and GST

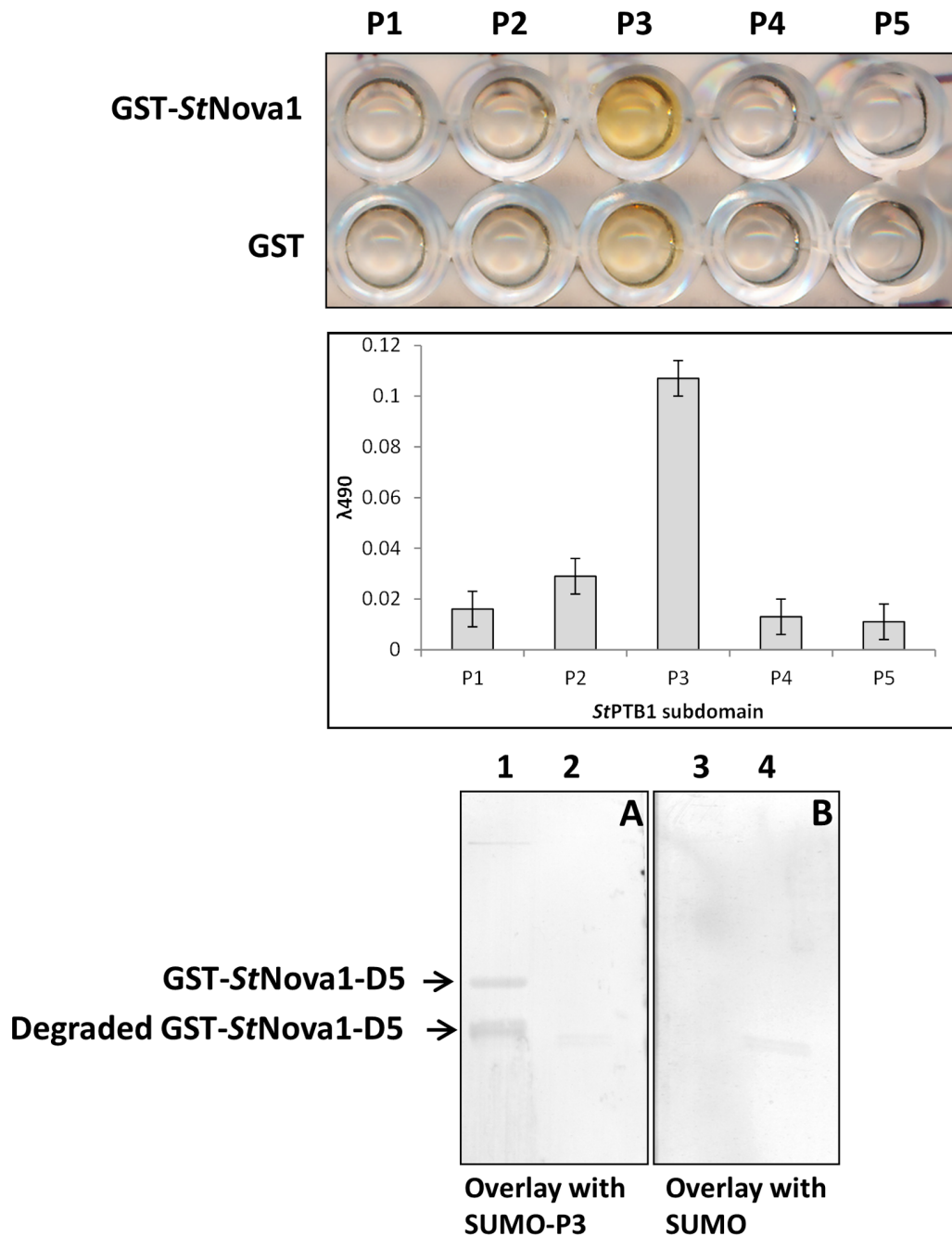


Figure 7. Mapping protein-protein interaction site on *StPTB1*. (Top) Phage ELISA using *StPTB1* single domain phage clones (P1 to P5) against GST-*StNova1* protein and GST control (Middle) Bar diagram showing quantification of phage binding (Bottom) **A**. SUMO-P3 protein overlay demonstrating binding to GST-D5 region (lane1), GST alone as a control shows no binding (lane2) **B**. SUMO protein (control) overlay shows no binding to either GST-D5 (lane 3) or GST alone (lane 4). doi:10.1371/journal.pone.0064783.g007

protein were blotted onto membrane and incubated with either SUMO-P3 or SUMO protein. SUMO-P3 bound to the GST-D5 of *StNova1* (Fig. 7, bottom panel A, lane 1) but not to GST (Fig. 7, bottom panel A, lane 2). No binding was observed with SUMO alone (Fig. 7, bottom panel B, lanes 3 and 4). This experiment affirmed the specific interaction between the P3 region of *StPTB1* and D5 of *StNova1*.

In parallel experiments using subdomain phage libraries, after two rounds of panning, 20 randomly selected colonies were sequenced. D5 in *StNova1* was exclusively represented as the

interaction partner for both *StPTBs* (Table 1) and P3 in *StPTB1* as the predominant interaction partner for *StNova1* (Table 1). These results clearly demonstrate the specificity of interaction between linker regions of both the proteins and also verify that the respective RNA-binding domains in the two proteins are not involved in the interaction.

Table 1. Domain interaction phage panning: Series of *StNova1* and *StPTB1* overlapping regions were cloned into phagemid vector and domain phage library was constructed for each proteins.

<i>StNova1</i> sub-domains	Frequency	
	<i>StPTB1</i>	<i>StPTB6</i>
D1	0	0
D2	0	0
D3	0	0
D4	0	0
D5	20	20
D6	0	0

<i>StPTB1</i> sub-domains	Frequency <i>StNova1</i>	
	<i>StPTB1</i>	<i>StPTB6</i>
P1	0	0
P2	2	2
P3	18	18
P4	0	0
P5	0	0

The *StPTB* proteins were screened against *StNova1* domain phage library and *StNova1* was screened against *StPTB1* domain phage library. After four rounds of panning 20 randomly selected clones were DNA sequenced. Table 1 shows the frequency of peptide selectively binding to *StPTB1* and *StNova1* proteins. doi:10.1371/journal.pone.0064783.t001

Identifying the Binding Site within *StNova1* Linker (D5) Region

To narrow down the binding sequence within D5, the linker region was further subdivided into three overlapping subdomains S1, S2 and S3 (Fig. 8A) and a secondary phage domain library was constructed using these regions. After two rounds of panning against *StPTB* proteins, twenty randomly selected colonies were DNA sequenced. In the case of *StPTB1*, out of twenty colonies, fourteen colonies represented S2 and six colonies were for S1, whereas with *StPTB6*, 15 colonies represented S2 and five colonies were for S1 (Fig. 8B). Additionally, phage-ELISA performed with homogeneous phage particles of the individual sub-domains showed that the S2 subdomain has the strongest binding affinity towards *StPTB* proteins that is comparable to the binding of the entire D5 domain (Fig. 8C). In comparison, subdomain S3 displays no affinity and subdomain S1 shows weak binding. Further validation of these interactions came from protein overlays assays (Fig. 9) in which the strongest binding by the PTBs again was observed against the GST fused S2 subdomain. No interaction was observed with S1 and S3.

The combination of DIPP and phage display screening thus suggest that the D5 region of *StNova1* and specifically the S2 subdomain within may be most important for binding with *StPTB* proteins. Interestingly, contained within the S2 subdomain is the glycine/proline-rich sequence **GMNYGPPNGIGGRYP** (Fig. 10A). This stretch of sequence has two motifs GxxxGxP (underlined) and GxxGxxP (double underlined) that are similar to the motif present in the peptide VNVQRYRMDGVLGPWP-GYNLE isolated in phage-panning experiments against *StPTB6*. Given the important role of Gly and Pro residues in protein-protein interactions [25] we checked the involvement of some of these residues in the S2 subdomain through three deletion mutants (Fig. 10A). In mutants S2Mu1 and S2Mu2, the deletions targeted

specific Gly and Pro residues within the GxxGxxP and GxxxGxP sequences respectively, whereas in S2Mu3, only the specific Gly residues in the entire S2 subdomain were deleted. The deletion mutants were cloned into phagemid vector and single clone phage particles were generated. Phage ELISA was performed with each of the homogeneous population of phages displaying the mutant subdomains against SUMO tagged *StPTB1*, *StPTB6* and SUMO protein as control (Fig. 10B). In comparison with the wild-type S2, a reduction in binding affinity was observed for all three mutants in the order Mu2 < Mu3 < Mu1.

Discussion

In RNA-binding proteins, the RRM domain is essential for both recognition and binding to specific target RNA molecules. One of the better characterized RNA-binding proteins in mammals, the PTB protein, is involved in numerous aspects of RNA metabolism including post-transcriptional processing and other distinct functions: mobility, stability, translation, and AS [2]. A newly discovered aspect of plant PTBs is their unique role as chaperones that transport RNA complexes [13]. Particularly well-studied is the role of mammalian PTB-proteins in AS. To a large extent, it is this biological context that has driven the structure and function analysis of PTB proteins and delineated the structural preferences of the RRMs in interactions with RNA. Simultaneously, structural studies have also underscored the mobility and dynamic nature of the linker regions in correctly orienting the RRMs towards specific substrates [26,27,28,29].

It has become increasingly evident, however, that in addition to direct RNA interactions, the function of PTB proteins can also be fine-tuned by interactions with other proteins. Thus, in studies on the splicing repressor domain in mammalian PTB protein, regions outside of RRM2 and proximal to the C-terminal linker were demonstrated to serve as sites for recruitment of other proteins that were independent of the RNA-binding activity [30]. On the other hand, in the case of the splicing repression of α -tropomyosin mediated by interactions of PTB with the protein Raver1, crystallographic studies show that the peptide motif [S/G][I/L]LGxxP in Raver1 interacts specifically with a hydrophobic groove in RRM2 but not with the linker following it [8,21]. The Raver1 protein contains four such motifs referred to as PTB-Raver1 interacting motifs (PRIs) that are essential for its function [8] and are also conserved in a related protein of unknown function, Raver2 [7]. Interestingly, however, in the yeast 2-hybrid analysis of both full-length and truncated constructs of Nova1 and Nova2 that interact with brain-enriched PTB (br-PTB), the spacer domain of Nova-2 (amino acids 230–407) was clearly established as the protein-protein interaction site that was sufficient and distinct from the RNA-binding sites [5,7,10].

In the present work we have validated the interaction between two RNA-binding proteins in potato, *StPTB1/StPTB6* and *StNova1*, through multiple techniques. The significant sequence homology of these plant proteins to their mammalian counterparts and the conserved domain architecture, which are also suggestive of conserved molecular mechanisms, is indeed revealed in our studies. We demonstrate that it is the ~65 amino-acid linker region (D5) between the KH2 and KH3 domains of *StNova1* that specifically binds to the *StPTBs* (Fig. 6). This region is analogous to the much longer spacer domain of mammalian Nova-2 comprising ~180 amino acids that is also sufficient for binding to brPTB [5]. More importantly, within the linker region of *StNova1*, the binding site can be narrowed further to the peptide sequence AAYLGM-NYGPPNGIGGRYPNNRYQ in subdomain S2 (Fig. 8). Reciprocally, it is the linker region (P3) comprising amino acid residues

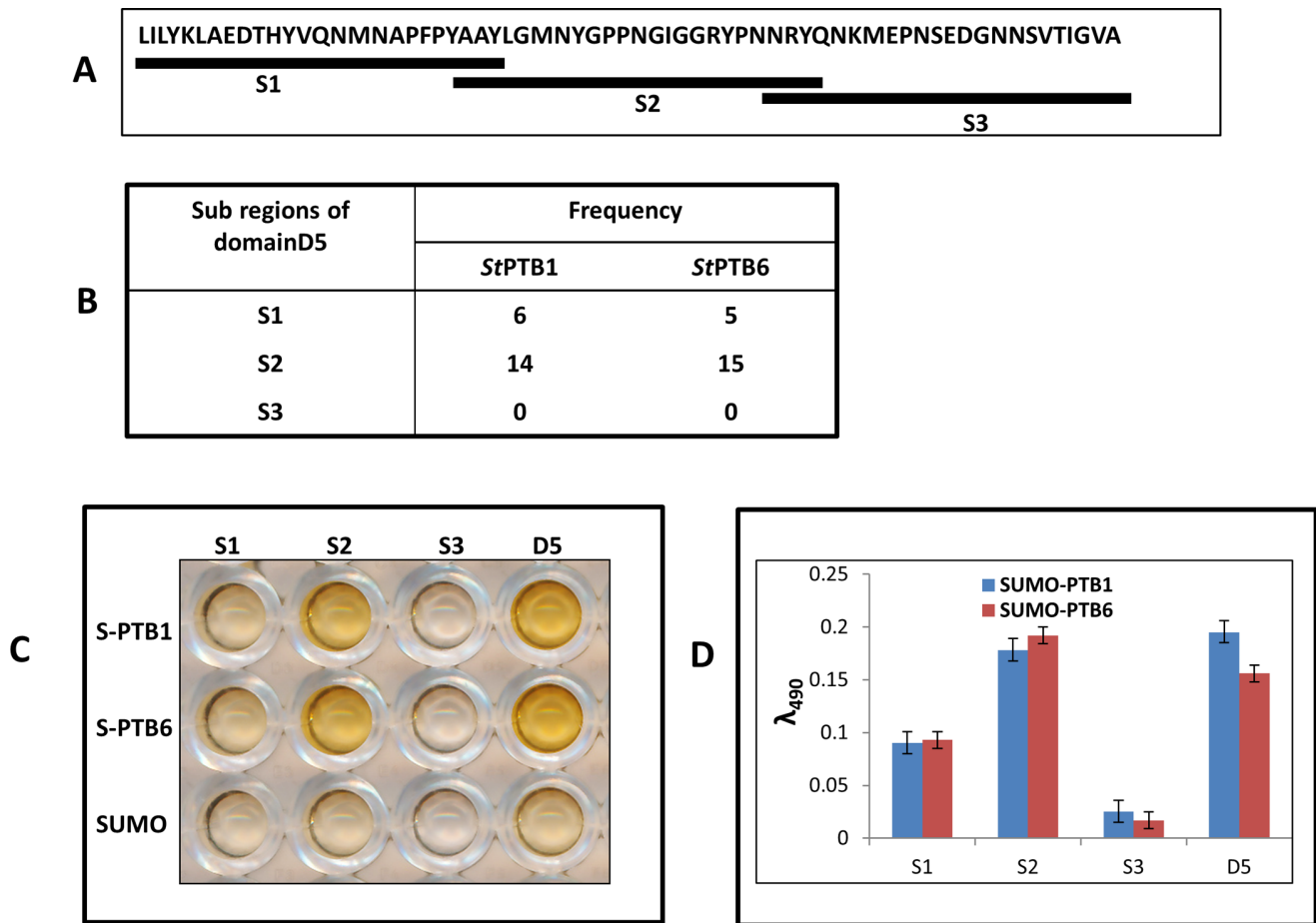


Figure 8. Determination of specific binding subdomain within *StNova1* D5. (A) Schematic showing the subdivision of D5 domain into three overlapping regions S1, S2 and S3 (B) Determination of specific binding region within D5 domain of *StNova1* by phage ELISA using phage clones of sub-domains S1, S2, S3 and domain D5 (C) Phage ELISA using phage clones of Domain D5 and sub-domains S1, S2, S3 against SUMO-*StPTB1*, SUMO-*StPTB6* and SUMO control (D) Bar diagram showing quantification of phage binding. doi:10.1371/journal.pone.0064783.g008

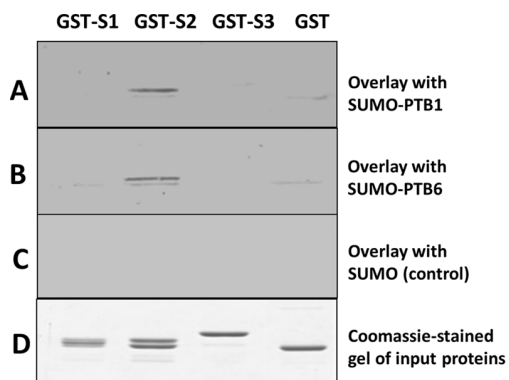


Figure 9. Binding of D5 sub-domains S1, S2 and S3 to *StPTB* proteins by protein overlay. GST-fused S1, S2 and S3 domains separated on SDS-PAGE were blotted on to PVDF membrane and probed with (A) SUMO-*StPTB1* protein (B) SUMO-*StPTB6* protein and (C) SUMO protein. Specific binding to the S2 domain is seen with the SUMO-PTB proteins (A and B) but not with SUMO (C). (D) Coomassie stained SDS PAGE of pure proteins. Binding was detected using anti-SUMO primary antibody, alkaline-phosphatase (AP) labeled secondary antibody and the AP substrate kit for color development as described in Materials & Methods. doi:10.1371/journal.pone.0064783.g009

176–280 of *StPTB1* (Fig. 5) that binds to *StNova1* (Fig. 7, top panel) specifically in the D5 region (Fig. 7, bottom panel). Thus, the linker regions in both proteins, *StPTB1/StPTB6* and *StNova1*, constitute the “hotspots” of the protein-protein interaction interface while leaving their respective RRM domains free to simultaneously bind RNA and participate in other functions.

Hotspots are defined as essential amino acids that constitute a small subset of interface residues that are necessary and sufficient for binding between two proteins [31,32]. In addition to predictive computational methods and alanine scanning mutagenesis to locate these sites, screening with a population of random peptides, such as those displayed on the surface of phage, has proven to be a powerful method for identifying “consensus” sequence motifs that mimic protein-protein interaction surfaces [22]. In many instances, peptide ligands isolated from phage-peptide libraries show conservation of such motifs with native interacting proteins, facilitating the identification of natural candidate proteins [33,34,35]. Along these lines, it is interesting that screening with a random, untrained phage-peptide library yields a single dominant peptide, VNVQRYRMDGVLGPWPGYNLE, that binds to *StPTB6* (Fig. 4). Intriguingly, this peptide contains the sequence ‘GVLGPWP’ which is strikingly similar to the peptide motif ‘[S/G][I/L]LGxxP’ in Raver1 protein that is essential for interaction with mammalian PTB. A similar motif, GIGGRYP,

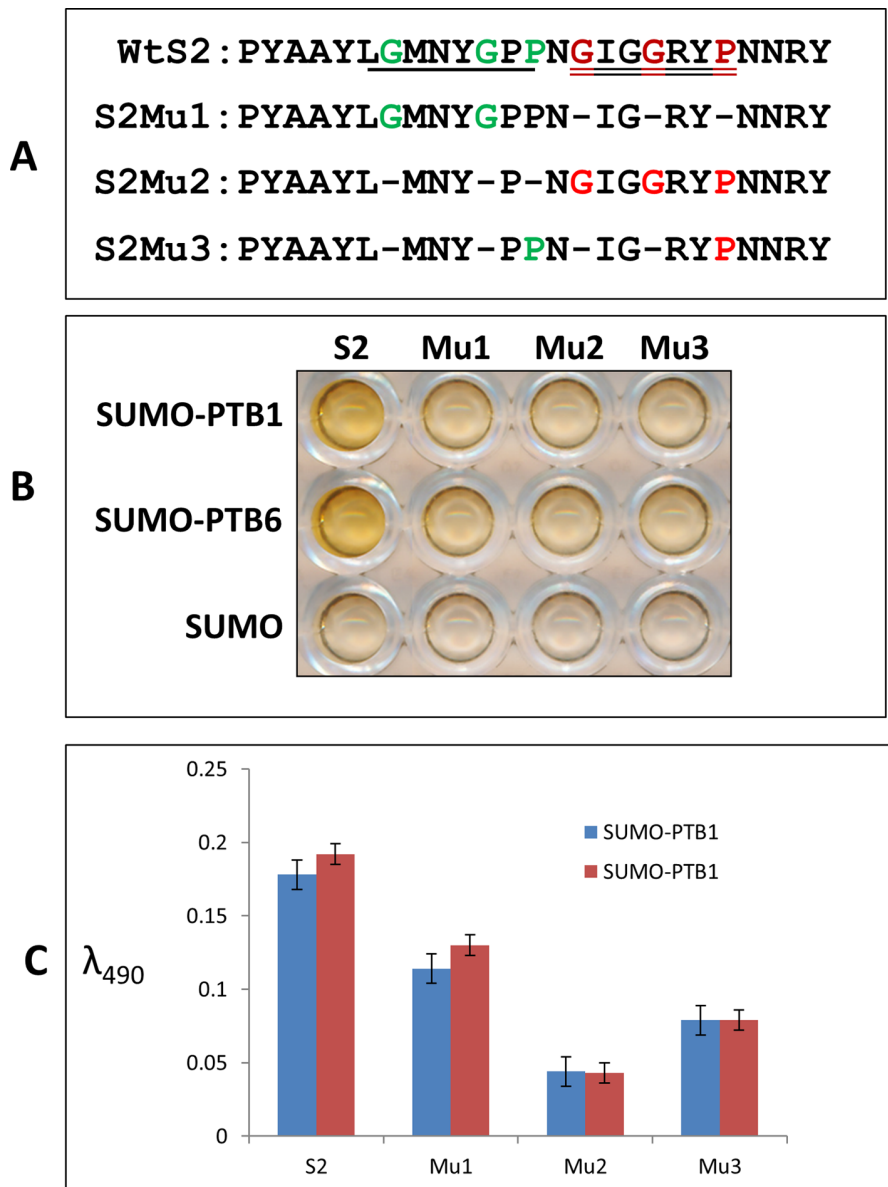


Figure 10. Mutational analysis to show that the glycine and proline rich motif of S2 region of *SfNova1* is involved in binding with *SfPTB* proteins. (A) Peptide sequence for S2 and deletion mutant S2Mu1, S2Mu2 and S2Mu3. The motifs 'GxxxGxP' and 'GxxGxxP' in the S2 peptide are underlined. (B) Phage ELISA showing that the deletion mutants of S2 have less affinity towards *SfPTB* proteins compared to S2. (C) Bar diagram showing quantification of phage binding. doi:10.1371/journal.pone.0064783.g010

occurs in the sequence AAYLGMNYGPPNGIGGRYPNNRYQ in the S2 region of *SfNova1* and specifically interacts with *SfPTB* (Fig. 8). The significant decrease in binding efficiency that occurs upon deletion of the Gly and Pro residues within the GIGGRYP and GMNYGPP sequences, (Fig. 10) illustrates the critical dependence of just a few amino acids that contribute to the energetics of interactions within a hotspot region. It is also interesting to note that the sequence alignment of the spacer region between the KH2 and KH3 domains of Nova-like proteins from other plant species shows high similarity (Fig. S1). This suggests that, in general, the Gly-rich sequence within the spacer region of plant Nova-like proteins may also serve as a site for protein-protein interaction. Perhaps, more importantly, the conservation of a protein interaction motif, GxxGxxP in two

unrelated and non-homologous proteins, mammalian Raver1 and plant *SfNova1*, is illustrative of the plasticity of residues in the hotspot to adapt to different structural contexts imposed by dissimilar sequences, while retaining the ability to perform similar functions of molecular recognition [25,36]. However, in the case of human Nova the linker region does not contain any similar motifs and potential interactions with other proteins may depend on determinants specified by the sequence. It is also worth noting that the amino acid sequence of the linker region (P3) between RRM2 and RRM3 in other plant PTB-like proteins is highly conserved (Fig. S2), suggesting also a possible conservation of function as a protein interaction interface in these proteins.

In conclusion, our studies indicate that the interaction between the RNA-binding proteins *SfPTB1/6* and *SfNova1* is mediated

through linker regions that are distinctly separated from the RRM. While the molecular details of this interaction and the specific involvement of the RRM will have to await atomic level resolution of the structure with and without RNA, one can gain some understanding by drawing upon the extensive structural investigations of mammalian PTBs in complex with RNA and proteins and their mechanistic implications in AS [2,27,29]. Such studies provide strong evidence for the formation of ternary complexes driven by cooperative interactions involving all the RRM and the linker regions. Therefore, given the multiple RRM and separate (or overlapping) protein interaction surfaces in *St*PTB and *St*Nova1, it is entirely reasonable that they can provide the scaffolding for macromolecular assembly of specific proteins and RNA for multidimensional regulation of various processes. Included among these functions is the potential ability to facilitate long-distance RNA transport. To this extent, identification of specific protein-binding regions outside of the RRM suggests independent functional interaction of these domains within the PTB/Nova complex in plants. Support for this comes from the characterization of a large ribonucleoprotein complex in pumpkin that binds phloem-mobile RNAs. Mediated by the 50 kDa PTB protein, RBP50, this complex contains six RNAs and sixteen other proteins [13]. Although no Nova-like protein was identified in this complex, a KH-domain protein very similar in sequence to *St*Nova1 was identified in the phloem sap of pumpkin [37], suggesting a role for Nova proteins in long-distance transport of mRNAs. *St*PTB1 and -6 are the potato orthologs of the pumpkin RBP50 and have been implicated in interactions with full-length phloem-mobile mRNAs that function in plant development [38]. One of these, *St*BEL5, is a phloem-mobile RNA that is involved in regulating tuber formation [39]. Coupled to the fact that both *St*PTB1 and -6 can also bind to CU-rich sequences in the 3' untranslated region of *St*BEL5, a plausible argument can be

made for the existence of a RBP50-like complex in potato similarly involved in long-distance mobile RNA transport. We are currently pursuing studies to establish the identity of this complex and its components.

Supporting Information

Figure S1 Sequence alignment of linker region between KH2 and KH3 domain of plant Nova-like proteins.

Sequence identification number appears in brackets following the species name.
(DOCX)

Figure S2 Sequence alignment of linker region between RRM2 and RRM3 domain of plant PTB like proteins.

Sequence identification number appears in brackets following the species name.
(DOC)

Table S1 List of forward and reverse primers.

(DOC)

Acknowledgments

We thank Dr. Yasufumi Yamamoto for providing the random 21-mer phage peptide library and to Dr. Sung Ki Cho for reconfirming the yeast two-hybrid data.

Author Contributions

Conceived and designed the experiments: SS AGR DH. Performed the experiments: SS NMB. Analyzed the data: SS AGR DH. Contributed reagents/materials/analysis tools: SS AGR. Wrote the paper: SS AGR DH.

References

- Oberstrass FC, Auweter SD, Erat M, Hargous Y, Henning A, et al. (2005) Structure of PTB bound to RNA: specific binding and implications for splicing regulation. *Science* 309: 2054–2057.
- Auweter SD, Allain FH (2008) Structure-function relationships of the polypyrimidine tract binding protein. *Cell Mol Life Sci* 65: 516–527.
- Spellman R, Rideau A, Matlin A, Gooding C, Robinson F, et al. (2005) Regulation of alternative splicing by PTB and associated factors. *Biochem Soc Trans* 33: 457–460.
- Sawicka K, Bushell M, Spriggs KA, Willis AE (2008) Polypyrimidine-tract-binding protein: a multifunctional RNA-binding protein. *Biochem Soc Trans* 36: 641–647.
- Polydorides AD, Okano HJ, Yang YY, Stefani G, Darnell RB (2000) A brain-enriched polypyrimidine tract-binding protein antagonizes the ability of Nova to regulate neuron-specific alternative splicing. *Proc Natl Acad Sci U S A* 97: 6350–6355.
- Kleinhenz B, Fabienke M, Swiniarski S, Wittenmayer N, Kirsch J, et al. (2005) Raver2, a new member of the hnRNP family. *FEBS Lett* 579: 4254–4258.
- Henneberg B, Swiniarski S, Sabine B, Illenberger S (2010) A conserved peptide motif in Raver2 mediates its interaction with the polypyrimidine tract-binding protein. *Exp Cell Res* 316: 966–979.
- Joshi A, Coelho MB, Kotik-Kogan O, Simpson PJ, Matthews SJ, et al. (2011) Crystallographic analysis of polypyrimidine tract-binding protein-Raver1 interactions involved in regulation of alternative splicing. *Structure* 19: 1816–1825.
- Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, et al. (2010) Regulation of alternative splicing by histone modifications. *Science* 327: 996–1000.
- Huttelmaier S, Illenberger S, Grosheva I, Rudiger M, Singer RH, et al. (2001) Raver1, a dual compartment protein, is a ligand for PTB/hnRNP1 and microfilament attachment proteins. *J Cell Biol* 155: 775–786.
- Gromak N, Rideau A, Southby J, Scadden AD, Gooding C, et al. (2003) The PTB interacting protein raver1 regulates alpha-tropomyosin alternative splicing. *EMBO J* 22: 6356–6364.
- Fujisaki K, Ishikawa M (2008) Identification of an Arabidopsis thaliana protein that binds to tomato mosaic virus genomic RNA and inhibits its multiplication. *Virology* 380: 402–411.
- Ham BK, Brandom JL, Xoconostle-Cazares B, Ringgold V, Lough TJ, et al. (2009) A polypyrimidine tract binding protein, pumpkin RBP50, forms the basis of a phloem-mobile ribonucleoprotein complex. *Plant Cell* 21: 197–215.
- Wachter A, Ruhl C, Stauffer E (2012) The Role of Polypyrimidine Tract-Binding Proteins and Other hnRNP Proteins in Plant Splicing Regulation. *Front Plant Sci* 3: 81.
- Ruhl C, Stauffer E, Kahles A, Wagner G, Drechsel G, et al. (2012) Polypyrimidine tract binding protein homologs from Arabidopsis are key regulators of alternative splicing with implications in fundamental developmental processes. *Plant Cell* 24: 4360–4375.
- Stauffer E, Westermann A, Wagner G, Wachter A (2010) Polypyrimidine tract-binding protein homologues from Arabidopsis underlie regulatory circuits based on alternative splicing and downstream control. *Plant J* 64: 243–255.
- Wang S, Okamoto T (2009) Involvement of polypyrimidine tract-binding protein (PTB)-related proteins in pollen germination in Arabidopsis. *Plant Cell Physiol* 50: 179–190.
- Butler NM, Hannapel DJ (2012) Promoter activity of polypyrimidine tract-binding protein genes of potato responds to environmental cues. *Planta* 236: 1747–1755.
- Chen H, Rosin FM, Prat S, Hannapel DJ (2003) Interacting transcription factors from the three-amino acid loop extension superclass regulate tuber formation. *Plant Physiol* 132: 1391–1404.
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248–254.
- Rideau AP, Gooding C, Simpson PJ, Monie TP, Lorenz M, et al. (2006) A peptide motif in Raver1 mediates splicing repression by interaction with the PTB RRM2 domain. *Nat Struct Mol Biol* 13: 839–848.
- Sidhu SS, Fairbrother WJ, Deshayes K (2003) Exploring protein-protein interactions with phage display. *ChemBiochem* 4: 14–25.
- Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14: 1188–1190.
- Lewis HA, Chen H, Edo C, Buckanovich RJ, Yang YY, et al. (1999) Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding domains. *Structure* 7: 191–203.

25. Moreira IS, Fernandes PA, Ramos MJ (2007) Hot spots-A review of the protein-protein interface determinant amino-acid residues. *Proteins-Structure Function and Bioinformatics* 68: 803–812.
26. Conte MR, Grune T, Ghuman J, Kelly G, Ladas A, et al. (2000) Structure of tandem RNA recognition motifs from polypyrimidine tract binding protein reveals novel features of the RRM fold. *EMBO J* 19: 3132–3141.
27. Simpson PJ, Monic TP, Szendroi A, Davydova N, Tyzack JK, et al. (2004) Structure and RNA interactions of the N-terminal RRM domains of PTB. *Structure* 12: 1631–1643.
28. Vitali F, Henning A, Oberstrass FC, Hargous Y, Auweter SD, et al. (2006) Structure of the two most C-terminal RNA recognition motifs of PTB using segmental isotope labeling. *EMBO J* 25: 150–162.
29. Clerie C, Hall KB (2009) The domains of polypyrimidine tract binding protein have distinct RNA structural preferences. *Biochemistry* 48: 2063–2074.
30. Robinson F, Smith CW (2006) A splicing repressor domain in polypyrimidine tract-binding protein. *J Biol Chem* 281: 800–806.
31. DeLano WL (2002) Unraveling hot spots in binding interfaces: progress and challenges. *Curr Opin Struct Biol* 12: 14–20.
32. Ofran Y, Rost B (2007) Protein-protein interaction hotspots carved into sequences. *Plos Computational Biology* 3: 1169–1176.
33. DeLano WL, Ultsch MH, de Vos AM, Wells JA (2000) Convergent solutions to binding at a protein-protein interface. *Science* 287: 1279–1283.
34. Smothers JF, Henikoff S (2000) The HP1 chromo shadow domain binds a consensus peptide pentamer. *Curr Biol* 10: 27–30.
35. Kay BK, Kasanov J, Knight S, Kurakin A (2000) Convergent evolution with combinatorial peptides. *FEBS Lett* 480: 55–62.
36. Lise S, Archambeau C, Pontil M, Jones DT (2009) Prediction of hot spot residues at protein-protein interfaces by combining machine learning and energy-based methods. *BMC Bioinformatics* 10.
37. Lin MK, Lee YJ, Lough TJ, Phinney BS, Lucas WJ (2009) Analysis of the pumpkin phloem proteome provides insights into angiosperm sieve tube function. *Mol Cell Proteomics* 8: 343–356.
38. Mahajan A, Bhogale S, Kang IH, Hannapel DJ, Banerjee AK (2012) The mRNA of a Knotted1-like transcription factor of potato is phloem mobile. *Plant Mol Biol* 79: 595–608.
39. Banerjee AK, Chatterjee M, Yu Y, Suh SG, Miller WA, et al. (2006) Dynamics of a mobile RNA of potato involved in a long-distance signaling pathway. *Plant Cell* 18: 3443–3457.