Protein Targeting to Starch Synthase 1, a Functional Protein of Starch Biosynthesis in Wheat (Triticum Aestivum L.)

Vinita Sharma
National Agri-Food Biotechnology Institute: National Agri-Food Biotechnology Institute
Vikas Fandade
National Agri-Food Biotechnology Institute
Prashant Kumar
National Agri-Food Biotechnology Institute: National Agri-Food Biotechnology Institute
Afsana Parveen
National Agri-Food Biotechnology Institute: National Agri-Food Biotechnology Institute
Akansha Madhawan
National Agri-Food Biotechnology Institute
Manik Bathla
national agri food biotechnology instittute
Ankita Mishra
National Agri-Food Biotechnology Institute
Himanshu Sharma
National Agri-Food Biotechnology Institute
Vikas Rishi
National Agri-Food Biotechnology Institute
Santosh Satbhai
Indian Institute of Science Education and Research Mohali
Joy Roy (joykroy@nabi.res.in)
National Agri-Food Biotechnology Institute https://orcid.org/0000-0001-5441-3538

Research Article

Keywords: Wheat grain, Starch, Starch synthase, GBSSI, Carbohydrate-binding domain, coiled-coil domain, protein-protein interaction.

Posted Date: November 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1033032/v1

Version of Record: A version of this preprint was published at Plant Molecular Biology on March 24th, 2022. See the published version at https://doi.org/10.1007/s11103-022-01260-1.

Abstract

In cereal endosperm, native starch comprising amylose and amylopectin is synthesized by the coordinated activities of several pathway enzymes. Amylose in starch influences its physio-chemical properties resulting in several human health benefits. The **G**ranule-**B**ound **S**tarch **S**ynthase I (GBSSI) is the most abundant starch-associated protein. GBSSI lacks dedicated Carbohydrate-binding module (CBM). Previously, Protein Targeting Starch Synthase 1 (PTST1) was identified as a crucial protein for the localization of GBSSI to the starch granules in *Arabidopsis*. The function of its homologous protein in the wheat endosperm is not known. In this study, TaPTST1, an AtPTST1 homolog, containing a CBM and a coiled-coil domain was identified in wheat. Protein-coding nucleotide sequence of TaPTST1 from Indian wheat variety 'C 306' was cloned and characterized. Homology modelling and molecular docking suggested the potential interaction of TaPTST1 with glucans and GBSSI. The TaPTST1 expression was higher in wheat grain than the other tissues, suggesting a grain-specific function. *In vitro* binding assays demonstrated different binding affinities of TaPTST1 for native starch, amylose, and amylopectin. Furthermore, the immunoaffinity pull-down assay revealed that TaPTST1 directly interacts with GBSSI, and the interaction is mediated by a coiled-coil domain. The direct protein-protein interaction was further confirmed by bimolecular fluorescence complementation assay (BiFC) *in planta*. Based on our findings we postulate a functional role for TaPTST1 in starch metabolism by targeting GBSSI to starch granules in wheat endosperm.

Key Message

TaPTST1, a wheat homolog of AtPTST1 containing CBM can interact with GBSSI and regulate starch metabolism in wheat endosperm.

Introduction

Starch is a plant-derived carbohydrate and a significant source of calories for humans and livestock. In wheat, grain starch accounts for nearly 70% of the dry weight and is the key determinant of its nutritional quality and yield (Shewry and Hey 2015). The plant diurnal dynamics enable plants to form two types of starch, namely transient and storage starch. The transient starch is synthesized in leaves during the daytime and gets metabolized as a carbon source for non-photosynthetic tissues during the night. The storage starch gets deposited in the storage tissue like endosperm in cereal seeds and tubers of different crops and serves as the carbon reserve for germination (Bahaji et al. 2014). Starch is semi-crystalline granules, composed of two types of glucose polymers amylose and amylopectin. Amylose is mainly a linear polymer of glucose linked by α-1,4-glycosidic linkage, whereas amylopectin is a highly branched molecule (Pérez and Bertoft 2010).

The starch biosynthesis in cereal grain is resultant of the coordinated activities of different enzymes including ADP-glucose pyrophosphorylase (AGPase), soluble starch synthase (SSs), granule-bound starch synthase (GBSS) starch branching and debranching enzymes, and starch phosphorylase. The

detailed starch biosynthesis pathway in cereals is reviewed elsewhere (Keeling and Myers 2010; Tetlow 2018). Briefly, the first committed step towards the starch biosynthesis is the AGPase reaction resulting in the ADP-glucose, a substrate for starch synthases, that is transported to amyloplast by Bt1 transporter (Ghosh and Preiss 1966; Beckles et al. 2001). Two isoforms of GBSS are expressed in a tissue-specific manner. GBSSI in the amyloplast solely elongates the linear glucose chain by transferring glucose moiety using ADPG as the substrate. The other isoform GBSSII is responsible for amylose biosynthesis in leaves contributing to the transitory starch (Vrinten and Nakamura 2000). Soluble starch synthase (SSs) is involved in the biosynthesis of amylopectin. SSI is involved in the elongation of the short-chain of glucans of DP 6-10, while SSII is responsible for the synthesis of chain DP 12-20 (Delvallé et al. 2005; Zhang et al. 2008). An earlier study in *Arabidopsis* reported the involvement of SSIII in the synthesis of longer glucan chains in leaves (Zhang et al. 2005). SSIV is required for the initiation of granules rather than the elongation of glucan chains (Roldán et al. 2007).

The proportion of amylose and amylopectin in starch largely influences its physio-chemical properties and nutritional quality by virtue of structural differences between amylose and amylopectin. The health benefits associated with starch rich in amylose have been studied and discussed elsewhere (Birt et al. 2013; Li et al. 2019; Lim et al. 2020).

GBSS is the only starch synthase responsible for the synthesis of longer amylose chain. The knockout or loss of function mutants of GBSSI is reported to synthesize amylose-free (waxy) starch in different species, like hexaploid wheat (Nakamura et al. 1998), rice (Wang et al. 1995), potato (Hovenkamp-Hermelink et al. 1987) and maize (Shure et al. 1983).

In cereals, amylose biosynthesis lags behind the amylopectin biosynthesis during seed development (Tomlinson and Denyer 2003; Radchuk et al. 2009). The amylose biosynthesis genes are mainly upregulated in the mid and later stages of seed development, unlike the expression of branching enzymes and soluble starch synthases which are higher during early stages of grain filling. The temporal arrangement of the amylose and amylopectin biosynthesis is hypothesized to be crucial since amylose biosynthesis occurs within the amylopectin matrix (Tatge et al. 1999; Tomlinson and Denyer 2003).

However, the relationship between the overexpression of GBSSI and amylose content is not linear in maize and potato (Tsai 1974; Flipse et al. 1996). GBSSI overexpression in wheat did not affect amylose content substantially (Sestili et al. 2012), suggesting the involvement of other factors associated with amylose biosynthesis. GBSS is one of the most abundant starch-bound proteins closely associated with starch granules. The enzyme is highly unstable in the soluble fraction (Seung et al. 2015; Hebelstrup et al. 2017). The enzyme lacks any dedicated CBM for its localization to the starch granules, a requisite to the amylose biosynthesis. A non-catalytic carbohydrate-binding protein PTST1 was identified as critical to the localization of GBSS in *Arabidopsis* leaves. The c-terminal catalytic domain of GBSS interacts with a protein PTST via coiled-coil domain and the interaction is crucial for the localization of enzyme to the starch granules. PTST was identified as a non-catalytic protein pertaining to CBM48 and the coiled-coil domain in *Arabidopsis* (Seung et al. 2015). The mechanism of starch biosynthesis in leave and

endosperm differs (Tetlow et al., 2017). In rice, a GBSS binding protein, namely GBP was identified to affect the quality and yield of rice (Wang et al. 2020). A PTST1 homolog was found to be involved in targeting HvGBSSI and normal starch biosynthesis in barley endosperm (Zhong et al. 2019). However, there is no report of GBSS binding proteins and its targeting to starch granules in wheat endosperm and leaves. Despite the availability of whole-genome sequences of *Triticum aestivum*, there are many genes with no assigned functions. In this study, we used molecular and biochemical approaches to identify and characterize the homolog of AtPTST1 in wheat endosperm. The direct protein-protein interaction between wheat GBSSI and PTST was identified *in vitro* and *in planta*. Our study suggests TaPTST1 as a GBSS binding protein with a possible role in starch metabolism by ushering GBSSI to starch granules in wheat endosperm.

Materials And Methods

Plant materials

The Indian bread wheat variety ‘C 306’ was planted during the Rabi season 2019-2020 at the research farm of the National Agri-Food Biotechnology Institute, Mohali, India. The primary spikes from the individual plant in three biological replicates were harvested at 7, 14, and 21 DPA (days post anthesis) and stored at -80°C, till further use. For the bimolecular fluorescence complementation assay (BiFC) assay, *Nicotiana. benthamiana* wild-type plants were grown at 25 °C with a photoperiod of 10hrs Light/14hrs dark in the plant growth chamber.

Phylogenetic and bioinformatic analysis

The amino acid sequence of AtPTST1 was used to search EnsemblPlants database (https://plants.ensembl.org/Triticum_aestivum/Info/Index). The amino acid sequences of orthologous genes were retrieved with the highest similarity with AtPTST1. The alignment of the sequence was created using MEGA X (Kumar et al. 2018). The neighbour-joining method was used to construct the phylogenetic tree with the following parameters: p-distance model, pairwise deletion, and bootstrap (1,000 replicates). The alignment used for phylogenetic tree construction is provided in DataS1. The coiled-coil domain was predicted by the COILS/PCOILS server (https://embnet.vital-it.ch/). The domain analysis was performed using the SMART webserver (Letunic et al. 2021).

Cloning and sequencing of TaPTST1

The cDNA libraries were prepared from the Indian wheat variety ‘C 306’ from the pooled samples of total RNA isolated from wheat grains at 7, 14 and 21 DAP. Full-length TaPTST1 protein-coding nucleotide sequence of 879 bp was amplified from cDNA by using primers (Table S1) and cloned to pJET1.2 cloning vector. A total of 30 clones were sequenced by the Sanger sequencing method ABI3730XL instrument. Sequences were aligned to ‘Chinese spring’ reference sequences retrieved from the ensemble plant using multialign (Corpet 1988).

Homology modelling and docking

Homology model of wheat PTST_AMPK domain was built using I-TASSER webserver, and TOME V3 Server (Pons and Labesse 2009; Yang and Zhang 2015). The model was validated using various validation tools like ERRAT plot, Ramachandran plot, ProSA, ProQ, and verify-3D. The COACH server was used to predict protein-ligand binding sites (Yang et al. 2013a; Yang et al. 2013b). For homology modelling of CC structure of TaPTST1 Swiss model was used and molecular docking of TaPTST1 coiled-coil domain and GBSSI was performed by the HDOCK webserver (Yan et al. 2017; Yan et al. 2020).

Quantitative real-time PCR

Total RNA was extracted from the grains collected at different seed developmental stages using TRIzol method (Simms et al. 1993)(Simms et al. 1993). RNA concentration was measured on the Nano Droplite (Thermo Fisher Scientific, USA) and the integrity of RNA was checked on the 1.5% agarose gel. Genome-specific primers of the genes were designed using Prime 3.0 software. cDNA was prepared using 'SuperScript™ III First-Strand Synthesis System' following the manufacture's protocol. The relative quantitative gene expression analysis was conducted in three technical replicates on the CFX96 thermocycler (BIO-RAD, USA) system. For normalization of gene expression, the wheat *tubulin gene* was used as an internal control. All the primers used in the study are listed in Table S1.

Cloning and recombinant protein expression in *E. coli*

The full-length CDS of *TaPTST1* was amplified using forward and reverse primers (Table S1) and CDS of GBSSI was previously cloned (Mishra et al. 2021). For each of the genes, 5 clones were sequenced and aligned with sequence retrieved from the EnsemblPlants database.

A 1605-bp nucleotide fragment coding for mature GBSSI protein but lacking 70 amino acids predicted transit peptide was cloned in pET28a expression vector using EcoRI and XhoI restriction endonucleases. For the full-length PTST expression, the expression vector was constructed using BamHI and SacI restriction enzymes. The N-terminal fragment of wheat PTST was cloned into the pET28a expression vector using EcoRI and XhoI enzymes to create PTST-N1-209 with N-terminal His-tag. Recombinant proteins with N-terminal His-tag were expressed in *E. coli* Rosetta cells harbouring each of the IPTG-inducible expression construct i.e., pET28a-TaGBSSI, pET28a-TaPTST1, and pET28a-TaPTST-N1-209. Cells were induced using 0.2 mM IPTG and grown in LB medium at 16 °C with constant agitation for 16 hours. 0.5 M sorbitol was used in the culture media to promote protein solubilization (Prasad et al. 2011). Cell lysis and protein extraction was performed using an earlier described method (Spriestersbach et al. 2015). Proteins were purified by Ni-NTA affinity chromatography using standard protocol followed by desalting. GBSSI and TaPTST1 were eluted at the concentration of 100 and 150mM imidazole in 100mM Tricine buffer. Purified proteins were detected by immunoblotting with anti-His antibodies.

Starch binding assay

The starch binding assay was performed using the protocol described previously (Seung et al. 2015). Briefly, 1 µg of the recombinant TaPTST1 was incubated in a binding medium with purified wheat starch, amylose, and amylopectin from potato starch (Sigma Aldrich) for 30 min at 20 °C with end-over-end shaking. The supernatant was collected by centrifugation of the mixture for 2 min at 13 000 g, and the pellet was washed three times with wash buffer. The pellet, final wash, and supernatant were heated at 95 °C into the SDS-sample loading buffer, followed by SDS-PAGE and immunoblotting with anti His-tag rabbit monoclonal antibodies (RM146 Thermofisher scientific) at 1:3000 dilutions. Goat Anti-Rabbit IgG (H + L)-HRP conjugated secondary antibodies were used at the same dilution for the detection of expressed proteins.

Immunoaffinity pull-down assay with recombinant protein

Purified recombinant GBSSI protein (10 µM) was incubated with protein-A sepharose beads immobilized with equimolar concentration of anti-wheat GBSSI antibodies in 100 µl of binding buffer containing 100mM tricine (pH 7.5) and 0.1% BSA for 4 hours. GBSSI antibodies were raised in rabbits against the peptide CRGKTKEKIYGPDAGTDYED (ABGENEX Pvt. Ltd.). The recombinant TaPTST1 was added to the reaction and incubated for 4 hours. The resin was washed five times in wash buffer, all steps were carried out at 4 °C. The resin was boiled in an SDS-sample loading buffer and the proteins were visualized using SDS-PAGE followed by immunoblotting. Immunoblotting was performed with the 6x-His tag antibodies. The pull-down assay was performed for the full-length PTST peptide as well as for the N-terminal truncated peptide.

Generation of BiFC vector constructs andtransient expression in tobacco

The full coding sequence of wheat PTST was amplified and cloned in plant compatible BiFC vector pSPYCE, using BamHI and XmaI restriction sites, enabling the construction of C-terminal fusion protein with eYFP C-terminus. GBSSI was cloned in pSPYNE173 using XhoI and XmaI restriction sites to synthesise fusion protein with eYFP N-terminus. Both fusion proteins expressed under 35S promoter were used to transform *Agrobacterium tumefaciens* (strain GV3101). BiFC experiments were performed according to the method described earlier (Thakur et al. 2021). Briefly, positive transformants were grown in LB medium with the selection markers rifampicin, gentamycin, and kanamycin at 28 °C for 24 hrs. Cells were pelleted by centrifugation and resuspended in a resuspension medium and kept for 6 hrs before the infiltrations. The cell suspensions of both the expression constructs and blank vector control were infiltrated into the intercellular spaces between abaxial epidermal cells of intact *N. benthamiana* leaves (4–5-week-old) using a 1 ml syringe. Plants were kept for 48 hours under the same growth condition in growth chambers prior to microscopy.

Microscopy and imaging

For the microscopic analysis, tobacco leaf discs were cut after 48 hours of infiltration. Using water as the imaging medium, lower epidermis cells were analysed at room temperature. The fluorescent YFP-signals were detected using a Carl Zeiss confocal microscope (LSM880) (Karl Zeiss, Jena, Germany). YFP

signals were monitored using argon laser for excitation at 514 nm and wavelength window of 518 to 557 nm for emission signals. ImageJ software was used for the processing of the images.

Results

Identification of AtPTST1 homologs in wheat

Homologs of AtPTST1 were identified using BLASTp and AtPTST1 amino acid sequence as a query at NCBI and UniProt database. There was no sequence with significant similarity in the *Triticum* species. The BLASTp search at ensemble plant database (Ensembl Plants release 47) found three peptides of 292 amino acid each, TraesCS6D02G090900.1, TraesCS6A02G102300.1 and TraesCS6B02G130700.1 with the 54.5, 55.1, and 60.9% identity with AtPTST1 respectively. The identified protein was found to be encoded by the three homeologues gene located at chromosome 6A, 6B and 6D.

Sequencing of all the three homeologues in the ‘C 306’ Indian wheat variety identified no sequence variation compared to the ‘Chinese spring’ reference genome. The nucleotide and amino acid alignments of all three TaPTST1 homeologues of ‘C 306’with Chinese spring are shown in supplementary data (Figure S1 & S2). The sequence similarity among the nucleotide sequences of three homeologues encoded by A, B, and D genome was A to B (90%), B to D (90%), and A to D (93%). The peptide encoded by the three genomes of wheat showed a high sequence identity of A to B (94.52%), B to D (94.18%), and A to D (92.18%). The TaPTST1 was encoded by 8 exons with transcript lengths of 1346, 1321, and 1282 bp of A, B, and D genome homeologues respectively. All the transcripts resulted in 879 bp long coding sequences encoding for 292 amino acids long peptide.

The candidate protein was found to have 72.88% identity with rice PTST (XP_015627711.1). The molecular phylogenetic analysis revealed the conservation of the protein sequences among the different species (Figure 1). The PTST1 from monocots and dicots are grouped into two different clades. All three Homeologues of TaPTST1 showed closed sequence similarity and grouped into the same clade and showed the highest identity with barley PTST1 sequence among all the monocots.

Domain analysis of TaPTST1

The amino acid sequence of candidate TaPTST1 was analysed using the Pfam and cdd database. The protein was found to have a 28 amino acid long coiled-coil domain (108-135) and an AMPKbeta-like C-terminal domain (cd02859). The AMPKbeta-like C- terminal domain was found to have carbohydrate-binding residues, suggesting a probable site that may interact directly with starch granules in amyloplast. The c-terminal domain was found to have conserved residues among the different species including AtPTST (Figure 2 a). The coiled-coil prediction using a coils web server (https://embnet.vital-it.ch/software/COILS_form.html) identified the strong probability of the residues from 110-140 to form coiled-coil structures (Figure 2 b). The coiled-coil structures are the potential interaction site for GBSS binding in AtPTST (Seung et al. 2015). Similar coiled-coil structures were found in the wheat-GBSSI

(Figure 2 c). The presence of the potential binding sites among both the proteins is indicative of the probable interactions between the TaGBSSI and TaPTST1.

Homology Modelling and structural analysis

A detailed amino acid analysis for secondary structure predictions using the PSIPRED tool showed the strong potential of residues at N-terminal to form coiled structures. PSIPRED analysis showed the three helical structures in N-terminal one from residues 51-86, second from 88-117, and the third helix of 18 residues from 135-152 (Figure S3).

The unavailability of decoded structural information of any homolog of the PTST type of proteins is a limitation to access the structural aspects. To identify the residues of PTST which are directly involved in interaction with glucans, the TaPTST1 secondary and tertiary structural homology model was constructed. TaPTST1 Psi-blast in uniport did not result in any significant sequence similarity to any PDB structure. Therefore, AMPK domain and N-terminal CC structures were modelled separately. The homology model of AMPK domain of TaPTST1 was generated by i-Tasser server. The built model had the C-score of -4.26 with the estimated TM-score of 0.26±0.08. Additionally, homology model of TaPTST1 was also built using AAKB2_RAT (PDB: 4Y0G) as a template with the TOME V3 Server. The selected template showed highest identity percentage to the query sequence and the homology protein model was built with the Atome score of 90.74. Both the homology models showed similar structures. TaPTST1 was then analysed for potential ligand binding sites and identified W232, K263, W270, N286 consensus alpha-D glucopyranose binding residues in C-terminal CBM. The modelled PTST structure docked with GLC is depicted in (Figure 3 a). The identified residues were surface exposed and may be potential glucans-interacting residues.

The homology model of TaGBSSI without predicted signal peptide was generated using decoded rice GBSSI crystal (PDB: 3UVE). The CC-structure model generated by Swiss model server (Figure 3b) was docked with TaGBSSI at Hawkdock server predicted coiled-coil mediated interaction between the proteins. Further, analysis of docked-molecule identified a coiled-coil helical structure in the C-terminal domain of TaGBSSI (Figure 3c), where, some key surface-exposed residues K365, E368, K372, E375 were identified as the possible interacting residues (Figure 3d).

Homeologues specific expression profile of GBSSI and PTST

We investigated the expression profile of TaPTST1 encoding transcripts in six different tissues i.e., root, stem, leaf, spike, seedling and grain using gene expression values (TPM -Transcripts Per Million) of hexaploid wheat cultivar ‘Chinese Spring’ from publicly available wheat expression database. The three transcripts of homeologues of TaPTst1 was found to be expressed substantially in all the tested tissues. Interestingly, the TaPTst1 showed highest expression in grain, while the expression of all homeologues was comparable in other tissues. Among the three homeologues TaPTST1_6D (TraesCS6D02G090900) showed highest expression in all tissues (Figure 4 a).

To further analyse the expression profile of TaPTST1 and GBSSI homeologues at different seed developmental stages, i.e., 7DAA, 14DAA, 21DAA, and 28DAA, q-RT PCR was performed.

Among the three homeologues of GBSSI, GBSSI_4AL was found to be null for transcript expression, while the expression of GBSSI_7A was consistent throughout the seed development. The expression of GBSSI_7D was found to be higher after the 21 DAA. The 6D and 6A homeologuess of wheat TaPTST1 showed higher expression at the later seed developmental days, while the expression of PTST_6B was lower than other homeologuess (Figure 4 b). The results indicated expression of the transcripts GBSSI and TaPTST1 in wheat grain throughout the grain filling stages. The highest expression of TaPTST1 in grain may be indicative of its probable function in wheat grain. GBSSI is found to be the most abundant starch-associated protein in wheat grain. Although, it lacks any dedicated carbohydrate-binding domain for its association with starch granules.

Direct binding of TaPTST1 with starch granules

The TaPTST1 was predicted to have a CBM module that could bind to starch. We analysed the PTST binding affinity against wild-type wheat starch, purified amylose, and amylopectin. The amylose and amylopectin were assayed separately to identify the specific interaction of TaPTST1 with two different polymers. His_6-TaPTST1 was incubated with any of the native starch, amylose, and amylopectin. We found that The TaPTST1 can interact with wheat starch well when assayed with wheat starch. Although, the protein interacts with less affinity to purified amylose fraction of starch and showed no binding with amylopectin (Figure 5). The interaction of TaPTST1 with starch suggested the probable role of the protein in starch biosynthesis.

The C-terminal domain is essential for carbohydrate-binding

The C-terminal AMPKbeta-like domain was predicted to have carbohydrate-binding residues and serve as the carbohydrate-binding module. To understand the significance of the domain in TaPTST1, a truncated peptide TaPTST1-209 lacking the AMPKbeta-like domain was expressed and purified. Interestingly, the TaPTST1-209 fails to interact with the starch granules *in vitro,* suggesting the crucial role of the domain as the carbohydrate-binding module in TaPTST1 (Figure 5). Additionally, no interaction between TaPTST1-209 with starch also rule out any possibility of interaction between his_6-tag and starch.

Identification of direct interaction between TaPTST1 and GBSSI

Domain analysis and docking results suggested the possible interaction between TaPTST1 and TaGBSSI mediated by the coiled-coil domain. We investigated the direct interaction between the proteins by immunoaffinity pull-down assay using recombinant proteins expressed in *E. coli*. N-terminal His-tagged GBSSI was incubated with N-terminal his tagged TaPTST1. GBSSI was recovered using protein-A sepharose beads coated with anti-TaGBSSI antibodies. Additionally, an immunoaffinity pull-down assay was also performed with TaPTST1-209 lacking the AMPK domain. Both TaPTST1 and TaPTST1-209 were co-eluted with GBSSI in a substantial amount, indicating a direct interaction between two

recombinant proteins. The TaPTST1-209 was able to interact with GBSSI, indicating that the AMPK domain has no role to play in the GBSSI binding of TaPTST1 (Figure 6).

Direct interaction between TaPTST1 and GBSSI in planta

Further, to confirm the direct protein-protein interaction between TaPTST1 and GBSSI in plants, BiFC assays were performed. The pSPYCE::TaPTST1 and pSPYNE173::GBSSI constructs encoding TaPTST1_YFP^{C} and GBSSI_YFP^{N} respectivelywere transiently expressed in *Nicotiana benthamiana* leaf epidermal cells. YFP protein was expressed under the control of constitutive CaMV 35S promoter as YFP control for the experiment and YFP signals were detected as expected. The YFP fluorescence signal was detected in chloroplasts when both plasmids were co-expressed in *N. benthamiana* leaves. The leaves co-infiltrated with either pSPYCE::TaPTST1 and pSPYNE173 or pSPYNE173::GBSSI/ pSPYCE resulted in non-detectable YFP signals (Figure 7). These results confirmed that both the proteins are localized to chloroplast and TaPTST1 directly interacts with GBSSI in leaves.

Therefore, the interaction between the GBSSI and TaPTST1 *in vitro* and *in planta* and starch binding affinity of TaPTST1 suggested the probable role of TaPTST1 in starch biosynthesis by targeting GBSSI to starch granules.

Discussion

Understanding the starch metabolism pathways in cereals, mostly rice and wheat is at the centre to improve grain's nutritional and processing quality by targeting one or more pathway enzymes. GBSSI is the most abundant starch-associated protein in cereal endosperm and the enzyme is dependent on unknown protein for interaction with glucans, as it lacks any dedicated Carbohydrate-binding domain (CBM). Previously, PTST1-mediated targeting of GBSSI was reported in *Arabidopsis* (Seung et al. 2015). A homolog of AtPTST1 was identified and found crucial for the development of starchy endosperm in barley (Zhong et al. 2019). In this study, we identified and characterized TaPTST1, a homolog of AtPTST1 in hexaploid wheat.

TaPTST1 is a functional protein and can interact with starch via its CBM domain

Expression analysis of starch metabolism genes and related transporters and transcription factors in wheat grain at different grain developmental stages showed different patterns (Singh et al. 2015). Based on the highest expression of TaPTST1 wheat grain, the potential role of the protein in grain may be postulated. The identified TaPTST1 showed significant sequence similarity with known PTST sequences from other plant species. However, monocots and dicots were grouped in two different clades in the phylogenetic tree, which may be due to the early divergence of monocots from their dicot relatives (Schuyler et al. 1983).

The AMPK domain-containing proteins have a glycogen-binding domain (GBD) in mammals (McBride et al. 2009). In plants, the GBD is known as CBMs that directly interacts with carbohydrates like starch

(Christiansen et al. 2009). Different CBM containing proteins have been identified in plants (Seung et al. 2015; Seung et al. 2017; Seung et al. 2018; Zhong et al. 2019; Wang et al. 2020) and their role was established in starch metabolism. The CBM was found to be located on the N or C-terminal of these proteins. The TaPTST1 peptide contains a C-terminal AMPK domain serving as a carbohydrate-binding module and showed close sequence similarity with barely PTST1 (Zhong et al. 2019). Some of the residues are conserved (marked by * in Figure 3) across the analysed sequences, suggesting the conserved role of the residues. We demonstrated the direct binding of TaPTST1 with starch and amylose with different affinities that may be attributed to the structural differences between the molecules. TaPTST1 showed no affinity for amylopectin, unlike AtPTST1, which showed binding with both amylose-free and wild-type starch(Seung et al. 2015). The TaPTST1 lacking C-terminal AMPK failed to bind the starch *in vitro*, suggesting that the AMPK may serve as a CBM. These findings demonstrated that TaPTST1 is a novel CBM containing protein with no assigned function in wheat and can interact with starch by its C-terminal CBM.

Potential role of TaPTST1 in starch biosynthesis by targeting GBSSI to starch granules

In cereal endosperm, the site of amylose biosynthesis is pre-existing, semi-crystalline amylopectin matrix, permeable to small molecules (Tatge et al. 1999). GBSSI is one of the most abundant starch-bound proteins in wheat grain. Although it lacks any CBM for its localization to starch granule, that is essential for the functioning of the enzyme. However, the mechanism of GBSSs entry into the matrix of the starch granule remains elusive. The TaGBSSI has a C-terminal coiled-coil structure that was previously found as a potential structure for interaction with other coiled-coil containing proteins via electrostatic interactions mediated by surface-exposed lysine and glutamate residues. Charged amino acid-mediated electrostatic interactions are shown to play a major role in coiled-coil interactions (Mason and Arndt 2004; Thompson et al. 2012).

In our study, the direct interaction between GBSSI and TaPTST1 identified by immunoaffinity pull-down assay indicated the potential role of TaPTST1 in targeting GBSSI to the starch granules. Further, the BiFC assay identified the interaction between TaPTST1 and GBSSI in the chloroplast of tobacco leaves. Taken together the findings, we postulate that TaPTST1 interacts with TaGBSSI protein via the coiled-coil domain and targets it to the starch granules in wheat endosperm. Understanding the mechanism and role mediated by TaPTST1 in starch biosynthesis is important to develop or select better crop varieties with enhanced nutrition properties.

Declarations

Acknowledgements

We would like to thank the Executive Director of the National Agri-Food Biotechnology Institute (NABI), Mohali, an Autonomous Institute of the Department of Biotechnology, Government of India for funds and support. VS acknowledges the Council of Scientific and Industrial Research for JRF and SRF fellowships for PhD work and Indian Institute of Science Education and Research for PhD program. We thank Dr.

Prafull Salvi for providing vectors for BiFC experiments. We also acknowledge DeLCON (DBT-electronic library consortium), Gurgaon, India, for the online journal access.

Conflict of interest

All authors declare that they have no conflict of interest.

Author's contributions

VS and JR conceived and designed the experiment, and manuscript writing. VS performed the experimental work and data analysis. VF, PK, AP, AM, MB helped in experimental work. AM and HS helped to manuscript writing. VR and SB helped to draft the final manuscript.

All authors read and approved the final manuscript.

References

1. Bahaji A, Li J, Sánchez-López ÁM, Baroja-Fernández E, Muñoz FJ, Ovecka M, Almagro G, Montero M, Ezquer I, Etxeberria E, Pozueta-Romero J (2014) Starch biosynthesis, its regulation and biotechnological approaches to improve crop yields. Biotechnol. Adv. 32
2. Beckles DM, Smith AM, Ap Rees T (2001) A cytosolic ADP-glucose pyrophosphorylase is a feature of graminaceous endosperms, but not of other starch-storing organs. Plant Physiol 125: . https://doi.org/10.1104/pp.125.2.818
3. Birt DF, Boylston T, Hendrich S, Jane JL, Hollis J, Li L, McClelland J, Moore S, Phillips GJ, Rowling M, Schalinske K, Paul Scott M, Whitley EM (2013) Resistant starch: Promise for improving human health. Adv. Nutr. 4
4. Christiansen C, Hachem MA, Glaring MA, Viksø-Nielsen A, Sigurskjold BW, Svensson B, Blennow A (2009) A CBM20 low-affinity starch-binding domain from glucan, water dikinase. FEBS Lett 583: . https://doi.org/10.1016/j.febslet.2009.02.045
5. Corpet F (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 16: . https://doi.org/10.1093/nar/16.22.10881
6. Delvallé D, Dumez S, Wattebled F, Roldán I, Planchot V, Berbezy P, Colonna P, Vyas D, Chatterjee M, Ball S, Mérida Á, D'Hulst C (2005) Soluble starch synthase I: A major determinant for the synthesis of amylopectin in *Arabidopsis* thaliana leaves. Plant J 43: . https://doi.org/10.1111/j.1365-313X.2005.02462.x
7. Flipse E, Keetels CJAM, Jacobsen E, Visser RGF (1996) The dosage effect of the wildtype GBSS allele is linear for GBSS activity but not for amylose content: Absence of amylose has a distinct influence on the physico-chemical properties of starch. Theor Appl Genet 92: . https://doi.org/10.1007/BF00222961
8. Ghosh HP, Preiss J (1966) Adenosine diphosphate glucose pyrophosphorylase. A regulatory enzyme in the biosynthesis of starch in spinach leaf chloroplasts. J Biol Chem 241: .

https://doi.org/10.1016/S0021-9258(18)99747-4

9. Hebelstrup KH, Nielsen MM, Carciofi M, Andrzejczak O, Shaik SS, Blennow A, Palcic MM (2017) Waxy and non-waxy barley cultivars exhibit differences in the targeting and catalytic activity of GBSS1a. J Exp Bot 68: . https://doi.org/10.1093/jxb/erw503
10. Hovenkamp-Hermelink JHM, Jacobsen E, Ponstein AS, Visser RGF, Vos-Scheperkeuter GH, Bijmolt EW, de Vries JN, Witholt B, Feenstra WJ (1987) Isolation of an amylose-free starch mutant of the potato (Solanum tuberosum L.). Theor Appl Genet 75: . https://doi.org/10.1007/BF00249167
11. Keeling PL, Myers AM (2010) Biochemistry and genetics of starch synthesis. Annu Rev Food Sci Technol 1: . https://doi.org/10.1146/annurev.food.102308.124214
12. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35: . https://doi.org/10.1093/molbev/msy096
13. Letunic I, Khedkar S, Bork P (2021) SMART: Recent updates, new developments and status in 2020. Nucleic Acids Res 49: . https://doi.org/10.1093/nar/gkaa937
14. Li H, Gidley MJ, Dhital S (2019) High-Amylose Starches to Bridge the “Fiber Gap”: Development, Structure, and Nutritional Functionality. Compr. Rev. Food Sci. Food Saf. 18
15. Lim SM, Page A, Carragher J, Muhlhausler B (2020) Could High-Amylose Wheat Have Greater Benefits on Diabesity and Gut Health than Standard Whole-wheat? Food Rev. Int. 36
16. Mason JM, Arndt KM (2004) Coiled coil domains: Stability, specificity, and biological implications. In: ChemBioChem
17. McBride A, Ghilagaber S, Nikolaev A, Hardie DG (2009) The Glycogen-Binding Domain on the AMPK β Subunit Allows the Kinase to Act as a Glycogen Sensor. Cell Metab 9: . https://doi.org/10.1016/j.cmet.2008.11.008
18. Mishra A, Sharma V, Rahim MS, Sonah H, Pal D, Mantri S, Sharma TR, Roy J (2021) Genotyping-by-sequencing based QTL mapping identified a novel waxy allele contributing to high amylose starch in wheat. Euphytica 217: . https://doi.org/10.1007/s10681-021-02861-5
19. Nakamura T, Vrinten P, Hayakawa K, Ikeda J (1998) Characterization of a granule-bound starch synthase isoform found in the pericarp of wheat. Plant Physiol 118: . https://doi.org/10.1104/pp.118.2.451
20. Pérez S, Bertoft E (2010) The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review. Starch/Staerke 62
21. Pons JL, Labesse G (2009) @TOME-2: A new pipeline for comparative modeling of protein-ligand complexes. Nucleic Acids Res 37: . https://doi.org/10.1093/nar/gkp368
22. Prasad S, Khadatare PB, Roy I (2011) Effect of chemical chaperones in improving the solubility of recombinant proteins in Escherichia coli. Appl Environ Microbiol 77: . https://doi.org/10.1128/AEM.05259-11
23. Radchuk V V., Borisjuk L, Sreenivasulu N, Merx K, Mock HP, Rolletschek H, Wobus U, Weschke W (2009) Spatiotemporal profiling of starch biosynthesis and degradation in the developing barley

grain1[W]. Plant Physiol 150: . https://doi.org/10.1104/pp.108.133520

24. Roldán I, Wattebled F, Mercedes Lucas M, Delvallé D, Planchot V, Jiménez S, Pérez R, Ball S, D'Hulst C, Mérida Á (2007) The phenotype of soluble starch synthase IV defective mutants of *Arabidopsis* thaliana suggests a novel function of elongation enzymes in the control of starch granule formation. Plant J 49: . https://doi.org/10.1111/j.1365-313X.2006.02968.x
25. Schuyler AE, Dahlgren RMT, Clifford HT (1983) The Monocotyledons: A Comparative Study. Bull Torrey Bot Club 110: . https://doi.org/10.2307/2996198
26. Sestili F, Botticella E, Proietti G, Janni M, D'Ovidio R, Lafiandra D (2012) Amylose content is not affected by overexpression of the Wx-B1 gene in durum wheat. Plant Breed 131: . https://doi.org/10.1111/j.1439-0523.2012.02004.x
27. Seung D, Boudet J, Monroe J, Schreier TB, David LC, Abt M, Lu KJ, Zanella M, Zeeman SC (2017) Homologs of PROTEIN TARGETING TO STARCH control starch granule initiation in *Arabidopsis* leaves. Plant Cell 29: . https://doi.org/10.1105/tpc.17.00222
28. Seung D, Schreier TB, Bürgy L, Eicke S, Zeeman SC (2018) Two plastidial coiled-coil proteins are essential for normal starch granule initiation in *Arabidopsis*. Plant Cell 30: . https://doi.org/10.1105/tpc.18.00219
29. Seung D, Soyk S, Coiro M, Maier BA, Eicke S, Zeeman SC (2015) PROTEIN TARGETING TO STARCH Is Required for Localising GRANULE-BOUND STARCH SYNTHASE to Starch Granules and for Normal Amylose Synthesis in *Arabidopsis*. PLoS Biol 13: . https://doi.org/10.1371/journal.pbio.1002080
30. Shewry PR, Hey SJ (2015) The contribution of wheat to human diet and health. Food Energy Secur. 4
31. Shure M, Wessler S, Fedoroff N (1983) Molecular identification and isolation of the Waxy locus in maize. Cell 35: . https://doi.org/10.1016/0092-8674(83)90225-8
32. Simms D, Cizdziel P, Chomczynski P (1993) TRIzol: a new reagent for optimal single-step isolation of RNA. Focus (Madison)
33. Singh A, Kumar P, Sharma M, Tuli R, Dhaliwal HS, Chaudhury A, Pal D, Roy J (2015) Expression patterns of genes involved in starch biosynthesis during seed development in bread wheat (*Triticum aestivum*). Mol Breed 35: . https://doi.org/10.1007/s11032-015-0371-9
34. Spriestersbach A, Kubicek J, Schäfer F, Block H, Maertens B (2015) Purification of His-Tagged Proteins. In: Methods in Enzymology
35. Tatge H, Marshall J, Martin C, Edwards EA, Smith AM (1999) Evidence that amylose synthesis occurs within the matrix of the starch granule in potato tubers. Plant, Cell Environ 22: . https://doi.org/10.1046/j.1365-3040.1999.00437.x
36. Tetlow IJ (2018) Starch biosynthesis in crop plants. Agronomy 8: . https://doi.org/10.3390/agronomy8060081
37. Thakur T, Gandass N, Mittal K, Jamwal P, Muthamilarasan M, Salvi P (2021) A rapid, efficient, and low-cost BiFC protocol and its application in studying in vivo interaction of seed-specific transcription factors, RISBZ and RPBF. Funct Integr Genomics. https://doi.org/10.1007/s10142-021-00801-z

38. Thompson KE, Bashor CJ, Lim WA, Keating AE (2012) Synzip protein interaction toolbox: In vitro and in vivo specifications of heterospecific coiled-coil interaction domains. ACS Synth Biol 1: . https://doi.org/10.1021/sb200015u
39. Tomlinson K, Denyer K (2003) Starch synthesis in cereal grains. Adv. Bot. Res. 40
40. Tsai CY (1974) The function of the Waxy locus in starch synthesis in maize endosperm. Biochem Genet 11: . https://doi.org/10.1007/BF00485766
41. Vrinten PL, Nakamura T (2000) Wheat granule-bound starch synthase I and II are encoded by separate genes that are expressed in different tissues. Plant Physiol 122: . https://doi.org/10.1104/pp.122.1.255
42. Wang W, Wei X, Jiao G, Chen W, Wu Y, Sheng Z, Hu S, Xie L, Wang J, Tang S, Hu P (2020) GBSS-BINDING PROTEIN, encoding a CBM48 domain-containing protein, affects rice quality and yield. J Integr Plant Biol 62: . https://doi.org/10.1111/jipb.12866
43. Wang Z -Y, Zheng F -Q, Shen G -Z, Gao J -P, Snustad DP, Li M -G, Zhang J -L, Hong M -M (1995) The amylose content in rice endosperm is related to the post-transcriptional regulation of the waxy gene. Plant J 7: . https://doi.org/10.1046/j.1365-313X.1995.7040613.x
44. Yan Y, Tao H, He J, Huang SY (2020) The HDOCK server for integrated protein–protein docking. Nat Protoc 15: . https://doi.org/10.1038/s41596-020-0312-x
45. Yan Y, Zhang D, Zhou P, Li B, Huang SY (2017) HDOCK: A web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. Nucleic Acids Res 45: . https://doi.org/10.1093/nar/gkx407
46. Yang J, Roy A, Zhang Y (2013a) Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics 29: . https://doi.org/10.1093/bioinformatics/btt447
47. Yang J, Roy A, Zhang Y (2013b) BioLiP: A semi-manually curated database for biologically relevant ligand-protein interactions. Nucleic Acids Res 41: . https://doi.org/10.1093/nar/gks966
48. Yang J, Zhang Y (2015) I-TASSER server: New development for protein structure and function predictions. Nucleic Acids Res 43: . https://doi.org/10.1093/nar/gkv342
49. Zhang X, Myers AM, James MG (2005) Mutations affecting starch synthase III in *Arabidopsis* alter leaf starch structure and increase the rate of starch synthesis. Plant Physiol. 138
50. Zhang X, Szydlowski N, Delvallé D, D'Hulst C, James MG, Myers AM (2008) Overlapping functions of the starch synthases SSII and SSIII in amylopectin biosynthesis in *Arabidopsis*. BMC Plant Biol 8: . https://doi.org/10.1186/1471-2229-8-96
51. Zhong Y, Blennow A, Kofoed-Enevoldsen O, Jiang D, Hebelstrup KH (2019) Protein Targeting to Starch 1 is essential for starchy endosperm development in barley. J Exp Bot 70: . https://doi.org/10.1093/jxb/ery398

Figures

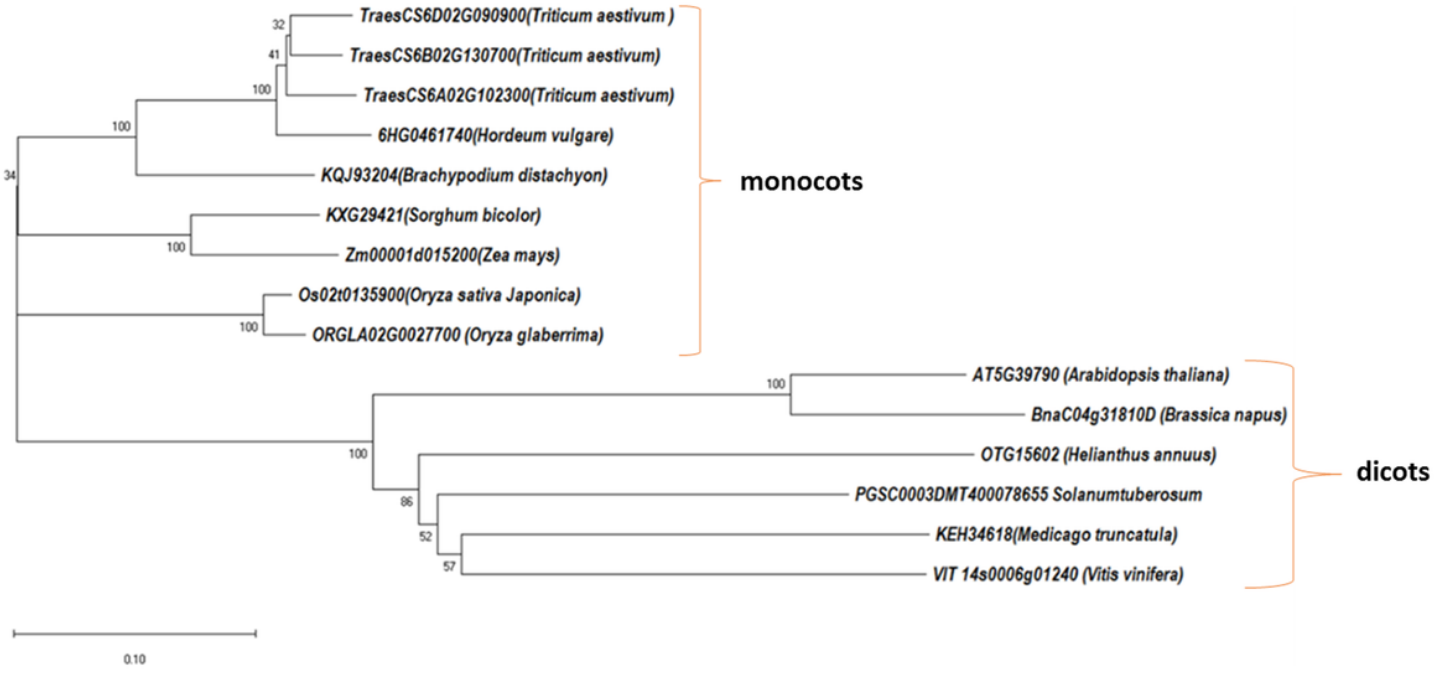

Figure 1

The neighbour-joining phylogenetic tree of PTST protein from MAFFT alignment. Tree was constructed with the p-distance model, pairwise deletion, and bootstrap (1,000 replicates). Bootstrap values are shown over the branches.

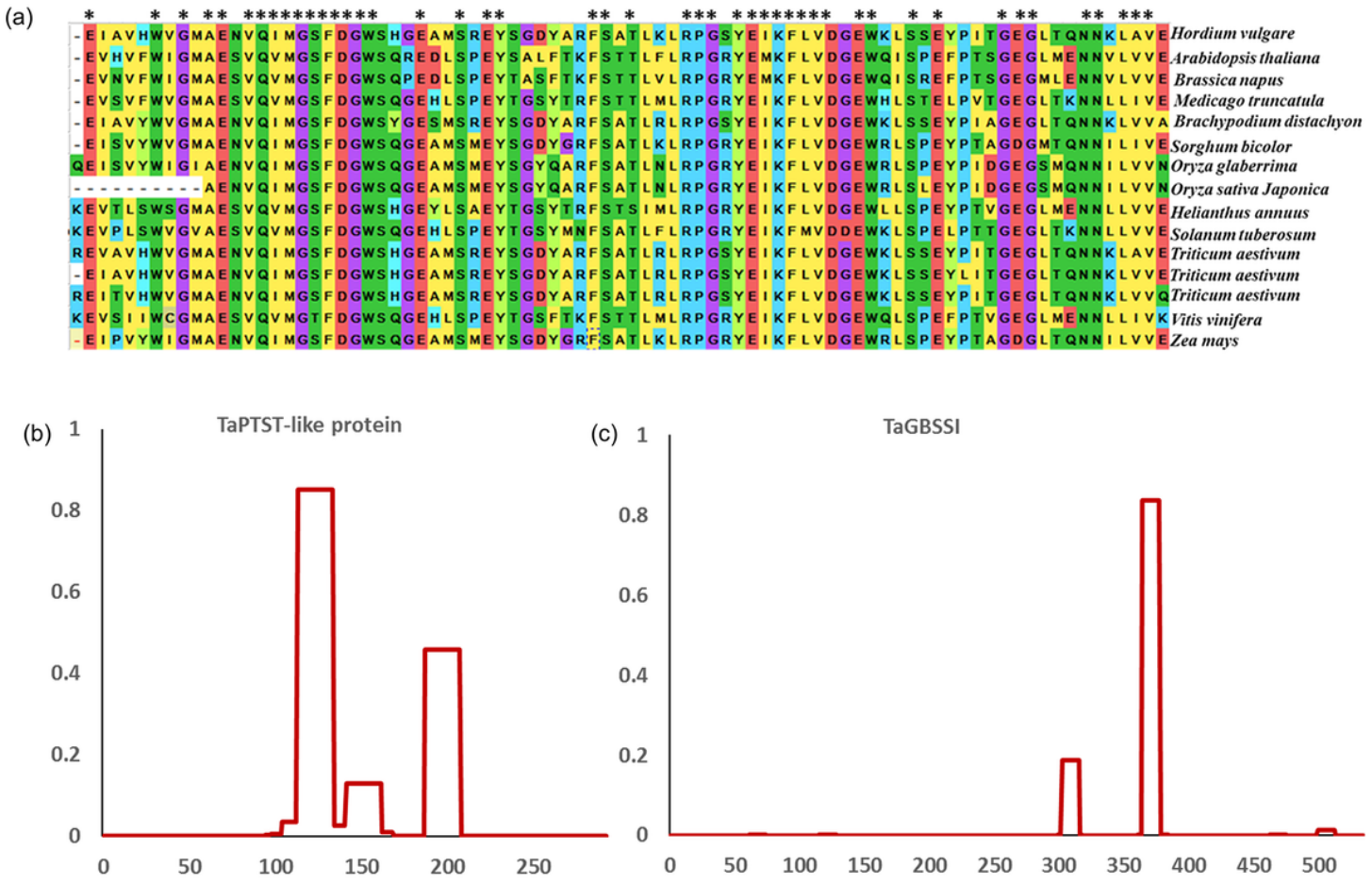

Figure 2

Domain analysis of TaPTST1 (a) multiple sequence alignment of CBM domains of known PTST1 protein. all conserved residues are marked by * mark. Coiled-coil prediction of (b) TaPTST1 (c) TaGBSSI. Probability scores are represented on the x-axis.

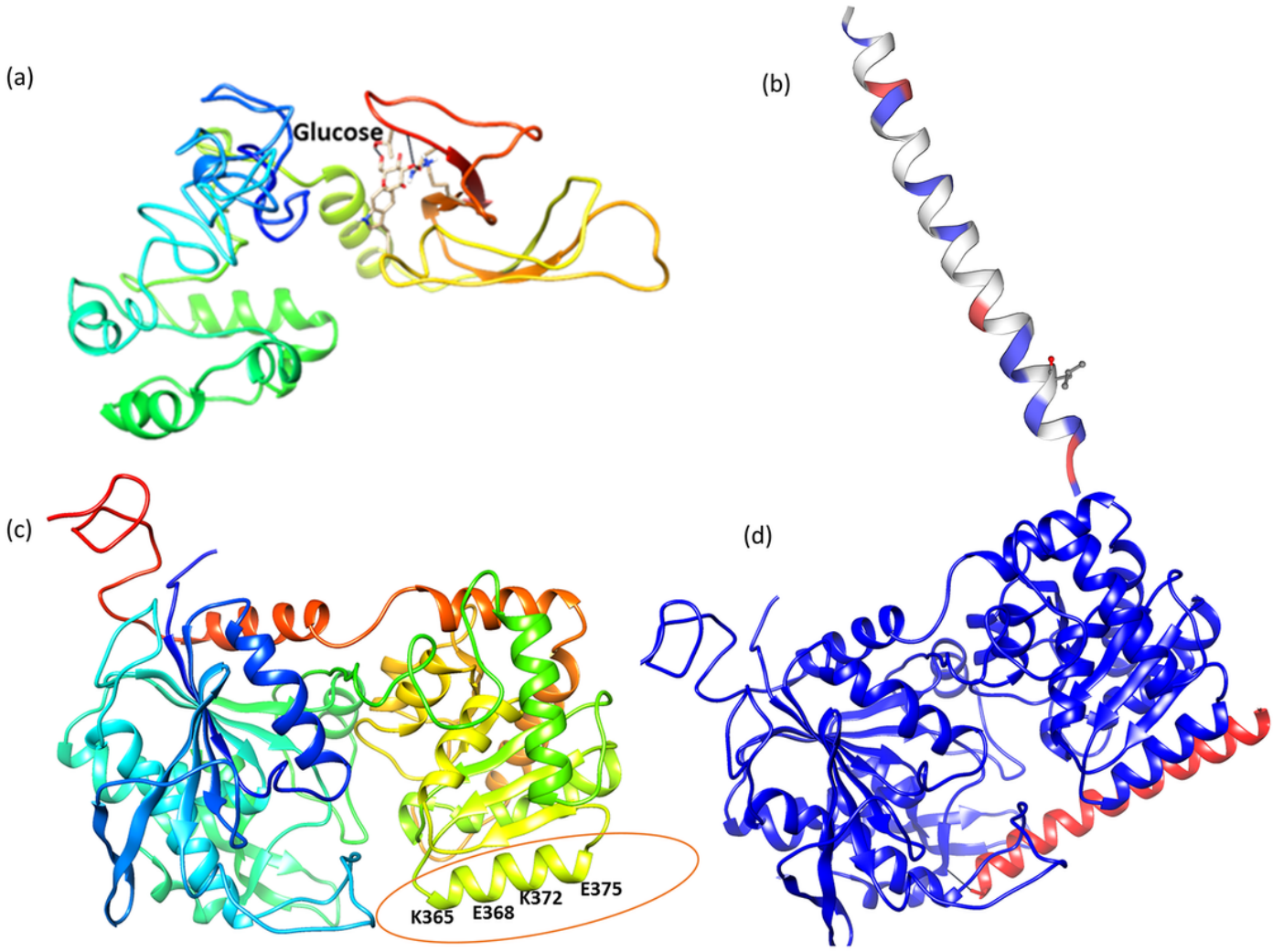

Figure 3

Homology modelling and molecular docking of (a) TaPTST1 with alpha-D glucopyranose. (b) Homology model of TaPTST1 coiled-coil domain. (c) The homology model of TaGBSSI. the coiled-coil structure with surface exposed charged residues is highlighted by outline. (d) The TaGBSSI in complex with coiled-coil domain of TaPTST1.

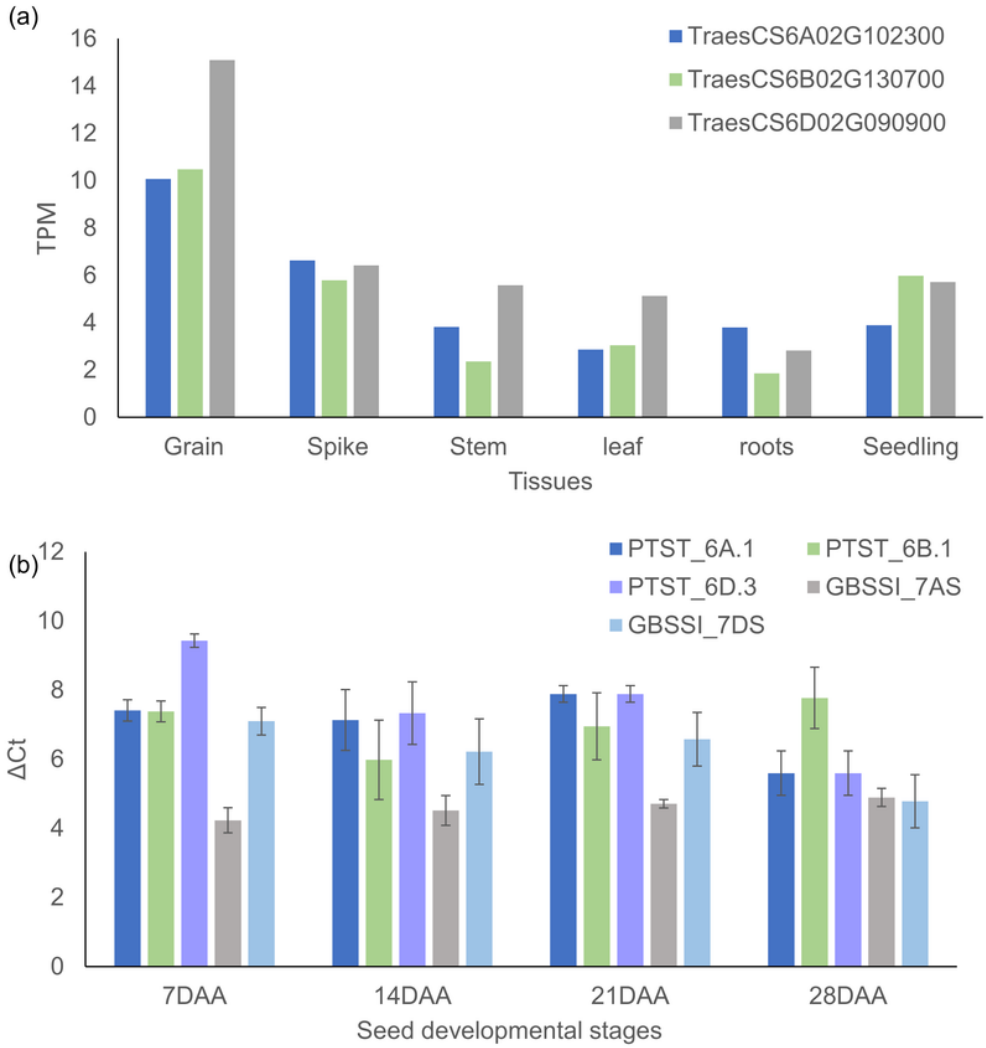

Figure 4

Expression analysis of TaPTST1 and GBSSI (a) expression of three homeologues of TaPTST1 in different tissues. (b) expression of homeologues of TaPTST1 and GBSSI in developing wheat grain. Where TPM stands for transcript per million.

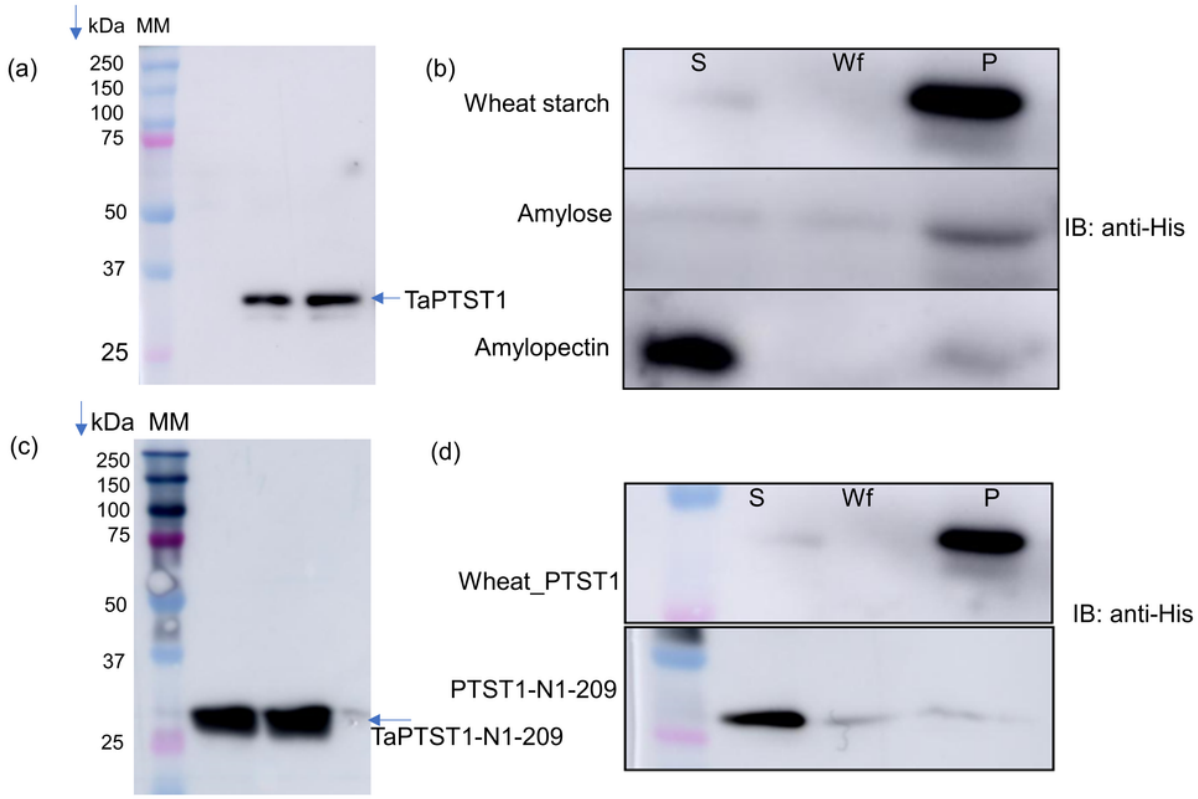

Figure 5

C-terminal AMPK is essential for the carbohydrate-binding of TaPTST1. (a) purified recombinant TaPTST1 with N-terminal His6-tag (b) Binding of His6-TaPTST1 to wild-starch, amylose, and amylopectin. Unbound proteins are represented as a soluble fraction (S), Wf is the final wash, and starch-bound proteins are shown as a pellet (P). (c) The truncated peptide His6-TaPTST1-209 lacking C-terminal AMPK. (d) Binding of His6-TaPTST1 His6-TaPTST-209 with wild-type starch in vitro.

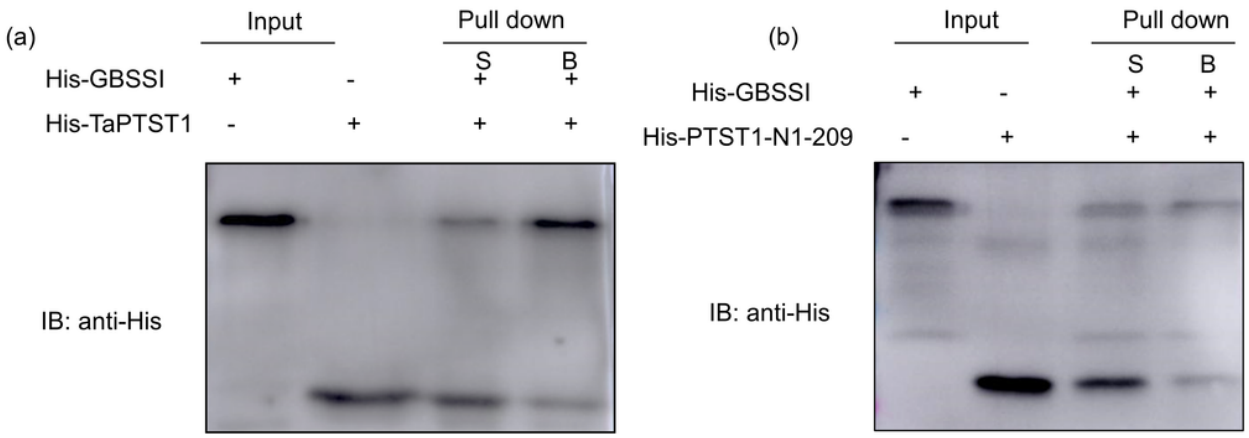

Figure 6

In vitro immunoaffinity pull-down assay using GBSSI antibodies. (a) with recombinant His6-TaPTST1 (b) recombinant His6-TaPTST1-209. Inputs represent the proteins used in the assay and S and B denote soluble and bound proteins. Immunoblotting was done using His6 antibodies.

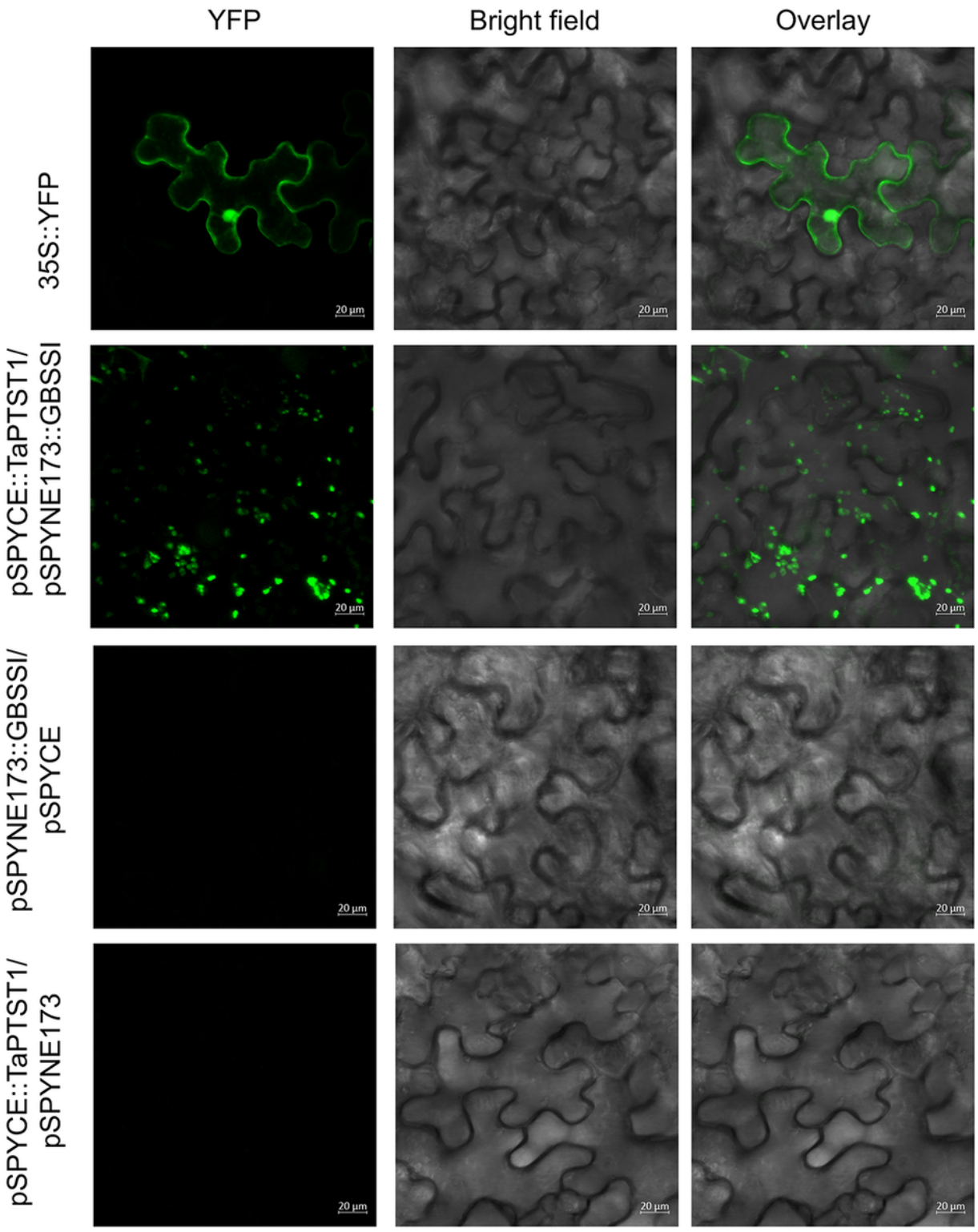

Figure 7

GBSS1 interacts with TaPTST1 in tobacco leaves as confirmed by BiFC. Confocal microscopy images of individually or co-expressed pSPYCE::TaPTST1 and pSPYNE173::GBSSI in N. benthamiana. Full-length YFP was used as a positive control. scale bar = 20 µm.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- DataS1.txt
- supplementarymaterial.docx