

Regular paper

Detection of single nucleotide polymorphisms in 24 kDa dimeric α -amylase inhibitors from cultivated wheat and its diploid putative progenitors

Ji-Rui Wang¹, Yu-Ming Wei¹, Ze-Hong Yan, You-Liang Zheng**Triticeae Research Institute, Sichuan Agricultural University, Dujiangyan 611830, China*

Received 25 January 2005; received in revised form 27 February 2005; accepted 1 March 2005

Available online 19 March 2005

Abstract

Seventeen new genes encoding 24 kDa family dimeric α -amylase inhibitors had been characterized from cultivated wheat and its diploid putative progenitors. And the different α -amylase inhibitors in this family, which were determined by coding regions single nucleotide polymorphisms (cSNPs) of their genes, were investigated. The amino acid sequences of 24 kDa α -amylase inhibitors shared very high coherence (91.2%). It indicated that the dimeric α -amylase inhibitors in the 24 kDa family were derived from common ancestral genes by phylogenetic analysis. Eight α -amylase inhibitor genes were characterized from one hexaploid wheat variety, and clustered into four subgroups, indicating that the 24 kDa dimeric α -amylase inhibitors in cultivated wheat were encoded by multi-gene. Forty-five cSNPs, including 35 transitions and 10 transversions, were found, and resulted in a total of ten amino acid changes. The cSNPs at the first site of a codon cause much more nonsynonymous (92.9%) than synonymous mutations, while nonsynonymous and synonymous mutations were almost equal when the cSNPs were at the third site. It was observed that there was Ile105 instead of Val105 at the active region Val104-Val105-Asp106-Ala107 of the α -amylase inhibitor by cSNPs in some inhibitors from *Aegilops speltoides*, diploid and hexaploid wheats.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Single nucleotide polymorphism; 24 kDa dimeric α -amylase inhibitor; *Triticum*; *Aegilops*

1. Introduction

Single nucleotide polymorphisms (SNPs) are single base-pair substitutions that occur within and outside genes, and they are the most common form of sequence variation [1]. A substantial private and public effort has been undertaken to characterize the SNPs responsible for genetic diversity. The SNPs are identified in expressed sequence tag (EST) sequences, thus the polymorphisms can be directly used to map functional and expressed genes, rather than DNA sequences derived from conventional RAPD and AFLP techniques, which are typically not genes [2–4]. SNPs are found in both the coding and non-coding regions of the genome, and randomly distributed markers as well as markers

clustered in genes can be discovered [5,6]. The majority of SNPs in coding regions (cSNPs) are single base substitutions, which may or may not result in amino acid substitutions. Some cSNPs alter a functionally important amino acid residue, and these are of interest for their potential links with phenotype [7]. Other cSNPs are useful for their potential links to functional cSNPs via linkage disequilibrium mapping [5,7]. It is very important to identify, map and determine the allele frequency of a useful set of SNPs, which are evenly distributed along the genome. Several methods had been carried out in the SNPs mining, including the re-sequencing of PCR amplicons with or without pre-screening, electronic SNP (eSNP) discovery in shotgun genomic libraries, and SNP discovery in EST libraries [8].

α -Amylase [α -(1,4)-glucan-4-glucanohydrolases] is a family of enzymes that hydrolyze α -D-(1,4)-glucan linkages in amylase, amylopectin, glycogen and phytoglycogen, and plays an important role in the carbohydrate metabolism of

* Corresponding author. Tel.: +86 835 2882007; fax: +86 835 2882336.
E-mail address: grmb@sicau.edu.cn (Y.-L. Zheng).

¹ The first two authors are equally contributed to this paper.

many autotrophic and heterotrophic organisms [9]. Heterotrophic organisms use α -amylase primarily to digest starch in their food sources [10]. Chitinase and β -1, 3 glucanase enzymes, lectins, arcelins, vicilins, systemins and enzyme inhibitors are the proteins which could help plant to defense insect pests and pathogens [11–13]. The enzyme inhibitors act on the key insect gut digestive hydrolases, the α -amylases and proteinases. Several kinds of α -amylase and proteinase inhibitors in seeds and vegetative organs act to regulate the numbers of phytophagous insects [14–16]. α -Amylase inhibitors are attractive candidates for the control of seed weevils as these insects are highly dependent on starch as an energy source [17].

It is well established that there were multiple forms of proteins active on exogenous or endogenous α -amylases in wheat kernel, and proteinaceous dimeric α -amylase inhibitors with inhibitory activity were against the α -amylase of various origins [18,19]. It was known that the bulk of wheat albumins consisted of a few amylase isoinhibitor families very likely phylogenetically related and coded by a small number of parental genes [20]. In the absence of dissociating agents, wheat inhibitors active on exogenous α -amylases can be fractionated by gel filtration into three isoinhibitor families with apparent molecular weight close to 60 kDa, 24 kDa and 12 kDa [18]. The two inhibitor families with molecular weight 12 kDa (0.28 family) and 24 kDa (0.19 family) had been extensively investigated [10,21–23].

In cereal seeds, α -amylase inhibitors proteins with 120–130 amino acids, which include trypsin inhibitors as well as α -amylase inhibitors, were known as one great family on the basis of the homology of their amino acid sequence [24]. Usually, each α -amylase inhibitor has ten Cys residues and forms five disulfide bonds. In this family, the dimeric α -amylase inhibitor 0.19 has been well characterized. Using site-directed mutagenesis and X-ray crystallographic analysis, the structure–function relationships of α -amylase inhibitor 0.19 had been investigated, such as in diabetes control and nutritional and/or toxicological aspects of food [25].

For weevil control, α -amylase inhibitors could be used through plant genetic engineering. However, many insects have several α -amylases that differ in specificity, and the successful utilization of a food source is dependent on the presence of a α -amylase for which there is no specific inhibitor [10]. The structure and function of dimeric α -amylase inhibitor genes with different cSNPs by direct sequencing, which is the most direct way to identify SNP polymorphisms [8], from cultivated wheat varieties and its diploid putative progenitor were investigated.

2. Materials and methods

2.1. Plant material

The genomic constitution of cultivated wheat (*T. aestivum*) is AABBDD. It was thought that the AA, BB

and DD genomes of cultivated wheat were derived from *T. urartu*, *Ae. speltooides* and *Ae. tauschii*, respectively [26–29]. In this study, the cultivated wheat and its diploid putative A-, B-, and D-genome progenitor species were chosen to characterize 24 kDa family α -amylase inhibitors. There were 5 *T. monococcum* ($2n=2x=14$, A^mA^m), 5 *T. urartu* ($2n=2x=14$, A^uA^u), 3 *Ae. tauschii* ($2n=2x=14$, DD), and 3 *Ae. speltooides* ($2n=2x=14$, SS) accessions and 1 cultivated wheat variety (Table 1).

2.2. Data of the α -amylase inhibitors

The EST data were retrieved from NCBI (<http://www.ncbi.nlm.nih.gov>), the GrainGenes database (<http://wheat.pw.usda.gov>) and BBSRC cereal functional genomics (<http://www.cerealsdb.uk.net/index.htm>). Almost all the EST sequences of α -amylase inhibitor genes were obtained from three hexaploid wheat varieties, Chinese Spring, Cheyenne and Butle86.

The protein data were obtained from Protein Data Bank (<http://www.rcsb.org/pdb/>) and ExPASy Proteomics Server (<http://au.expasy.org/>).

2.3. DNA isolation and PCR amplification

Genomic DNA was extracted from leaves of single plants about 2 weeks old with a modified CTAB protocol as described in Yan et al. [30].

Based on the alignment of 72 α -amylase inhibitor gene sequences, two primers, P1 (5'-CTATGTATGCTCGTGGC-GAC-3') and P2 (5'-ACTCATTT/CGCTTGACTAGGC-3'), were designed to amplify the complete ORFs of dimeric α -amylase inhibitors. PCR amplification was performed with a GeneAmp PTC-200 cyclor (MJ Research, USA) in 50 μ L

Table 1
List of accessions used for the amplification and sequencing of the α -amylase inhibitor genes

Species	Genome	Accession	GenBank number
<i>T. monococcum</i>	A ^m	As261	AY729685
		PI 428149	No sequence obtained
		PI 428152	No sequence obtained
		PI 428154	No sequence obtained
		PI 428156	No sequence obtained
<i>T. urartu</i>	A ^u	PI 428197	AY856086
		PI 428208	No sequence obtained
		PI 428257	No sequence obtained
		PI 428328	AY856087
		PI 487268	No sequence obtained
<i>Ae. speltooides</i>	S	PI 487231	AY856089
		PI 487232	AY856090
		PI 487233	AY856091
<i>Ae. tauschii</i>	D	As69	AY729679
		PI 499262	AY856085
		PI 486266	AY856084
<i>T. aestivum</i>	ABD	Chuannong16 ^a	AY729672–AY729678 AY834214

^a Sequenced by cloning the fragments into the plasmid.

volume, which consisted of 200–300 ng of genomic DNA, 200 μ M of each dNTPs, 4 μ M of each primers, 1 U *Taq* polymerase, 1.5 mM Mg^{2+} and 1 $\times$ PCR buffer. The cycling parameters were 95 °C for 5 min to pre-denature, followed by 40 cycles of 95 °C for 1 min, 60 °C for 30 s and 72 °C for 1 min, and a final extension at 72 °C for 5 min.

2.4. Sequence analysis of α -amylase inhibitor

Amplification products were separated in 0.8% agarose gels. The desired DNA fragments were recovered from gels and ligated to the pBluescript SK (+) T-vector plasmid (Stratagene), and then the positive clones were screened and sequenced. The analysis of full-length sequence and the construction of subsequent nucleotide sequence were carried out under DNAMAN 4.0, and multiple sequence alignment software Clustal W (<http://www.ebi.ac.uk/clustalw>) was used for the SNP evaluation. The α -amylase inhibitor ORFs were translated into amino acid sequences using ORF Finder program in the NCBI (<http://www.ncbi.nlm.nih.gov>). The phylogenetic trees were constructed by MEGA, version 2 [31].

2.5. Structure predictions

The three secondary structure elements of greatest interest are helical region known as α -helix (H), extended region known as β -stand, which join together to form β -sheet (E), and coil/loop (C or '.'). The method of Gibrat was used to predict the structures of α -amylase inhibitors for α -helix, β -sheet, random coil and β -turn conformations [32]. For the method of Gibrat in which:

$$\text{Info}(k, i) = \sum_{j=-8}^8 I(k, i+j)$$

where k designates the given state (α -helix, β -sheet, β -turn, and coil); j is the directional position apart from residue i ; $\text{Info}(k, i)$ is the information that amino acid residue i behaves on state k .

Fold recognition, disulfide bridges analysis and structural alignments were performed with software ANTHEPROT [31,33].

To compare the secondary structure of α -amylase inhibitors, 3D-PSSM protein fold recognition (threading) server (<http://www.embl-heidelberg.de/predictprotein>) was used to assume the predicted secondary structure of these 24 kDa family α -amylase inhibitors.

3. Results

3.1. Alignment of the dimeric α -amylase inhibitor gene sequences

To retrieve the dimeric α -amylase inhibitor gene sequences from the available database, the coding region

of a dimeric α -amylase inhibitor gene, BE423110, was chosen to perform a Blast search against the GenBank non-redundant DNA database, Bavarian State Research Center for Agronomy, the GrainGenes database and BBSRC cereal functional genomics. Seventy-two dimeric α -amylase inhibitor genes, gene fragments and pseudo-genes were retrieved, and used in the nucleotide alignment through DNAMAN and Clustal W. Although the α -amylase inhibitor genes were much conserved in its length, there were about 40 candidate SNPs in those sequences. It is necessary to get further information of the SNPs in α -amylase inhibitor genes. By analyzing the amino acid sequences of dimeric α -amylase inhibitors from the PDB and ExPASy database, it was found that there was no intron in the α -amylase inhibitor genes. Thus, it is possible to isolate the complete coding sequences of α -amylase inhibitor genes by direct PCR amplification.

3.2. Isolation the coding region of dimeric α -amylase inhibitor

Using the primers P1 and P2, genomic PCR amplifications were conducted, and one desire DNA band was detected in each accession of *Triticum* and *Aegilops* (figure not shown). It indicated that the PCR products consisted of only one sequence by direct sequencing of the amplification products from diploid species. This has been further supported by the clone sequencing. However, it was obvious that there were a number of sequences in the amplification products from one cultivated wheat variety by both direct sequencing and clone sequencing. Therefore, it is necessary to clone the PCR amplification fragments of common wheat cultivars for sequence analysis. After cloning the fragments into the plasmid T-vector pBluescript SK (+), eight positive clones from cultivated wheat were sequenced.

A number of 24kDa family dimeric α -amylase inhibitor genes from cultivated wheat and its diploid putative progenitors had the similar sequence with the other amylase inhibitor genes from the database. Thus, 38 representative α -amylase inhibitor genes were chosen in the further analysis. Thirty-six out of 38 dimeric α -amylase inhibitor genes were 375 bp, while only 2 genes (1 from *T. monococcum* As261 and 1 from *Ae. speltooides* PI 487232) were 376 bp (Fig. 1). However, no amplification fragments were obtained from 3 *T. urartu* (i.e. PI 428208, PI 428257 and PI 687268) and 4 *T. monococcum* (i.e. PI 428149, PI 428152, PI 428154 and PI 428156) accessions. Among the new α -amylase inhibitor genes, 8 genes were from cultivated wheat, 1 from *T. monococcum*, 2 from *T. urartu*, 3 from *Ae. tauschii* and 2 from *Ae. speltooides* (Table 1). These α -amylase inhibitor gene sequences have been deposited in the GenBank (<http://www.ncbi.nlm.nih.gov>).

AY856091	AGTGGTCCCTTGGATGTGCTATCCAGGGTATGCTTCAAGGTGCCC GCCTCCCTAGCTGTCGT	63
AY856085	AGTGGTCCCTTGGATGTGCTATCCGGGGCAGGCCTTCCAAGGTGCCC GCCTCCCTAGCTGTCGT	63
AY856087	AGTGGTCCCTTGGATGTGCTATCCGGGGCAGGCCTTCCAAGGTGCCC GCCTCCCTAGCTGTCGT	63
AY729672	AGTGGTCCCTTGGATGTGCTATCCGGGGCAGGCCTTCCAAGGTGCCC GCCTCCCTAGCTGTCGT	63
AY834214	AGTGGTCCCTTGGATGTGCTATCCGGGGCAGGCCTTCCAAGGTGCCC GCCTCCCTAGCTGTCGT	63
AY729685	AGTGGTCCCTTGGATGTGCTATCCAGGGTATGCTTCCAAGGTGCCC GCCTCCCTAGCTGTCGT	63
AY856091	CCAATTGCTGAGGCTCCAGTGCAATGCGACGCCAGGTGCCCGA. GGC TGTCC TAAGGGA CTGCTG	125
AY856085	CCAATTGCTGAGGCTCCAGTGCAATGCGACGCCAGGTGCCCGA. GGC TGTCC TAAGGGA CTGCTG	125
AY856087	CCAATTGCTGAGGCTCCAGTGCAATGCGACGCCAGGTGCCCGA. GGC TGTCC TAAGGGA CTGCTG	125
AY729672	CCAATTGCTGAGGCTCCAGTGCAATGCGACGCCAGGTGCCCGA. GGC TGTCC TAAGGGA CTGCTG	125
AY834214	CCAATTGCTGAGGCTCCAGTGCAATGCGACGCCAGGTGCCCGA. GGC TGTCC TAAGGGA CTGCTG	125
AY729685	CCAATTGCTGAGGCTCCAGTGCAATGCGACGCCAGGTGCCCGA. GGC TGTCC TAAGGGA CTGCTG	126
AY856091	CCAGCAGCTCGCCGACATCAGCGAGTGGTGCAAGGTGCTGTCCTCTACAGCATGTTGGACAG	188
AY856085	CCAGCAGCTCGCCGACATCAGCGAGTGGTGCAAGGTGCTGTCCTCTACAGCATGTTGGACAG	188
AY856087	CCAGCAGCTCGCCGACATCAGCGAGTGGTGCAAGGTGCTGTCCTCTACAGCATGTTGGACAG	188
AY729672	CCAGCAGCTCGCCGACATCAGCGAGTGGTGCAAGGTGCTGTCCTCTACAGCATGTTGGACAG	188
AY834214	CCAGCAGCTCGCCGACATCAGCGAGTGGTGCAAGGTGCTGTCCTCTACAGCATGTTGGACAG	188
AY729685	CCAGCAGCTCGCCGACATCAGCGAGTGGTGCAAGGTGCTGTCCTCTACAGCATGTTGGACAG	189
AY856091	CATGTA TAAGGAGCATGGCGGTCAGGAGGGACAGGCGGGACAGGTCGTTCCCAAGCTGCCG	251
AY856085	CATGTA TAAGGAGCATGGCGGTCAGGAGGGACAGGCGGGACAGGTCGTTCCCAAGCTGCCG	251
AY856087	CATGTA TAAGGAGCATGGCGGTCAGGAGGGACAGGCGGGACAGGTCGTTCCCAAGCTGCCG	251
AY729672	CATGTA TAAGGAGCATGGCGGTCAGGAGGGACAGGCGGGACAGGTCGTTCCCAAGCTGCCG	251
AY834214	CATGTA TAAGGAGCATGGCGGTCAGGAGGGACAGGCGGGACAGGTCGTTCCCAAGCTGCCG	251
AY729685	CATGTA TAAGGAGCATGGCGGTCAGGAGGGACAGGCGGGACAGGTCGTTCCCAAGCTGCCG	252
AY856091	GAGGGAAGGTGTGAAGCTGACGGCGCGAGCA TCA CAGCGGTCTG CAGGCTACC CATCGTCAT	314
AY856085	GAGGGAAGGTGTGAAGCTGACGGCGCGAGCA TCA CAGCGGTCTG CAGGCTACC CATCGTCAT	314
AY856087	GAGGGAAGGTGTGAAGCTGACGGCGCGAGCA TCA CAGCGGTCTG CAGGCTACC CATCGTCAT	314
AY729672	GAGGGAAGGTGTGAAGCTGACGGCGCGAGCA TCA CAGCGGTCTG CAGGCTACC CATCGTCAT	314
AY834214	GAGGGAAGGTGTGAAGCTGACGGCGCGAGCA TCA CAGCGGTCTG CAGGCTACC CATCGTCAT	314
AY729685	GAGGGAAGGTGTGAAGCTGACGGCGCGAGCA TCA CAGCGGTCTG CAGGCTACC CATCGTCAT	315
AY856091	TGATGCTGCCGGA GATGGAGCGTATGTC TGCAAGGATG TGGC CGCATACCCGAGCCTAG	375
AY856085	TGATGCTGCCGGA GATGGAGCGTATGTC TGCAAGGATG TGGC CGCATACCCGAGCCTAG	375
AY856087	TGATGCTGCCGGA GATGGAGCGTATGTC TGCAAGGATG TGGC CGCATACCCGAGCCTAG	375
AY729672	TGATGCTGCCGGA GATGGAGCGTATGTC TGCAAGGATG TGGC CGCATACCCGAGCCTAG	375
AY834214	TGATGCTGCCGGA GATGGAGCGTATGTC TGCAAGGATG TGGC CGCATACCCGAGCCTAG	375
AY729685	TGATGCTGCCGGA GATGGAGCGTATGTC TGCAAGGATG TGGC CGCATACCCGAGCCTAG	376

Fig. 1. The alignment of six representative nucleotide sequences of the mature protein in 24 kDa α -amylase inhibitors. AY834214 was isolated from hexaploid wheat, which had similar amino acid sequence to inhibitor 0.53 with amino acid substitution at only one position. AY729672 and AY856085 were isolated from hexaploid wheat and *Ae. tauschii*, respectively, which had the same amino acid with inhibitor 0.19. AY856087 was isolated from *T. urartu*. AY856091 was isolated from *Ae. speltoides*. There was an extra A at position 104 (dot) of AY729684 from *T. monoccoccum*, and as a result, the α -amylase inhibitor gene could not code the correct mature proteins.

3.3. Evolutionary relationships

To investigate the evolutionary relationships between the dimeric α -amylase inhibitor genes in this study and the representative full genes from the database, evolutionary distances were calculated and subjected to construct the neighbour-joining trees showing the evolutionary relationships of the dimeric α -amylase inhibitor genes. It was found that the α -amylase inhibitor genes were divided into six subgroups (Fig. 2). However, the deduced amino acid sequences were divided into four subgroups (Fig. 3). In the amino acid phylogenetic tree, 16 and 11 dimeric α -amylase inhibitors were clustered together with α -amylase inhibitor 0.19 (ExPASy database: P01085) and α -amylase inhibitor 0.53 (ExPASy database: P01084) in high bootstrap values, respectively. The remaining 9 dimeric α -amylase inhibitors were clustered into two subgroups.

In nucleotide sequences, the dimeric α -amylase inhibitor genes from the A^u and A^m genomes were not clustered into one subgroup, and they were divided into 2 subgroups, one for those from the A^u genome and one for those from the A^m genome (Fig. 2). Furthermore, the dimeric α -amylase inhibitor genes from the A^u genome were more closely related to those from D genome other than those from the A^m and S genomes.

The deduced amino acid sequences of dimeric α -amylase inhibitor from *Ae. tauschii* were of high coherence (99.73%), and 2 inhibitors (AY856085 and AY856086) had the same sequence with the α -amylase inhibitor 0.19. The genes controlling the inhibitor 0.19 were located on chromosome 3D [34]. These findings are in agreement with the presence of the 0.19 α -amylase inhibitors in *Ae. tauschii* and hexaploid wheat, and with the absence of this inhibitor in *durum* wheats, which were lack of the D genome [35].

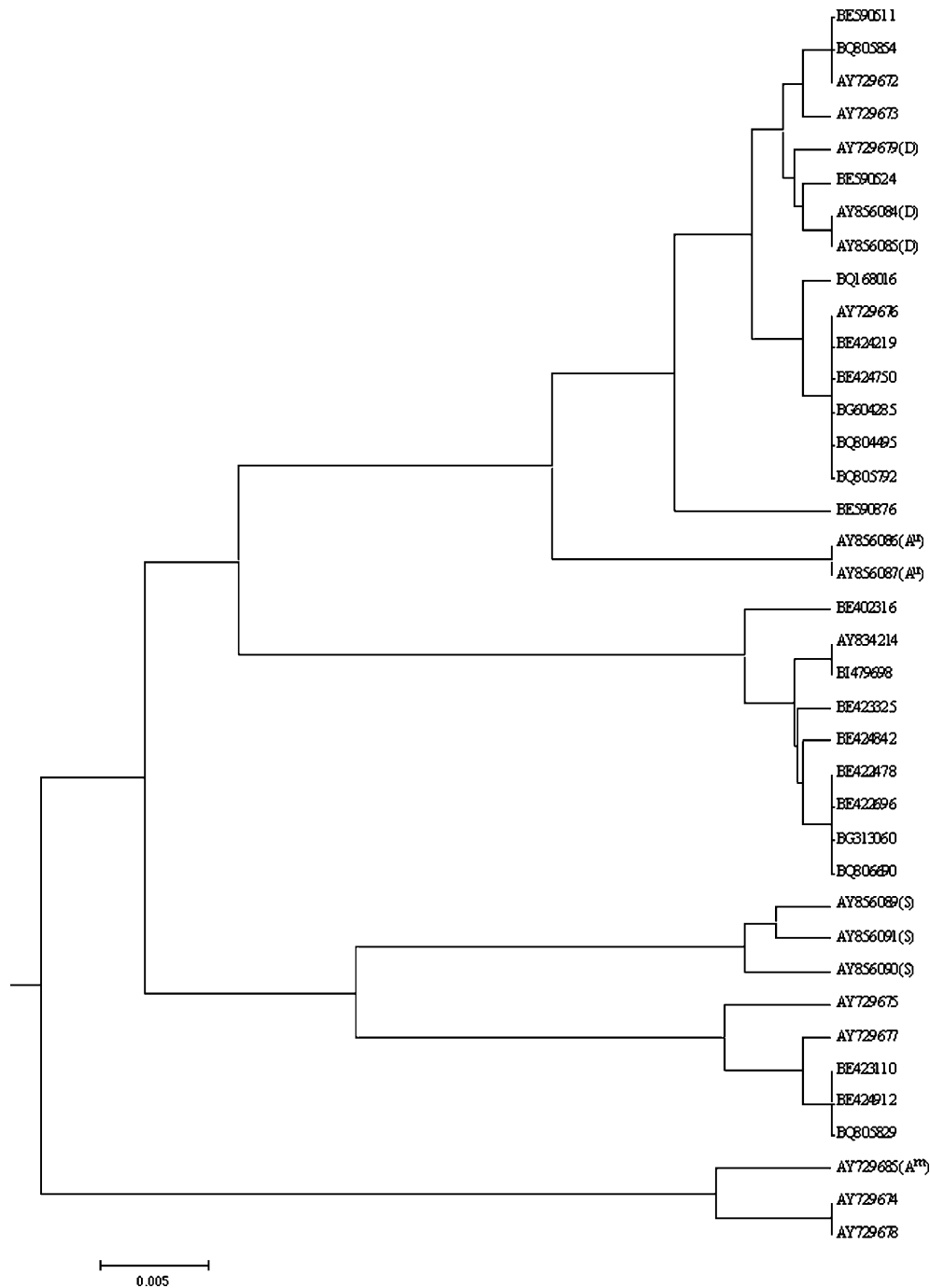


Fig. 2. Phylogenetic tree based on the alignment of proteinaceous dimeric α -amylase inhibitor genes. Genes obtained from diploid species were signed by their genome in the “()”, such as (D) for the D genome of *Ae. tauschii*, and others were obtain from hexaploid wheats.

Therefore, inhibitor 0.19 has been used as a marker for the presence of bread wheat (hexaploid) in *durum* wheat (tetraploid) food products [35].

Eight different α -amylase inhibitor genes from one hexaploid wheat variety Chuannong16 were clustered into four subgroups with three genes close to those from *Ae. tauschii*, two close to those from *Ae. speltoides*, two close to those from *T. monococcum*, and one closely related to those

from *T. urartu* (Figs. 2 and 3). It suggested that the 24 kDa family proteinaceous dimeric α -amylase inhibitors were encoded by multi-gene in cultivated wheat.

3.4. cSNPs analyses

Sequence alignment was performed with the α -amylase inhibitor sequences both from the database and from our

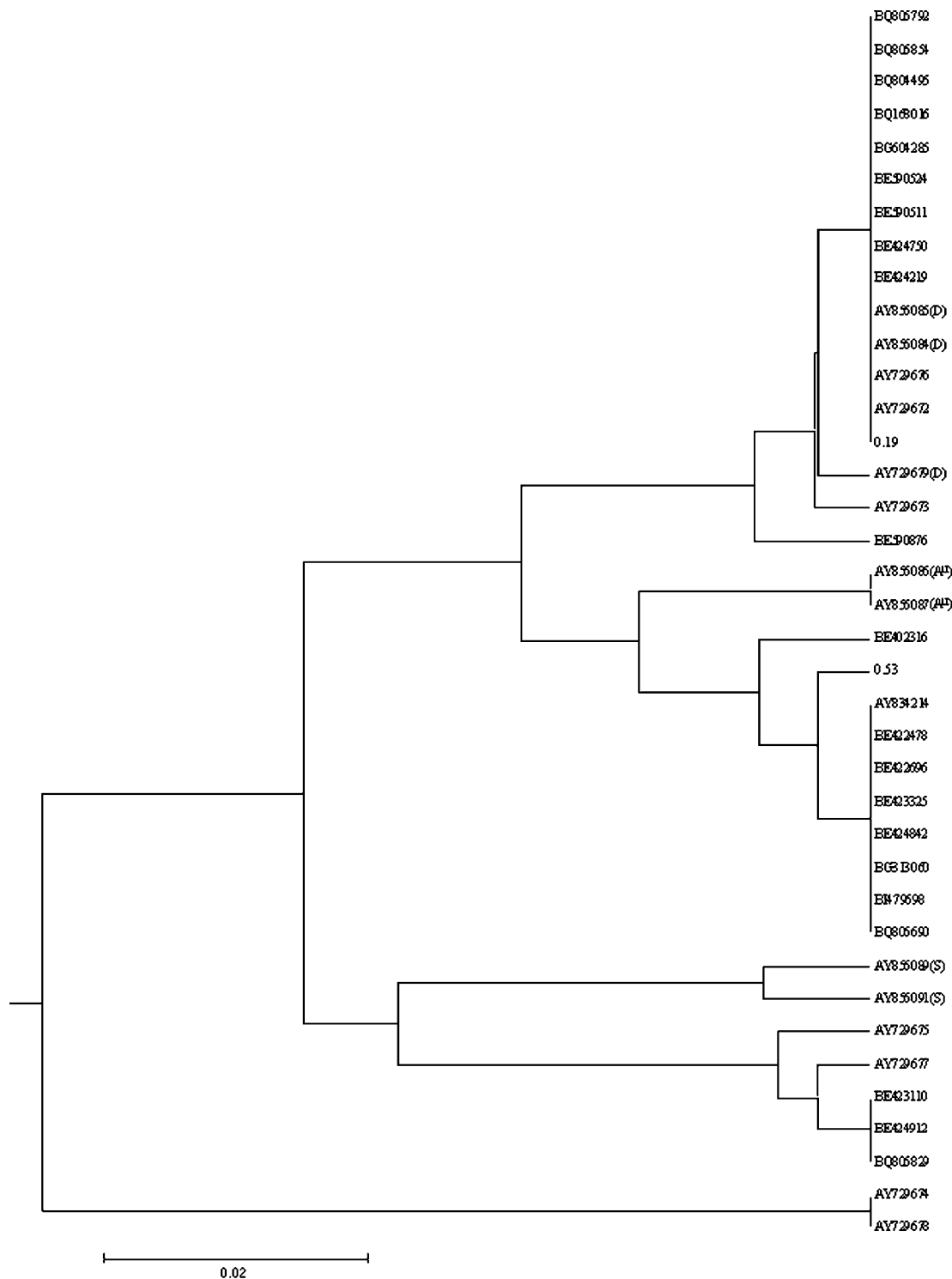


Fig. 3. Phylogenetic tree based on the alignment of proteinaceous dimeric α -amylase inhibitors from PDB and ExPASy, and the deduced amino acid sequences of α -amylase inhibitors clones in this study.

experiment to detect the candidate cSNPs. The frequency of cSNPs in the α -amylase inhibitor genes was 1 out of 9 bases, and there were 45 candidate cSNPs in the 38 α -amylase inhibitor genes (Table 2). The high frequency of cSNPs in wheat was due to the high similarity of EST sequences in the A, B and D genomes of cultivated wheat [36,37]. Similar

results had been obtained in the kunitz-type α -amylase inhibitor and α -amylase/subtilisin inhibitor genes in barley [38,39].

In human genome, cSNPs were classified into transitions (C–T and A–G) and transversions (A–C, A–T, C–G and G–T) according to their substitution types, and the transitions

Table 2

Variation of single nucleotides in the α -amylase inhibitor conserved sequences and appearing time of the nucleotide on the SNP position of 50 sequences

Position	A	G	C	T
9			19	31
24	15	35		
28			35	15
30		35		15
36			42	8
37	7		43	
42		36		14
45	7		43	
54			4	46
55	3	47		
56		27	23	
74	27	23		
104	6	44		
109			3	47
113		8	42	
121		2	48	
142			38	12
147			37	13
156			2	48
162			44	6
164	8	42		
165			37	13
172			6	44
175	8	42		
185			30	20
210			31	19
212			38	12
213	38		12	
226	23	27		
235	31	19		
238	12	38		
245	19	31		
256	3	47		
278	3	47		
289	23	27		
295			38	12
300	15	35		
314	15	35		
330	42	8		
332	8	42		
341	8	42		
351	6	44		
359	8	42		
366		8	42	
367	23	27		
370			37	13

were more common than the transversions among the sequence changes [40]. Since, 41% of the substitutions were found to occur at a CpG site, a dinucleotide known for its high mutability [41]. In the cerealsDB (http://www.cerealsdb.uk.net/wheat_snips), 47,874 putative SNPs had been found in 397,673 EST sequences, among which 26,013 SNPs were transitions, 14,291 SNPs were transversions and 8004 SNPs were indels. The distribution and proportion of transitions and transversions in the cSNPs of α -amylase inhibitor genes were investigated (Table 2, Fig. 4). As expected, most of the cSNPs (78%) in dimeric α -amylase

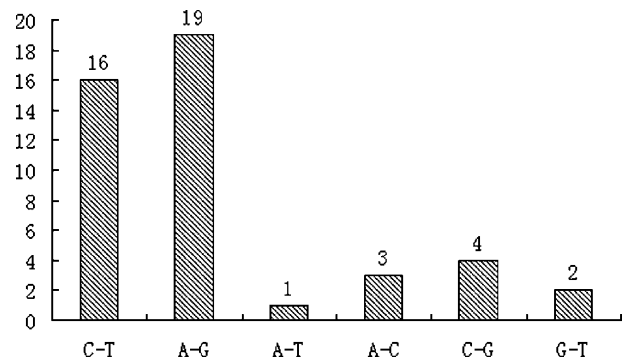


Fig. 4. Distribution of cSNPs by the type of nucleotide substitution transitions (C–T; A–G) and transversions (C–A; C–G; T–A; T–G).

inhibitor genes were transitions, and the rest (22%) were transversions.

3.5. Amino acid changes resulted from cSNPs

The relationship between cSNPs and the amino acid changes in α -amylase inhibitors was summarized in Table 3. It was clear that the 45 cSNPs resulted in ten amino acid variations. One potential way to obtain information about the functional constraint is via evolutionary comparisons, as it is generally believed that conserved residues are more likely to be functionally significant than nonconserved residues [7]. The position of each cSNP in the sequence, whether the predicted change was synonymous (silent) or nonsynonymous (replacement), was determined. Thirty-nine percent of the cSNPs were found to occur at the third codon position, and as expected, half of these was synonymous (Fig. 5). A number of changes were also identified in codon positions 1 and 2 (Fig. 5), and these accounted for most of the nonsynonymous changes. In all,

Table 3

The variation of amino acids caused by the nucleotide changes in the encoding regions

Number	Position ^a	Substitution	Amino acid	Frequency ^b
1	10	TAT-CAG ^c	Try-Gln	0.70
2 ^d	13	AAG-CAG	Lys-Gln	0.86
3	19	GGC-GCC	Gly-Ala	0.54
4 ^c	23	CTG-GTG-TTG	Val-Leu	
5	25	AAG-AGG	Lys-Arg	0.54
6 ^{d,e}	47	CCA-CGA-CAA	Asp-Asn-His	
7	70	CGT-CGC	Val-Ala	0.62
8	71	GCA-GTC ^c	Gln-Ser	0.76
9	82	AAG-ACG	Ser-Arg	0.62
10	100	CAG-CAA	Lys-Arg	0.70
11 ^d	105	CGT-CAT	Ile-Val	0.70
12	117	GGG-GGA	Gly-Asp	0.88

^a From the first codon.

^b Frequency of the most frequent allele.

^c Linkage equilibrium between two adjacent loci, and these two cSNPs have the same frequency.

^d Variation at the most important residue positions.

^e Amino acid does not change by cSNP.

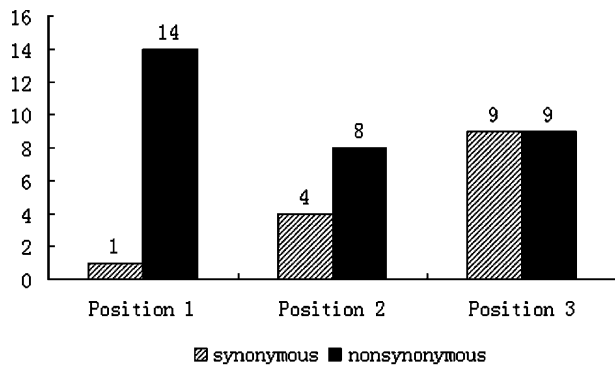


Fig. 5. Distribution of candidate cSNPs by codon position for synonymous (white) and nonsynonymous (black) amino acid substitutions.

70.1% of the candidate cSNPs resulted in nonsynonymous changes.

The amino acid sequences of dimeric α -amylase inhibitors deduced from the retrieved EST sequences and the nucleotides sequences in this study had the conserved amino acid sequences with the α -amylase inhibitors from PDB and ExPASy. Vittozzi et al. [42] reported that no detectable amylase inhibitory activity was found in the extracts from the diploid wheats by biochemical analysis. Since amylase inhibitors were absent in diploid *Triticum* species, but really present in tetraploid and hexaploid *Triticum* species, protein inhibitors of insect amylases could be considered as one of the advantages acquired by wheat through its evolution [42]. By comparison of the dimeric α -amylase inhibitor genes from diploid wheat with those from other species, one insertion with A at position 104 was found in one α -amylase inhibitor gene from *T. monococcum* accession As261, and as a result, the α -amylase inhibitor gene from *T. monococcum* could not encode the mature protein (Fig. 1). And

no PCR fragments were detected in the other 4 *T. monococcum* accessions and 3 *T. urartu* accessions.

3.6. Relationships between the structure and function of the dimeric α -amylase inhibitors and cSNPs in its genes

Nucleotide sequences affected the proteins by the determination of amino acid sequences. Sequence alignment indicated that all the α -amylase inhibitors from *T. monococcum*, *T. urartu*, *Ae. speltoideis*, *Ae. tauschii*, and hexaploid wheat and the sequences retrieved from the database were highly homologous and belong to a 24 kDa α -amylase inhibitor family including inhibitors 0.19 and 0.53. There were only a 7 amino acid difference between inhibitors 0.19 and 0.53, including Cys52, which made inhibitor 0.53 different with all other cereal type inhibitors [17,43–45].

From the structure of α -amylase inhibitors 0.19, it was known that not only the 10 Cys residues were of importance, but also Asp110, Lys116, Asn29, Glu35, Ser94, Leu90, Trp51, His47 and Gln13 were important to form the structure of inhibitors [25]. And three inhibitor spots of interest were proposed using the modeled complex of human salivary α -amylase with α -amylase inhibitor 0.19. The first was residue His47, the second concerned Ser49, and the third region of interest was the sequence Val104–Val105–Asp106–Ala107 [17]. However, only Gln13, His47 and Val105 were changed by cSNPs in some α -amylase inhibitors (Fig. 6). Most cSNPs did not occur at the conservative site, which ensures the α -amylase inhibitors keeping their activity structure to combine with the α -amylase. It was observed that the cSNP (A–C) at position 37 made the amino acid change from Gln to Lys in some inhibitors. All the dimeric α -amylase inhibitors had the Ser49, which was closely packed by Lys352 and Asp356 of

AY834214	SGPW	MCYP	GQAF	QVP	ALP	PCR	PL	LK	LQ	NGS	QVPE	AVL	RDC	CQQL	ADIS	EWCR	CGAL	YSML	D	62
AY729672	SGPW	MCYP	GQAF	QVP	ALP	PCR	PL	LK	LQ	NGS	QVPE	AVL	RDC	CQQL	ADIS	EWCR	CGAL	YSML	D	62
AY729674	SGPW	MCYP	GQAF	QVP	ALP	PCR	PL	LK	LQ	NGS	QVPE	AVL	RDC	CQQL	ADIS	EWCR	CGAL	YSML	D	62
AY729675	SGPW	MCYP	GQAF	QVP	ALP	PCR	PL	LK	LQ	NGS	QVPE	AVL	RDC	CQQL	ADIS	EWCR	CGAL	YSML	D	62
AY856085	SGPW	MCYP	GQAF	QVP	ALP	PCR	PL	LK	LQ	NGS	QVPE	AVL	RDC	CQQL	ADIS	EWCR	CGAL	YSML	D	62
AY856087	SGPW	MCYP	GQAF	QVP	ALP	PCR	PL	LK	LQ	NGS	QVPE	AVL	RDC	CQQL	ADIS	EWCR	CGAL	YSML	D	62
AY856091	SGPW	MCYP	GQAF	QVP	ALP	PCR	PL	LK	LQ	NGS	QVPE	AVL	RDC	CQQL	ADIS	EWCR	CGAL	YSML	D	62
BE423110	SGPW	MCYP	GQAF	QVP	ALP	PCR	PL	LK	LQ	NGS	QVPE	AVL	RDC	CQQL	ADIS	EWCR	CGAL	YSML	D	62
0.19	SGPW	MCYP	GQAF	QVP	ALP	PCR	PL	LK	LQ	NGS	QVPE	AVL	RDC	CQQL	ADIS	EWCR	CGAL	YSML	D	62
0.53	SGPW	MCYP	GQAF	QVP	ALP	PCR	PL	LK	LQ	NGS	QVPE	AVL	RDC	CQQL	ADIS	EWCR	CGAL	YSML	D	62
0.28	SGPW	MCYP	GQAF	QVP	ALP	PCR	PL	LK	LQ	NGS	QVPE	AVL	RDC	CQQL	ADIS	EWCR	CGAL	YSML	D	64
AY834214	SMYKEH	GVSE	EGQAGT	GAPPS	CRREVV	KLTAAS	ITAV	CKL	P	IV	VDAS	SGDG	AYVCK	DVAAY	PDPA					124
AY729672	SMYKEH	GVSE	EGQAGT	GAPPS	CRREVV	KLTAAS	ITAV	CKL	P	IV	VDAS	SGDG	AYVCK	DVAAY	PDPA					124
AY729674	SMYKEH	GVSE	EGQAGT	GAPPS	CRREVV	KLTAAS	ITAV	CKL	P	IV	VDAS	SGDG	AYVCK	DVAAY	PDPA					124
AY729675	SMYKEH	GVSE	EGQAGT	GAPPS	CRREVV	KLTAAS	ITAV	CKL	P	IV	VDAS	SGDG	AYVCK	DVAAY	PDPA					124
AY856085	SMYKEH	GVSE	EGQAGT	GAPPS	CRREVV	KLTAAS	ITAV	CKL	P	IV	VDAS	SGDG	AYVCK	DVAAY	PDPA					124
AY856087	SMYKEH	GVSE	EGQAGT	GAPPS	CRREVV	KLTAAS	ITAV	CKL	P	IV	VDAS	SGDG	AYVCK	DVAAY	PDPA					124
AY856091	SMYKEH	GVSE	EGQAGT	GAPPS	CRREVV	KLTAAS	ITAV	CKL	P	IV	VDAS	SGDG	AYVCK	DVAAY	PDPA					124
BE423110	SMYKEH	GVSE	EGQAGT	GAPPS	CRREVV	KLTAAS	ITAV	CKL	P	IV	VDAS	SGDG	AYVCK	DVAAY	PDPA					124
0.19	SMYKEH	GVSE	EGQAGT	GAPPS	CRREVV	KLTAAS	ITAV	CKL	P	IV	VDAS	SGDG	AYVCK	DVAAY	PDPA					124
0.53	SMYKEH	GVSE	EGQAGT	GAPPS	CRREVV	KLTAAS	ITAV	CKL	P	IV	VDAS	SGDG	AYVCK	DVAAY	PDPA					124
0.28	SMYKEH	GVSE	EGQAGT	GAPPS	CRREVV	KLTAAS	ITAV	CKL	P	IV	VDAS	SGDG	AYVCK	DVAAY	PDPA					123

Fig. 6. Alignment of the primary structures of the mature protein of the 24 kDa family α -amylase inhibitor. We chose 11 representative inhibitors to carry out the alignment. Inhibitors 0.19, 0.28 and 0.53 (SwissProt database: accession No. P01083) were retrieved from ExPASy; AY856085 was from *Ae. tauschii*, AY856091 from *Ae. speltoideis*, AY856087 from *T. urartu*, AY834214, AY729672, AY729674 and AY729675 from *T. aestivum*, and BE423110 from NCBI.

human salivary α -amylase and by Trp51 and Cys52 of its own [17]. It meant that this position was conserved to interest the structure. It was of importance that only α -amylase inhibitors closely related to inhibitor 0.19 from D genome of *Ae. tauschii* and common wheat had the His47, which was replaced by Asp or Asn in all the other inhibitors. Franco et al. [17] proposed that His47 was situated near the Glu349 of human salivary α -amylase. This could explain that all the inhibitors in this family shared very high sequence coherence but show dramatically different abilities to inhibit human salivary α -amylase.

The electron density for residues 1–4 and 69–77 are absent in the four subunits of α -amylase inhibitors 0.19, probably because of the intrinsic flexibility of these segments [25]. Thus, we conducted the structure predictions without these two absent segments using five methods (i.e. GORI, Gibrat, DPM, Homologue and Predato). Finally, the method of Gibrat was chosen in the further investigation because the results from this method was more close to the 3D structure in X-ray analysis by Oda et al. [25]. Ten amino acid sequences of α -amylase inhibitors, which represented all the 24 kDa α -amylase inhibitors (Fig. 6), were selected in the structure prediction. In a phylogenetic tree based on amino acid sequences, α -amylase inhibitor 0.19, AY729672 (a sequence from *T. aestivum*) and AY856085 (a sequence from *Ae. tauschii*) were clustered into one subgroup. Moreover, AY856085 had the same amino acid sequence with α -amylase inhibitor 0.19. Only the α -amylase inhibitor 0.19, AY856085 and AY729672 had four major α -helices, two short antiparallel β -strands and one small α -helix, while the rest seven inhibitors had only four α -helices. In spite of the high similarity of the main secondary structures among the ten inhibitors, the variations were found in the proportions (percentages) of the α -helix, β -sheet and coil structures (Fig. 7). The differences between the proportions (percentages) of the α -helix, β -sheet and coil structures

were detected by u -test. For α -helix and coil structures, the u values between the highest and lowest proportions were 1.80 and 1.48, respectively. They were not significant ($u_{0.05}=1.96$). It indicated that there was no difference between the proportions of the α -helix and coil structures. For the β -sheet structures, the u values between AY834214 and AY856091, and AY834214 and BE423110 were significant ($u=2.08>u_{0.05}=1.96$), indicating that the proportion of β -sheet structures in AY834214 were quite different with those of AY856091 and BE423110. Although there was only one amino acid difference at position 52 between inhibitor 0.53 and AY834214 (Fig. 6), the predicted secondary structures by both ANTHEPROT and PredictProtein server showed that there were 1, 5 and 6 variations in the α -helix, β -sheet and coil structures, respectively. These results suggested that the cSNPs in the nucleotide sequence had affected the structure of the 24 kDa α -amylase inhibitor family by changing the amino acid sequences.

Each inhibitor of the 24 kDa α -amylase inhibitor family consisted of four same subunits [25]. Therefore, one amino acid change in one subunit would result in a 4 times change in the inhibitor. In this study, each subunit in a different subgroup of this α -amylase inhibitor family had more than one different amino acid. AY729674 and AY729678 were most different from all the other inhibitors, and they had 14 mutative amino acids with both inhibitors 0.19 and 0.53.

4. Discussion

The presence of a high homology between 24 kDa family proteinaceous dimeric α -amylase inhibitors in cultivated wheat and the putative donors of A, B and D genomes that the dimeric α -amylase inhibitors might

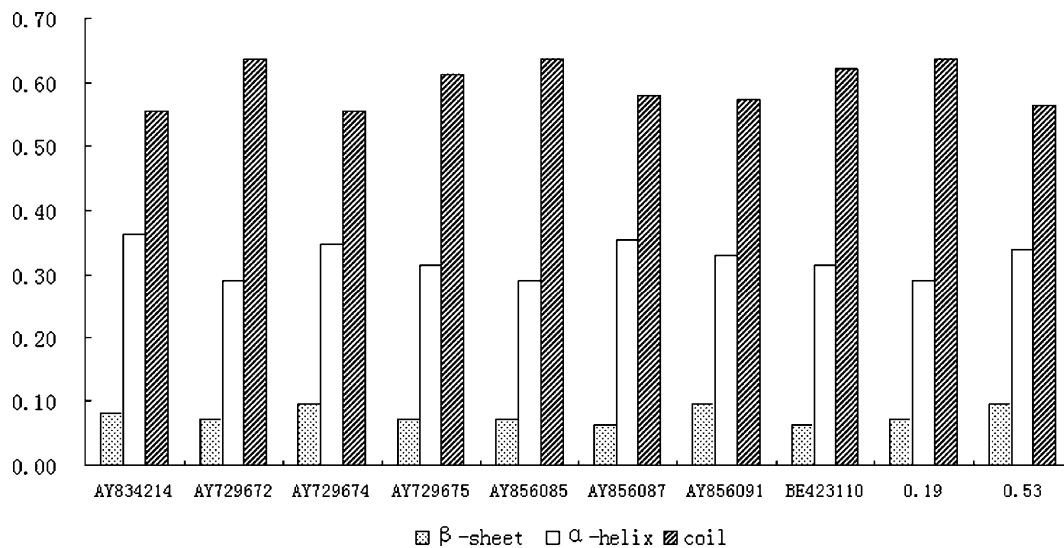


Fig. 7. The secondary structure of the 24 kDa family α -amylase inhibitors. The proportion of the α -helix, β -sheet and coil structures of α -amylase inhibitors was obtained by ANTHEPROT 5.0.

derive from a very limited number of ancestral genes coding for peptide units with molecular weight 12 kDa and the distribution of the inhibitor fractions in diploid *Aegilops* species indicated that there was a duplication of these ancestral genes followed by a divergence of the duplicated genes through mutation [20,22]. Furthermore, structure modeling sheds light on inhibitor specificity in this family that might be used for future inhibitor design in cases for which no inhibitors were available.

In this study, the nucleotide sequences of the dimeric α -amylase inhibitors from diploid wheat and *Aegilops* accessions were obviously clustered into different subgroups with other sequences from cultivated wheat in high bootstrap values (Fig. 2). However, there was no dimeric α -amylase inhibitor activity detected in diploid wheat (i.e. *T. monococcum* and *T. urartu*) [35,42]. There were three possible reasons. Firstly, no PCR amplification products were detected in 70% of the diploid wheat accessions in this study. The appearance or disappearance of an inhibitor might be the result of any mutation capable of affecting the polymerization properties of the basic peptide units of the α -amylase inhibitors [20,22]. Second, some sequences were 376 bps, which could not be deduced in correct amino acid sequences. Third, α -amylase inhibitors from some diploid wheat had Ile105 instead of Val105. The sequence Val104-Val105-Asp106-Ala107 was one of the regions of interests the α -amylase inhibitor spots. This mutation from Val105 to Ile105 might result in no α -amylase inhibitor activity in diploid wheat due to that the sequence Val104-Val105-Asp106-Ala107 occupied a cavity on the surface of human salivary α -amylase with side chains forming, largely hydrophobic, interactions with the enzyme [17].

In barley, most of SNP polymorphisms (71%) were transitions, and the remaining (29%) were transversions [46]. In this study, the similar proportion of SNP was found in wheat 24 kDa family α -amylase inhibitors. It was thought that a nonrandom nucleotide substitution mutation was the result of 5-methylcytosine deamination reactions that are known to occur at the CpG islands [41,47]. It was found that the cSNPs in dimeric α -amylase inhibitors genes resulted in ten amino acid variations. One potentiality to obtain the information about the functional constraint is via evolutionary comparisons, as it is generally believed that conserved residues are more likely to be functionally significant than are nonconserved residues [7]. In this study, it was observed that 70.1% of the candidate cSNPs of dimeric α -amylase inhibitors genes resulted in nonsynonymous changes as expected.

The relationships between the cSNPs of α -amylase inhibitor genes and the predicted structure of the α -amylase inhibitors in *Ae. tauschii*, *Ae. speltoides*, *T. urartu*, *T. monococcum* and *T. aestivum* were undertaken in this study. It was obvious that the structural variations of the α -amylase inhibitors were determined by cSNPs in their genes, which encompass proteins with mainly alpha, beta and mixed folds

in its secondary structures. The twelve (ten were determined by cSNP) amino acid changes in this α -amylase inhibitor family might affect the activities of the inhibitors to the α -amylases from various sources. Different α -amylase inhibitors exhibit different specificities against α -amylases from diverse sources. Some of these inhibitors specially inhibit insect enzymes, in that they inhibit α -amylase from insects strongly but inhibit mammalian α -amylase only weakly or not at all [48]. Thus, these α -amylase inhibitors could distinguish the α -amylases from various sources. In some cases, the α -amylase inhibitors act only against mammalian α -amylases or, on the contrary, just against insect α -amylases. In the latter case, α -amylase inhibitor provides a highly specific potential weapon in plant defense [17]. Therefore, the determination of the specific inhibition of a dimeric α -amylase inhibitor is the important first step towards the discovery of an inhibitor, which could be useful for generating insect-resistant transgenic plants.

The different structure of inhibitors affected the specificity and activities against different mammalian and insect α -amylase [49,50]. In addition, different amino acid residues determined by cSNPs significantly affected the charge between inhibitors 0.19 and 0.53 [45]. The net charge of one subunit of the inhibitor 0.19 was increased by three basic charge units with seven amino acid substitutions meaning that six more basic charge units increase per dimeric form protein [45]. This charge difference might result in the relative mobility of gel electrophoresis and the differential inhibitory activities of inhibitors 0.19 and 0.53, even 0.27, 0.36 and 0.38.

All the cereal type α -amylase inhibitors had 10 Cys [43]. In this study, the deduced proteins of the inhibitors in the 24 kDa α -amylase inhibitor family also had 10 Cys. However, there was 9 Cys in inhibitor 0.53, which was sequenced by using CNBr- and trypsin-peptides of carboxymethyl (CM)-protein (Fig. 6) [44]. There was a 7 amino acid difference between inhibitors 0.19 and 0.53. At position 52 of dimeric α -amylase inhibitors, only inhibitor 0.53 was Pro instead of Cys (Fig. 6). Cys52 in inhibitor 0.19 might form a disulfide bridge with Cys54, which is a free Cys residue in inhibitor 0.53. This difference was the main reason that inhibitor 0.19 has 5 times lower selective inhibitory activity towards human salivary α -amylase than inhibitor 0.53 did, and it was found that inhibitor 0.19 has higher specific inhibitory activity [45]. Eleven inhibitors, which were deduced from EST sequences from the database and the cloning sequences in this study, had the same amino acid sequences, and they had only one different amino acid (Cys52) with inhibitor 0.53. This result was not in agreement with Maeda et al. [44]. It was difficult to further validate the nucleotide sequence of inhibitor 0.53 due to that the wheat accessions in the characterization of the amino acid sequence of inhibitor 0.53 had not been described in detail by Maeda et al. [44]. Moreover, Maeda et al. [44] thought that several amino acids were ambiguously identified in the region of residues 43 to 60.

The crystal structure of α -amylase inhibitor 0.19 was determined by the multiple-isomorphous replacement method coupled with density modification and noncrystallographic symmetry averaging and then refined by simulated annealing using diffraction data to 2.06 Å resolutions [25]. The asymmetric unit has four molecules (A–D) of inhibitor 0.19, each composed of 124 amino acid residues. All 10 Cys residues in this family were highly conserved, which would form disulfide bonds with the pairs of Cys6–Cys52, Cys20–Cys41, Cys28–Cys83, Cys42–Cys99, and Cys54–Cys115 as α -amylase inhibitor 0.19 (Fig. 6). Because the disulfide bonds were a main factor to affect the secondary structure of proteins, we could believe that the α -amylase inhibitors in this family had the similar primary structure. Since the motilities and the activities of some inhibitors were of diversity, it suggested that the 3D structures might have much difference. It was obvious that α -amylase inhibitor 0.19 had a high α -helix content, by CD spectroscopy analysis, and five α -helix domains were found in the structure of α -amylase inhibitor 0.19 [51].

To evaluate the differences of 24 kDa α -amylase inhibitors determined by the cSNPs in their genes, it is necessary to express the representative genes in each subgroup. A synthetic gene encoding α -amylase inhibitor 0.19 was introduced into vector for expression in *Escherichia coli* [52]. In SDS-PAGE and Western blotting analyses of the *E. coli* cell lysate, the expression product could not be detected. Okuda et al. [52] proposed that the putative hairpin-like structure of mRNA in the 5'-terminal coding region of α -amylase inhibitor 0.19 might interrupt efficient expression, when the hairpin structure was eliminated, the resulting gene could be overexpressed in *E. coli*.

The α -amylase inhibitors have long been proposed as possibly important weapons against pests whose diets make them highly dependent on α -amylase activity. In vitro and in vivo trials, including those made under field conditions, have now fully confirmed the potential of raising the possibility and significantly increased yields. Thus, it is of importance to identify the specificity of α -amylase inhibitor. The inhibition spectra of inhibitors so far characterized are remarkably variable and include some with the desirable combination of activity against insect enzymes and inactivity towards mammalian enzymes [50]. Several different classes of plant proteins have been shown to be insecticidal towards a range of economically important insect pests when tested in artificial diets or transgenic plants [15]. Possible environmental side effects of the transgenic technology will thereby be minimized, helping to achieve consumer acceptance for the transgenic crop.

Acknowledgements

We wish to thank the two anonymous reviewers for helpful comments that improved the manuscript. This work was supported by the National High Technology Research

and Development Program of China (863 program 2003AA207100) and the FANEDD project (200357) from Ministry of Education, China. Y.-L. Zheng was supported by the Program for Changjiang Scholars and Innovative Research Team in University.

References

- [1] D.G. Wang, J.B. Fan, C.J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, T. Topaloglou, E. Hubble, E. Robinson, M. Mittmann, M.S. Morris, N. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T.J. Hudson, E.S. Lander, Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome, *Science* 280 (1998) 1077–1082.
- [2] A. Rafalski, Novel genetic mapping tools in plants: SNPs and LD-based approaches, *Plant Sci.* 162 (2002) 329–333.
- [3] D.L. Remington, J.M. Thornsberry, Y. Matsuoka, L.M. Wilson, S.R. Whitt, J. Doebley, S. Kresovich, M.M. Goodman, E.S. Buckler, Structure of linkage disequilibrium and phenotypic associations in the maize genome, *Proc. Natl. Acad. Sci. U. S. A.* 98 (2001) 11479–11484.
- [4] M.I. Tenaillon, M.C. Swakins, A.D. Long, R.L. Gaut, J.F. Doebley, B.S. Gaut, Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.), *Proc. Natl. Acad. Sci. U. S. A.* 98 (2001) 9161–9166.
- [5] F.S. Collins, M.S. Guyer, A. Chakravarti, Variations on a theme: cataloging human DNA sequence variations, *Science* 278 (1997) 1580–1581.
- [6] R.M. Harding, S.M. Fullerton, R.C. Griffiths, J. Bond, M.J. Cox, J.A. Schneider, D.S. Moulin, J.B. Clegg, Archaic African and Asian lineages in the genetic ancestry of modern humans, *Am. J. Hum. Genet.* 60 (1997) 772–789.
- [7] K. Garg, P. Green, D.A. Nickerson, Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags, *Genome Res.* 9 (1999) 1087–1092.
- [8] A. Rafalski, Applications of single nucleotide polymorphisms in crop genetics, *Plant Biol.* 5 (2002) 94–100.
- [9] E.A. MacGregor, S. Janecek, B. Svensson, Relationship of sequence and structure to specificity in the α -amylase family of enzymes, *Biochim. Biophys. Acta* 1546 (2001) 1–20.
- [10] M.C. Silva, M. Da, M.F. Grossi, S. De, M.J. Chrispeels, R.C. Togawa, G. Neshich, Analysis of structural and physico-chemical parameters involved in the specificity of binding between α -amylases and their inhibitors, *Protein Eng.* 13 (2000) 167–177.
- [11] C.A. Ryan, G. Pearce, Systemin: a polypeptide signal for plant defensive genes, *Annu. Rev. Cell Dev. Biol.* 14 (1998) 1–17.
- [12] C.A. Ryan, Protease inhibitors in plants: genes for improving defenses against insects and pathogens, *Annu. Rev. Phytopathol.* 28 (1990) 425–449.
- [13] M.P. Sales, I.R. Gerhardt, M.F. Grossi-de-Sá, J. Xavier-Filho, Do legumes storage proteins play a role in defending seeds against bruchids? *Plant Physiol.* 124 (2000) 515–522.
- [14] M.J. Chrispeels, M.F. Grossi-de-Sá, T.J.V. Higgins, Genetic engineering with α -amylase inhibitors seeds resistant to bruchids, *Seed Sci. Res.* 8 (1998) 257–263.
- [15] A.M.R. Gatehouse, J.A. Gatehouse, Identifying proteins with insecticidal activity: use of encoding genes to produce insect-resistant transgenic crops, *Pest. Sci.* 52 (1998) 165–175.
- [16] A.V. Konarev, Interaction of insect digestive enzymes with plant protein inhibitors and host-parasite co-evolution, *Euphytica* 92 (1996) 89–94.
- [17] O.L. Franco, D.J. Rigden, F.R. Melo, J.C. Bloch, C.P. Silva, M.F. Grossi de Sá, Activity of wheat α -amylase inhibitors towards bruchid

- α -amylases and structural explanation of observed specificities, Eur. J. Biochem. 267 (8) (2000) 1466–1473.
- [18] R. Deponte, R. Parlamenti, V. Petrucci, V. Silano, M. Tomasi, Albumin α -amylase inhibitor families from wheat flour, Cereal Chem. 53 (1976) 805–820.
 - [19] J. Mundy, J. Hejgaard, I. Svendsen, Characterization of a bifunctional wheat inhibitor of endogenous (α -amylase and subtilisin, FEBS Lett. 167 (1984) 210–214.
 - [20] V. Buonocore, T. Petrucci, V. Silano, Wheat protein inhibitors of α -amylase, Phytochemistry 16 (1977) 811–820.
 - [21] J.E. Baker, Interaction of partially-purified amylases from larval anagasta kuehniella (lepidoptera: pyralidae) with amylase inhibitors from wheat, Comp. Biochem. Physiol. 9311 (2) (1989) 239–246.
 - [22] V. Buonocore, M.G.D. Biasi, P. Giardina, E. Poreio, V. Silano, Purification and properties of an α -amylase tetrameric inhibitor from wheat kernel, Biochim. Biophys. Acta 831 (1985) 40–48.
 - [23] L. Carrano, G. Nitti, V. Buonocore, C. Caporale, E. Poerio, An effective purification procedure of amylase and trypsin inhibitors from wheat flour; isolation of a new water-soluble, Plant Sci. 65 (1989) 25–31.
 - [24] S. Strobl, P. Muhlhahn, R. Bernstein, R. Wilschek, K. Maskos, M. Wunderlich, R. Huber, R. Glockshuber, T.A. Holak, Determination of the three-dimensional structure of the bifunctional α -amylase/trypsin inhibitor from ragi seeds by NMR spectroscopy, Biochemistry 34 (26) (1995) 8281–8293.
 - [25] Y. Oda, T. Matsunaga, K. Fukuyama, T. Miyazaki, T. Morimoto, Tertiary and quaternary structures of 0.19 α -amylase inhibitor from wheat kernel determined by X-ray analysis at 2.06 Å resolution, Biochemistry 36 (44) (1997) 13503–13511.
 - [26] N.K. Blake, B.R. Leffeldt, A. Hemphill, X. Shan, L.E. Talbert, DNA sequence analysis suggests a monophyletic origin of the wheat B genome, Proc. 9th Wheat Genet. Symp. 2 (1998) 14–16.
 - [27] J. Dvorak, P. Terlizzi, H.B. Zhang, P. Resta, The evolution of polyploid wheats: identification of the A genome donor species, Genome 36 (1993) 21–31.
 - [28] O. Raskina, A. Belyyev, E. Nevo, Repetitive DNAs of wide emmer wheat *Triticum dicoccoides* and their relation to S-genome species: molecular-cytogenetic analysis, Genome 45 (2) (2002) 391–401.
 - [29] Z. Lianquan, L. Dengcai, Y. Zehong, L. Xiuji, Z. Youliang, Z. Yonghong, Rapid changes of microsatellite flanking sequence in the allopolyploidization of new synthesized hexaploid wheat, Sci. China, Ser. C, Life Sci. 47 (6) (2004) 553–561.
 - [30] Z. Yan, Y. Wan, K. Liu, Y.L. Zheng, D.W. Wang, Identification of a novel HMW glutenin subunit and comparison of its amino acid sequence with those of homologous subunits, Chin. Sci. Bull. 47 (3) (2002) 220–225.
 - [31] J.R. Wang, Z.H. Yan, Y.M. Wei, Y.L. Zheng, A novel HMW glutenin subunit gene Eel.5 from *Elytrigia elongata* (Host) Nevski, J. Cereal Sci. 40 (2004) 289–294.
 - [32] J.F. Gibrat, J. Garnier, B. Robson, Further developments of protein secondary structure prediction using information theory. New parameters and consideration of residue pairs, J. Mol. Biol. 198 (1987) 425–443.
 - [33] G. Deléage, C. Combet, C. Blanchet, C. Geourjon, An integrated protein sequence analysis software with client/server capabilities, Comput. Biol. Med. 31 (4) (2001) 259–267.
 - [34] C. Aragoncillo, L. Rodriguez, P. Carbonero, F. Garcia-Olmedo, Chromosomal control of non-gliadin proteins from 70% ethanol extract of wheat endosperm, Theor. Appl. Genet. 45 (1975) 322–326.
 - [35] C. Bedetti, A. Bozzini, V. Silano, V. Luciano, Amylase protein inhibitors and the role of *Aegilops* species in polyploid wheat speciation, Biochim. Biophys. Acta 362 (2) (1974) 299–307.
 - [36] G.T. Marth, I. Korf, M.D. Yandell, R.T. Yeh, Z. Gu, H. Zakeri, N.O. Stizel, L. Hillier, P.Y. Kwok, W.R. Gish, A general approach to single-nucleotide polymorphism discovery, Nat. Genet. 23 (1999) 452–456.
 - [37] L. Picoult-Newberg, T.E. Ideker, M.G. Pohl, S.L. Taylor, M.A. Donaldson, D.A. Nickerson, M. Boyce-Jacino, Mining SNPs from EST databases, Genome Res. 9 (1999) 167–174.
 - [38] V.G. Battershell, R.J. Henry, High performance liquid chromatography of α -amylases from germinating wheat and complexes with the α -amylase inhibitor from barley, J. Cereal Sci. 12 (1990) 73–81.
 - [39] S.J. Jarrett, R.J. Marschke, M.H. Symons, C.E. Gibson, G.P. Fox, R.J. Henry, α -amylase/subtilisin inhibitor levels in Australian barleys, J. Cereal Sci. 25 (1997) 261–266.
 - [40] D.N. Cooper, M. Krawczak, The mutational spectrum of single base-pair substitutions causing human genetic disease: patterns and predictions, Hum. Genet. 85 (1990) 55–74.
 - [41] D.N. Cooper, H. Youssoufian, The CpG dinucleotide and human genetic disease, Hum. Genet. 78 (1988) 151–155.
 - [42] L. Vittozzi, G. Morisi, V. Silano, Continuous automated assay of α -amylase inhibitors, J. Sci. Food Agric. 27 (5) (1976) 449–453.
 - [43] E. Poerio, C. Caporale, L. Carrano, P. Pucci, V. Buonocore, Assignment of the five disulphide bridges in an α -amylase inhibitor from wheat kernel by FAB/MS and Edman degradation, Eur. J. Biochem. (1991) 199595–199600.
 - [44] K. Maeda, The complete amino-acid sequence of the endogenous α -amylase inhibitor in wheat, Biochim. Biophys. Acta 871 (1983) 250–256.
 - [45] K. Maeda, S. Kakabayashi, H. Matubara, Complete amino acid sequence of an α -amylase inhibitor in wheat kernel (0.19-inhibitor), Biochim. Biophys. Acta 828 (1985) 213–221.
 - [46] V.D. Solemimani, B.R. Baum, D.A. Johnson, Efficient validation of single nucleotide polymorphisms in plants by allele-specific PCR, with an example from barley, Plant. Mol. Biol. Report. 21 (2003) 281–288.
 - [47] R. Holliday, W.G. Girgg, DNA methylation mutation, Mutat. Res. 285 (1993) 61–67.
 - [48] F. Payan, Structural basis for the inhibition of mammalian and insect α -amylase by plant inhibitors, Biochim. Biophys. Acta 1696 (2004) 171–180.
 - [49] G.H. Feng, M. Richardson, M.S. Chen, K.J. Kramer, T.D. Morgan, G.R. Reeck, α -Amylase inhibitors from wheat: a sequences and patterns of inhibition of insect and human α -amylases, Insect Biochem. Mol. Biol. 26 (1996) 419–426.
 - [50] O.L. Franco, D.J. Rigden, F.R. Melo, M.F. Grossi-de-Sá, Plant α -amylase inhibitors and their interaction with insect α -amylases. Structure, function and potential for crop protection, Eur. J. Biochem. 269 (2002) 397–412.
 - [51] T. Petrucci, A. Rab, M. Tomasi, V. Silano, Further characterization studies of the α -amylase protein inhibitor of gel electrophoretic mobility 0.19 from wheat kernel, Biochim. Biophys. Acta 420 (1976) 288–297.
 - [52] M. Okuda, T. Satoh, N. Sakurai, K. Shibuya, H. Kaji, T. Samejima, Overexpression in *Escherichia coli* of chemically synthesized gene for active 0.19 α -amylase inhibitor from wheat kernel, J. Biochem. 122 (5) (1997) 918–926.