Molecular Cloning And Characterization of An Alpha-Amylase Inhibitor (Tkaai) Gene From *Trichosanthes Kirilowii* Maxim

Yipeng Zhang (zypmail.hzau.edu.cn@webmail.hzau.edu.cn)
HUAZHONG Agricultural University https://orcid.org/0000-0002-7232-654X

Keyue Wang
HUAZHONG Agricultural University

Qiyuan Huang
HUAZHONG Agricultural University

Shaohua Shu
HUAZHONG Agricultural University https://orcid.org/0000-0001-5230-7707

Research Article

Keywords: Trichosanthes kirilowii Maxim, Alpha-amylase inhibitors, Gene cloning, Enzymatic reaction, Diabetes

Posted Date: January 27th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1194994/v1

Abstract

Trichosanthes kirilowii Maxim., taxonomically belongs to the Cucurbitaceae and Trichosanthes genus, and whose whole fruit, fruit peel, seed and root are widely used in traditional Chinese medicines (TCM). A ribosome-inactivating protein with RNA N-glycosidase activity called Trichosanthrip was isolated and purified from the seeds of *T. kirilowii* in the previous research. To further explore the biological functions of Trichosanthrip, the cDNA of *T. kirilowii* AAI (*TkAAI*) was cloned through rapid-amplification of cDNA ends and its sequence was analyzed, as well as the heterologous protein was expressed in *Escherichia coli* and its alpha-amylase activity was further measured under optimized conditions. The full-length cDNA of *TkAAI* was 613 bp, the speculated open reading frame sequence encodes 141 amino acids with a molecular weight of 16.14 kDa. Phylogenetic analysis demonstrated that the AAI_SS domain sequence of TkAAI revealed significant evolutionary homology with the of 2S albumin derived from the other plants in Cucurbitaceae. In addition, *TkAAI* was assembled into pET28a with *eGFP* to generate a prokaryotic expression vector and induced to express in *E. coli*. The TkAAI-eGFP infusion protein was proven to exhibit alpha-amylase inhibitory activity against porcine pancreatic amylase (PPA) in a suitable reaction system. Analysis of gene expression patterns proved that the relative expression level of *TkAAI* in seeds is highest. Above results forecast that the *TkAAI* might play a crucial role during the development of *T. kirilowii* seeds and provide insights into the possibility of *T.kirilowii*-derived medicine to treat diabetes related diseases.

Introduction

Trichosanthes kirilowii Maxim. is a perennial herbaceous liana of the Trichosanthes genus in the Cucurbitaceae family, and its distribution ranges from the eastern Himalayas in India and southern China through southern Japan, Malaysia, and tropical Australia. The fruit, seed, and root of the plant have been commonly used as traditional Chinese medicines. Moreover, it is considered as one of the 50 fundamental herbs [1]. The bioactive substances of *T. kirilowii* vary widely, such as trichosanthin, a ribosome inactivating protein extracted from the root tube, was reported to have properties of immunomodulatory, anti-tumor, and anti-human immunodeficiency virus (HIV) properties [2, 3]. Moreover, it was found that *T. kirilowii* seeds are energy-rich diet containing 62% oil, up to 30% proteins, 2.5% mono and oligosaccharides [4].

Alpha-amylases (a-1,4-glucan-4-glucanohydrolases; EC 3.2.1.1) are hydrolytic enzymes and widely distributed in nature [5], which catalyze the cleavage of the α-1,4 glycosidic linkages found in starch and other oligosaccharides, hydrolysates are molecules such as dextrin and reducing sugars [6]. Cereal seeds stock large amounts of the substrate for these enzymes, making them vulnerable to the attack by pests and herbivores [7]. Nonetheless, many plant species have developed defense systems against these attacks, which include the expression of a set of seed inhibitors acting on a range of amylases of different origins [7]. Alpha-amylase inhibitors are a kind of glycoside hydrolase inhibitors which was first isolated from wheat in 1933 [5]. Natural alpha-amylase inhibitors are numerous, mainly including peptidoglycan, alkaloids, glycosides, glycopeptides, polysaccharides, oligosaccharides and other

compounds [8], among which flavonoids are a kind of substance with significant inhibitory effect on alpha-amylase activity, and their inhibitory potential related to the number of hydroxyl groups in compound molecules [9].

Diabetes as a type of metabolic disorder and global disease, the number of cases increased year by year, and the patients showed a trend of younger. Targeting enzymes in the metabolic pathway of diabetes and designing specific enzyme inhibitors is a research direction for the treatment of diabetes in the future [10]. One of the vital enzymes associated with diabetes is alpha-amylases. For human beings, alpha-amylase inhibitors effectively restrain the activity of amylase in saliva and pancreatic to inhibit the hydrolysis and digestion of carbohydrates in food, reducing the absorption of glucose [11]. In addition, alpha-amylase inhibitors play a role in reducing blood fat levels and controlling food intake [12]. Consequently, it's used in clinical medicine for the prevention and treatment of hyperglycemia, hyperlipidemia and diabetes mellitus type II, etc [13]. Otherwise, alpha-amylase inhibitors likewise show a wide range of applications value in agriculture. AAI has proven to have defense related activity in botanical systems. Through the genetic engineering technology to carry *AAI* into crop cells or genome to generate genetically modified (gm) crops, resistance of plants could be improved to defend from diseases and pests and reduce the loss of crop yield [14, 15].

Recently, trichosanthrip, a novel RIP with a molecular mass of 10,964.617 Da, was purified from the seeds of *T. kirilowii*, and its enzyme activity was assayed[16]. Function analysis showed that the protein belongs to the AAI-LTSS superfamily which widely exist in the seeds of plants can inhibit the amylase activity of a protein, lipid transport and storage, and so on. In this paper, the alpha-amylase inhibitor gene from *T. kirilowii* (*TkAAI*, GenBank accession no. AJO68010.1.) was cloned and the expression pattern of *TkAAI* in different tissues of *T. kirilowii* was investigated. The TkAAI protein sequences, structure, and function were characterized by bioinformatics analyses. Protein was obtained via prokaryote expression in *Escherichia coli* and alpha-amylase activities were assayed. To our knowledge, this is the first alpha-amylase inhibitor gene characterized from *T. kirilowii*. The present study revealed sectional biological functions and mechanisms of trichosanthrip/TkAAI protein in *T. kirilowii* seeds from the perspective of biochemistry, molecular biology and bioinformatics. Providing a meaningful insight into the characterization and application value of this kind of multifunctional protein.

Materials And Methods

Plant materials

The leaves, stems, roots, flowers, fruits and seeds of *T. kirilowii* were collected from the medicinal botanical garden of Huazhong Agricultural University and immediately frozen in liquid nitrogen and stored at -80 ^{o}C for RNA extraction.

RNA isolation and cloning of *TkAAI*

Total RNA was isolated via the total RNA Kit (Promega). The extracted RNA was quantified using a NanoDrop 2000 Spectrophotometer (Thermo Scientific, San Jose, California, USA) and the integrity was further analyzed by 1% agarose gel electrophoresis. Prior to cDNA synthesis, total RNA was treated with a DNase to remove DNA contamination from the samples. The first cDNA fragment was synthesized by reverse transcription using the Oligo(dT) primer, AA7-F(5'-CCTACCGCACCACTATCACCA-3'), and 2 mg total RNA as the template according to the instructions of the PrimeScript@1stStrand cDNA Synthesis Kit (TaKaRa, Japan). In order to isolate the *TkAAI* specific fragments from *T. kirilowii*, a 326 bp fragment was amplified from cDNA prepared from seeds. Two Primers AA7-F (5'-CCTACCGCACCACTATCACCA-3') and AA7-R (5'- GAGACCATTTAGAAGTCGCATCG -3') which contained a conserved sequence were designed based on the homology of the *TkAAI* from *M. charantia* (gi|21327880|), *Cucurbita cv.* (gi|459404|). The degenerate PCR reaction was conducted in a total volume of 20 μL mixture containing 12.7 μL of ddH_2O, 2.0 μL of 2.5 mM dNTP mixture, 2.5 μL of Ex Taq buffer (Takara), 0.3 U Ex Taq (Takara), 0.5 mM of each primer, and 1.0 μL of cDNA (50 ng/μL). PCR was performed using the thermal cycle profile of 94 ^{o}C 3 min, 35 cycles of 94 ^{o}C 30 s, 55 ^{o}C 30 s and 72 ^{o}C 1 min, and a final extension of 10 min at 72 ^{o}C. One amplified product was recovered using the DNA rapid purification kit (Axygen). The purified PCR products were ligated into pMD-18T vector (Takara) and then transformed to competent *E. coli* DH 5α cells for sequencing. RACE-PCR was processed to obtain the 3'end and 5'end of the cDNA sequence of *TkAAI* via Terminal deoxynucleotidyltransferase as previously described [17, 18]. 3' RACE: The cDNA first strand was synthesized by reverse transcription using Q_T (5'-CCAGTGAGCAGAGTGACGAGGACTCGAGCTCAAGCTTTTTTTTTTTTTTTT- 3'), GSP-5F (5'-AAGAAGGCGGGTCCTTTGAT- 3') and Q_0 (5'- CCAGTGAGCAGAGTGACG-3') were designed to amplify 3' RACE products. 5' RACE: The cDNA first strand was synthesized by reverse transcription using AA7-R (5'-GAGACCATTTAGAAGTCGCATCG -3'), add poly A to the purified cDNA using TdT and dATP, QT(5'-CCAGTGAGCAGAGTGACGAGGACTCGAGCTCAAGCTTTTTTTTTTTTTTTT-3'), Q_0 (5'-CCAGTGAGCAGAGTGACG- 3') and GSP-3R (5'- ATCAAAGGACCCGCCTTCTT-3') were designed to amplify 5' RACE product. The RACE-PCR products were analyzed on 1.2% (w/v) agarose gels, purified, and cloned into the pMD-18T vector (Takara), and sequenced (Sangon, Shanghai). After the complete sequences of the 5'-end and the 3'-end were obtained, the full cDNA length of *TkAAI* was established by assembling analysis.

Characterization

The physical and chemical properties of the deduced TkAAI protein were determined using the Expasy ProtParam Tool (https://web.expasy.org/protparam/). The subcellular localization was predicted through TargetP-2.0 (http://www.cbs.dtu.dk/services/TargetP/). The sequence analysis was performed via the InterProScan online website(http://www.ebi.ac.uk/interpro/). The predicted TkAAI protein sequences were aligned with their orthologs and other members of TkAAI super family using the BLASTX program, from the database of NCBI (https://www.ncbi.nlm.nih.gov/). Multiple sequence alignments were viewed and edited using the ClustalW (https://www.genome.jp/tools-bin/clustalw) and ESPript (https://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi). For the phylogenetic analysis, TkAAI amino acids

sequences were aligned with their orthologs in other seed plant species using ClustalW and MAFFT software [19]. The evolutionary history was inferred by using MEGA7 with the Neighbor-Joining analysis method (https://robetta.bakerlab.org/)[20]. Phylogenetic tree was constructed online by itol (https://itol.embl.de/). TkAAI protein tertiary structure via SWISS-MODLE (https://swissmodel.expasy.org/interactive) and Robetta (https://robetta.bakerlab.org/). The accuracy of the tertiary structure models is determined by Savesv6.0 (https://saves.mbi.ucla.edu/).

Gene expression pattern analysis of *TkAAI* in various organs

Real-time PCR amplification and analysis was performed on CFX96™ Real-Time PCR System (Bio-Rad, California, USA) via SYBR® Premix Ex Taq™ (Tli RnaseH Plus) (TaKaRa, Japan). Primers (GSP-F 5'-AAGAAGGCGGGTCCTTTGAT -3', GSP-R 5'- TCTGCTCCTCCCTAGCAATCT -3') used for real-time PCR were designed to amplify 100-120 bp fragments from full-length cDNA of *TkAAI*. A constitutively expressed housekeeping gene, GAPDH (GAPDH-F 5'- TGCACTACCAACTGCCTAGC -3', GAPDH-R 5'-CCTTCACCAAGTCATCCCCC- 3') from cucumber, was used for normalization in the quantification of the gene expression in different tissues. qPCR was performed at a final volume of 20 μl containing 1 μL cDNA, 0.5 μL of each primer diluted to 10 mM, 10 μL SYBR® Premix Ex Taq™, and 7 μL ddH2O. The thermal cycle condition used in real-time PCR was: 94 ^{o}C for 1 min, followed by 40 cycles of 94 ^{o}C for 10 s, and 56 ^{o}C for 20 s. Following the real-time PCR cycles, the specificity of the SYBR green PCR signal was confirmed by melting curve analysis. Data analysis was performed according to the instructions of the manufacturer of the quantitative real-time PCR instrument CFX96™ management software. The expression level for each sample was calculated as $2^{-\Delta\Delta C(t)}$ where Ct represents the cycle number when the fluorescence signal in each reaction reaches the threshold. All samples were repeated three times.

Protein expression and purification

Firstly, amplify the *TkAAI* coding sequence and reporter gene *eGFP* fragments. Prokaryotic expression plasmid pET28a was linearized by PCR. Primers for *TkAAI*, (TkAAI-F 5'- CTTTAAGAAGGAGATAT-ACCATGGCAAGACTCACAGGTATCATTG- 3', TkAAI-R 5'- TCCTCGCCCTTGCTCACCATTGCT-GCTGCTGCTGCTGCGAAGGCGCATCGGTCT -3') eGFP (eGFP-F 5'- GCCCAGACCGATGCG-CCTTCGCAGCAGCAGCAGCAGCAATGGTGAGCAAGGGCGAGGAGCTG -3', eGFP-R 5'- CTTC-CTTTCGGGCTTTGTTAGTGGTGGTGGTGGTGGTGCTTGTACAGCTCGTCCATGCCGAG -3') and pET28a (Vector-R 5'- GCATGGACGAGCTGTACAAGCACCACCA-CCACCACCACTAACAAAGCCC-GAAAGGAAGCTGAGT-3', Vector-F 5'-ATACCTGTGAGTCTTGCCATGGTATATCTCCTTCTTAA-AGTTAAACAAAATTATTTCTAGAG-3') were used to amplified these three sequences, purified the PCR products using QIAquick PCR Purification Kit (Cwbio, China), three purified TkAAI and eGFP fragments were assembled into linearized pET28a to generate *TkAAI-eGFP*-pET28a plasmid. The recombinant vector was transformed into *E. Coli* BL21(DE3) strain for expression of TkAAI-eGFP fusion protein. Then *E. coli* BL21(DE3) cells harbouring the expression vector *TkAAI-eGFP*-pET28a were cultured in LB broth containing 50 μg/ml kanamycin overnight, the cells were cultured for 12 h with 0.2 mM isopropyl-b-D-thiogalacto-pyranoside (IPTG) in 16 ^{o}C for induction, harvested and then disrupted by sonication on ice.

After centrifugation (12,000 g, 10 min at 4 ^{o}C), the resulting supernatant was purified by Ni–NTA affinity column chromatography (TransGen, China) and His-tagged recombinant protein was eluted with a linear gradient of 100-300 mM imidazole in 50 mM 2-Amino-2-(hydroxymethyl)-1,3-propanediol (Tris-HCL) buffer (pH 7.9).

Western blotting analysis

Purified protein was analysis by 10% SDS–PAGE (80 V, 2.5 h) and transferred to the polyvinylidene difluoride (PVDF) membrane (100 V, 1.5 h). Then immersed the PVDF membrane into 1×TBST (3% (w/v) skim milk) for 2 hours at room temperature. Washed the membrane by 1×TBST for three times, incubated the membrane with primary antibody diluted in 1×TBST (3% (w/v) skim milk, 1:5,000) and shaken for 2 hours, washed the membrane as mentioned before. Incubated the membrane with anti-eGFP secondary antibody diluted in 1×TBST (3% (w/v) skim milk, 1:10,000) and shaken for 1 hour, washed the membrane as mentioned before. Specific binding was detected with omnipotent imaging system (Bio-Rad).

Alpha-amylase inhibitory activity assays

The purified protein was concentrated to 0.2 mg/mL by molecular weight cutoff (MWCO) with 10kDa ultrafiltration tube (Merk). 20 μL (4 μg) of TkAAI-eGFP solution and 20 μL of porcine pancreatic amylase (PPA) solution (1 units/ml) were incubated at 37 °^{o}C for 10 minutes, then reaction was started by the adding 50 μL 1% (w/v) soluble starch solution to 20 mM phosphate buffer (pH 6.9, containing 6.7 mM sodium chloride), and accurately incubated for 8 min at 37 °^{o}C. The reaction was immediately stopped by adding 50 μL NaOH (1 M) solution, filled in 50 μL DNS color reagent solution (Solarbio, Beijing) and heated up to 100 °^{o}C for color reaction for 10 min. After the reaction, cooled the mixture to room temperature and added 100 μL distilled water. The 200 μL samples were taken to measure the absorbance at 540 nm. The formula for calculating the inhibition activity of alpha-amylase inhibition was shown as:

$$Inhibition(\%) = 1 - \left(\frac{A_s - A_b}{A_t - A_c}\right) \times 100\%$$

A_s is the absorbance of the mixture of TkAAI-eGFP, starch solution, alpha-amylase solution and DNS color reagent solution. A_b is the absorbance of the mixture of TkAAI-eGFP, starch solution, phosphate buffer (replace alpha-amylase solution) and DNS color reagent solution. A_t is the absorbance of the mixture of phosphate buffer (replace TkAAI-eGFP), starch solution, alpha-amylase solution and DNS color reagent solution. A_c is the absorbance of the mixture of phosphate buffer (replace TkAAI-eGFP and alpha-amylase solution), starch solution, and DNS color reagent solution.

Results

Cloning the full-length cDNA of *TkAAI*

The total cDNA was extracted from *T. kirilowii* seeds. The splicing of RACE PCR products shows full-length sequence of the cDNA consisting of 585 bp nucleotides with a 423 bp open reading frame, encoding a 141 amino acids polypeptide (Supplementary Fig. 1). The deduced amino acid sequence of the cDNA showed that it encoded a polypeptide of approximately 16.14 kDa, with a calculated isoelectric point of 5.36. This gene was designated as *TkAAI* and the cDNA sequence was deposited in GenBank with accession number AJO68010.1.

Characterization

Blastp manifested that TkAAI belonged to TkAAI super-LTSS family, and was specifically classified as alpha-amylase inhibitors (AAIs) and seed storage (SS) proteins (Fig. 1a). ProtParam speculated that Arg was the richest amino acid in composition with frequency 14.2%. TkAAI did not contain His, Trp, Pyl and Sec. The instability index (II) was computed to be 59.44 which indicated the protein was unstable. SOPMA was employed to analysis the secondary structure of TkAAI and the result revealed that the protein contained 60.99% α-helices, 9.220% extended strand and 29.79% random coils (Fig. 1b). TargetP-2.0 predicated that from the first to 19th amino acid residues were presumed to be signal peptide region and the cleavage site was located in 22th amino residues, which meant that TkAAI protein was located in secretion pathway and secreted outside the cells in plants (Fig. 1c). In the meanwhile, the TkAAI protein contained five alpha-amylase binding sites including 41th, 94th, 95th, 96th and 103th amino acid residues (Fig. 1a). These results showed that TkAAI protein shared the typical characteristics of alpha-amylase inhibitors' proteins. Sequence analysis was performed through InterProScan and the result indicated that TkAAI contained Bifunctional inhibitor (trypsin-alpha amylase inhibitor)/plant lipid transfer protein/seed storage helical domain (55-139 amino acid residues) and belonged to Protease inhibitor/seed storage/LTP family (55-139 amino acid residues). In the meanwhile, AAI_SS conserved domain was founded in TkAAI (44-136 amino acid residues) (Supplementary Fig. 2). The comparison with NCBI database blastp further demonstrated that TkAAI had extremely high homology with the 2S albumin protein in Alpha-Amylase Inhibitors (AAIs) and Seed Storage (SS) Protein subfamily. Four 2S albumin from different species which had relatively high homology to TkAAI were selected for amino acids sequence alignment, the result revealed that TkAAI protein had 70.21–83.69% sequence identity to C.maxima2S (XP_022993226.1, *Cucurbita maxima*), C.moschata2S (XP_022942593.1, *Cucurbita moschata*), C.sativus2S (XP_011650534.1, *Cucumis sativus*) and Bh2S1 albumin-like (XP_038904176.1, *Benincasa hispida*) (Fig. 1d). The gene bank accession numbers and names referred in this article of 2S seed storage albumins selected from the 30 species were shown in Supplementary Table 1. The phylogenetic tree was clustered into three large branches, and phylogenetic analysis revealed that *TkAAI* was closely related to 2S albumin from other species in the Cucurbitaceae family, which gather in one of the big branches. This involved species phylogenetically related (Fig. 1e). The sequence was uploaded to SWISS-MODEL database employing comparative modeling, a valid three-dimensional model of TkAAI (Fig. 2b) was generated via Brazil Nut 2S albumin Ber e 1 as a template, identity value reached 37.37%. Nonetheless, GMQE (Global Model Quality Estimate) value was only 0.44, which means that the modelling of the protein was not sufficiently reliable. Therefore, the Robetta servers were employed to predict the tertiary structures of the TkAAI (Fig. 2a). To estimate the quality of protein models,

PROCHECK, Errat and Verify_3D of SAVESv6.0 servers was simultaneously employed to validate the accuracy and quality of the structures. Errat evaluated the three-dimensional structure of proteins based on crystallography to identify incorrect regions of protein, the overall quality factor of Robetta (100.00)'s protein model was apparently higher than the models from SWISS-MODEL (91.01). In addition, VERIFY3D indicated that the percentage of residues with averaged 3D-1D score >= 0.2 of protein models from Robetta was 70.21%, which was also higher than the models from SWISS-MODEL (44.12%) (Mahtarin et al., 2020). Besides, the results of 8 evaluations from PROCHECK likewise supported that the models from Robetta were significantly superior models from SWISS-MODEL in passage and error rates.

Gene expression pattern analysis of *TkAAI* in various organs

The mRNA expression level of *TkAAI* in *T. kirilowii* was investigated via real-time PCR. The results verified that *TkAAI* was detected in stems, fruits, roots, flowers and seeds. *TkAAI* expressed extensively in stems and fruits. Transcripts were relatively more abundant in seeds (**P < 0.01) (Fig. 3), followed were detected in stems and fruits. Consequently, the following investigation of the *TkAAI* expression could be focused on different developing period of seeds. In the early stage of seed germination, the activity of alpha-amylase was enhanced to catalyze the hydrolysis of starch stored in the endosperm to produce small molecular reducing sugars [21]. The decomposition of these sugars provides a direct source of energy for the elongation of seeds' coleoptile [22]. Hence, we deductive that *TkAAI* directly affects the activity of alpha-amylase in endosperm to regulate the formation, maturation and germination of *T. kirilowii* seeds.

Optimal design of heterologous expression of *TkAAI*

As a type of reporter gene, *eGFP* was introduced to downstream of *TkAAI*. In the meanwhile, 6-Histidine tag was introduced to C-terminus to purify fusion protein. In *E. coli* expression system, the adjustment of induction temperature (^{o}C), induction time (h) and inducer IPTG concentration (mM) all affected the expression of TkAAI protein. Hence, the present experiment adopts three significant factors affecting prokaryotic expression, using SPSS Statistics 19 to design (final concentration of IPTG (0.1 mM, 0.5 mM, 1mM), induction temperature (16 ^{o}C, 28 ^{o}C, 37 ^{o}C), induction time (4 h, 8 h, 12 h)) three factors and three-level orthogonal experiment, with a total of 9 treatments. The three-factor and three-level design and the combination of experimental treatments was shown in Table 1.

Table 1

L9 (3^3) Orthogonal design table and experimental results, relative fluorescence intensity values are shown as mean + SD. Fluorescence analysis of TkAAI-eGFP production by using Duncan's multiple range test

Combination	A Induction temperature (°C)	B Induction time(h)	C Concentration of IPTG (mM)	Relative fluorescence intensity
1	16	4	0.1	0.45 ± 0.081^{c}
2	16	8	1	0.76 ± 0.034^{b}
3	16	12	0.5	0.95 ± 0.043^{a}
4	28	4	0.5	0.31 ± 0.0043^{d}
5	28	8	0.1	0.28 ± 0.042^{d}
6	28	12	1	0.19 ± 0.016^{e}
7	37	4	1	0.069 ± 0.038^{f}
8	37	8	0.5	0.027 ± 0.032^{f}
9	37	12	0.1	0.035 ± 0.0057^{f}
S1	2.16	0.83	0.77	
S2	0.78	1.06	1.29	
S3	0.13	1.18	1.02	
K1	0.72	0.28	0.26	
K2	0.26	0.35	0.43	
K3	0.04	0.39	0.34	
R	0.68	0.12	0.17	

The results of the range analysis demonstrated that the primary and secondary order of factors affecting the relative fluorescence intensity of TkAAI-eGFP bacterial solution was induction temperature, IPTG concentration, and induction time, which indicated that induction temperature played a significant role in expression efficiency of the TkAAI-eGFP prokaryotic expression system. After inducing 12 hours under 0.5 mmol final IPTG concentration and 16 °C inducing expression temperature, the relative fluorescence intensity of TkAAI-eGFP fusion protein was the highest, and significantly higher than other treatment groups (*P<0.05). The relative fluorescence intensities of the induced bacterial solution under 16 °C (Trials1, 2, 3) were also significantly varied from other trails with different temperature. Simultaneously, observation of the bacterial liquid by super resolution laser confocal microscope (Nikon) could be found

the fluorescence intensity of the cells treated with IPTG was obvious higher than the blank control. It indicated that IPTG could significantly up-regulate the expression level of protein in *E. coli* (Supplementary Fig. 3).

The molecular weight of TkAAI-eGFP infusion protein was estimated to 44 kDa according to the encoding cDNA. The results of SDS-PAGE analysis demonstrated that there were obvious target bands in the 44 kDa region in lane 2 and lane 3, and after purification by Ni-NTA column, a certain purity target band could be detected (Fig. 4a).

Western blotting analysis

Western blotting displayed that non-specificity binding in lane 1 (negative control), and in lane 2 to 6, specific binding bands aggregation appeared in the range of molecular weights from 27 kDa to 44 kDa. Interestingly, the predicted molecular weight of eGFP protein was 27 kDa. Accordingly, we infer that the fusion protein may undergo protease cleavage leading to the separation of TkAAI and eGFP peptides, or the mechanical fracture of the fusion protein caused by the external force (Shear force generated by ultrasonication) applied during the extraction of the separated protein, which causes the fragmented eGFP to accumulate in the 27 kDa region (Fig. 4b).

Alpha-amylase inhibition assay

Considering the identical enzyme activity will have certain disparity under different temperature, solution pH and other reaction systems. Accordingly, the main purpose of the experiment was to qualitatively investigate the designated condition of solution PH 6.9, 37 ^{o}C, and the total reaction time was 8 minutes, the alpha-amylase inhibitory activity of TkAAI-eGFP fusion protein (4 μg) on porcine pancreatic amylase (PPA) (1 unit/ml). The experimental result manifested that in the above-mentioned reaction system, the TkAAI-eGFP fusion protein showed 60.1±15.5% alpha-amylase inhibitory activity against PPA (Supplementary Table 2).

Discussion

The full-length cDNA of *TkAAI* was cloned from the seeds of *T. kirilowii* by RACE PCR. The gene expression pattern indicated that the transcript of *TkAAI* in seeds was the most abundant and the expression levels were significantly higher than the other tissues. Blast alignment annotates that 40-136 amino acids in the TkAAI protein sequence were AAI_SS domains, annotation display members of AAI_SS subfamily involve cereal-type alpha-amylase inhibitors and seed storage protein. TkAAI protein contained multiple functional domains, which played a significant role in plant response to different stresses such as insects, microorganisms and environmental factors.

NJ phylogenetic tree analysis revealed a high ortholog of TkAAI with 2S albumin. Moreover, 2S albumin was reported to be a member of the alcohol-soluble superfamily [23]. Most of these small molecules with a molecular weight of about 12-15 kDa come from dicotyledonous plants. 2S albumin had 8 or more

conservative cysteine patterns Cys-(X)n-Cys-(X)n-CysCys-(X)n-Cys-X-Cys-(X)n-Cys-(X)n-Cys-(X)n, we further confirmed this feature in the alignment of amino acid sequences (Fig. 1d). The conservative cysteine pattern of TkAAI is Cys_{43}-(X)n-Cys_{55}-(X)n-$Cys_{86}Cys_{87}$-(X)n-Cys_{97}-X-Cys_{99}-(X)n-Cys_{132}-(X)n-Cys_{139}. However, only six cysteine residues in this domain may have formed three pairs of disulfide bonds (Fig. 2d). The conservation of these cysteine sequences and disulfide bond pairs may be related to the specific binding of lipids in plant seeds, thereby protecting the development and reproduction of plant embryo sacs [24]. Robetta server was employed to guarantee the accuracy of the tertiary structure modeling of the tertiary structure of TkAAI [25],

If the expression rate of heterologous protein was too high and there was no molecular chaperone to assist protein folding in the expression system, the newly formed protein may not fold correctly to form a reasonable conformation, resulting in the formation of insoluble and inactive inclusion bodies, which was not conducive to the follow-up study of the biological activity of the protein [26]. Reports manifested that the low temperature within a condition of a certain range, the folding process of the protein came to correct and the rate of the formation of inclusion body may be reduced [27]. We confirmed the favorable effect of low temperature (16 ^{o}C) on the soluble expression of the TkAAI-eGFP fusion protein, together with appropriate IPTG concentrations ranging around 0.5 mM, which was rapidly detected via determining the relative green fluorescence intensity signal of the bacterial solution under different treatments.

With the rapid development of maturity of molecular cloning technology, plenty of plant-derived alpha-amylase genes have been reported in recent years. [28–30]. Interestingly, the tertiary structure of *Amaranthus hypochondriacus* (AAI) contained three conserved cysteine residues which were the same as TkAAI, and cysteine residues were reported to maintain the stability of these types of proteins [31]. However, the difference was embodied in the number of remaining conservative disulfide bonds between AhAI (1) and TkAAI (2), which revealed the conservation and tendency of variation of AAI during evolution [24]. It was worth mentioning that the methanolic extract of *T. kirilowii* revealed a high anti-amylase activity [32]. Results of enzymes assay and previous reports demonstrated that *T. kirilowii*, as a traditional Chinese medicine (TCM), possessed the ability to produce strong alpha-amylase inhibitory substances. Generally, alpha-amylase inhibitors naturally-derived from the plants had few negative effects on humans, so the prospect of alpha-amylase inhibitors peptides could be developed as a treatment for diabetes mellitus type II by alleviating abnormal absorption of carbohydrates.[33]

To our knowledge, *TkAAI* was the first alpha-amylase inhibitor gene extracted from *T. kirilowii*. Inhibitory activity of TkAAI against PPA was preliminarily measured. Interestingly, the alpha-amylase inhibition mechanism of the cereal-type inhibitor superfamily had been reported [34]. Ragi bifunctional alpha-amylase/trypsin inhibitor (RBI) competed with the substrate for the amylase active site (Asp_{185}, Glu_{222} and Asp_{287}) in alpha-amylase yellow mealworm alpha-amylase (TMA). Due to the high homology of TMA and PPA in binding to substrate and binding to inhibitor sequences [35], we speculated that TkAAI, which also belonged to cereal-type inhibitors and contained a similar domain with RBI, had a similar competitive

inhibitory pattern to PPA. Subsequent experiments were necessary to be designed for more mammalian alpha-amylases to further explore the inhibitory pattern between TkAAI and alpha-amylase. In the meanwhile, it was likewise worth designing the enzymatic kinetic experiments expanded around the multi-site mutants of TkAAI, such as before mentioned five alpha-amylase binding sites including Gln_{41}, Asp_{94}, Glu_{95}, Gln_{96} and Glu_{103} amino acid residues [36]. It was also worth mentioning that plant hormones such as gibberellin (GA) and abscisic acid (ABA) jointly regulated the expression and activity of alpha-amylase genes by antagonizing or promoting effects to control the process of seeds from embryogenesis to germination [37], thereby whether *TkAAI* and plant hormones co-regulate the processes of *T. kirilowii* seeds at above period was also an interesting direction of research.

Declarations

Author Contributions All authors contributed to the study conception and design. ZYP participated in design of the study, performed the cloning and identify of gene, analyzed the data and drafted manuscript; HQY helped to RNA extraction; WKY contributed to enzymatic reaction. SSH helped designed the research, and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements The study was supported by the Fundamental Research Funds for the Central Universities (Grant numbers 2011QC085).

Data availability The datasets generated during the current study are available from the corresponding author on reasonable request.

Conflict of interest The authors have no financial or proprietary interests in any material discussed in this article.

Ethical approval The present article does not involve intervention on a population of humans and/or animals directly and all authors agree with the policies in the ethical Statement.

References

1. Yang J, Zhou C, Yuan G, Li D (2012) Effects of geographical origin on the conjugated linolenic acid of trichosanthes kirilowii maxim seed oil. JAOCS. J Am Oil Chem Soc 89:401–407. https://doi.org/10.1007/s11746-011-1928-3
2. Multiple PW (1994) Pang-Chui Shaw'*, Wah-Lun Chan 2, Hin-Wing Yeung ",3 and Tzi-Bun Ng" Departments of ~Biochemistry, 2Physiology and 3Chinese Medicinal Material Research Centre, The Chinese University of Hong Kong, Shatin, NT., Hong Kong. 55:253–262
3. Shaw PC, Lee KM, Wong KB (2005) Recent advances in trichosanthin, a ribosome-inactivating protein with multiple pharmacological properties. Toxicon 45:683–689. https://doi.org/10.1016/j.toxicon.2004.12.024

4. Huang Y, He P, Bader KP et al (2000) Seeds of Trichosanthes kirilowii, an energy-rich diet. Zeitschrift fur Naturforsch - Sect C J Biosci 55:189–194. https://doi.org/10.1515/znc-2000-3-409
5. Franco OL, Rigden DJ, Melo FR, Grossi-de-Sá MF (2002) Plant α-amylase inhibitors and their interaction with insect α-amylases: Structure, function and potential for crop protection. Eur. J. Biochem
6. de Sales PM, de Souza PM, Simeoni LA et al (2012) α-amylase inhibitors: A review of raw material and isolated compounds from plant source. J Pharm Pharm Sci 15:141–183. https://doi.org/10.18433/j35s3k
7. Capocchi A, Muccilli V, Cunsolo V et al (2013) A heterotetrameric alpha-amylase inhibitor from emmer (Triticum dicoccon Schrank) seeds. Phytochemistry 88:6–14. https://doi.org/10.1016/j.phytochem.2012.12.010
8. Coelho CMM, Bellato C, de Santos M et al (2007) Effect of phytate and storage conditions on the development of the ' hard-to-cook'. J Sci Food Agric 1243:1237–1243. https://doi.org/10.1002/jsfa
9. Williams LK, Zhang X, Caner S et al (2015) The amylase inhibitor montbretin A reveals a new glycosidase inhibition motif. Nat Chem Biol. https://doi.org/10.1038/nchembio.1865
10. Alam F, Shafique Z, Amjad ST, Bin Asad MHH (2019) Enzymes inhibitors from natural sources with antidiabetic activity: A review. Phyther Res 33:41–54. https://doi.org/10.1002/ptr.6211
11. Xiao-Ping Y, SONG C-Q YUANP, MAO R-G (2011) α-Glucosidase and α-Amylase Inhibitory Activity of Common Constituents from Traditional Chinese Medicine Used for Diabetes Mellitus. Chin J Nat Med 8:349–352. https://doi.org/10.3724/sp.j.1009.2010.00349
12. Obiro WC, Zhang T, Jiang B (2008) The nutraceutical role of the Phaseolus vulgaris α-amylase inhibitor. Br. J. Nutr
13. Tormo MA, Gil-Exojo I, De Tejada AR, Campillo JE (2006) White bean amylase inhibitor administered orally reduces glycaemia in type 2 diabetic rats. Br J Nutr 96:539–544. https://doi.org/10.1079/BJN20061836
14. Shade RE, Schroetter HE, Pueyo JJ et al (1994) Transgenic pea seeds expressing the α-amylase inhibitor of the common bean are resistant to bruchid beetles. Bio/Technology 12:793–796. https://doi.org/10.1038/nbt0894-793
15. Lüthi C, Álvarez-Alfageme F, Ehlers JD et al (2013) Resistance of αaI-1 transgenic chickpea (Cicer arietinum) and cowpea (Vigna unguiculata) dry grains to bruchid beetles (Coleoptera: Chrysomelidae). Bull Entomol Res. https://doi.org/10.1017/S0007485312000818
16. Shu S, Xie G, zhen, Guo X, liang, Wang M (2009) Purification and characterization of a novel ribosome-inactivating protein from seeds of Trichosanthes kirilowii Maxim. Protein Expr Purif 67:120–125. https://doi.org/10.1016/j.pep.2009.03.004
17. Scotto-Lavino E, Du G, Frohman MA (2007) 3' End cDNA amplification using classic RACE. Nat Protoc 1:2742–2745. https://doi.org/10.1038/nprot.2006.481
18. Scotto-Lavino E, Du G, Frohman MA (2007) 5' end cDNA amplification using classic RACE. Nat Protoc 1:2555–2562. https://doi.org/10.1038/nprot.2006.480

19. Wong KM, Suchard MA, Huelsenbeck JP (2008) Alignment uncertainty and genomic analysis. Science (80-) 319:473–476. https://doi.org/10.1126/science.1151532
20. Tamura K, Peterson D, Peterson N et al (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739. https://doi.org/10.1093/molbev/msr121
21. Stanley D, Rejzek M, Naested H et al (2011) The role of α-glucosidase in germinating barley grains. Plant Physiol 155:932–943. https://doi.org/10.1104/pp.110.168328
22. Daussant J, Miyata S, Mitsui T, Akazawa T (1983) Enzymic Mechanism of Starch Breakdown in Germinating Rice Seeds. Plant Physiol 71:88–95. https://doi.org/10.1104/pp.71.1.88
23. Millst ENC, Trust W, Campus G (2002) Classification o f plant proteins. 30:906–910
24. Zhang D (2009) Homology between DUF784, DUF1278 domains and the plant prolamin superfamily typifies evolutionary changes of disulfide bonding patterns. Cell Cycle 8:3428–3430. https://doi.org/10.4161/cc.8.20.9674
25. Kim DE, Chivian D, Baker D (2004) Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res 32:526–531. https://doi.org/10.1093/nar/gkh468
26. Baneyx F, Mujacic M (2004) Recombinant protein folding and misfolding in Escherichia coli. Nat Biotechnol 22:1399–1407. https://doi.org/10.1038/nbt1029
27. Donovan RS, Robinson CW, Glick BR (2000) Optimizing the expression of a monoclonal antibody fragment under the transcriptional control of the escherichia coli lac promoter. Can J Microbiol 46:532–541. https://doi.org/10.1139/cjm-46-6-532
28. Yu J, Li Y, yang, Xiang M et al (2017) Molecular cloning and characterization of α-amylase/subtilisin inhibitor from rhizome of Ligusticum chuanxiong. Biotechnol Lett 39:141–148. https://doi.org/10.1007/s10529-016-2227-8
29. Bhide AJ, Channale SM, Yadav Y et al (2017) Genomic and functional characterization of coleopteran insect-specific α-amylase inhibitor gene from Amaranthus species. Plant Mol Biol 94:319–332. https://doi.org/10.1007/s11103-017-0609-5
30. Bunyatang O, Chirapongsatonkul N, Bangrak P et al (2016) Molecular cloning and characterization of a novel bi-functional α-amylase/subtilisin inhibitor from Hevea brasiliensis. Plant Physiol Biochem 101:76–87. https://doi.org/10.1016/j.plaphy.2016.01.014
31. Lehmann K, Schweimer K, Reese G et al (2006) Structure and stability of 2S albumin-type peanut allergens: Implications for the severity of peanut allergic reactions. Biochem J 395:463–472. https://doi.org/10.1042/BJ20051728
32. Uchholz TB, Chen C, Zhang XY, Melzig MF (2016) Pancreatic lipase and α-amylase inhibitory activities of plants used in Traditional Chinese Medicine (TCM). Pharmazie 71:420–424. https://doi.org/10.1691/ph.2016.6569
33. Quintero-Soto MF, Chávez-Ontiveros J, Garzón-Tiznado JA et al (2021) Characterization of peptides with antioxidant activity and antidiabetic potential obtained from chickpea (Cicer arietinum L.) protein hydrolyzates. J Food Sci 86:2962–2977. https://doi.org/10.1111/1750-3841.15778

34. Strobl S, Maskos K, Wiegand G et al (1998) A novel strategy for inhibition of α-amylases: Yellow meal worm α-amylase in complex with the Ragi bifunctional inhibitor at 2.5 Å resolution. Structure 6:911–921. https://doi.org/10.1016/S0969-2126(98)00092-6
35. Strobl S, Gomis-Rüth FX, Maskos K et al (1997) The α-amylase from the yellow meal worm: complete primary structure, crystallization and preliminary X-ray analysis. FEBS Lett 409:109–114. https://doi.org/10.1016/S0014-5793(97)00451-1
36. Rydberg EH, Li C, Maurus R et al (2002) Mechanistic analyses of catalysis in human pancreatic α-amylase: Detailed kinetic and structural studies of mutants of three conserved carboxylic acids. Biochemistry 41:4492–4502. https://doi.org/10.1021/bi011821z
37. Gómez-Cadenas A, Zentella R, Walker-Simmons MK, Ho THD (2001) Gibberellin/abscisic acid antagonism in barley aleurone cells: Site of action of the protein kinase PKABA1 in relation to gibberellin signaling molecules. Plant Cell. https://doi.org/10.1105/tpc.13.3.667

Figures

Figure 1

Characterization of TkAAI. **a** Conserved domains of TkAAI. **b** Calculated secondary structure of TkAAI protein, blue area indicates alpha helix, red area indicates extended strand and purple area indicates random coil. **c** Signal peptide and subcellular localization analysis of TkAAI. **d** Alignment and of the amino acids sequences between TkAAI and 2S albumin from other four cucurbitaceous species. The conserved cysteine sequence of 2S albumin was marked with red triangle (Wilding et al., 2009). **e** Phylogenetic relationship of 2S albumin and 2S albumin-like. The Phylogenetic tree analysis was performed by the MEGA 7.0 program with the neighbor joining method and with 1000 replicates. The phylogenetic tree was constructed based on the alignment of amino acids sequences of 30 orthologs proteins, different color indicated to different species, and TkAAI was marked in red

Figure 2

Tertiary structure of TkAAI. **a** TkAAI tertiary structure constructed by Robetta server ab initio modeling. **b** TkAAI tertiary structure constructed through homologous modeling approaches employing SWISS-MODLE. **c** Structural aligned of predicted TkAAI tertiary structure models acquired from Robetta (Green) and SWISS-MODLE (Blue) (RMSD 4.374 Å). **d** Conserved cysteine pattern in tertiary structure of TkAAI. The protein model was constructed by Robetta. Eight sites of cysteine have been labeled and there are three disulfide bonds in the yellow region

Figure 3

Relative expression level of *TkAAI* in different tissues of *T. kirilowii*. The expression profile of *TkAAI* was analyzed by real-time PCR, a constitutively expressed housekeeping gene, GADPH gene was adopted as the internal control for real-time PCR. The error bars represent the mean ± SE values (n = 12). Significant difference was shown as **P < 0.01

Figure 4

Detection of heterologous expression of TkAAI. **a** 10% SDS-PAGE gel electrophoresis of TkAAI-eGFP. Lane M indicated marker. Lane 1 indicated total protein from *E. coli* carrying pET28a. Lane 2 indicated total protein from *E. coli* carrying TkAAI-eGFP-pET28a with IPTG induction. Lane 3 indicated supernatant of ultrasonic treated bacterial solution from *E. coli* carrying TkAAI-eGFP-pET28a with IPTG induction. Lane 4 indicated flow-through fluid. Lane 5 indicated 10mM Imidazole washing solution. Lane 6, 7, and 8 indicated Ni-NTA column purified TkAAI-eGFP. **b** Western blotting detection of TkAAI-eGFP by eGFP antibody. Lane M indicated marker. Lane 1 indicated total protein from *E. coli* carrying pET28a. Lane 2 indicated supernatant of ultrasonic treated bacterial solution from *E. coli* carrying TkAAI-eGFP-pET28a without IPTG induction. Lane 3 indicated total protein from *E. coli* carrying TkAAI-eGFP-pET28a with IPTG induction. Lane 4 indicated supernatant of ultrasonic treated bacterial solution from *E. coli* carrying TkAAI-eGFP-pET28a with IPTG induction. Lane 5 and 6 indicated Ni-NTA column purified TkAAI-eGFP

Supplementary Files

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

- Supplementarymaterial.docx