Purification, developmental expression, and in silico characterization of α -amylase inhibitor from *Echinochloa frumentacea*

Priyankar Panwar¹ · A. K. Verma¹ · Ashutosh Dubey¹

Received: 22 September 2017 / Accepted: 23 April 2018 / Published online: 27 April 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Barnyard (*Echinochloa frumentacea*) and finger (*Eleusine coracana*) millet growing at northwestern Himalaya were explored for the α -amylase inhibitor (α -AI). The mature seeds of barnyard millet variety PRJ1 had maximum α -AI activity which increases in different developmental stage. α -AI was purified up to 22.25-fold from barnyard millet variety PRJ1. Semi-quantitative PCR of different developmental stages of barnyard millet seeds showed increased levels of the transcript from 7 to 28 days. Sequence analysis revealed that it contained 315 bp nucleotide which encodes 104 amino acid sequence with molecular weight 10.72 kDa. The predicted 3D structure of α -AI was 86.73% similar to a bifunctional inhibitor of ragi. In silico analysis of 71 α -AI protein sequences were carried out for biochemical features, homology search, multiple sequence alignment, phylogenetic tree construction, motif, and superfamily distribution of protein sequences. Analysis of multiple sequence alignment revealed the existence of conserved regions NPLP[S/G]CRWYVV[S/Q][Q/R]TCG[V/I] throughout sequences. Superfam analysis revealed that α -AI protein sequences were distributed among seven different superfamilies.

Keywords α -Amylase inhibitor · Barnyard millet · Bifunctional inhibitor · In silico analysis · Expression analysis

Introduction

α -AI is extensively found in many plant seeds and tubers, being particularly abundant in cereals and legumes (Franco et al. 2002; Mehrabadi et al. 2010). These molecules play a key role in plant defense toward pests and pathogens (Franco et al. 2000), which cause severe damages to field crops and stored grains (Koiwa et al. 1997; Pereira et al. 1999; Franco et al. 2002). As these inhibitors could show different specificities against α -amylases from different sources and inhibitors with a wide specificity spectrum are strongly favoured for insect control (Franco et al. 2000). Several kinds of α -amylase and proteinase inhibitors in seeds and vegetative organs act to regulate the numbers of phytophagous insects (Konarev 1996; Chrispeels et al. 1998; Gatehouse and Gatehouse 1998). α -AI are attractive candidates for the control of seed weevils as these insects are highly dependent on starch as an energy source (Franco et al. 2000). In

cereal seeds, α -AI are proteins with 120–130 amino acids based on their primary and tertiary structures, disulphide bonds positions and reactive site localization α -AI may be classified into six families (lectin-like, knotting-like, Kunitz-like, γ -purothionin-like, thaumatin-like and cereal-type) (Richardson 1981; Strobl et al. 1995). Many insects have several α -amylases that differ in specificity, and successful utilization of a food source is dependent on the expression of α -amylase for which there is no specific inhibitor (Silva et al. 2000). The α -AI have long been proposed as possible important weapons against pests whose diets make them highly dependent on α -amylase activity. In vitro and in vivo trials using α -AI, including those made under field conditions, have now fully confirmed their potential for increasing yields by controlling insect populations (Franco et al. 2000). α -AIs have been isolated and characterized in white bean, kidney bean, sorghum, green gram, black gram, wheat, little millet and finger millet (Mulimani and Supriya 1993; Veronique et al. 1997; Kokiladevi et al. 2005; Heidari et al. 2005; Sivakumar et al. 2006; Barrett and Udani 2011) However, there is no report on presence of α -AI in barnyard millet.

Barnyard millet (*Echinochloa frumentacea*) is important millet grown in the arid and semiarid region of the world. Barnyard millet is a good source of protein, fat,

✉ A. K. Verma
akv72@rediffmail.com

¹ Department of Biochemistry, College of Basic Sciences and Humanities, G. B. Pant University of Agriculture and Technology, Pantnagar, Uttarakhand, India

carbohydrates, and crude fiber (Kumar and Parameshwaran 1998; Veena et al. 2005; Devi et al. 2014) apart from mineral (Panwar et al. 2016) and vitamins (Veena et al. 2005). It also contains phytochemicals, such as phenolic acids, flavonoids and tannins (Kulkarni et al. 1992; Panwar et al. 2016) which serve as good source of natural antioxidants.

The present study, α -AI, was investigated in the barnyard (*E. frumentacea*) and finger (*Eleusine coracana*) millet. The α -AI genes and protein from barnyard millet were purified. We isolated the α -AI and performed a comparative in silico analysis of the reference nucleotide and protein sequences of α -AI from different organisms available in NCBI and ExPASy databases. The sequences were utilized for study of variation in their biochemical features, multiple sequence alignment, and sequence homology, phylogenetic tree construction, distribution of motifs, and superfamily using various bioinformatics tools and software.

Materials and methods

Plant materials

Seeds of different varieties of barnyard millet (VL21, VL29, VL172, VL207, and PRJ1) were procured from Vivekananda Parvatiya Krishi Anushandhan Sansthan, Almora, (ICAR) and Uttarakhand University of Horticulture and Forestry, Bharsar, Uttarakhand, India.

α -AI activity

Seeds (1 g) were finely ground in pestle mortar and mixed with 10 ml 1% NaCl (w/v) by shaking for 1 h and kept overnight at 4 °C. The mixture was centrifuged and final volume of supernatant was made with 25 ml with 1% NaCl solution. α -AI activity was measured in the supernatant. The enzyme was assayed using the procedure of Bernfeld (1955). One unit of amylase was defined as the amount of enzyme which yields 1 mol of glucose at 37 °C for 1 min at pH 6.7. One unit of inhibitory activity is that which reduce the activity of amylase by one unit under the assay conditions. Seeds of different developmental stages of Barnyard millet variety PRJ1 (viz. 7, 14, 21, and 28 days and matured) was harvested and α -AI was isolated and assayed.

Purification of α -AI

Purification of α -AI was done as described by McEwan et al. (2010) with slight modification. Barnyard millet variety PRJ1 seeds (50 g) were ground with 200 ml of 0.02M phosphate buffer (6.9 pH) containing 1% (w/v) polyvinyl-pyrrolidone. The mixture was stirred at room temperature (25 °C) for 4 h and then filtered through muslin cloth.

The residue was re-extracted and combined filtrate was centrifuged at 12,000g for 20 min and collected the supernatant. The obtained supernatant was treated with ammonium sulfate salt to 80% saturation for overnight at 4 °C. The precipitated protein was collected by centrifugation at 10,000 rpm for 20 min. The supernatant was decanted and the pellet was dissolved in a minimum amount of phosphate buffer (20 mM, pH 6.9). The ammonium sulfate fractions were dialyzed against 10 mM phosphate buffer (pH 6.9) at 4 °C for 48 h with repeated changes of buffer. The protein was concentrated in PEG-6000 at 4 °C until. α -AI activity and protein content was measured in concentrated fraction using starch as substrate (Bernfeld 1955) and Bradford method (Bradford 1976), respectively. Concentrated sample obtained after dialysis was loaded into a DEAE-cellulose column (28 × 1.25 cm), equilibrated with 0.02M phosphate buffer (pH 6.9). Stepwise elution was carried out with a linear gradient of 0.1–0.5 M NaCl in 0.02 M phosphate buffer. Fractions (2 ml) were collected at 0.42 ml/min elution rate. Total protein was monitored by the absorbance at $\lambda_{280\text{nm}}$ on a spectrophotometer. All fractions were analyzed for α -AI activity, active fractions were pooled, and protein content was measured in each fraction by the Bradford method (1976). Active fractions were concentrated and were used for carrying out size-exclusion chromatography using Sephadex G 75 columns. The standard molecular weight marker protein catalases (240 kDa), Bovine serum albumin (66 kDa), ovalbumin (43 kDa), and lysozyme (13.5 kDa) were used for the determination of molecular weight. The SDS-PAGE of purified protein samples was analyzed using the method described by Laemmli (1970).

Semi-quantitative PCR-based expression profiling of α -AI

Different stages of seeds (100 mg) of barnyard millet variety PRJ1 were ground in with liquid nitrogen. The ground seeds were immediately mixed with 1 ml of RNA-XPress Reagent and transferred to a 2.0 ml collection tube and RNA isolation was performed as per the protocol given in the RNA—press kit procured from Himedia Pvt. Ltd. Complementary DNA (cDNA) was synthesized using reverse transcriptase (RT) by ProtoScript® First Strand cDNA Synthesis kit (New England Biolabs).

Eleusine coracana alpha-amylase–trypsin bifunctional inhibitor gene, partial cds (DQ494211.1) sequence was retrieved from NCBI (<http://www.ncbi.nlm.nih.gov/>) for primer designing (forward ACAATCCGCTGGACAGCT GC and reverse CCGAGTAAGGAGAGGCAGAA). For semi-quantitative analysis, tubulin gene was selected as endogenous internal standard. The tubulin primer (forward CTCCAAGCTTTCTCCCTCCT and reverse GCATCATCA CCTCCTCCAAT) used as internal control was designed

from Genbank database (Acc. no. CX265249). The primers were custom synthesized by SBS Gentech Co. Ltd., Biochem Lifesciences, New Delhi.

Semi-quantitative PCR analysis of α -AI in different developmental stages of seeds of barnyard millet was carried by made the reaction mixture contain 200 ng of genomic DNA, 200 μ M of each dNTPs, 1 μ M each primer, 1U *Taq* polymerase, 1.5 mM Mg^{2+} and 1X PCR buffer. Amplification was carried in temperatures: 95 °C for 5 min, followed by 30 cycles of 94 °C for 45 s, 60 °C for 45 s, 72 °C for 2 min and final extension for 10 min at 72 °C. The PCR amplification was resolved on 1% (w/v) agarose gel.

In silico analysis of α -AI sequence

Amplified PCR products were sequenced by automated DNA sequencer at the DNA Sequencing Facility, SciGenome Labs private LTD, Cochin, Kerala, India. The sequence analysis was done using NCBI database by employing BLASTN algorithm (<http://www.ncbi.nlm.nih.gov/GenBank/html>) (Altschul et al. 1990). Nucleotide sequence was translated into protein sequence using ExpASY translate tool (<http://web.expasy.org/translate/>) and used for in silico characterization. The tertiary structure of α -AI was predicted using the homology modelling approach at SWISS-MODEL workspace (<http://swissmodel.expasy.org/>) (Guex and Peitsch 1997; Arnold et al. 2006).

In silico analysis of different classes of α -AI

Seven classes of α -AI were analyzed separately using reference protein sequence entries for these α -AI classes from online databases. Representative protein sequences were used as probes to the BLAST database from NCBI (<http://www.ncbi.nlm.nih.gov/>). The protein sequences in FASTA format from refseq entries, which were shown to exhibit α -AI activities, were selected for further in silico study. The evolutionary history was inferred using the neighbor-joining method (Saitou and Nei 1987). The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method and were in the units of the number of amino acid substitutions per site (Zuckerandl and Pauling 1965). All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA5 (Tamura et al. 2011). For domain search, the Pfam site (<http://www.sanger.ac.uk/software/pfam/search.html>) was used. Superfam server (<http://supfam.cs.bris.ac.uk/SUPERFAMILY/hmm.html>) was used for the analysis of superfamily present of the protein sequences. Domain analysis was used MEME (<http://meme.sdsc.edu/meme/meme.html>) (Bailey and Elkan 1994). The conserved protein motifs

deduced by MEME were characterized for biological function analysis using BlastP (<http://www.ncbi.nlm.nih.gov/>) and domains were studied with Interproscan providing the best possible match based on highest similarity score.

Physiochemical data were produced from various tools in the EXPASY Proteomic server (Clustal W, ProtParam, Protein calculator, Compute pI/Mw, ProtScale) (Kyte and Doolittle 1982). The molecular weights (kDa) of the various α -AIs were calculated by the addition of average isotopic masses of amino acid in the protein and deducting the average isotopic mass of one water molecule. The pI of the enzyme was calculated using pKa values of amino acid according to Bjellqvist et al. (1993).

Results and discussion

α -AI quantity

α -AI in seeds of barnyard millet varieties varied from 36.27 ± 0.33 to 24.82 ± 0.63 U/g, while in seeds of finger millet varieties, it ranged from 24.81 ± 0.30 to 20.04 ± 0.19 U/g. It indicates that barnyard millet contains higher α -AI activity compared to finger millet. In barnyard millet varieties, α -AI was distinctly higher in variety PRJ1 (36.27 ± 0.33 U/g) followed by other VL172 (27.68 ± 0.23 U/g), VL29 (26.73 ± 0.58 U/g), VL207 (26.72 ± 0.69 U/g) and VL21 (24.82 ± 0.63 U/g), respectively. α -AI in seeds of finger millet varieties was markedly higher in variety VL315 (24.81 ± 0.30 U/g) followed by four other varieties VL146 (22.90 ± 0.45 U/g), VL324 (21.95 ± 0.54 U/g), VL149 (21.00 ± 0.38 U/g) and VL204 (20.04 ± 0.19 U/g), respectively. Results revealed that α -AI content in seeds of barnyard millet varieties showed inhibitory activity from 60.32 ± 0.46 to $41.26 \pm 0.84\%$, while α -AI content in seeds of finger millet varieties exhibited inhibitory activity from 41.27 ± 0.27 to $33.34 \pm 0.18\%$ (Fig. 1). Duncan post hoc analysis of α -AI in seeds of barnyard millet and finger millet showed nine groups. No significant difference was observed between varieties falling within a single group depicted with the same superscript at $P \leq 0.05$, i.e., group a (PRJ1), group b (VL172), group c (VL29 and VL207), group d (VL21 and VL315), group f (VL146), group g (VL324), group h (VL149), and group i (VL204). Kokiladevi et al. (2005) studied the effect of α -AI activity crude protein extracts from the seeds of *Vigna sublobata* on larval α -amylase of insect (*Callosobruchus analis*) and exhibited that inhibitory activity 110.01% for *Vigna umbellata*, 80.50% for *V. sublobata*, and 70.03 for *V. Glabracens*. Barnyard millet variety PRJ1 contained the highest α -AI; hence, it was chosen for the purification and characterization of α -AI.

Fig. 1 α -AI in seeds of different varieties of barnyard millet (VL21, VL29, VL172, VL207, and PRJ1) and finger millet (VL146, VL149, VL204, VL315, and VL324). Values are expressed as mean $\pm$ standard deviation ($n=3$). Means with different letters (a, b, c, d, e and f) were significantly different at the level of $P<0.05$

Table 1 Screening of α -AI at different developmental stages of maturation of barnyard millet seed

Seed stage	Inhibition (U/g)	% inhibition	Specific inhibitory activity (U/mg)
7 days	18.61 $\pm$ 0.13	30.37 $\pm$ 0.04	4.65
14 days	25.67 $\pm$ 0.39	41.38 $\pm$ 0.09	6.42
21 days	31.40 $\pm$ 0.05	48.16 $\pm$ 0.14	7.85
28 days	32.55 $\pm$ 0.12	52.78 $\pm$ 0.28	8.14
Matured seed	35.98 $\pm$ 0.07	59.09 $\pm$ 0.09	8.99

Screening of α -AI in different developmental stages of seed of barnyard millet (PRJ1), i.e., 7, 14, 21, and 28 days and matured seed. Values are mean values $\pm$ standard deviation of three determinations ($n=3$)

Screening of α -AI at different developmental stages of seed

Barnyard millet variety PRJ1, which showed the highest α -AI activity, was subjected to the screening of α -AI activity

at different developmental stages of their seed. α -AI activity of seeds increased from 7 days after seed set to maturation stage (Table 1). α -AI 35.98 $\pm$ 0.07 U/g was found highest in matured seed with specific inhibitory activity 8.99 U/mg followed by other, 28 days (32.55 $\pm$ 0.12 U/g), 21 days (31.40 $\pm$ 0.05 U/g), 14 days (25.67 $\pm$ 0.39 U/g), and 7 days (18.61 $\pm$ 0.13 U/g), respectively. In matured seeds, α -AI activity was found 59.09 $\pm$ 0.09%.

Richardson (1991) reported inhibitors as a storage protein because of presence in a large amount in dry/matured seed. α -AI represents about 6% of the protein in soybean (Rackis and Anderson 1964), up to 10% of the total protein of barley seeds (Mikola and Kirsii 1972), over 10% of the soluble proteins in potato tubers (Weder 1981) and about 20% of the soluble proteins of seeds of the Brazilian Carolina tree *Ade-nanthera pavonina* (Richardson et al. 1986). Hence, matured seeds of barnyard millet variety PRJ1 contained high α -AI protein and showed highest α -AI activity. Kokiladevi et al. (2005) reported, during the development of *Vigna sublobata* seed the α -AI activity was detected after 10 days of seed development and gradually increased until maturation of seeds, i.e., 15 days (45.23 IU/g), 20 days (93.27 IU/g), 25 days (126.25 IU/g), 30 days (151.39 IU/g), and dry seed (155.20 IU/g), respectively; similar pattern was observed in seeds of barnyard millet in our investigation.

Purification of α -AI

α -AI from matured seeds of barnyard millet variety PRJ1 was extracted with the specific activity 27.87 U/mg in 0.02 M phosphate buffer (pH 6.9) containing 1% polyvinylpyrrolidone. Specific activity of α -AI after 80% ammonium sulfate precipitation was 50.77 U/mg with 1.82 purification fold. After ion-exchange chromatography, specific activity of α -AI was found 271.03 U/mg with 9.73-fold purification (Table 2). After gel-filtration chromatography specific activity of α -AI was found 620.14 U/mg with 22.25-fold purification which was approximately 8% in yield (Fig. 2). Shivaraj and Pattabiraman (1981) also characterized two α -AI from finger millet (*Elusine coracana* Geartm) after subjecting to 55% ammonium sulfate precipitation, dialysis against 2 mM sodium acetate buffer, pH 5.0, CM-cellulose

Table 2 Purification table of α -AI protein from Barnyard millet seed (variety PRJ1)

Purification steps	Volume (ml)	Total inhibitory activity (U/ml)	Protein (mg/ml)	Specific activity (U/mg)	Fold purification	Yield (%)
Crude extract	121	25.69	0.92	27.87	1	100
Ammonium sulfate precipitation	6.5	363.64	7.16	50.77	1.82	76
Ion-exchange	2	397.60	1.47	271.03	9.73	26
Size exclusion	5	267.90	0.43	620.14	22.25	8

Fig. 2 Purification profile of α -AI with Sephadex G-75 (gel-filtration chromatography). Marker protein mixture contained catalase (240 kDa), Bovine serum albumin (66 kDa), ovalbumin (43 kDa), and lysozyme (14.3 kDa)

Fig. 3 SDS-PAGE of purified α -AI from barnyard millet seed (PRJ1) lane 1, crude protein; lane 2, ammonium sulfate precipitate at 80%; lane 3, ion-exchange chromatography; lane 4, gel-filtration chromatography; lane 5 molecular weight marker

and Sephadex G-50 and found 14.3 and 16.5 kDa protein. SDS-PAGE electrophoresis analysis of the purified α -AI showed the presence of a single band in Coomassie Blue R-250 dye (Fig. 3). Its apparent molecular mass was approximately 14.3 kDa. The molecular mass was also confirmed by gel-filtration chromatography (Fig. 2). In the gel filtration, applied sample resolved the protein in a single peak (fraction number 37–50) which contained α -AI activity. The calculated elution volume (Ve) was 78.25 ml. The molecular weight of α -AI was determined from the standard calibration curve, which was approximately 14 kDa (Fig. 4). α -AIs have been purified from various sources and found the molecular

Fig. 4 Molecular weight determination of α -AI (AAI) from the standard curve drawn between $\log_{10}$ molecular weight v/s elution volume (Ve)

weight between the range of 14.0 and 77.0 kDa (Table 3). The molecular weight of the α -AI of barnyard millet was found relatively low compared to *Phaseolus vulgaris* (Veronique et al. 1997), *Palo fierro* (Guzman-Partida et al. 2007), and *Phaseolus acutifolius* (Yamada et al. 2001), but was found identical with *Elusine coracana* (Shivaraj and Pattabiraman 1981), *Panicum miliaceum* (Nagaraj and Pattabiraman 1985), and *Vigna sublobata* (Kokiladevi et al. 2005).

Semi-quantitative PCR for expression analysis of α -AI

Semi-quantitative PCR of different developmental stages of barnyard millet seeds showed increased levels of the transcript from 7 to 28 days, while in matured seeds of barnyard

Table 3 α -AI purified from different organisms

Source organism	Purification techniques	Size of α -amylase inhibitor	Subunits	References
<i>Elusine coracana</i>	CM-Cellulose and Sephadex-G75	14.3 and 16.5 kDa	Monomer	Shivaraj and Pattabiraman (1981)
<i>Panicum miliaceum</i>	CM-Cellulose, DEAE-Cellulose and Sephadex-G75	14 kDa	Monomer	Nagaraj and Pattabiraman (1985)
<i>Colocasia esculenta</i>	DEAE-Sephacel and Sephadex G-100	17 and 19 kDa	Monomer	McEwan et al. (2010)
<i>Phaseolus vulgaris</i>	Chromatofocusing and Gel filtration	43 kDa	Dimer	Veronique et al. (1997)
<i>Palo fierro</i>	Affinity Chromatography	77 kDa	Dimer	Guzman-Partida et al. (2007)
<i>Vigna sublobata</i>	Sephadex-G75 and RP-HPLC	14 kDa	Monomer	Kokiladevi et al. (2005)
<i>Triticum aestivum</i>	FPLC	–	Dimer	Heidari et al. (2005)
<i>Phaseolus acutifolius</i>	DEAE-Sephacel and ConA-Sepharose	35 kDa	Dimer	Yamada et al. (2001)

Fig. 5 α -AI gene expression profiling in different developmental stages (7, 14, 21, and 28 days and matured seed) of seeds of barnyard millet (PRJ1). Sterile distilled water was used as a control. Tubulin was used as an internal control

millet, α -AI gene expression was not found (Fig. 5). Variation in the band intensity would, therefore, represent differences in cDNA concentration. Intensity of bands was revealed by relative densitometry values in gel on densitometry analysis (Alpha EaseFC software, Alpha Innotech Corporation, USA) found increasing order of intensity, i.e., in 7 days (3.3%), 14 days (7.8%), 21 days (18.3%), 28 days (31.0%), matured seed (0.0%) and in tubulin gene (39.5%).

Sequencing and characterization of α -AI gene sequence

The amplified PCR product was sequenced to found 315 bp. NCBI nucleotide BLAST (<http://www.ncbi.nlm.nih.gov/>) revealed 99% homology with *E. coracana* alpha-amylase–trypsin bifunctional inhibitor gene, partial cds (gbIDQ494211.1). The translated product of DNA sequence contained 104 amino acids, which was retrieved from ExPASy translate tool (<http://web.expasy.org/translate/>) (Fig. 6).

The 3D structure model of α -AI was constructed in SWISS-MODEL workspace using already existing α -AI structure (*Tenebrio molitor* Larval Alpha-Amylase complex with the bi-functional inhibitor of rafi) as a template. 1–98 amino acids were used for the construction of the 3D structure and 86.73% similarity was found with the template sequence (1tmqB). Predicted 3D structure of α -AI

```

acaatccgctggacagctgccgctggtacgtggccaaacgtgcctgcggcgtggggcccccac
N P L D S C R W Y V A K R A C G V G P H
ctctgacgcaggagatgaaggcgctagtagcgggcagctggaggccatcccgattac
L L T Q E M K A R S S G Q L E A I P D Y
tgccgctgcgaggcggtgcgcatactcatggacggcggtggtcacgtcgagcgggcagcac
C R C E A V R I L M D G V V T S S G Q H
gagggggcgctcctctgaggacgtccgggatgccaggcagggtcaaaggaggttcgca
E G R L L E D L P G C P R Q V Q R E F A
cccaaatgtaactgaggcggtgcaacctgtccaccatccatggaggcccttctgccc
P K L V T E A E C N L S T I H G G P S A
tctccttactcgg
S P Y S

```

Fig. 6 Translated protein sequence (104 amino acids) form α -AI nucleotide sequence (315) bp. Conserved sequences are highlighted. Two conserved motifs ‘N-P-L-P-[S/G]-C-R-W-Y-V-V-[S/Q]-[Q/R]-T-C-G-[V/I]’ (amino acids, 1–17) and ‘L-E-D-L-P-G-C-P-R-Q-V-Q-R-E-F-A-[A/P/R]-K-L-V-T-[E/P]-[G/A]-E-C-N-L-S-T-[I/V]-H-[G/N]’ (amino acids 65–96), were found in α -AI of barnyard millet

of barnyard millet is shown in (Fig. 7). Very less *E* value (1.64068e–42) in sequence identity indicates that the 1tmqB and α -AI of barnyard millet has a very similar structure.

Bioinformatics analysis for characterization of α -AI

Amino acid composition of the translated product of α -AI gene from barnyard millet was analyzed by ProtParam to find the molecular weight (10.72 kDa), theoretical pI (6.77), and negatively charged as it contained 12 negatively charged amino acids and 11 positively charged amino acids (Table 4). Aliphatic index of α -AI was 83.93. Aliphatic index of protein measures the relative volume occupied by aliphatic side chains of the amino acids: alanine, valine, leucine, and isoleucine. Globular proteins with high aliphatic index have high thermostability and an increase in aliphatic

Fig. 7 Three-dimensional structural model of α -AI from barnyard millet. The 3D structure model of α -AI was constructed in SWISS-MODEL workspace using α -AI structure of ragi as a template. 1–98 amino acids were used for the construction of the 3D structure and 86.73% similarity was found with the template sequence

index protein thermostability increases (Ikai 1980; Rawlings et al. 2006). The amino acid composition of α -AI protein sequence was in (Table 5). Ala, Arg, Gly, Glu, Gly, Leu, Pro, and Ser were major amino acids constituting about 57.68% α -AI. Six cysteine residues present in the sequence revealed that α -AI bear disulfide bonds, which were believed to be essential for conformational stability and catalytic activity (Hung et al. 2003) (Fig. 6).

Alignment of homologous sequences was performed with Mega 5.0 (Tamura et al. 2011) showed two conserved motifs, namely, 'N-P-L-P-[S/G]-C-R-W-Y-V-V-[S/Q]-[Q/R]-T-C-G-[V/I]' (amino acids, 1–17) and 'L-E-D-L-P-G-C-P-R-Q-V-Q-R-E-F-A-[A/P/R]-K-L-V-T-[E/P]-[G/A]-E-C-N-L-S-T-[I/V]-H-[G/N]' (amino acids, 65–96), were found in α -AI of barnyard millet and these sequences are homologous to α -AI of *E. coracana*, *Sorghum bicolor* and *Zea mays*. (Fig. 8). Superfam analysis of protein sequence of α -AI of barnyard millet revealed their similarity to proteinase/ α -AI family (Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin super family) with *Amaranthus hypochondriacus*, *Hordeum vulgare*, *Triticum aestivum*, *S. bicolor*, and *Secale cereal*.

In silico analysis of protein sequences of α -AI

Seventy-one protein sequences of α -AI were retrieved from NCBI protein database. The accession numbers of

Table 5 Amino acid composition of α -AI from Barnyard millet

S. no.	Amino acid	Number	Com-position (%)
1	Ala (A)	8	7.69
2	Arg (R)	8	7.69
3	Asn (N)	2	1.92
4	Asp (D)	4	3.84
5	Cys (C)	6	5.77
6	Gln (Q)	5	4.8
7	Glu (E)	8	7.69
8	Gly (G)	9	8.65
9	His (H)	3	2.88
10	Ile (I)	3	2.88
11	Leu (L)	10	9.62
12	Lys (K)	3	2.88
13	Met (M)	2	1.92
14	Phe (F)	1	0.96
15	Pro (P)	8	7.69
16	Ser (S)	9	8.65
17	Thr (T)	4	3.84
18	Trp (W)	1	0.96
19	Tyr (Y)	3	2.88
20	Val (V)	7	6.73

retrieved sequences along with species names are listed in Table 6. Analysis of multiple sequence alignment revealed the presence of conserved regions throughout sequences. In sequences, a conserved site 'MASKSS[I/C][T/D]LLLA AVL[A/V]S[V/I]' was observed toward their N-terminus, followed by one more sequence 'LV[T/A][S/P]G[H/Q]CN[V/L][M/A]T[V/I]HN[A/T/V][P/R]Y'. Evolutionary relationships between different sequences were studied using phylogenetic tree constructed by Neighbor-joining method (Fig. 9). Overall, seven major clusters were observed in the phylogenetic tree. Cluster '1' represents 12 sequences of *A. hypochondriacus*, *E. frumentacea*, *H. vulgare*, *T. aestivum*, *S. bicolor*, and *S. cereale*. The biochemical features of this cluster are listed in Table 6. The amino acid residues in sequences of this cluster ranged from 25 to 171 with a molecular weight range 2.71–18.52 kDa. Isoelectric point

Table 4 Characterization of α -AI from barnyard millet

Source organ-ism	No. of Amino acids	Molecular weight	Theoretical pI	No. of negatively charged amino acids	No. of positively charged amino acids	Instability index	Aliphatic index	GRAVY
<i>Echinoch-loa frumen-tacea</i>	104	10,720.5	6.77	12	11	43.55	83.93	0.007

Fig. 8 Phylogenetic tree constructed by NJ method based on α -AI protein sequences similarity

(pI) for this group was placed between 3.92 and 7.50. The instability index is used to measure in vivo half-life of a protein (Guruprasad et al. 1990). The proteins which have been reported as in vivo half-life of less than 5 h showed an

instability index greater than 40, whereas those having more than 16 h half-life (Rogers et al. 1986) have an instability index of less than 40. Instability index of this cluster placed between 25.76 and 57.42. Aliphatic index of protein measures the relative volume occupied by aliphatic side chains of the amino acids: alanine, valine, leucine, and isoleucine. Globular proteins with high aliphatic index have high thermostability and an increase in aliphatic index increases protein thermostability (Ikai 1980; Rawlings et al. 2006). Aliphatic index of cluster one ranged from 21.25 to 93.47. The variation among α -AI in this group in terms of physiological parameters like a number of negatively charged and positively charged amino acids and hydropathicity (GRAVY). Cluster '2' represents seven α -AI sequences of *H. vulgare*, *Ricinus communis*, and *T. aestivum*. The amino acids of this group ranged between 44 and 690 with molecular weight 4.80–75.75 kDa. The theoretical pI of this group was found from 5.23 to 7.56. Instability index of this cluster ranged between 25.65 and 57.61. Aliphatic index of cluster two ranged from 66.10 to 92.42 represents high thermostability of this group. Cluster '3' represents *S. bicolor*, *H. vulgare*,

Table 6 List of source organism of retrieved α -AI protein sequences (with accession number)

S. no.	Source organism	Accession no.	Total sequences
1	<i>Oryza sativa</i>	NP_001059199.1, P29421	2
2	<i>Zea mays</i>	NP_001232812.1, NP_001106233.1, NP_001105061.1, NP_001148385.1, NP_001147967.1, NP_001148116.1	6
3	<i>Sorghum bicolor</i>	XP_002461684.1, XP_002461685.1, XP_002459322.1, P81367, XP_002459323.1, XP_002461687.1, XP_002459556.1, P81368	8
4	<i>Corynebacterium matruchotii</i>	ZP_07404539.1	1
5	<i>Streptomyces ghanaensis</i>	ZP_06575095.1	1
6	<i>Streptomyces avermitilis</i>	NP_822072.1	1
7	<i>Streptomyces roseosporus</i>	ZP_04713226.1, ZP_06588930.1	2
8	<i>Arabidopsis thaliana</i>	NP_973763.1, NP_001117596.1, NP_177496.3	3
9	<i>Medicago truncatula</i>	XP_003628867.1, XP_003628864.1	4
10	<i>Ricinus communis</i>	XP_002531545.1, XP_002530403.1, XP_002528537.1	3
11	<i>Hordeum vulgare</i>	P16968, P13691, P32936, P34951, P01086, P07596, P16969, P28041, P11643	9
12	<i>Triticum aestivum</i>	P10846, Q43691, P16852, P16347, P17314, P16850, P01083, P01085, Q43723, P16159, P16851, P01084	12
13	<i>Eleusine coracana</i>	P01087, 1B1U, 1003192A, 1BIP	4
14	<i>Secale cereale</i>	P83048	1
15	<i>Phaseolus vulgaris</i>	Q41114, P81870	2
16	<i>Coix lachryma-jobi</i>	P15326	1
17	<i>Delonix regia</i>	P86366, P86367	2
18	<i>Streptomyces tendae</i>	P01092	1
19	<i>Streptomyces aureofaciens</i>	P04082	1
20	<i>Amaranthus hypochondriacus</i>	P80403	1
21	<i>Streptomyces griseosporus</i>	P20078, P01093	2
22	<i>Streptomyces olivaceoviridis</i>	P20596, P09921	2
23	<i>Streptomyces rochei</i>	P07512	1
24	<i>Echinochloa frumentacea</i>	AAI-BM	1

Fig. 9 Phylogenetic tree of α -AI protein sequences from different source organisms

Zea mays, *Oryza sativa*, and *E. coracana*. This cluster was further divided into two subclusters. Subcluster 1 includes 6 sequences of *S. bicolor*. Subcluster 2 includes sequences of *S. bicolor*, *H. vulgare*, *Zea mays*, *Oryza sativa*, and *E. coracana*. The amino acid residues of this cluster ranged 105–185 and molecular weight 11.47–19.62 kDa with pI values placed between 6.45 and 8.96. Instability index of cluster 3 ranged between 41.62 and 73.58, indicated that in vivo half-life of proteins of this cluster are less than 5 h and aliphatic index of this group ranged from 66.95 to 99.01. Cluster 4 consists α -AI sequences of *Delonix regia* and *T. aestivum*. The amino acid residues of this cluster ranged 19–121 and molecular weight ranged 2.07–13.89 kDa and pI values placed between 9.08 and 9.99. Instability index of this cluster ranged between 50.14 and 55.23 and Aliphatic index ranged from 41.05 to 95.12. Cluster 5 represents the bacterial α -AI sequences of *Streptomyces* (*S. ghanaensis*, *S. avermitilis*, *S. roseosporus*, *S. griseosporus*, *S. olivaceoviridis*, *S. tendae*, *S. rochei*, and *S. aureofaciens*) and one sequence of *R. communis*. These α -AI contained amino acids ranged 36–120 and their molecular weight ranged from 3.93 to 12.62 kDa. The pI of this group was ranged from 4.11 to 7.73. Instability index of these bacterial sequences ranged 15.97–38.69, indicated that in vivo half-life of these proteins are more than 16 h. Aliphatic index of cluster 5 ranged from 56.67 to 90.19. Cluster 6 represents the α -AI sequences of *Arabidopsis thaliana*, *Corynebacterium matruchotii*, *Coix lacryma-jobi*, *E. coracana*, and *Zea mays*. The sequences of cluster 6 contained amino acids in the range of 95–205 with molecular weight ranged from 9.33 to 20.62 kDa and their pI value ranged in 5.21–9.89. Instability index of this cluster ranged between 9.45 and 56.43 and Aliphatic index ranged from 26.82 to 113.67 represents high thermostability variation in this group. Cluster 7 represents 13 α -AI sequences of *Delonix regia*, *H. vulgare*, *T. aestivum*, *Oryza sativa*, *R. communis*, *P. vulgaris*, and *Medicago truncatula*. The amino acids residues in this group ranged from 22 to 769 with molecular weight ranged 2.31–78.59 kDa and theoretical pI values ranged from 3.84 to 9.42.

Superfam tools for super family analysis of sequences of α -AI revealed seven different super families (Table 7). All bacterial α -AI sequences of *Streptomyces ghanaensis* (ZP_06575095.1), *S. avermitilis* (NP_822072.1), *S. roseosporus* (ZP_04713226.1 and ZP_06588930.1), *S. griseosporus* (P20078 and P01093), *S. olivaceoviridis* (P20596 and P09921), *S. tendae* (P01092), *S. rochei* (P07512), *S. aureofaciens* (P04082), and *C. matruchotii* (ZP_07404539.1) were belong to α -AI tendamistat super family. α -AI sequences of *A. hypochondriacus* (P80403), *A. thaliana* (NP_973763.1, NP_001117596.1 and NP_177496.3), *D. regia* (P86366 and P86367), *E. coracana* (P01087, 1B1U, 1003192A, and 1BIP), *E. frumentacea* (AAI-BM), *H. vulgare* (P16968, P13691, P32936, P34951, P01086, P16969,

P28041, and P11643), *Oryza sativa* (NP_001059199.1), *S. cereal* (P83048), *S. bicolor* (XP_002461684.1, XP_002461685.1, XP_002459322.1, XP_002459323.1, XP_002461687.1, XP_002459556.1, P81367, and P81368), *T. aestivum* (P10846, Q43691, P16852, P17314, P16850, P01083, P01085, Q43723, P16159, P16851, and P01084), and *Zea mays* (NP_001232812.1, NP_001106233.1, NP_001105061.1, and NP_001148385.1) were belong to Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin super families. α -AI of *M. truncatula* (XP_003628867.1 and XP_003628864.1), *P. vulgaris* (Q41114 and P02873), *R. communis* (XP_002531545.1, XP_002530403.1 and XP_002528537.1) and *Zea mays* (NP_001147967.1 and NP_001148116.1) were found in Concanavalin A-like lectins/glucanases super family. In protein kinase-like (PK-like) superfamily, α -AI of *R. communis* (XP_002531545.1, XP_002530403.1 and XP_002528537.1) and *Zea mays* (NP_001147967.1 and NP_001148116.1) was found. α -AI sequences of *H. vulgare* (P07596), *Oryza sativa* (P29421), and *T. aestivum* (P16347) were found in STI-like superfamily. MEME analysis with given parameters results in frequently observed 20 motifs (Table 8). A set of 29 amino acids sequences 'WWWCYPGMAIPHNPLPSCR-WYVKQQTTCGI' representing motif 1 was conserved and uniformly observed in 32 α -AI protein sequences, revealing their identity with domain AAI_LTSS super family (Table 9).

Conclusions

The mature seeds of barnyard millet variety PRJ1 contain higher α -AI activity compared to finger millet and the α -AI activity is increased during the developmental stages of seed up to 7 days. α -AI gene from barnyard millet variety PRJ1 contains 315 bp nucleotides and translated product contains 104 amino acids with MW 10.72 kDa. Bioinformatic analysis reveals that it has more similar to the bifunctional inhibitor of rafi. Similarity analysis of various α -AI showed a conserved region in all the sequences and is distributed in seven superfamilies. Hence, phylogenetic clustering and variation among biochemical features of different α -AI might contribute in further classification and selection of suitable α -AI for biotechnological application. Conserved sequences in motifs may be utilized for designing of specific degenerate primers for identification and isolation of type and class of α -AI. Apart from this sequence information, structural information is also required to identify new α -AI. Thus, in silico analysis might be used for future genetic engineering of α -AI.

Table 7 ProtParam results of α -AI sequences from different source organisms

Serial no.	Accession no.	Source organism	No. of amino acids	Molecular weight	Theoretical pI	Total no. of negatively charged residues (Asp + Glu)	Total no. of positively charged residues (Arg + Lys)	Instability index	Aliphatic index	GRAVY
1	NP_001059199.1	<i>Oryza sativa</i> Japonica Group	148	15767.3	7.48	14	15	42.11	87.97	0.099
2	NP_001232812.1	<i>Zea mays</i>	155	16342.1	8.07	12	14	60.78	88.26	0.132
3	NP_001106233.1	<i>Zea mays</i>	155	16301.9	8.07	12	14	60.91	88.90	0.144
4	NP_001105061.1	<i>Zea mays</i>	185	19626.7	8.29	14	17	58.77	85.03	0.077
5	XP_002461684.1	<i>Sorghum bicolor</i>	145	15522.1	7.46	11	12	50.92	92.90	0.334
6	XP_002461685.1	<i>Sorghum bicolor</i>	148	15565.2	8.55	10	14	56.25	87.03	0.253
7	XP_002459322.1	<i>Sorghum bicolor</i>	105	11471.2	7.38	12	13	73.58	66.95	-0.237
8	XP_002459323.1	<i>Sorghum bicolor</i>	147	15563.2	7.46	11	12	46.56	89.59	0.265
9	XP_002461687.1	<i>Sorghum bicolor</i>	151	15769.3	6.54	12	12	63.29	99.01	0.291
10	XP_002459556.1	<i>Sorghum bicolor</i>	143	14898.5	8.04	9	11	47.84	91.54	0.487
11	ZP_07404539.1	<i>Corynebacterium matruchotii</i> ATCC 14266	170	18288.7	5.21	19	16	35.22	60.29	-0.405
12	ZP_06575095.1	<i>Streptomyces ghanaensis</i> ATCC 14672	115	12020.6	7.73	9	10	24.87	67.91	-0.028
13	NP_822072.1	<i>Streptomyces avermitilis</i> MA-4680	119	12552.1	6.54	11	11	38.69	81.18	-0.004
14	ZP_04713226.1	<i>Streptomyces roseosporus</i> NRRL 11379	101	10220.5	5.60	6	5	33.36	80.40	0.193
15	NP_973763.1	<i>Arabidopsis thaliana</i>	205	20621.4	6.50	8	8	44.78	75.27	0.199
16	ZP_06588930.1	<i>Streptomyces roseosporus</i> NRRL 15998	83	8531.4	5.00	6	4	33.21	67.11	0.00
17	NP_001117596.1	<i>Arabidopsis thaliana</i>	134	14607.2	8.14	8	10	50.12	93.88	0.158
18	NP_177496.3	<i>Arabidopsis thaliana</i>	152	16509.5	8.67	8	12	40.02	96.25	0.384
19	NP_001148385.1	<i>Zea mays</i>	213	21654.2	8.12	9	11	56.43	84.88	0.302
20	XP_003628867.1	<i>Medicago truncatula</i>	265	29260.9	4.91	23	15	25.84	90.11	0.028
21	XP_003628864.1	<i>Medicago truncatula</i>	300	33616.0	5.36	30	22	33.63	92.83	-0.070
22	NP_001147967.1	<i>Zea mays</i>	749	77272.2	6.69	67	65	49.34	87.33	0.050
23	NP_001148116.1	<i>Zea mays</i>	769	78596.8	6.55	67	64	49.76	87.36	0.092
24	XP_002531545.1	<i>Ricinus communis</i>	681	76273.8	5.96	77	66	29.98	90.18	-0.109
25	XP_002530403.1	<i>Ricinus communis</i>	346	38765.4	9.42	27	36	38.51	85.06	-0.149
26	XP_002528537.1	<i>Ricinus communis</i>	690	75756.3	6.05	73	65	36.17	92.42	-0.082

Table 7 (continued)

Serial no.	Accession no.	Source organism	No. of amino acids	Molecular weight	Theoretical PI	Total no. of negatively charged residues (Asp + Glu)	Total no. of positively charged residues (Arg + Lys)	Instability index	Aliphatic index	GRAVY
27	P16968	<i>Hordeum vulgare</i>	146	15816.0	5.36	13	12	27.72	66.10	-0.127
28	P01093	<i>Streptomyces griseus</i>	78	8121.9	4.22	9	4	15.97	68.85	-0.012
29	P09921	<i>Streptomyces olivaceoviridis</i>	73	7424.1	4.42	7	1	28.95	64.25	-0.036
30	P13691	<i>Hordeum vulgare</i>	152	16429.3	5.36	15	13	40.03	83.95	0.186
31	P20596	<i>Streptomyces olivaceoviridis</i>	75	7626.3	4.27	8	1	28.45	62.53	-0.092
32	P81367	<i>Sorghum bicolor</i>	118	12499.5	8.96	8	14	57.42	71.19	-0.070
33	P01084	<i>Triticum aestivum</i>	124	13185.1	5.23	12	10	32.71	78.71	0.038
34	P32936	<i>Hordeum vulgare</i>	149	16526.1	5.77	13	12	52.11	73.36	-0.057
35	P16851	<i>Triticum aestivum</i>	145	15459.7	6.86	12	12	25.76	77.31	-0.049
36	P34951	<i>Hordeum vulgare</i>	143	15178.7	6.68	9	9	57.59	86.71	0.318
37	P01086	<i>Hordeum vulgare</i>	148	16135.6	7.50	11	12	48.51	84.46	0.075
38	P07596	<i>Hordeum vulgare</i>	203	22163.8	7.77	21	22	60.41	64.43	-0.469
39	P29421	<i>Oryza sativa subsp. japonica</i>	200	21417.4	8.66	19	22	57.03	81.90	-0.156
40	P01087	<i>Eleusine coracana</i>	122	13138.3	8.07	10	12	43.55	83.93	0.007
41	P16969	<i>Hordeum vulgare</i>	147	15964.5	8.12	18	20	41.62	95.51	0.019
42	P83048	<i>Secale cereale</i>	25	2712.9	3.92	4	1	32.18	62.40	-0.260
43	P07512	<i>Streptomyces rochei</i>	76	8129.8	4.49	8	2	32.63	51.32	-0.250
44	P16159	<i>Triticum aestivum</i>	143	15782.3	5.31	12	10	56.21	81.19	0.015
45	Q43723	<i>Triticum aestivum</i>	121	13831.2	9.23	14	21	53.23	95.12	-0.130
46	P15326	<i>Coix lachryma-jobi</i>	133	14305.2	6.07	11	10	9.45	58.80	-0.456
47	Q41114	<i>Phaseolus vulgaris</i>	240	26596.5	5.17	25	18	31.91	82.00	-0.223
48	P81870	<i>Phaseolus vulgaris</i>	30	3195.6	3.84	4	1	21.78	113.67	0.283
49	IB1U	<i>Eleusine coracana</i>	122	13138.3	8.07	10	12	43.55	83.93	0.007
50	1003192A	<i>Eleusine coracana</i>	95	9331.5	9.89	2	11	48.72	77.37	0.099
51	IBIP	<i>Eleusine coracana</i>	122	13128.2	8.07	10	12	43.55	83.93	0.013
52	P01085	<i>Triticum aestivum</i>	124	13337.3	6.66	11	11	39.41	77.98	0.015
53	P20078	<i>Streptomyces griseus</i>	120	12624.2	4.91	13	9	22.29	80.50	0.225
54	P01083	<i>Triticum aestivum</i>	153	16799.5	7.45	14	15	25.65	78.30	0.022
55	P81368	<i>Sorghum bicolor</i>	116	12776.1	8.55	8	12	47.47	83.19	0.157
56	P28041	<i>Hordeum vulgare</i>	145	15499.7	5.86	14	12	38.64	78.07	-0.090

Table 7 (continued)

Serial no.	Accession no.	Source organism	No. of amino acids	Molecular weight	Theoretical PI	Total no. of negatively charged residues (Asp + Glu)	Total no. of positively charged residues (Arg + Lys)	Instability index	Aliphatic index	GRAVY
57	P16850	<i>Triticum aestivum</i>	145	15516.7	7.50	12	13	35.86	79.38	-0.031
58	P17314	<i>Triticum aestivum</i>	167	18108.1	7.43	11	12	48.69	93.47	0.213
59	P111643	<i>Hordeum vulgare</i>	171	18525.5	6.07	12	11	50.48	86.78	0.177
60	P80403	<i>Amaranthus hypochondriacus</i>	32	3592.1	6.08	3	3	37.70	21.25	-0.613
61	P16347	<i>Triticum aestivum</i>	180	19633.0	6.77	22	21	56.01	70.44	-0.476
62	P13867	<i>Zea mays</i>	206	22074.7	8.16	16	19	30.19	53.54	-0.257
63	P04082	<i>Streptomyces aureofaciens</i>	36	3938.2	4.11	5	2	31.89	56.67	-0.317
64	P01092	<i>Streptomyces tendae</i>	104	10759.1	6.01	8	7	27.87	90.19	0.179
65	P16852	<i>Triticum aestivum</i>	27	3070.3	4.37	2	1	47.59	57.78	-0.233
66	Q43691	<i>Triticum aestivum</i>	121	13889.2	9.08	15	21	55.52	95.12	-0.155
67	P02873	<i>Phaseolus vulgaris</i>	246	27207.2	5.03	25	18	32.97	84.76	-0.192
68	P86366	<i>Delonix regia</i>	19	2071.3	9.99	1	3	50.14	41.05	-1.274
69	P86367	<i>Delonix regia</i>	22	2312.4	5.51	4	3	31.83	26.82	-1.268
70	P10846	<i>Triticum aestivum</i>	44	4796.6	7.56	2	3	57.61	77.50	0.107
71	AAI-BM	<i>Echinocloa frumentacea</i>	104	10720.5	6.77	12	11	43.55	83.93	0.007

Table 8 Distribution of Super families among α -AI from different source organisms using superfam server

Serial no.	Family	Super family	Sequences accession no. (range of amino acids)
1	Alpha-amylase inhibitor tendamistat	Alpha-amylase inhibitor tendamistat	ZP_04713226.1 (28–96), ZP_07404539.1 (11–65), ZP_06575095.1 (41–109), NP_822072.1 (33–103), ZP_06588930.1 (10–78), P01093 (2–72), P09921 (2–71), P20596 (2–73), P07512 (2–71), P20078 (34–104), P04082 (2–36), P01092 (31–103)
2	Proteinase/alpha-amylase inhibitors	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin	NP_001059199.1 (32–138), NP_001232812.1 (33–146), NP_001106233.1 (33–146), NP_001105061.1 (63–176), XP_002461684.1 (30–135), XP_002461685.1 (29–131), XP_002459322.1 (3–102), XP_002459323.1 (33–138), XP_002461687.1 (30–142), XP_002459556.1 (18–137), NP_973763.1 (36–104), NP_001117596.1 (32–113), NP_177496.3 (32–113), NP_001148385.1 (43–117), P16968 (19–136), P13691 (36–147), P81367 (7–110), P01084(6–117), P32936 (30–137), P16851 (13–140), P34951 (30–142), P01086 (24–138), P01087 (2–116), P16969 (29–138), P83048 (1–25), P16159 (30–137), Q43723 (25–114), 1B1U(2–116), 1003192A (1–93), 1BIP (2–116), P01085 (6–117), P01083 (36–140), P81368 (3–108), P28041 (29–142), P16850 (29–142), P17314 (25–160), P11643 (32–164), P80403 (1–32), P16852 (2–27), Q43691 (25–114), P10846 (6–44), AAI-BM
3	Legume lectins	Concanavalin A-like lectins/glucanases	XP_003628867.1 (25–258), XP_003628864.1 (25–259), NP_001147967.1 (28–256), NP_001148116.1 (34–266), XP_002531545.1 (49–277), XP_002530403.1 (42–251), XP_002528537.1 (34–274), Q41114 (23–223), P02873 (24–228)
4	Protein kinases, catalytic subunit	Protein kinase-like (PK-like)	NP_001147967.1 (375–688), NP_001148116.1 (393–704), XP_002531545.1 (339–631), XP_002528537.1 (333–622)
5	Kunitz (STI) inhibitors	STI-like	P07596 (24–201), P29421 (23–195), P16347 (1–178)
6	Family 19 glycosidase	Lysozyme-like	P15326 (10–133)
7	Osmotin, thaumatin-like protein	Osmotin, thaumatin-like protein	P13867 (2–206)

Table 9 MEME analysis of α -AI sequences

Motifs	Motifs presents in no. of sequences	Motifs width	Meme.xml (peptide) best possible match	Domain
1	32	29	WWWCYPGMAIPHNPLPSRWYVKQQT CGI	AAI_LTSS super family
2	22	32	LEDLPGCPREMQRDFAKTLVTEGECNLM TIHG	Local conserved domains
3	32	15	MKRRCCQELAKIPQY	Local conserved domains
4	32	15	CRCEALRIMMDGMYT	Local conserved domains
5	10	50	PAPECVHYYQDWRYTFVTNDCAITYSVTVEYQDGQ EVP-CRILNPGDITTF	Local conserved domains
6	20	21	HHQLILSAAVLLSVFAAAAAT	Local conserved domains
7	4	50	WRHRYEIIAGVASALHYLHHEWEQ RVIHRDIKSSNVMLDENY-NARLGDFG	PKc_like super family
8	4	50	RSEFLAELSHIAGLRHRNLVRLQGWCYKKGEILLVYDYMPNG-SLDRHLFD	PKc_like super family
9	8	26	KNHIVAVEFDTFMDEQFHDVNNHNIK	Lectin_L-type superfamily
10	2	50	QQAMRCVAKLMPCQPYIHL SIPPPLCCNPMKQIAEKDVS-CLCTAFKHPD	AAI_LTSS super family
11	3	50	DVRISFRAYTTTCVQSTEW HIDSELVSGRRHVITGPVKDPSPSGRE-NAFRI	STI super family
12	6	38	HEFSYKELSFATKGFDEKNLIGQGDFGVVYKGILPKHN	PKc_like super family
13	6	21	GFSGATQGSYETHDILWWSFS	Lectin_L-type super family
14	2	49	EKYSGAEVHEYKLMACGDWCQDLGVFRDLKGGAWFLGATEPY-HVVVFKK	STI super family
15	4	15	AECPWILGGGTMP SK	Local conserved domains
16	7	16	PGYGTNGNYVLGLVLC	Local conserved domains
17	5	25	PGCRKEVMKLTAA SIPAVCKLPIN	AAI_LTSS super family
18	4	42	EGRILEAVDPRLGNEYEEGEARRVLLLGLACSHPEPALRPGM	Local conserved domains
19	4	35	FYKAPFQIRDSTTG NVMSFDNFTMNITTHRQANS	Lectin_L-type super family
20	2	50	MLHRNWDLSFKFPNFFG PYRNDIINYQGDAVESNGTIQLTK-IENDINMPY	Lectin_L-type super family

Acknowledgements This work is highly acknowledged to the Department of Science and Technology (DST) for provided facilities under DST-FIST programme. We sincerely thank Director Research, Dean CBSH, G. B. P. U. A. and T., Pantnagar for providing necessary requirements for the research work.

Author contributions PP contributed in all the experiment under the guidance of AKV and AD. All authors' proof reads the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410

- Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22:195–201
- Bailey TL, Elkan C (1994) University of California SDDoCS, engineering. Fitting a mixture model by expectation maximization to discover motifs in bipolymers: Citeseer. UCSD technical report CS94-351
- Barrett ML, Udani JK (2011) A proprietary alpha-amylase inhibitor from white bean (*Phaseolus vulgaris*): a review of clinical studies on weight loss and glycemic control. *J Nutr* 24:1–10
- Bernfeld P (1955) Amylases, alpha and beta. In: Colowick SP, Kaplan NO (eds) *Methods in enzymology*, vol 1. Academic Press, New York, pp 149–158
- Bjellqvist B, Hughes GJ, Pasquali C, Paquet N, Ravier F, Sanchez JC, Frutiger S, Hochstrasser D (1993) The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis* 14:1023–1031
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Ann Biochem* 72:248–254
- Chrispeels MJ, Grossi-de-Sá MF, Higgins TJV (1998) Genetic engineering with α -amylase inhibitors seeds resistant to bruchids. *Seed Sci Res* 8:257–263

- Devi PB, Vijayabharathi R, Sathyabama S, Malleshi NG, Priyadarisini VB (2014) Health benefits of finger millet (*Eleusine coracana* L.) polyphenols and dietary fiber: a review. *J Food Sci Technol* 51(6):1021–1040
- Franco OL, Ridgen DJ, Melo FR, Bloch C Jr, Silva CP, Grossi-de-a MF (2000) Activity of wheat α -amylase inhibitors towards bruchid α -amylases and structural explanation of observed specificities. *Eur J Biochem* 267:2166–2173
- Franco OL, Ridgen DJ, Melo FR, Grossi-de- Sa MF (2002) Plant α -amylase inhibitors and their interaction with insect α -amylases: structure, function and potential for crop protection. *Eur J Biochem* 269:397–412
- Gatehouse AMR, Gatehouse JA (1998) Identifying proteins with insecticidal activity: use of encoding genes to produce insect resistant transgenic crops. *Pest Sci* 52:165–175
- Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *Electrophoresis* 18:2714–2723
- Guruprasad K, Reddy BVB, Pandit MW (1990) Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng* 4:155–161
- Guzman-Partida AM, Jatomea FO, Robels BMR, Ortega NM, Vazquez MM (2007) Characterization of α -amylase inhibitor from *Palo fierro* seeds. *Plant Physiol Biochem* 45(3):711–715
- Heidari R, Zareae S, Heidarizadeh M (2005) Extraction, purification and inhibitory effect of alpha-amylase inhibitor from wheat (*Triticum aestivum* var. *Zarrin*). *Pak J Nutr* 4(2):101–105
- Hung CH, Huang CC, Tsai WS, Wang HL, Chen YL (2003) Purification and characterization of a trypsin inhibitor from *Brassica campestris* seeds. *J Yuanpei Univ Sci Technol* 10:13–22
- Ikai A (1980) Thermostability and aliphatic index of globular proteins. *J Biochem* 88:1895–1898
- Koiwa H, Bressan RA, Hasegawa PM (1997) Regulation of protease inhibitors and plant defence. *Trends Plant Sci* 2:379–384
- Kokiladevi E, Manickam A, Thayumanavan B (2005) Characterization of alpha-amylase inhibitor in *Vigna sublobata*. *Botanical. Bull Acad Sin* 46:189–196
- Konarev AV (1996) Interaction of insect digestive enzymes with plant protein inhibitors and host-parasite co-evolution. *Euphytica* 92:89–94
- Kulkarni LR, Naik RK, Katarki PA (1992) Chemical composition of minor millets. *Karnataka J Agric Sci* 5(3):255–258
- Kumar KK, Parameshwaran PK (1998) Characterization of storage protein from selected varieties of foxtail millet (*Setaria italica* (L) Beauv). *J Food Agric* 77:535–542
- Kyte J, Doolittle RF (1982) A simple method for displaying the hydrophobic character of a protein. *J Mol Biol* 157:105–132
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685
- McEwan R, Madivha RP, Djarova T, Oyediji OA, Opoku AR (2010) Alpha α -amylase inhibitor of madumbe (*Colocasia esculenta*): isolation, purification and selectivity toward α -amylases from various sources. *Afr J Biochem Res* 4(9):220–224
- Mehrabadi M, Bandani AR, Saadati F (2010) Inhibition of Sunn pest, *Eurygaster integriceps*, α -amylases by α -amylase inhibitors (T- α AI) from Triticale. *J Insect Sci* 10:179 (**available online: insectscience.org/10.179**)
- Mikola J, Kirsi M (1972) Difference between endospermal and embryonal trypsin inhibitor in barley, wheat and rye. *Acta Chem Scand* 26:787–795
- Mulimani VH, Supriya D (1993) Alpha amylase inhibitors in sorghum (*Sorghum bicolor*). *Plant Foods Hum Nutr* 44:261–266
- Nagaraj RH, Pattabiraman TN (1985) Purification and properties of an α -amylase inhibitor specific for human pancreatic amylase from proso (*Panicum miliaceum*) seeds. *J Biosci* 7:257–268
- Panwar P, Dubey A, Verma AK (2016) Evaluation of nutraceutical and antinutritional properties in barnyard and finger millet varieties grown in Himalayan region. *J Food Sci Technol* 53(6):2779–2787
- Pereira PJB, Lozanov V, Patthy A, Huber R, Bode W, Pongor S, Strobl S (1999) Specific inhibition of insect α -amylases: yellow meal worm α -amylase in complex with the *Amaranth* α -amylase inhibitor at 2.0 Å resolution. *Structure* 7:1079–1088
- Rackis JJ, Anderson RL (1964) Isolation of four soybean inhibitors by DEAE-cellulose chromatography. *Biochem Biophys Res Commun* 15(3):230–235
- Rawlings ND, Morton FR, Barrett AJ (2006) MEROPS: the peptidase database. *Nucleic Acids Res* 34:270–272
- Richardson M (1981) Plant protein inhibitors of enzymes: their role in animal nutrition and plant defense. *J Biol Educ* 15:178–182
- Richardson M (1991) Protein' inhibitors of enzymes. *Food Chem* 6:235–253
- Richardson M, Campos FAP, Xavier-Filho J, Macedo MLR, Maia GMC, Yarwood A (1986) The amino acid sequence and reactive (inhibitory) site of the major trypsin iso inhibitor (DE5) isolated from seeds of the Brazilian Carolina tree (*Adenanthera pavonina* L.). *Biochem Biophys Acta* 872:134–140
- Rogers S, Wells R, Rechsteiner M (1986) Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* 234:364
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Shivaraj B, Pattabiraman TN (1981) Purification and properties of two alpha-amylase inhibitors from ragi (*Eleusine coracana*) grains. *Indian J Biochem Biophys* 17:181–185
- Silva MC, Da M, Grossi de sa MF, Chrispeels MJ, Togawa RC, Neshich G (2000) Analysis of structural and physico-chemical parameters involved in the specificity of binding between α -amylases and their inhibitors. *Protein Eng* 13:167–177
- Sivakumar S, Mohan M, Franco OL, Thayumanavan B (2006) Inhibition of insect pest α -amylases by little and finger millet inhibitors. *Pestic Biochem Physiol* 85:155–160
- Strobl S, Muhlhahn P, Bernstein R, Wilschek R, Maskos K, Wunderlich M, Huber R, Holk TA (1995) Determination of three-dimensional structure of the bifunctional alpha-amylase/trypsin inhibitor from ragi seeds by NMR spectroscopy. *Biochemistry* 34(2):8281–8291
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739
- Veena B, Chimmad BV, Naik RK, Shantakumar G (2005) Physico-chemical and nutritional studies in barnyard millet. *Karnataka J Agric Sci* 18(1):101–105
- Veronique LBA, Coralie BG, Francoise P, Pierre R (1997) Characterization and functional properties of the α -amylase inhibitor (α -AI) from kidney bean (*Phaseolus vulgaris*) seeds. *Biochim Biophys Acta* 1343:31–40
- Weder JKP (1985) Chemistry of legume protease inhibitors and their use in taxonomy. *Plant Foods Hum Nutr* 35:183–194
- Yamada T, Hattori K, Ishimoto M (2001) Purification and characterization of two α -amylase inhibitors from seeds of tepary bean (*Phaseolus acutifolius* A. Gray). *Phytochemistry* 58:59–66
- Zuckerkindl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. *Evol Genes Proteins* 97:166