MALAYSIAN JOURNAL OF BIOCHEMISTRY & MOLECULAR BIOLOGY

The Official Publication of The Malaysian Society For Biochemistry & Molecular Biology (MSBMB)
http://mjbmb.org

IN SILICO ANALYSIS OF GENE-ENCODED INVERTASE INHIBITOR (*Sininh*) FROM SEVERAL SUGARCANE VARIETIES

Maulidi Firlandiana[1], Giyanto[2], Wilhelmus Terang Arga Sanjaya[1,3], Dwi Andreas Santosa[1,3,*]

[1]*Department of Soil Science and Land Resources, Faculty of Agriculture, IPB University, 16680 Bogor, Indonesia*
[2]*Department of Plant Protection, Faculty of Agriculture, IPB University, 16680 Bogor, Indonesia*
[3]*Biotech Center, IPB University, 16680 Bogor, Indonesia*

*Corresponding Author: dsantosa@indo.net.id

History
Received: 8 December 2021
Accepted: 13 June 2022

Keywords:

Invertase inhibitor; Sugarcane; Sucrose accumulation; Post-translation

Abstract

Invertase inhibitor gene fragments have been isolated from various cultivated plants including sugarcane. This article will review the structure of the invertase inhibitor gene (*sininh*) at 650 bp from multiple sugarcane varieties in Indonesia (PS 881, PS 882, PSJT 941, PS 862, BL/Bulu Lawang, KK/Kidang Kencana). Moreover, the primary structure, physicochemical properties, the secondary structure, 3D structure, and subcellular localization of invertase inhibitor protein (*Sininh*) were predicted utilizing bioinformatics tools. This information is beneficial for future research, especially in controlling sucrose accumulation at the post-translation level in sugarcane plants. The prediction of the *Sininh* consisting of 124-145 residues has different characteristics. The predicted *Sininh* sequences from PS 882 and BL varieties have the potency for translocation due to peptide signals. In addition, there are some Cys residues in both varieties associated with the formation of disulfide bonds for protein structure stability. Sustainable, functional conserved domain Plant invertase/pectin methylesterase inhibitors found in the *Sininh* protein variety BL/Bulu Lawang where they inhibit the activity of pectin methylesterase (PMEs) invertase with a complex formation. N-glycosylation motifs are also found in the protein *Sininh* varieties PS 881, PSJT 941, PS 862, BL/Bulu Lawang, KK/Kidang Kencana, which is related to potential stability and interaction with other proteins. These studies build the foundation for studying the structural aspects and the mechanism of the inactivation of invertase via its inhibitory proteins at the molecular level.

INTRODUCTION

Sugarcane (*Saccharum officinarum L.*) is one of the world's largest sugar-producing food crops capable of growing in tropical and subtropical regions [1]. Sugars result from the accumulation of sucrose, which sucrose is the primary end product of the process of carbon metabolism of photosynthesis [2]. Sucrose accumulated from the source tissue (leaves) will be transported through phloem to the sink tissue (roots, stems, reproductive organs, vegetative storage organs) in apoplast or symplast [3]. Sucrose begins to accumulate on the stem during the elongation process until the elongation stops. At the ripening stage, sucrose concentration increases higher than non-sucrose (i.e., glucose and fructose). This condition indicates that sucrose metabolism also changes with the development of plants. Invertase is capable of hydrolyzing sucrose (disaccharides) into glucose and fructose (monosaccharide) [4]. Generally, sucrose metabolism involves several enzymes, including sucrose-phosphate synthase (SPS), sucrose synthase (SUSY), and invertase (CINV). Sucrose-phosphate synthase (SPS) becomes the catalyst for converting UDP-glucose and fructose-6-phosphate into sucrose-6-phosphate, while sucrose-phosphatase (SPP) converts sucrose-6-phosphate

into sucrose. Sucrose synthase (SUSY) plays a role in synthesizing sucrose, where SUSY can convert UDP-glucose and fructose into sucrose [5]. After going through the metabolic process, sucrose that has reached the point of saturation will be hydrolyzed by the enzyme invertase to lower the sugar content of sugarcane. This phenomenon was very detrimental, especially for farmers and sugar industry players. Historically, the increase in sugar yields was done with conventional breeding programs but only increased sugarcane yields, not sucrose content. In recent decades, an increase in yields in sucrose has been possible due to the potential for many genetically explored sources. Increased sugar yields become possible with technology and genetic resources because it is right on the target of related genes. Therefore, it is hoped that the genetic manipulation approach can be more effective in producing new varieties that are more specific according to the desired traits [4]. Although the enzyme invertase is vital in supporting plant growth and development, uncontrolled invertase activity will decrease the sucrose content in sugarcane stems. As reported by [6, 7], invertase activity was negatively correlated with sucrose content. One approach that can be expanded is to examine the mechanism of inhibition of invertase activity using the role of protein inhibitors. The presence of endogenous invertase inhibitors protein in sugarcane can suppress invertase activity at the post-translational level [8]. Protein inhibitors were chosen because they have several advantages, including being available naturally in the cells of organisms and being easy to control gene expression. In addition, protein inhibitors are the class of proteins functional most abundantly in an organism, one of which is in plants and animals [9].

Invertase inhibitor proteins are small proteins with a molecular weight (Mr) of about 15 – 23 kDa [3]. In high-level plants, invertase inhibitors have the same molecular weight. For example, potato plants have acid invertase inhibitors measuring 19.5 kDa with filtration gel. As for the sugar beet plant, root beet, and sweet potato consecutively amounted to 17.8 kDa, 18.1 kDa, and 22.9 kDa, respectively [10]. This invertase inhibitor protein is adequate not only for the invertase of potatoes but also in other plant invertase. It shows that invertase in all plants has similar characteristics to potatoes [11]. In 1994, [12] tested the inhibitory activity of invertase inhibitors in tomato plants (18 kDa) which actively inhibits invertase at pH 5.0. It was also able to inhibit invertase activity in potato plants. The invertase inhibitor encoding gene was also successfully characterized from tobacco plants [13]. In 2004, [14] successfully cloned and demonstrated inhibition activity in corn crops where invertase activity decreased with the recombinant gene encoding inhibitor in corn (*ZM-INVINH*). Identifying the *ZN-INVINH* gene is the first discovery to characterize the invertase inhibitor encoding gene in monocot plants. In sugarcane, cloning and description of the invertase inhibitor encoding gene (*SlNINH*) were first performed [15] in 2014. In 2018, further research was conducted that invertase inhibitor genes (*ShINH1*) were expressed in the leaves and stem of sugarcane than in the roots and flowers [8].

This study aims to isolate and characterize the protein invertase inhibitor gene in various sugarcane varieties in Indonesia. Furthermore, this study will underlie further study on how these invertase inhibitor proteins avoid sucrose content reduction on sugarcane at the post-translational level.

MATERIALS AND METHODS

Materials

Sugarcane (*S. officinarum* L.) was supplied by Jati Tujuh Co., a sugarcane plantation in Majalengka district, West Java province, Indonesia, during the 2018 harvest. A total of 6 sugarcane samples consisting of PS 881, PS 882, PSJT 941, PS 862, BL (Bulu Lawang), and KK (Kidang Kencana) were grown until the generative phase in Cikabayan Farm Research Station, Bogor Agricultural University (IPB University)

Sample Preparation

After five months, the young leaf shoots were taken to be carried out induction of embryogenic callus. The callus induction method was modified from the [16]. The explants were planted on the basal medium of MS (Murashige and Skoog) [17], added with 2.4-D (2,4-Dichlorophenoxyacetic acid) (3 mg/L) and kinetin (0.1 mg/L). The callus were stored in a dark room with a temperature of 25-28 ^{o}C for ±14 days.

Total RNA Isolation and cDNA Synthesis

Embryogenic calluses of sugarcane from various varieties were ground using liquid nitrogen. Total RNA was isolated using Total RNA Mini Kit (Plant) (RP100, Geneaid, New Taipei, Taiwan). RNA purity and concentration were measured by looking at the $A_{260/280}$ ratio using a spectrophotometer at an absorbance of 260 nm (Nanodrop 2000, Thermo Scientific, USA). cDNA synthesis was performed with an oligodT primer and reverTra Ace-α-® (FSK101, Toyobo, Osaka, Japan).

Amplification of *sininh* Gene

cDNA amplification was performed using the PCR method with KOD Fx Plus-Neo (Toyobo). Primer used refers to research [15] for inhibitor-specific genes invertase (*sininh*), among them primary forward 5' CTACCATCCACATCCAACTC 3' and primary reverse 5' CCAAGCACGCTGTATAGTAC 3' with Tm (melting temperature) 54 ^{o}C. Shortly, 1 µL cDNA template was added to 0.5 µL KOD Plus Neo, 2.5 µL dNTPs (2 mM), 1.5 µL $MgSO_4$ (25 mM), 2.5 µL KOD Buffer 10x, 0.5 µL (10 pmol) of each primer, and ddH_2O to final volume of 25 µL. The

PCR program was set for 35 cycles with the following details; denaturation stage at 98 ºC for 10 seconds, 54 ºC attachment stage for 1 minute, extension stage at 70 ºC for 30 seconds, last cycle added 7 minutes at 68 ºC for synthesis, and cycle ended at 4 ºC.

In silico Study

Identification of *sininh* Gene

Sequences of *sininh* gene from all sugarcane varieties are consensus-made in both forward and reverse sequences, then aligned with the application of BioEdit 7.2.5 [18]. The nucleotide sequence was then used to identify the invertase inhibitor (*sininh*) encoding gene using the BLAST (BlastN) program on the NCBI website *(https://www.ncbi.nlm.nih.gov/)*. The Neighbor-joining method constructed the phylogeny tree, while multiple sequence alignment was done using Clustal Muscle algorithms in Mega X software [19].

Analysis of Primary Structure Characterization of Predicted Invertase Inhibitor Protein (*Sininh*)

Nucleotide sequences of whole varieties are translated into amino acid sequences using https://web.expasy.org/translate/ and physiochemical analysis of *Sininh* protein predictions using http://web.expasy.org/protparam. Furthermore, the amino acid sequence of the varieties was identified using the BLAST (BlastX) program on the NCBI *(https://www.ncbi.nlm.nih.gov/)*. Peptide signals were detected using PrediSi and Signal IP 5.0. Analysis of domain predictions and motives was accessed on http://bioinf.cs.ucl.ac.uk/psipred/ and *http://myhits.isb-sib.ch/cgi-bin/motif_scan,* respectively. Moreover, we also used CDD (Conserved Domain Database) in NCBI to analyze sustainable domains on proteins [20]. Conserved domain protein searches were accessed on https://www.ncbi.nlm.nih.gov/cdd. Disulfide bonds were carried out using DiANNA 1.1 software (*http://clavius.bc.edu/~clotelab/DiANNA/*). Protein-specific motifs and active sites were analysed using *http://myhits.isb-sib.ch/cgi-bin/motif_scan* and *http://prosite.expasy.org* sites, respectively.

Subcellular Localization of Predicted *Sininh* Protein

Subcellular localization of *Sininh* protein prediction was accessed using PSORT II Prediction software (https://psort.hgc.jp/form2.html).

Secondary and 3D Structure of *Sininh* Protein

Creating a secondary structure model of the protein was carried out using the PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) program. The 3D structure model was built using the online software SWISS-MODEL *(https://swissmodel.expasy.org/interactive)*.

RESULTS AND DISCUSSION

Identification of *sininh* Gene Structure

The specific gene of the invertase inhibitor was successfully isolated and verified on the agarose gel 1%. PCR results showed 650 bp band for the *sininh* gene of all sugarcane varieties using a *sininh*-specific primer (Figure 1).

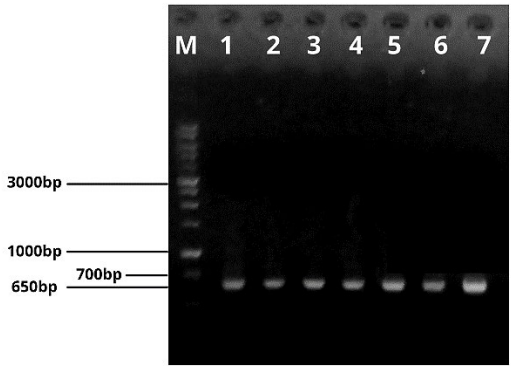


Figure 1. Visualization of 1 % agarose gel shows 650 bp amplicon invertase inhibitor (*sininh*) in various sugarcane varieties. Band number 1=PS 881; 2=PS 882; 4=PSJT 941; 5=PS 862; 6=BL; 7=KK

Gene sequences obtained for *sininh* from all six varieties showed the identity with *S. officinarum* (KP997207.1) of 99.66% (PS 881), 95.09% (PS 882), 99.66% (PSJT 941), 99.66% (PS 862), 97.0% (BL/Bulu Lawang), and 99.66% (KK/Kidang Kencana). As for *sininh* of the six varieties showed identification with *Saccharum sinence* (KP997206.1) of 99.35% (PS 881), 94.92% (PS 882), 99.35% (PSJT 941), 99.35% (PS 862), 95.84% (BL/Bulu Lawang), and 99.24% (KK/Kidang Kencana). Results showed that the *sininh* gene fragment is an invertase inhibitor encoding gene. Invertase inhibitor gene can potentially control sucrose accumulation at the post-translation level by the research that has been conducted by [15].

Various gene sequences of invertase inhibitors from multiple plant species, including sugarcane varieties, are aligned and made phylogeny trees using Neighbour-Joining (Figure 2). For the analysis of the phylogeny tree, six *sininh* gene fragments were identified, in addition to involving 16 invertase inhibitor genes from the *Solanaceae* family (consisting of potatoes, chilies, eggplant, and tobacco), two from the *Brassicaceae* family (Arabidopsis thaliana), two from the *Rosaceae* family, and 13 from the *Poaceae* family (sugarcane, corn, and sorghum). The results of the phylogeny tree showed that all six *sininh* genes had close

kinship with the PME/Invertase inhibitor-like protein of *Zea mays* (*ZmPMEI2,* NM 001155260.3), pectinesterase inhibitor 12 of *Zea mays* (*ZmPMEI3*, XM 020552657.3); invertase inhibitor genes of *Zea mays* clone 329934 (*ZmINH1*, EU969422.1), pectinesterase inhibitor 12 of *Sorghum bicolor* (*SbPMEI2*, XM 002454411.2); and invertase inhibitor genes from *Saccharum robustum* (*SrINVINH*, KP055631.1), *Saccharum spontaneum* (*SsINVINH1*, KP844455.1), *Saccharum hybrid* var. Q208 (*SlNVINH*, MG457818.1), *S. sinense* (*SsINVINH2*, KP997206.1), *S. officinarum* (*SoINVINH*, KP997207.1), *Saccharum barberi* var. Khadya (*SbINVINH2*, KU057162.1), and *S. barberi* (*SbINVINH1*, KU167101.1) [15].

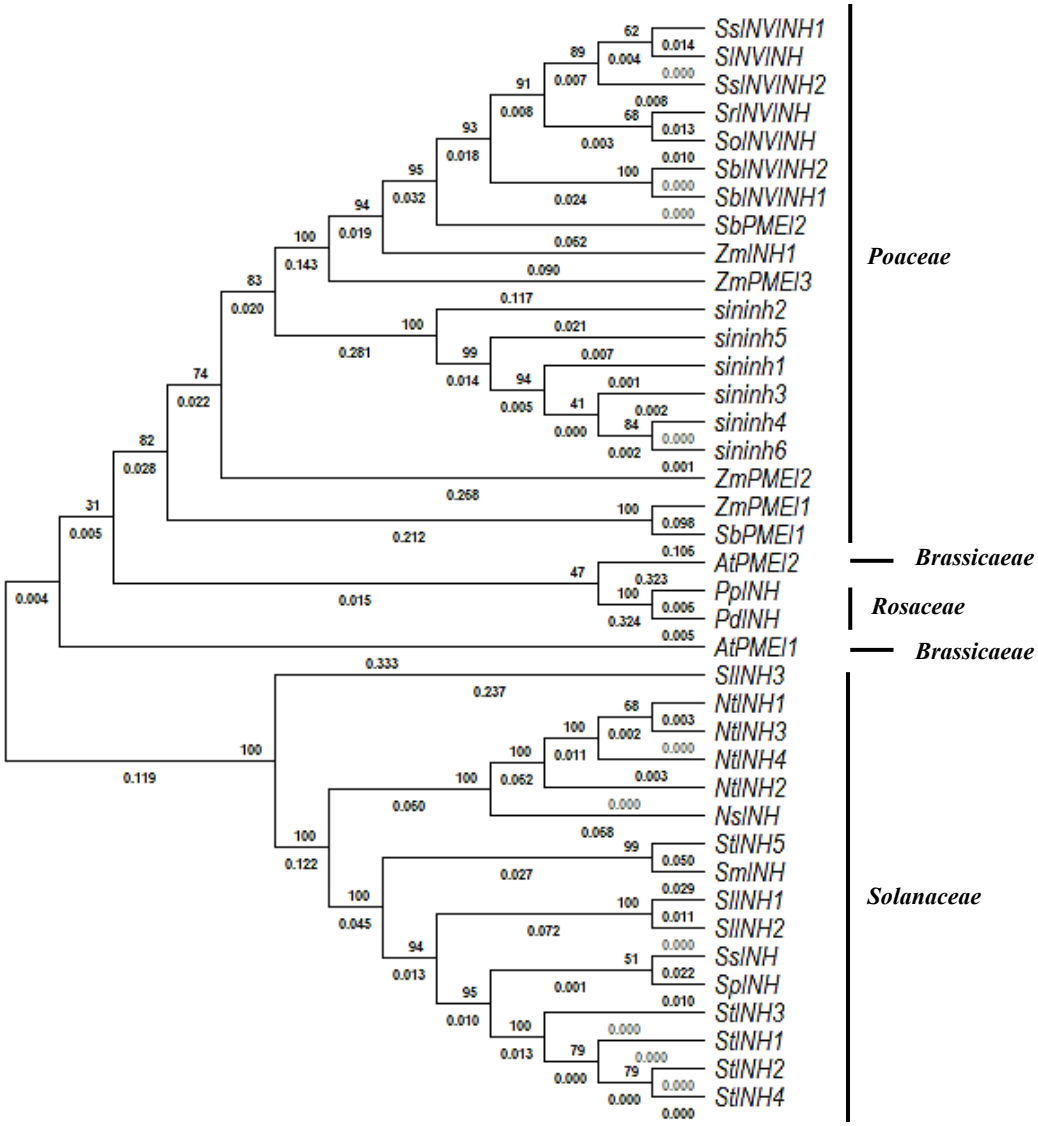


Figure 2. Phylogenetic relationship of several plant invertase inhibitor gene from *S. spontaneum* (*SsINVINH1*, KP844455.1); *Saccharum* hybrid (*SlNVINH*, MG457818.1); *S. sinense* (*SsINVINH2*, KP997206.1); *S. robustum* (*SrINVINH*, KP055631.1); *S. officinarum* (*SoINVINH,* KP997207.1); *S. barberi* (*SbINVINH1,* KU167101.1; *SbINVINH2,* KU057162.1); *Zea mays* (*ZmINH1,* EU969422.1); *ZmPMEI1,* XM 008660364.1; *ZmPMEI2*, NM 001155260.3; *ZmPMEI3,* XM 020552657.3); *Sorghum bicolor* (*SbPMEI1*, XM 002467406.2; *SbPMEI2,* XM 002454411.2); PS 881 (*sininh1*); PS 882 (*sininh2*); PSJT 941 (*sininh3*); PS 862 (*sininh4*); BL (*sininh5*); KK (*sininh6*); *Arabidopsis thaliana* (*AtPMEI1,* NM 130335.3; *AtPMEI2,* NM 118658.4); *Prunus persica* (*PpINH,* XM 007221481.2); *Prunus dulcis* (*PdINH,* XM 034351093.1); *Nicotiana tabaccum* (*NtINH1,* XM 016614445.1; *NtINH2,* AY145781.1; *NtINH3,* AY594179.1); *Nicotiana tomentosiformis* (*NtINH4,* XM 009591851.3); *Nicotiana sylvestris* (*NsINH,* XM 009765416.1); *Solanum lycopersicum* (*SlINH1,* NM 001329220.1; *SlINH2,* KC007465.1; *SlINH3,* NM 001247862.2); *Solanum tuberosum* (*StINH1,* MK405606.1; *StINH2,* GU321341.1; *StINH3,* MK405603.1; *StINH4,* JQ269669.1); *Solanum torvum* (*StINH5,* KC884746.1); *Solanum mauritianum* (*SmINH,* MK473854.1); *Solanum stoloniferum* (*SsINH,* MK405615.1); dan *Solanum pinnatisectum* (*SpINH,* MK405617.1)

Analysis of Primary Structure Characterization of Predicted *Sininh* Protein

The *sininh* genes in sugarcane varieties PS 881, PSJT 941, PS 862, and KK have the same number of amino acids and isoelectric points of 134 and 10.68, respectively, but they have different molecular weights. As for sugarcane varieties, Sininh protein from PS 882 and BL have a molecular weight of 12.76 kDa and 15.39 kDa, respectively. With almost the same number of amino acids, the molecular weight of the six sugarcane varieties is close to the molecular weight of the invertase inhibitor protein of the tobacco plant (*Nt-CIF*) by 16-20 kDa [21] and *Saccharum* hybrid var. Q208 of 18.17 kDa [8]. The isoelectric point (pI) is the pH of a solution at which the net charge of a protein becomes zero. At solution pH above the pI, the protein's surface is predominantly negatively charged, and therefore like-charged molecules will exhibit repulsive forces. Likewise, at a solution pH below the pI, the protein's surface is predominantly positively charged, and repulsion between proteins occurs. However, at the pI, the negative and positive charges are balanced, reducing repulsive electrostatic forces, and the attraction forces predominate, causing aggregation and precipitation [22].

The predicted amino acid sequences of *Sininh* proteins from various sugarcane varieties were aligned with several similar protein amino acid sequences with a similarity of 74.4% – 99.48%; these include ZmINVINH_*Z. mays*_invertase inhibitor (AKU19487.1), SbPMEI_*S. bicolor*_pectinesterase inhibitor 12 (XP_002454456.2), SbINVINH_*S. barberi*_invertase inhibitor (ANA08080.1), SoINVINH_*S. officinarum*_invertase inhibitor (ALG64884.1), SsINVINH_*S. sinense*_invertase inhibitor (ALG64883.1), and SspINVINH_*S. spontaneum*_invertase inhibitor (AKU19487.1).

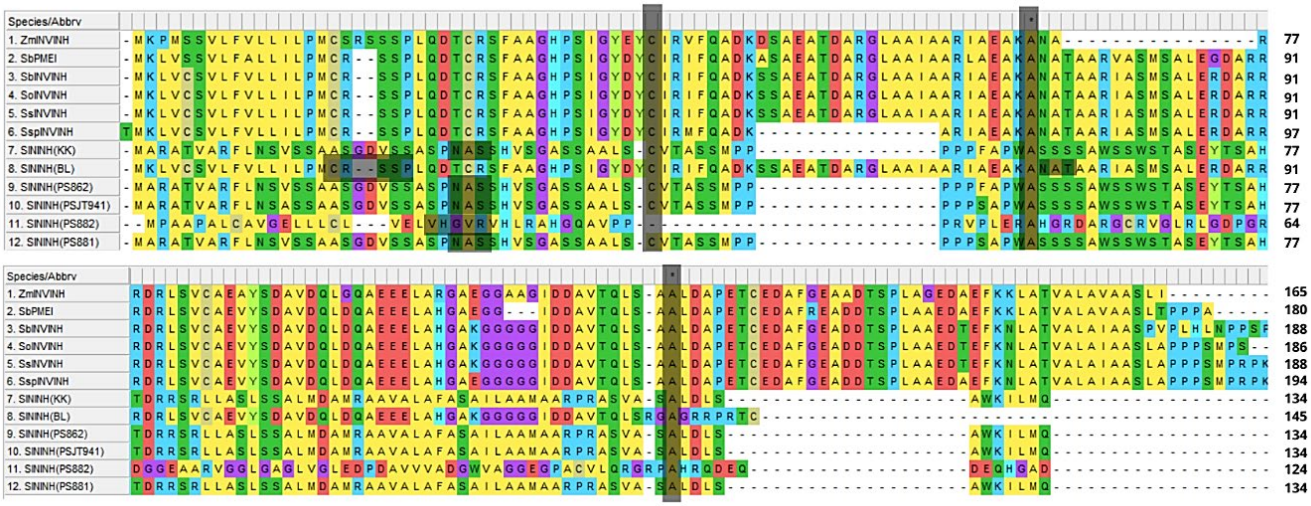


Figure 3. Comparison of predicted *Sininh* protein from several sugarcane varieties with plant invertase inhibitor-like protein. Multiple sequence alignment results showed the presence of Cys residual homology (except *Sininh* from PS882 sequence), N-glycosylation ASN site (NAS-NAT), and peptide signal sequence (CR—SSP; VHGVR)

Based on the results of the alignment of amino acid sequences in Figure 3, one homologous Cys residue was found except for the predicted *Sininh* protein sequence of the PS 882 variety. However, some Cys residues were also detected in the predicted *Sininh* protein sequences from PS882 and BL varieties. These Cys residues are engaged in disulfide bridge formation. Moreover, they were also found to contribute to the structural stabilization of the protein, and it has been observed that the well-conserved N-terminal end helical hairpin extension is not only crucial for the structural integrity and activity of the protein but also the conserved C-terminal domain is thought to contribute to the interface stabilization of the protein. It may affect the complex formation as well as stabilization of invertase and invertase inhibitors [23,3]. The result of the prediction of the disulfide bond is based on the position of the Cys sequence using the DiANNA 1.1 software (Table 1). If these two protein scores are compared (based on the table above), then the BL variety *Sininh* protein has a stronger bond with more bond numbers and a high score of up to 0.99680 in disulfide bonds 17 – 145.

Polypeptides that form after the translation stage will experience translocations at various locations inside and outside the cell. The translocation process is inseparable from several system components: signal peptide sequences, SRP (Signal Recognize Particle), and transmembrane systems. In the N-terminal part of the amino acid, there is a signal sequence that can connect with SRP so that it helps in the translocation of amino acids through the membrane, both the endoplasmic reticulum membrane and the cell

membrane. Proteins that have signal sequences in their N-terminal region have the opportunity to be exposed to the outside of the cell [24]. Signal sequences were only detected in predicted *Sininh* protein from PS 882 and BL varieties, based on signal peptide analysis of all six invertase inhibitor proteins. The *Sininh* protein of PS 882 cleavage site was found between headings 23 and 24 (between VHG and VR) with a probability of 0.7164. As for the *Sininh* protein in the BL variety, the cleavage site is between headings 19 and 20 (between CRS and SP) with a probability of 0.9509.

Table 1. Predicted disulfide bonding and state prediction of *Sininh*

Variety	Cystein Sequence Position (DiANNA Weighted matching)	Bond	Score
PS 882	8 - 53	AAPALCAVGEL - RDARGCRVGLR	0.01037
	16 - 103	GELLLCLVELV - GEGPACVLQRG	0.37268
BL	5 - 26	XMKLVCSVLFV - PLQDTCRSFAA	0.01037
	41 – 98	IGYDYCIRIFQ - DRLSVCAEVYS	**0.93681**
	17 - 145	LILPMCRSSPL - RRPRTCXXXXX	**0.99680**

Moreover, Table 2**.** shows that the predicted *Sininh* protein from PS 882 has a stable value (36.55) in the instability index (II). It is related to the expected disulfide bonding in the stability structure. Predicted *Sininh* protein from BL also shows a value near with stable index (43.70). The dipeptide composition-based Instability Index (II) is one of the protein's primary structure-dependent methods available for in vivo protein stability predictions. As per this method, proteins with an II value below 40 are stable proteins [25].

Table 2. Physiochemical characteristics of the primary structure of predicted *Sininh* protein

Variety	Nucleotide lenght (bp)	The number of amino acid	Molecule weight (Da)	Isoelectic Point (pH)	The number of atom	Aliphatic Index (AI)	Instability Index (II)
PS 881	616	134	13489.20	10.68	1873	77.61	50.61 (unstable)
PS 882	630	124	12759.45	5.81	1782	92.82	**36.55 (stable)**
PSJT 941	620	134	13461.15	10.68	1867	76.19	50.61 (unstable)
PS 862	629	134	13549.30	10.68	1882	77.61	50.61 (unstable)
BL	627	145	15387.48	7.60	2147	85.72	43.70 (unstable)
KK	655	134	13549.30	10.68	1882	77.61	50.61 (unstable)

Protein commonly consists of one or more submolecular parts, which are termed as the domain. A domain is a structural or functional module of protein, and it is usually evolutionarily conserved units. Differential association of domains provides a way to create new functions for organisms [26]. The interactions between domains can help in locating a protein at a specific subcellular site, recognizing protein post-translational modification or participating in signal transduction [27]. The interactions can also regulate the enzymatic activity, vigor, and substrate specificity [28]. Based on the domain prediction analysis of amino acid sequences of all sugarcane varieties, only protein sequences of BL varieties are predicted to have domains. This functional domain is Plant invertase/pectin methylesterase inhibitors, where they inhibit the activity of pectin methylesterase (PMEs) and invertase with a complex formation of non-covalent 1:1. The domain is located in the N-terminal region of PMEs predictions from DNA sequences [29], indicating that PMEs and their inhibitors are expressed as a single polyprotein and can then be processed. Both are connected by two main disulfide bridges, alpha-helix [30].

In Figure 3, small motifs of PKF (Proline, Lysine, and Phenylalanine) are not found in all amino acid sequences in both sugarcane varieties and other species. The small motif PKF is a critical sequence for invertase–inhibitor interaction. It has a target Cell Wall Invertase (CWI) through physical binding to substrate cleft in a pH-dependent manner in tobacco [23,31]. In sugarcane plants, N-glycosylation (ASN-glycosylation) motifs were found in the predicted sequences *Sininh* protein of the PS 881, PSJT 941, PS 862, and KK (at aa 27-30) and BL (in aa 73-76) varieties, which are also critical to the apoplastic/cell wall invertase inhibitor. Moreover, *N-glycosylation* plays a vital role in protein folding, glycan-dependent quality control processes in the endoplasmic reticulum, protein stability, and protein-protein interactions [32-34]. Asparagine *N-glycosylation* motifs are a significant post-translational modification in eukaryotes. It involves many enzymes. There are for lipid-linked production, nucleotide-conjugated sugar monomers,

enzymes for oligosaccharide synthesis and maturation [35,36]. The N-glycosylation motifs in *Sininh* protein can synthesize oligosaccharides such as sucrose from sugar monomers like fructose and glucose. The asn-Xaa-Ser/Thr (N-X-S/T) partial in which X represents any amino acid other than P (Proline), is the site of the acceptor or side-chain group of amino acid residues asparagine [37]. In addition to the N-glycosylation (ASN-glycosylation) motif, there are also other small motifs such as Casein kinase II phosphorylation site, N-myristoylation site, Protein kinase C phosphorylation site on the predicted *Sininh* protein at PS881, PSJT 941, PS 862, BL, and KK varieties but not for the PS 882 variety which only found small motif N-myristoylation site.

Subcellular Localization of Predicted *Sininh* Protein

The invertase enzyme is an extracellular protein [23]. According to [38], based on the discovery location, the invertase enzyme is then classified into CWINs (Cell wall invertases), VINs (Vacuolar invertases), and CINs / Cytoplasmic invertases. Acid invertases are either cell wall-bound (extracellular, insoluble) or localized in the vacuoles (vacuolar, soluble), whereas neutral invertases function in the cytosol, plastids, and mitochondria [39,40].

The predicted *Sininh* protein from the PS 882 and BL/Bulu Lawang varieties (Table 3) have a great chance of being extended to extracellular, including cell walls (44.4% and 55.6%, respectively). Therefore, it is predictable that the *Sininh* proteins of the PS 882 and BL varieties can work and potentially perform mechanisms in inhibiting the activity of CWINs (cell wall invertases). In addition, the predicted *Sininh* protein from the PS 882 variety also has a great chance of being supported in the cytoplasmic area (by 44.4%). The results show that the predicted *Sininh* protein of the PS 882 variety can work and perform mechanisms in inhibiting the activity of CINs (cytoplasmic invertases). The predicted *Sininh* protein of the PS 881, PSJT 941, PS 862, and KK varieties have a very low export probability (%) in extracellular/cell wall and cytoplasmic, making it less likely or less competent in inhibiting invertase activity. It is also related to the absence of peptide signal sites in the predicted *Sininh* proteins that play a role in helping the translocation of cellular proteins out of cells.

Based on the pH value of the overall *Sininh* protein prediction, the variety shows 6.5 [41]. Acid invertases have an optimum pH between 4.5-5.5 and neutral invertases optimally at pH 7.0-7.8. [42,39].

Table 3. Subcellular localization of predicted *Sininh* protein

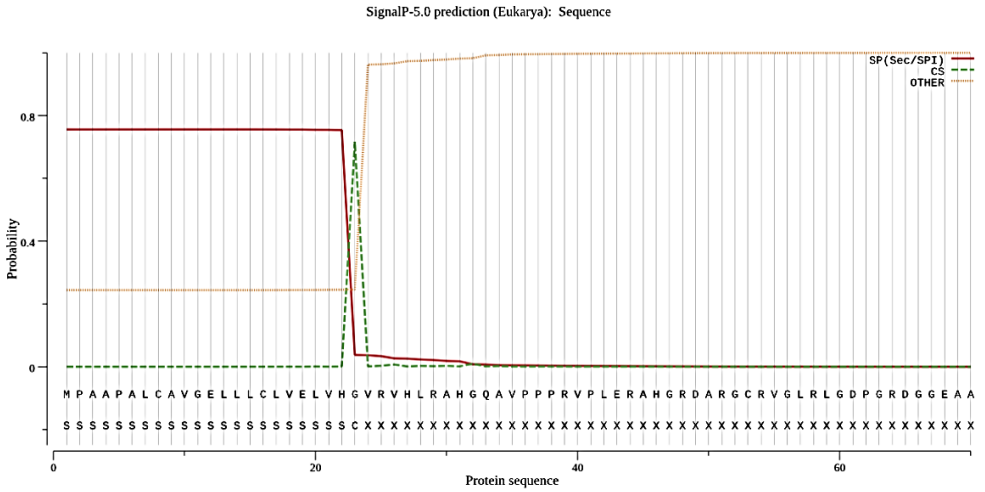
Variety	Export probability (%)						
	Mitochondrial	**Nuclear**	**Golgi**	**Cytoplasmic**	**Plasma membrane**	**Extracellular (include cell wall)**	**Cytoskeletal**
PS 881	**30.4**	**30.4**	13.0	8.7	8.7	4.3	4.3
PS 882	11.1	-	-	**44.4**	-	**44.4**	-
PSJT 941	**30.4**	**30.4**	13.0	8.7	8.7	4.3	4.3
PS 862	**30.4**	**30.4**	13.0	8.7	8.7	4.3	4.3
BL	22.2	-	-	**22.2**	-	**55.6**	-
KK	**30.4**	**30.4**	13.0	8.7	8.7	4.3	4.3

The Secondary and 3D Structure of Predicted *Sininh* Protein

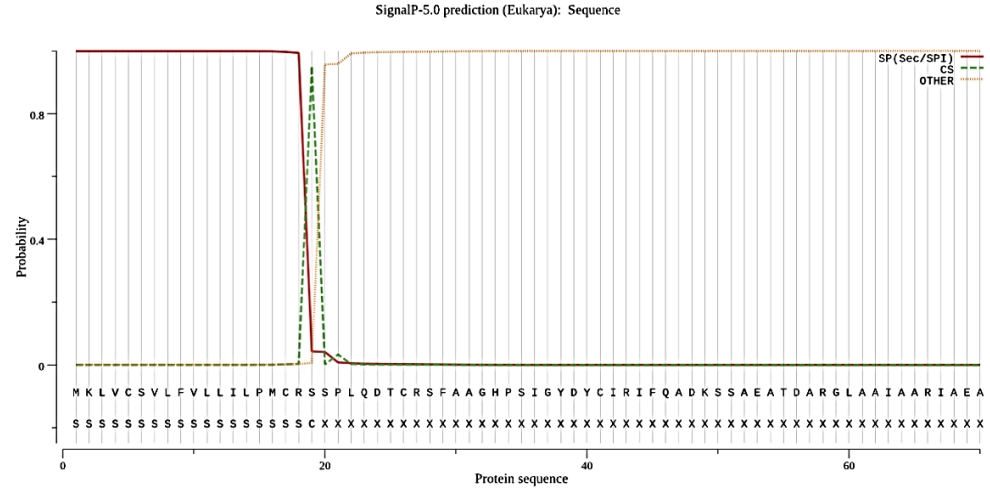
Based on Figure 4, the secondary structure of predicted *Sininh* protein in PS 881, PSJT 941, PS 862, and KK varieties has an almost balanced composition between α-helix with coils that are 55% (consisting of seven α-helix) and 45% in the absence of β-strand arrangement. The secondary structure of the predicted *Sininh* protein of BL variety also has no β-strand components. Still, it has a dominant α-helix composition of 77% (consisting of eight α-helix), followed by a 23% fewer coil. In contrast to the secondary structure of the PS 882 variety *Sininh,* which has mostly coil composition (62%), followed by α-helix (23%; consisting of three α-helix)) and β-strand (15%), which is less based on the results of online PSIPRED software.

The 3D structure of the predicted *Sininh* protein of various sugarcane varieties is shown in Figure 5. The structure of the predicted *Sininh* protein in PS 882 and BL varieties are different from other varieties; it corresponds to the conformation of amino acids. In the predicted *Sininh* proteins, the PS 882 and BL varieties have some Cys residues that can form disulfide bonds that play a role in the stability of the 3D structure. In Figures 5g and 5h, there is a comparison of 3D structures where the persisted *Sininh* protein PS 882 (Figure 5g) variety consists of the helix, coil, and strand structures, while BL (Figure 5h) consists only of helix and coil structures.

Studies related to invertase inhibition by specific proteinase inhibitors in sugarcane plants are still not widely known. Nevertheless, some research about these studies has been done. In 2014, a study from [15] was successfully



(a)



(b)

Figure 4. Peptide signals are found only in amino acid sequences of invertase inhibitors PS 882 and BL variety. (a) PS 882 = Cleavage site between pos. 23 and 24: VHG-VR. Probability: 0.7164; (b) BL = Cleavage site between pos. 19 and 20: CRS-SP. Probability: 0.9509

Figure 5. An overview of the secondary structure of predicted *Sininh* protein by PSIPRED software

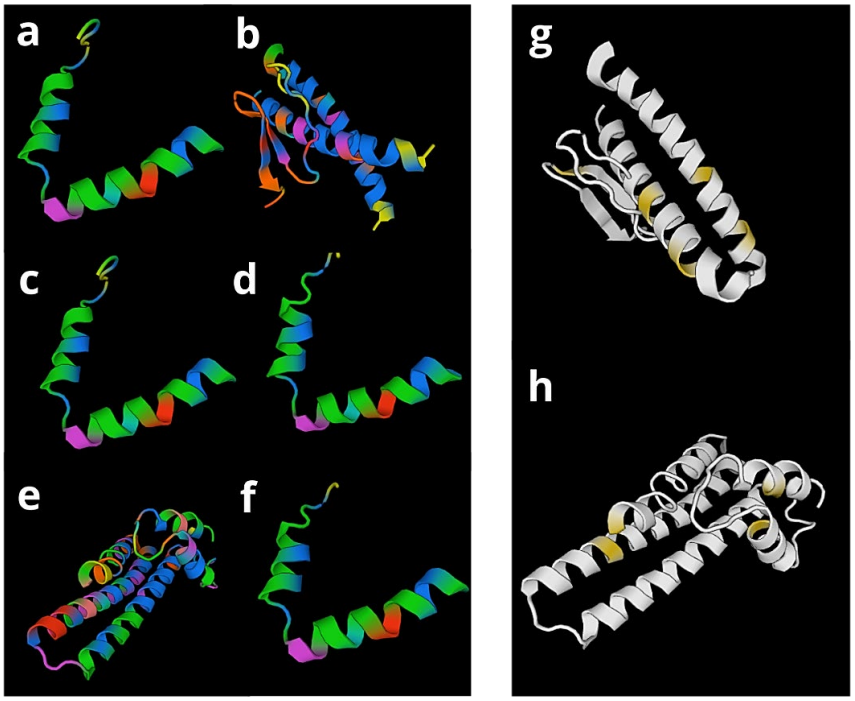


Figure 5. An overview of the 3D model of predicted *Sininh* protein by SWISS-MODEL. **(a)** PS 881; **(b)** PS 882; **(c)** PSJT 941; **(d)** PS 862; **(e)** BL/Bulu Lawang; **(f)** KK/Kidang Kencana; **(g and h)** PS 882 and BL/Bulu Lawang respectively with Cys region in yellow spot

cloned, characterized, and analyzed the expression profile of the invertase gene (soluble acid invertase) in *S. spontaneum*. The results showed that the *Sininh* protein was explicitly able to regulate in vivo invertase activity which regulates sucrose accumulation. In comparison, detailed structural and functional characterization studies on the expression, purification, and subcellular localization analysis of the invertase inhibitor (ShINH1) from sugarcane were reported [8]. However, characterization of the invertase inhibitor gene and biochemical studies of the protein has not yet been reported. These studies would be the foundation for studying the structural aspects and the mechanism of the inactivation of invertase via its inhibitory proteins at the molecular level. Moreover, it suggests the invertase inhibitor enzyme may be involved in invertase control in vitro.

CONCLUSION

Isolation of the *sininh* gene in various sugarcane varieties in Indonesia produced gene fragments along 616 bp - 655 bp, which identified 99.66% - 95.09%, with gene invertase inhibitors in sugarcane species *S. officinarum* and *S. sinense*. The encoded *Sininh* protein was predicted to contain amino acids ranging from 124 - 145 residues and the molecular weight of 12.76 kDa – 15.39 kDa. Peptide signals were found only in the predicted *Sininh* protein from PS 882 and BL/Bulu Lawang varieties. Disulfide bonds can also form in the predicted *Sininh* protein from BL variety with a score of 0.99680 in disulfide bonds 17 - 145. Based on domain prediction analysis on amino acid sequences of all sugarcane varieties, only *Sininh* protein sequences of BL variety were predicted to have a domain. This functional domain is Plant invertase/pectin methylesterase inhibitors, which inhibit the activity of pectin methylesterase (PMEs) and invertase with a complex formation of non-covalent. N-glycosylation motifs were also found in the predicted *Sininh* protein from PS 881, PSJT 941, PS 862, BL, and KK varieties, related to potential stability and interaction with other proteins. Subcellular localization of the predicted *Sininh* proteins was averaged mitochondrion and nuclear, except in PS 882 and BL varieties, which were cytoplasmic and extracellular, including cell walls. These studies build the foundation for studying the structural aspects and the mechanism of the inactivation of invertase via its inhibitory proteins at the molecular level.

ACKNOWLEDGEMENTS

The authors thank the authority of the Indonesia Center for Biodiversity and Biotechnology for their necessary and logistic support in the progress of the research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

REFERENCES

1. Jadoski, C. J., Toppa, E.V.B., Julianetti, A., Hulshof, T., Ono, E.O., and Rodrigues, J.D. (2010). Physiology development in the vegetative stage of sugarcane. *Appl Res and Agrotech*. 3, 177-185.
2. Chiou, T.J., and Bush, D.R. (1998). Sucrose is a signal molecule in assimilate partitioning. *Proc Natl Acad Sci USA.* 95, 4784-4788.
3. Rausch, T., and Greiner, S. (2004). Plant protein inhibitor of invertases. *Biochim et Biophysic Acta*. 1696(2), 253-261.
4. Chandra, A., Jain, R., and Solomon, S. (2012). Complexities of invertases controlling sucrose accumulation and retention in sugarcane. *Curr Sci*. 102(6), 857–866.
5. Wind, J., Smeekens, S., and Hanson, J. (2010). Sucrose: Metabolite and signaling molecule. *Phytochemistry.* 71(14-15), 1610-1614.
6. Wang, L., Zheng, Y., Ding, S., Zhang, Q., Chen, Y., and Zhang, J. (2017). Molecular cloning, structure, phylogeny and expression analysis of the invertase gene family in sugarcane. *BMC Plant Biol*. 17(1), 1-20.
7. Tana, B., Chanprame, S., Tienseree, N., and Tadakittiresarn, S. (2014). Realationship between invertase enzyme activities and sucrose accumulation in sugarcane (*Saccharum* spp.). *Kasetsart J (Nat. Sci)*. 48(6), 869-879.
8. Shivalingamurthy, A.G., Anangi, R., Kalaipandian, S., Glassop, D., King, G.F., and Rae, A.L. (2018). Identification and functional characterization of sugarcane invertase inhibitor (*ShINH1*): A potential candidate for reducing pre- and post-harvest loss of sucrose in sugarcane. *Front Plant Sci.* 9, 1-14.
9. Ryan, C.A. (1989). Proteinase inhibitor gene families: Strategies for transformation to improve plant defenses against herbivores. *BioEssays.* 10(1), 20-24.
10. Brancho, G.E., and Whitaker, J.R. (1990). Purification and partial characterization of potato Solanum tuberosum invertase and its endogenous proteinaceous inhibitor. *Plant Physiol.* 92, 386-394.
11. Pressey, R. (1967). Invertase inhibitor from potatoes: purification, characterization, and reactivity with plant invertases. *Plant Physiol.* 42, 1780-1786.
12. Pressey, R. (1994). Invertase inhibitors in tomato fruit. *Phytochemistry*. 36(3), 543-546.
13. Greiner, S., Krausgrill, S., and Rausch, T. (1998). Cloning of a tobacco apoplasmic invertase inhibitor. *Plant Physiol*. 116(2), 733-742.
14. Bate, N.J., Niu, X., Wang Y., Reimann, K.S., and Helentjaris, T.G. (2004). An invertase inhibitor from maize localize to the embryo surrounding region during early kernel development. *Plant Physiol.* 134, 246-254.
15. Prathima, P.T., Suparna, T.V., Anishma, S., Punnya, R., and Ramalakshmi, K. (2014). Cloning and characterization of a differentially regulated invertase inhibitor gene during sucrose accumulation in sugarcane. *Journal of Sugarcane Research*. 4(2), 21-28.
16. Gutierrez-Miceli, F.A., Rodriguez-Mendiola, M.A., Ochoa-Alejo, N., Mendez-Salas, R., Arias-Castro, C., and Dendooven, L. (2005). Sucrose accumulation and enzyme activities in callus culture of sugarcane. *Plant Physiology*. 49(3), 475-479.
17. Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Plant Physiology*. 15, 473-497.

18. Hall, T. (2011). BioEdit: An important software for molecular biology. *GERF Bulletin of Biosci.* 2(1), 60-61.

19. Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *MBE Brief Com*.

20. Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Lu, F., Matchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Geer, L.Y., and Bryant, S.H. (2017). CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res*. 45, 200-203.

21. Hothorn, M., Bonneau, F., Stier, G., Greiner, S., and Scheffzek, K. (2003). Bacterial expression, purification, and preliminary X-ray crystallographic charaterizatoin of the invertase inhibitor *Nt-CIF* from tobacco. *Acta Crystallography.* 59(12), 2279-2282.

22. Novak, P. and Havlicek, V. 2016. *Protein extracton and precipitation*. *In:* P Ciborawski and J Silberring (Ed). Proteomic profilling and analytical chemistry: The Crossroad (Second edition). *Academic Press.* London., pp. 51-62.

23. Hothorn, M., Wolf, S., Aloy, P., Greiner, S., and Scheffzek, K. (2004). Structural insights into the target specificity of plant invertase and pectin methylesterase inhibitory proteins. *The Plant Cell.* 16(12), 3437-3447.

24. Freudl, R. (2018). Signal peptides for recombinant protein secretion in bacterial expression systems. *Microb. Cell Fact*. 17(1), 1-10.

25. Gamage, D.G., Gunaratne, A., Periyannan, G.R., and Russell, T.G. (2019). Applicability of instability index for in vitro protein stability prediction. *Protein Peptide Lett.* 26, 339-347.

26. Bitard-Feildel, T., Kemena, C., Greenwood, J.M., and Bornberg-Bauer, E. (2015). Domain similarity based orthology detection. *BMC Bioinformatics.* 16(154), 1-11.

27. Li, X., Yang, L., Zhang, X., and Jiao, X. (2019). Prediction of protein-protein interaction based on domain. *Comput and Math Methods in Med*. 2019, 1-7.

28. Pawson, T., and Nash, P. (2000). Protein-protein interactions define specificity in signal transduction. *Genes Dev.* 14, 1027-1047.

29. Giovane, A., Balestrieri, C., Quagliuolo, L., Castaldo, D., and Servillo, L. (1995). A glycoprotein inhibitor of pectin methylesterase in kiwi fruit: Purification by affinity chromatography and evidence of a ripening-related precursor. *Eur J Biochem.* 233(3), 926-929.

30. Camardella, L., Carratore, V., Ciardiello, M.A., Servillo, L., Baletrieri, C., and Giovane, A. (2000). Kiwi protein inhibitor of pectin methylesterase; Amino-acid sequence and structural importance of two disulfide bridges. *Eur J Biochem*. 267, 4561-4565.

31. Hothorn, M., den Ende, W.V., Lammens, W., Rybin, V., and Scheffzek, K. (2010). Structural insights into the pH-controlled targeting of plant cell-wall invertase by a specific inhibitor protein. *PNAS*. 107(40), 17427-17432.

32. Moremen, K.W., Tiemeyer, M., and Nairn, A.V. (2012). Vertebrate protein glycosylation: diversity, syntesis, and function. *Mol Cell Biol.* 13, 448-462.

33. Hebert, D.N., Lamriben, L., Powers, E.T., and Kelly, J.W. (2014). The intrinsic and extrinsic effects of N-linked glycans on glycoproteostasis. *Nat Chem Biol.* 10, 902-910.

34. Datir, S.S., Mirikir, D., and RaviKumar, A. (2019). Sequence diversity and in silico structure prediction of the vacuolar invertase inhibitor gene from potato (*Solanum tuberosum* L.) cultivars differing in sugar content. *Food Chem*. 295, 403-411.

35. Strasser, R. (2016). Plant protein glycosylation. *Glycobiology*. 26(9), 926-939.

36. Nagashima, Y., von Schaewen, A., and Koiwa, H. (2018). Function of N-glycoslation in plants. *Plant Sci*. 274, 70-79.

37. Lowenthal, M.S., Davis, K.S., Formolo, T., Kilpatrick, L.E., and Phinney, W. (2016). Identification of novel N-glycosilation sites at non-canonical protein consensus motifs. *J Proteome Res*. 15, 2087-2101.

38. Wan, H., Wu, L., Yang, Y., Zhou, G., and Ruan, Y.L. (2017). Evolution of sucrose metabolism: The dichotomy of invertase and beyond. *Trends in Plant Science*. 23(2), 163-177.

39. Sturm, A., Hess, D., Lee, H.S., and Lienhard, S. (1999). Neutral invertase is a novel type of sucrose-cleaving enzyme. *Physiologia Plantarum*. 107(2), 159-165.

40. Slugina, M.A., Shchennikova, A.V., and Kochieva, E.Z. (2017). TAI vacuolar invertase orthologs: The interspecific variability in tomato plants (*Solanum* section Lycopersicon). *Mol Genet and Genomics.* 292(5), 1123-1138.

41. Anandakrishnan, R., Aguilar, B., and Onufriev, A.V. (2012). H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Research*. 40, 537-541.

42. Ruan, Y.L., Jin, Y., Yang, Y.J., Li, G.J., and Boyer, J.S. (2010). Sugar input, metabolism, and signaling mediated by invertase: Roles in development, yield potential, and response to drought and heat. *Mol Plant*. 3(6), 942-955.