

The starch-bound alpha-amylase/trypsin-inhibitors in *Avena*

Laura Gazza¹  · Gloria Gazzelloni¹ · Federica Taddei¹ · Arianna Latini^{1,4} · Vera Muccilli² · Michela Alfieri³ · Salvatore Conti¹ · Rita Redaelli³ · Norberto E. Pogna¹

Received: 21 June 2016 / Accepted: 29 July 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract Oat kernels exhibit an extra-soft texture, a trait recently demonstrated to be largely modulated by starch-bound tryptophan-rich 2S proteins, the vomindolines. In this study, fractionation by two-dimensional electrophoresis of starch-bound proteins in 25 oat (*Avena sativa*) cultivars and 11 diploid or tetraploid *Avena* species revealed novel 2S proteins called Avena α -amylase/trypsin-inhibitors (AATI) because of their sequence similarity with wheat α -amylase/trypsin inhibitors. Thirty-seven AATI polypeptides, about 14 kDa in size, were split into three families named AATI-1, AATI-2, and AATI-3 with different primary structures and isoelectric points. AATI-1 and AATI-2 proteins showed 55.5–60.0 % sequence similarity with wheat α -amylase inhibitors CM1, CM2, and CM16, which have been found to cause innate immunity responses

in celiac disease and non-celiac gluten sensitivity. Diploid A-genome and tetraploid AC-genome oat species possess three and five genes encoding for the AATI proteins, respectively, whereas hexaploid *A. sativa* exhibits 12 genes dispersed over the A-, C-, and D-genomes. Some AATI proteins expressed in hexaploid oats were assigned to the A-genome based on similarity to their counterparts in diploid species, contributing to further clarify the genetic origin of hexaploid oats. Moreover, AATI may interact with starch-bound vomindolines in determining the extra-soft texture of oat kernels and, due to their balanced amino acid compositions, may contribute to the biological value of oat proteins in a positive manner.

Keywords *Avena* · α -Amylase/trypsin-inhibitors · Kernel texture · Oat immunogenicity · Starch-bound proteins

Communicated by S. Hohmann.

N. E. Pogna Deceased 22 June 2016.

Electronic supplementary material The online version of this article (doi:10.1007/s00438-016-1238-4) contains supplementary material, which is available to authorized users.

✉ Laura Gazza
laura.gazza@crea.gov.it

¹ Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria (CREA-QCE), Via Manziana 30, 00189 Rome, Italy

² Department of Chemical Sciences, University of Catania, Via Andrea Doria 6, Catania, Italy

³ Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria (CREA-MAC), Via Stezzano 24, Bergamo, Italy

⁴ Present Address: ENEA Research Center Casaccia, Via Anguillarese, 301, 00123 Rome, Italy

Introduction

Amongst the cereal crops, oats (*Avena sativa* L.) are characterized by kernels with one-third more protein (15–20 %), four times more lipids (5–9 %), and relatively high amounts of fibers (12–14 %), including about 5 % of β -glucan (Peterson 1992). Oat groats also contain avenanthramides, low-molecular-weight soluble phenolic compounds not present in other cereal species (Peterson 2001; Meydani 2009). Because of their unique kernel composition, oats are currently promoted as a regular component of the human diet with a broad range of health benefits. Indeed, most studies indicate that the consumption of oat meal and oat bran can decrease total plasma cholesterol and LDL-cholesterol levels, suppresses inflammation, relaxes arteries, and reduces atherosclerosis and colon cancer risk (Singh et al. 2013; Wani et al. 2014). In addition, several studies claim that

pure oats are safe for most celiac patients (Fric et al. 2011; Comino et al. 2015; La Vieille et al. 2016), and the current European regulations include oats in the list of gluten-free ingredients. Another unique trait of *A. sativa* is the extra-soft texture of its kernels. When analyzed by the single kernel characterization system (SKCS), oat groats exhibit negative values ranging from −37 to −14 (Gazza et al. 2015), much lower than those of durum wheat (>70), common wheat (from 10 to 70), rye and barley (about 70) (Galassi et al. 2012), or einkorn wheat (from −7 to 5) (Taddei et al. 2009). Grain softness of cereal kernels decreases milling costs and increases flour yield (Tsilo et al. 2011). In addition, upon milling, particle size of flour prepared from soft wheat kernels is remarkably small compared to that of hard-textured grains, resulting in enhanced digestibility (Knowlton et al. 1996; Wondra et al. 1995). On the other hand, the extra-soft oat groats are prone to breakage during mechanical harvesting, threshing, and dehulling, which negatively affects the size and quality of oat flakes. Recently, grain texture in *A. sativa* has been shown to be modulated by accumulation on the starch granules of tryptophan-rich proteins named vromindolines (Gazza et al. 2015). Upon fractionation of starch-bound proteins by two-dimensional A-PAGE × SDS-PAGE electrophoresis, vromindolines were found to occur in the 13–14 kDa region of the gel as three spots flanked by a cluster of three novel polypeptides never described before and provisionally called Avena- α -amylase-trypsin inhibitors (AATI). Here, AATI proteins and their encoding genes are investigated in 25 hexaploid oat cultivars and 11 diploid or tetraploid *Avena* species. The potential contribution of AATI polypeptides in determining the nutritional value, the extra-soft texture of oat groats and the genetic relationships of *Avena* species are also discussed.

Materials and methods

Plant material

Kernels and plants of 25 oat (*A. sativa*) cultivars (Abel, Adam, Alcudia, Astra, BD134, Bikini, Bullion, Cornell, Donata, Flavia, Genziana, Grafton, Harpoon, Hamel, Hinoat, Icon, Konradin, Krypton, Kynon, Lisbeth, Primula, TeoBd 40, Terra, Tropicale, and Vir2301) were analyzed in the present study. Plants of wild diploid (*A. canariensis*, *A. clauda*, *A. damascena*, *A. longiglumis*, *A. strigosa*, and *A. ventricosa*) or tetraploid oat species (*A. barbata*, *A. insularis*, *A. magna*, *A. maroccana*, and *A. murphy*), from the CREA-MAC (Bergamo, Italy) collection, were used as well (Supplementary Table 1). Seeds of oat cvs Donata and Primula were harvested at 7, 14, 21, 28, and 35 days post anthesis (DPA), immediately frozen in liquid nitrogen and stored at −80 °C until used for protein extraction.

Protein extraction, electrophoretic analysis, and Western blotting

Air-dried starch granules were obtained as described by Corona et al. (2001), and starch-bound proteins were extracted, reduced, and alkylated as described by Gazza et al. (2015). Alkylated proteins were fractionated by A-PAGE, SDS-PAGE, and two-dimensional A-PAGE × SDS-PAGE as described previously (Gazza et al. 2006). Protein extractions and electrophoretic fractionations were performed at least twice for each genotype.

Western blotting was performed as described previously by Gazza et al. (2015) in a Biorad (USA) Semi Dry Transfer Cell using nitrocellulose membranes (Hybond-C Extra, GE Healthcare, USA). The polyclonal antiserum was developed in rabbits immunized with the synthetic peptide MRKT-PASLPYEVSLR. After three washes with PBS buffer and incubation for 4 h with a 1:2500 dilution of goat anti-rabbit horseradish peroxidase conjugate (Promega, USA), blots were stained in a solution containing 50 ml PBS buffer, 10 ml methanol, 30 mg 4-chloro-1-naphthol, and 1 ml H₂O₂ (30 %). Peroxidase activity was also visualized by chemiluminescence using the Supersignal Detection Kit (Pierce, USA), following the manufacturer's instructions. Western blotting was performed at least twice for each genotype.

DNA and RNA extraction, PCR amplification, DNA cloning, and sequencing

Genomic DNA was isolated from young leaves using the cetyltrimethylammonium bromide (CTAB) method (Ausubel et al. 1994). Total RNA was extracted with TRIZOL Reagent (Invitrogen, USA) from 70 mg of seeds collected at 21 and 28 DPA and grinded in liquid nitrogen. PCR primers (10 pmol each reaction) used in this study are listed in Table 1. The samples denatured at 94 °C for 3 min before the addition of Taq (Jump Start REDAccuTaq LA DNA Polymerase, SIGMA, Italy) were submitted to 35 cycles of 1 min of denaturation at 94 °C, 1 min annealing at T_m (Table 1) and 1.5 min elongation at 72 °C. A final cycle with an extension of 7 min at 72 °C completed the reactions. PCR products were analyzed on 1.8 % agarose gels, stained with Gel Red (Aurogene, Italy), and visualized under UV. Gels were analyzed with the UVipro Bronze Gel Documentation System (Uvitec, UK), using the UVipro 12.5 software.

Amplification products were eluted from agarose gels with the Nucleospin-Extract kit (Macherey–Nagel, Germany), cloned as previously described (Gazza et al. 2006) and sequenced with the Beckman Coulter CEQ 8000 sequencer according to the manufacturer's instructions. DNA sequences were deposited in the GenBank sequence database at NCBI. DNA amplification, cloning, and sequencing were performed at least three times for each genotype.

Table 1 Primer pairs and annealing temperatures used for PCR or RACE amplifications

Name	Forward primer (5'–3')	Reverse primer (5'–3')	T_m (°C)	Amplicon size (bp)	Encoded protein
A1	ATGGCGTCCAAGTGCAATTGC	TCAGAATGTTTCTCTTTATTC	54	592	AATI-1
A1cod	ATGGCGTCCAAGTGCAATTGC	GTACTGAAATTTGTCCAAAGCTG	58	453	AATI-1
3RA2	GACTATGTGGAACAAGCG		54		
5RA2	CGCTTGTTCCACATAGTC		54		
A2	GTCATCGAACCGGTCTAG	TCTCTAGTGCCACTGAGAC	56	478/481	AATI-2
3RA3	CTGCCGTGAGTACGTGGC		60		
5RA3	CACACACATCCATGGATGC		60		
A3	CACCCAGCAGTGAAGTGC	ACACACACATCCATGGATGC	58	542/545	AATI-3

RACE (rapid amplification cDNA ends)

An aliquot (10 µg) of total RNA extracted from seeds harvested at 28 DPA was submitted to 3' and 5' RACE experiments performed with 3' and 5' RACE System kits (Invitrogen, USA) according to manufacturer's instructions. Primers used for cDNA extension are shown in Table 1. cDNA samples, denatured at 94 °C for 3 min before the addition of 2.5 U of Go Taq DNA polymerase (Promega, USA), were submitted to 35 cycles of 1 min of denaturation at 94 °C, 1 min annealing at T_m , and 1.5 min elongation at 72 °C, with the final extension of 7 min at 72 °C. RACE experiments were performed at least twice for each genotype. RACE fragments were then gel-purified, cloned, and sequenced as described above.

In-gel digestion, mass spectrometric analysis, database search, and protein identification

Selected protein spots from the two-dimensional gels were excised manually, transferred into 1.5 ml microcentrifuge tubes, washed and subjected to in-gel digestion with trypsin (Modified Porcine Trypsin, Promega) in 50 mM NH_4HCO_3 , and incubated at 37 °C overnight, as previously reported in Gazza et al. (2015). After in-gel digestion, the peptides were separated and analyzed on an Ultimate 3000 LC system (Dionex Corporation, USA) coupled online with a linear ion trap nano-electrospray mass spectrometer (LTQ, ThermoFisher Scientific, USA). Experimental parameters for chromatographic separation, mass spectrometric data acquisition, database search, and protein identification were previously reported in Galassi et al. (2012) and Gazza et al. (2015), following a bottom-up proteomic approach (Cunsolo et al. 2012).

For N-terminal sequencing, starch-bound proteins fractionated by A-PAGE were transferred onto a 0.2 µm PVDF membrane (Sequi-Blot™, BioRad, USA). Gels and membranes were pre-equilibrated for 20 min in a transfer CAPS buffer, pH 11.0, containing 10 % (v/v) methanol in distilled water, transferred between two double layers of 3 MM chromatography paper (Whatman, UK), and electro-blotted

at 50 V for 90 min in a wet apparatus Transfer Cell (BioRad, USA). After the transfer, membranes were stained for 10 min in a solution containing 50 % methanol and 16 % Coomassie Brilliant Blue R250 and destained with a solution of 50 % methanol in distilled water. N-terminal sequencing of the first ten amino acids of selected protein was performed by Primm Srl (Milan, Italy) using the Applied Biosystems Model H49 Procise® Protein Sequencer.

Results

Two-dimensional A-PAGE x SDS-PAGE fractionation of AATI proteins in *A. sativa*

When fractionated by acidic polyacrylamide gel electrophoresis at pH 3.1 (A-PAGE), proteins extracted from the starch granules appeared as two broad bands in all the 25 oat cultivars analyzed (Fig. 1a, lane 1). Each of these bands generated a cluster of three spots after separation by SDS-PAGE in the second dimension (Fig. 1b). In particular, the cluster of vromindolines VIN-1, VIN-2, and VIN-3 (Fig. 1b, circle), the tryptophan-rich proteins exerting a softening effect on kernel texture (Gazza et al. 2015) was flanked by three novel polypeptides never described before and marked AATI-1, AATI-2, and AATI-3 in Fig. 1b. AATI-1 and AATI-2 occurred as two strong spots with apparent molecular weights of 14 and 13.5 kDa, respectively, whereas AATI-3 appeared as a faint spot, 13 kDa in size (Fig. 1b). The relative intensities of AATI-1, AATI-2, and AATI-3 and their scattering in the two-dimensional gel showed a clear parallelism with those of VIN-1, VIN-2, and VIN-3.

Identification of AATI-1 proteins by mass spectrometry and western blotting

The peptides obtained by trypsin digestion of the AATI-1 spot excised from the gel of the A-PAGE x SDS-PAGE fractionation of oat cv. Donata were analyzed by RP-HPLC/n-ESI-MS/MS. Some of the identified peptides

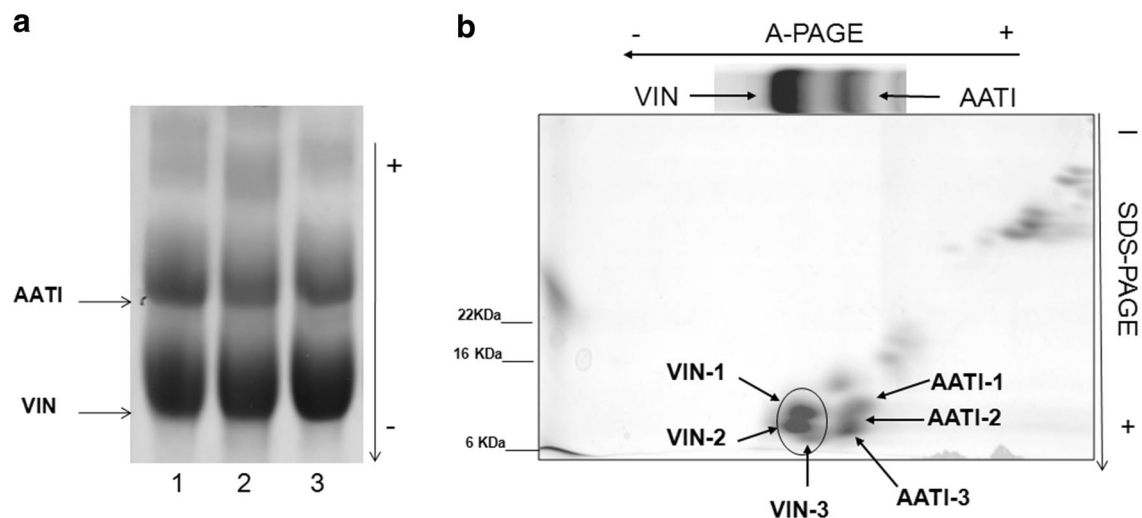


Fig. 1 **a** A-PAGE fractionation of reduced and alkylated starch-bound proteins from (1) *Avena sativa* cv. Donata (2) A-genome *Avena strigosa* and (3) C-genome *Avena ventricosa*; **b** Two-dimensional

A-PAGE x SDS-PAGE fractionation of reduced and alkylated starch-bound proteins from *Avena sativa* cv. Donata. Circled spots are vromindolines VIN-1, VIN-2, and VIN-3

could be arranged inside the deduced amino acid sequence of the 5' domain of the 636-bp-long cDNA clone JMT4-D065, GenBank CK780237 (Fig. 2a), and identified as an oat α -amylase/trypsin inhibitor. In addition, the remaining AATI-1 peptides could be assigned to the 3' domain of the same cDNA clone, translated backwards from the TAG stop codon at position 495–497. Therefore, primer pair “A1” designed on the 5' and 3' ends of the cDNA clone CK780237 (Table 1; Fig. 2a) was used for PCR amplification of genomic DNA of five single plants of oat cv. Donata. All plants produced a 592-bp-long amplicon identical to the JMT4-D065 clone, except for the insertion of an additional cytosine after the G nucleotide at position 315 (Fig. 2a, in bold). Translation of this amplicon resulted in a 151-amino-acid-long polypeptide here named *Avena* α -amylase/trypsin-inhibitor-1 (AATI-1), with an even number of ten cysteine residues, a CC dipeptide, and a CRC tripeptide typical of proteins belonging to the 2S family (Fig. 2b). Furthermore, the AATI-1 spot in the two-dimensional A-PAGE x SDS-PAGE patterns of starch-bound proteins extracted from cvs Donata and Primula was unique in showing a strong reaction with the polyclonal antiserum developed against the MRKTPASLPYEVSLR sequence (Fig. 2b, sequence in bold) specific to AATI-1 (data not shown). Finally, the N-terminal sequencing of the AATI-1 spot from oat cv. Donata produced the DNPFCFK-TAP polypeptide, suggesting that mature AATI-1 originates from post-translational cleavage of a 25-mer signal peptide.

The peptides obtained by trypsin digestion and RP-HPLC/n-ESI-MS/MS analysis of the AATI-1 spot covered about 80.5 % of mature AATI-1 (Fig. 2b, underlined sequences) and revealed a certain level of sequence

polymorphism. This polymorphism was due to the deletion of the P residue at position 35 of native AATI-1 or to single amino acid substitutions, such as L to M⁴⁰, T to I⁶³, L to I⁷⁹, S to P⁸³, L to V¹⁰⁸, M to T¹²⁸, T to I¹³⁷, or H to Y¹³⁸, suggesting that the AATI-1 spot contains three major polypeptides. Sequencing of *E. coli* clones transformed with PCR amplicons obtained from a single plant of oat cvs Donata, Primula and Hinoat using primer pair “A1cod” designed on the AATI-1 coding sequence (Table 1; Fig. 2b) confirmed this conclusion, and the three major polypeptides were named AATI-1.1, AATI-1.2, and AATI-1.3 (Fig. 3). The signal peptide of AATI-1.2 and AATI-1.3 was found to comprise 28 amino acid residues because of the insertion of the ²⁵ASA or ²⁵ASS sequences with respect to the 25-mer signal peptide of AATI-1.1 (Fig. 3). Moreover, cv. Hinoat released two clones only distinguishable from each other by a non-synonymous SNP, suggesting a gene duplication event that has resulted in two AATI-1 variants named AATI-1.1 and AATI-1.1* with a single amino acid substitution (P to S) at position 144 (Fig. 3).

After cleavage of the signal peptide, the AATI-1 peptides include 125 or 126 amino acids, the difference in length being due to the presence/absence of a P residue at position 10 of the mature protein. Similarity among AATI-1.1, AATI-1.2, and AATI-1.3 calculated over the native amino acid sequences (signal peptide included) was 89.5–96.7 % (Table 2). Alignment of mature proteins revealed that AATI-1.1 differs from AATI-1.2 and AATI-1.3 for ten amino acids (92.1 % similarity) and three amino acids (97.6 % similarity), respectively, whereas AATI-1.2 and AATI-1.3 differ from each other by seven amino acids (94.4 % similarity).

a

CTCGTGCCGAATTCGGCACGAGCCNNATCTAGCTAGAACAACATGGCGTCCAAGTGCAATTGCAGTCTCCTCCTCTTGCCGCCATCCT
 CGTCTCCGTTTTTCGTGCCTCCGCCGCGACAATCCCTGCTTCCCAAAACGGCGCCTACTCTGAGCAGTTTGCTCCAAAGGTGCCGTG
 ACTATGTGGAACAACAACCTGCGGCGTCGAACCTCCGGTCCGTTTACAATTTCTGTAAGAGAGCAGTACATGGTGAAGGAGCGTTG
 CTGCTGGGAGCTTGCCAACATTTGCGAGAAGTGTGCTGTGAGGCGT**TGG**TACCTTATGAGGAAGACGCCTGCTAGTCTTCCCTATG
 AAGTCAGCCTCAGGGGCATGCCCGGATGCCCAAGGGAGGCGCAGATGAACTTTGCCGAATACTCGTCATGCCGGGGCAGTGCAACT
 TGGCGACCACTACAACGTTCCGTACTGCCAGCTTTGGACAAATTTCACTACT**AG**AGATACAAATCTGGCTCTCATGAATAAGCAAAG
 CGTGTTCATCCACTGACGTGTGACATGCTTATCGACACATGTGCGCTCCAGGTTTCATGTGTGGTACTGTCTGGTATGATTAG**GAATAAA**
GAGAAACATTCTGAGGT⁶³⁶

b

M A S K C N C S L L L L A A I L V S V F ²⁰
ATG GCG TCC AAG TGC AAT TGC AGT CTC CTC CTC TTG GCC GCC ATC CTC GTC TCC GTT TTC

A A S A A D N P C F P K T A P T L S S L ⁴⁰
 GCT GCC TCC GCC GCG GAC AAT CCC TGC TTC CCA AAA ACG GCG CCT ACT CTG AGC AGT TTG

L Q R C R D Y V E Q Q T C G V E P S G P ⁶⁰
 CTC CAA AGG TGC CGT GAC TAT GTG GAA CAA CAA ACC TGC GGC GTC GAA CCC TCC GGT CCG

F T I S V R E Q Y M V K E E C C W E L A ⁸⁰
 TTT ACA ATT TCT GTA AGA GAG CAG TAC ATG GTG AAG GAG CGT TGC TGC TGG GAG CTT GCC

N I S Q K C R C E A L R Y L M R K T P A ¹⁰⁰
 AAC ATT TCG CAG AAG TGT CGC TGT GAG GCG TTG CGC TAC CTT ATG AGG AAG ACG CCT GCT

S L P Y E V S L R G M P G C P R E A Q M ¹²⁰
 AGT CTT CCC TAT GAA GTC AGC CTC AGG GGC ATG CCC GGA TGC CCA AGG GAG GCG CAG ATG

N F V R I L V M P G Q C N L A T T H N V ¹⁴⁰
 AAC TTT GTC CGA ATA CTC GTC ATG CCG GGG CAG TGC AAC TTG GCG ACC ACT CAC AAC GTT

R Y C* P A L D K F Q Y ¹⁵¹
 CGG TAC TGC CCA GCT TTG GAC AAA TTT CAG TAC TAG

Fig. 2 a cDNA sequence of clone JMT4-D065 (GenBank CK780237). The G nucleotide at position 315 and the stop codon TAG at position 495–497 are in bold. The 5' and 3' targets of the primer pair 'A1' (see Table 1) used for PCR amplification are *underlined*; **b** nucleotide and deduced amino acid sequences of the gene coding for AATI-1 obtained by PCR amplification of oat cv. Donata. The peptides obtained by trypsin digestion and RP-HPLC/n-ESI–

MS/MS analysis of the AATI-1 spot from cv. Donata are *underlined*. The MRKTPASLPYEVSLR sequence (in *bold*) was used to develop a polyclonal antiserum specific for AATI-1 proteins. The C nucleotide absent in the cDNA clone JMT4-D065 (GenBank CK780237) is *arrowed*. The 5' and 3' targets of primer pair 'A1cod' (see Table 1) used for PCR amplification are *double underlined*

Identification of AATI-2 and AATI-3 proteins by mass spectrometry

Trypsin digestion and subsequent mass spectrometric analysis of the AATI-2 spot, excised from the gel of the two-dimensional A-PAGE × SDS-PAGE fractionation of oat cv. Donata, allowed to identify nine different fragments, including the 20-mer DYVEQATCGFDTEGPPFMAR polypeptide. Based on this sequence, several primers were

designed to perform 3' and 5' RACE amplification of total RNA extracted from immature seeds of cv. Donata harvested at 28 DPA. In particular, amplification with primer "3RA2" for 3' RACE and primer "5RA2" for 5' RACE (Table 1) produced a DNA sequence coding for a protein named Avena α-amylase/trypsin-inhibitor-2 (AATI-2) harboured inside 478-bp-long cDNA amplicons (Fig. 4a). PCR amplification with primer "A2" designed on the 5' and 3' ends of this amplicon (Table 1; Fig. 4a) produced fragments

Fig. 3 Amino acid sequences of Avena- α -amylase-trypsin-inhibitor (AATI) proteins in *Avena sativa*. The arrows indicate the cleavage sites of the signal peptide. The amino acid domains, where AATI-1, AATI-2, and AATI-3 show higher sequence similarity, are underlined

Protein	Sequence			
AATI-1.1	MASKNCNS LL ¹⁰	LLAAILVSVF ²⁰	AASA--ADN ²⁷	PCFPKTAPTL ³⁷
AATI-1.1*	MASKNCNS LL	LLAAILVSVF	AASA--ADN ²⁷	PCFPKTAPTL ³⁷
AATI-1.2	MASKNCNS LF	LLAAVLLSIF	AAPAASADN ³⁰	PCFPKTA-IL ³⁹
AATI-1.3	MASKNCNS LL	LLAAILVSVF	AAAAASSADN ³⁰	PCFPKTAPTL ⁴⁰
	+	+	+	+
AATI-1.1	SSLLQRCRDY ⁴⁷	VEQQTCGV ⁵⁷	SGPFTISVRE ⁶⁷	QYMKERCCW ⁷⁷
AATI-1.1*	SSLLQRCRDY ⁴⁷	VEQQTCGV ⁵⁷	SGPFTISVRE ⁶⁷	QYMKERCCW ⁷⁷
AATI-1.2	SSMLQRCRDY ⁴⁹	VEQQTCGV ⁵⁹	SGPFIISVRE ⁶⁹	PYMKERCCW ⁷⁹
AATI-1.3	SSLLQRCRDY ⁵⁰	VEQQTCGV ⁶⁰	SGPFTISVRE ⁷⁰	QYMKERCCW ⁸⁰
	+	+	+	+
AATI-1.1	ELANISQKCR ⁸⁷	CEALRYLMRK ⁹⁷	TPASLPYEV ¹⁰⁷	LRGMPGCPRE ¹¹⁷
AATI-1.1*	ELANISQKCR ⁸⁷	CEALRYLMRK ⁹⁷	TPASLPYEV ¹⁰⁷	LRGMPGCPRE ¹¹⁷
AATI-1.2	EIANIPQKYR ⁸⁹	CEALRYLMRK ⁹⁹	TPASLPYEV ¹⁰⁹	VRGMPGCPRE ¹¹⁹
AATI-1.3	ELANISQKCR ⁹⁰	CEALRYLMRK ¹⁰⁰	TPASLPYEV ¹¹⁰	LRGMPGCPRE ¹²⁰
	+	+	+	+
AATI-1.1	AQMNFVRLV ¹²⁷	MPGQCNLATT ¹³⁷	HNVRYCPALD ¹⁴⁷	KFY ¹⁵¹
AATI-1.1*	AQMNFVRLV ¹²⁷	MPGQCNLATT ¹³⁷	HNVRYCSALD ¹⁴⁷	KFY ¹⁵¹
AATI-1.2	AQMNFVRLV ¹²⁹	TPGQCNLATT ¹³⁹	HNVRYCPALD ¹⁴⁹	KFY ¹⁵³
AATI-1.3	AQMNFVRLV ¹³⁰	TPGQCNLAT ¹⁴⁰	YNVRYCPALD ¹⁵⁰	KFY ¹⁵⁴
	+	+	++	+
AATI-2.1	MASNSSRN-L ⁹	LLFATVLLYV ¹⁹	FAPAATRSY ²⁹	CFPGMVGSPS ³⁹
AATI-2.2	MASNSSRNLL ¹⁰	LLFAAVLLSV ²⁰	FAAAATRGD ³⁰	CFQGMVGSPS ⁴⁰
AATI-2.2*	MASNSSRNLL ¹⁰	LLFAAVLLSV ²⁰	FAAA-TRGD ²⁹	CFQGMVGSPS ³⁹
AATI-2.3	MASNSSRN-L ⁹	LLFATVLLYV ¹⁹	FAPAAATRSY ²⁹	CFPGMVGSPS ³⁹
	+	++	+	+
AATI-2.1	DLIPSCRDYV ⁴⁹	EQATCGFDTE ⁵⁹	GPPFMARDDC ⁶⁹	CKQLAAIPQH ⁷⁹
AATI-2.2	DLIPSCRDYV ⁵⁰	EQETCGFDTE ⁶⁰	GPPFMAREDC ⁷⁰	CKQLAAIPQR ⁸⁰
AATI-2.2*	DLIPSCRDYV ⁴⁹	EQETCGFDTE ⁵⁹	GPPFMAREDC ⁶⁹	CKQLAAIPQR ⁷⁹
AATI-2.3	DLIPSCRDYV ⁴⁹	EQATCGFDTE ⁵⁹	GPPFMARDDC ⁶⁹	CKQLAAIPQH ⁷⁹
	+	+	+	+
AATI-2.1	CRCEALRCF ⁸⁹	GRRSRPDQIG ⁹⁹	GLIDLACGPR ¹⁰⁹	EPQRDFARIL ¹¹⁹
AATI-2.2	CRCEALRYF ⁹⁰	GRRSRPDQIG ¹⁰⁰	GLIDLPGCPR ¹¹⁰	EPQRDFARIL ¹²⁰
AATI-2.2*	CRCEALRYF ⁸⁹	GRRSRPDQIG ⁹⁹	GLIDLPGCPR ¹⁰⁹	EPQRDFARIL ¹¹⁹
AATI-2.3	CRCEALRCF ⁸⁹	GRRSRPDQIG ⁹⁹	GLIDMAGCPR ¹⁰⁹	EPQRDFARIL ¹¹⁹
	+	+	++	+
AATI-2.1	VTGQC�LAT ¹²⁹	IHNAPYCLAM ¹³⁹	DESQWH ¹⁴⁵	
AATI-2.2	VTGQC�LAT ¹³⁰	IHNAPYCLAM ¹⁴⁰	DESQWH ¹⁴⁶	
AATI-2.2*	VTGQC�LAT ¹²⁹	IHNAPYCLAM ¹³⁹	DESQWH ¹⁴⁵	
AATI-2.3	VTGQC�LAT ¹²⁹	IHNAPYCLAM ¹³⁹	DESQWH ¹⁴⁵	
	+	+	+	+
AATI-3.1	MASN--CRL ⁸	LLAALLVSF ¹⁸	AAAAATSGGE ²⁸	YCPSMGLPS ³⁸
AATI-3.2	MASN SKRVL ¹⁰	LMAALLVSF ²⁰	AAAGAT--GD ²⁸	YCPSMGLPS
AATI-3.2*	MASN SKRVL ¹⁰	LMAALLVSF ²⁰	AAAGAT--GD ²⁸	YCPSMGLPS
AATI-3.3	MASN--CRL ⁸	LLAALLVSF ¹⁸	AAATSGGE ²⁸	YCPSMGLPS
	++	+	++	+
AATI-3.1	RPLYGCREYV ⁴⁸	AQQTGCGARIL ⁵⁸	GAPSAPIET ⁶⁸	MERCCELEFSQ ⁷⁸
AATI-3.2	RPLDGCREYV	AQQTGCGTRIL	GAPSAPIEKL	MYQCCELEFSQ
AATI-3.2*	RPLDGCREYV	AQQTGCGTRIL	GAPSAPIEKL	MYQCCELEFSQ
AATI-3.3	RPLYGCREYV	AQQTGCGARIL	GAPSAPIEKL	MYQCCELEFSQ
	+	+	+	++
AATI-3.1	IQQHCRCAV ⁸⁸	RYLMGSRPES ⁹⁸	SGLMNLPGCP ¹⁰⁸	IEAQRDFARIL ¹¹⁸
AATI-3.2	IRQHCRCAQAL	RYLMGSDPET	SGLMKLP GCP	IEPQRDFARI
AATI-3.2*	IRQHCRCAQAL	RYLMGSDPET	SGLMKLP GCP	IEAQRDFARI
AATI-3.3	IQQHCRCEAL	RYLMGSHPER	SGLMNLPGCP	IEAQRDFARI
	+	+	+	+
AATI-3.1	LPTPRQC�LV ¹²⁸	IDYNTRYCLE ¹³⁸	MDKFM ¹⁴³	
AATI-3.2	LPTPRQC�LI	TEYNTRYCLE	MDKFS-	
AATI-3.2*	LPTPRQC�LV	TDYNTRYCLE	MDKFM-	
AATI-3.3	LPTPRQC�LV	TDYSTRYCLE	MDKFWQ ¹⁴⁴	
	+	+	+	+

Table 2 Percent similarity of amino acid sequences of native AATI-1, AATI-2, or AATI-3 proteins in diploid and tetraploid oat species as compared with their counterparts in *Avena sativa*

Encoded protein	GenBank	No. of amino acids		Molecular Mw(Da)/pI ^a	Species (genome)	% Similarity			
		Native	Mature			AATI-1.1	AATI-1.1*	AATI-1.2	AATI-1.3
AATI-1.1	KX014747	151	126	14,475/8.7	<i>A. sativa</i> (ACD)		99.3	90.7	96.7
AATI-1.1*	KX014748	151	126	14,465/8.7		99.3		90.0	96.0
AATI-1.2	KX014749	153	125	14,381/8.7		90.7	90.0		89.5
AATI-1.3	KX014750	154	126	14,483/8.7		96.7	96.0	89.5	
AATI-1.1 ^s	KX014751	151	126	14,475/8.7	<i>A. strigosa</i> (A)	100.0	99.3	90.7	96.7
AATI-1.1 ^l	KX014752	151	126	14,475/8.7	<i>A. longiglumis</i> (A)	99.3	98.7	90.0	96.0
AATI-1.1 ^c	KX014753	151	126	14,378/8.7	<i>A. canariensis</i> (A)	98.0	98.0	90.7	96.7
AATI-1.1 ^d	KX014754	151	126	14,378/8.7	<i>A. damascena</i> (A)	98.7	98.0	92.0	95.4
AATI-1.1 ^b	KX014755	151	126	14,475/8.7	<i>A. barbata</i> (AB)	100.0	99.3	90.7	96.7
AATI-1.2 ^b	KX014756	153	125	14,383/8.7		90.0	89.3	98.0	90.2
AATI-1.2 ⁱ	KX014757	153	125	14,407/8.7	<i>A. insularis</i> (AC)	90.7	90.0	97.4	92.2
AATI-1.3 ⁱ	KX014758	154	126	14,383/8.7		96.7	96.0	89.5	100.0
AATI-1.1 ^m	KX014759	150	125	14,404/8.6	<i>A. magna</i> (AC)	98.0	97.3	91.3	96.0
AATI-1.4 ^m	KX014760	150	125	14,412/8.6		96.0	95.3	92.0	95.3
Encoded protein	GenBank	No. of amino acids		Molecular Mw(Da)/pI ^a	Species (genome)	AATI-2.1	AATI-2.2	AATI-2.2*	AATI-2.3
		Native	Mature						
AATI-2.1	KU958672	145	120	13,337/5.8	<i>A. sativa</i> (ACD)		91.7	91.7	99.3
AATI-2.2	KU958683	146	120	13,499/5.3		91.7		99.8	91.3
AATI-2.2*	KX014762	145	120	13,499/5.3		91.7	99.8		91.0
AATI-2.3	KU958682	145	120	13,355/5.8		99.3	91.3	91.0	
AATI-2.1 ^l	KU958681	145	120	13,337/5.8	<i>A. longiglumis</i> (A)	100.0	91.7	90.4	99.3
AATI-2.4 ^v	KU958680	145	120	13,423/6.0	<i>A. ventricosa</i> (C)	95.9	91.7	91.7	95.2
AATI-2.4 ^p	KU958684	147	120	13,456/5.8	<i>A. clauda</i> (C)	96.5	93.8	93.8	95.9
AATI-2.3 ^b	KU958675	145	120	13,355/5.8	<i>A. barbata</i> (AB)	99.3	91.0	91.0	100.0
AATI-2.1 ⁱ	KU958674	145	120	13,397/5.8	<i>A. insularis</i> (AC)	99.3	92.4	92.4	98.6
AATI-2.1 ^m	KU958679	145	120	13,441/5.5	<i>A. magna</i> (AC)	98.6	92.4	92.4	97.9
AATI-2.1 ^{mr}	KU958676	145	120	13,397/5.8	<i>A. maroccana</i> (AC)	99.3	92.4	92.4	98.6
AATI-2.1 ^{mu}	KU958678	145	120	13,325/6.2	<i>A. murphy</i> (AC)	98.6	91.7	91.7	97.9
Encoded protein	GenBank	No. of amino acids		Molecular Mw(Da)/pI ^a	Species (genome)	AATI-3.1	AATI-3.2	AATI-3.2*	AATI-3.3
		Native	Mature						
AATI-3.1	KU958669	143	120	13,585/8.0	<i>A. sativa</i> (ACD)		83.7	86.5	93.7
AATI-3.2	KU958670	143	118	13,438/7.4		83.7		97.2	85.8
AATI-3.2*	KX014761	143	118	13,429/7.4		86.5	97.2		87.9
AATI-3.3	KU958671	144	121	13,797/7.4		93.7	85.8	87.9	
AATI-3.1 ^s	KU958668	143	120	13,654/8.0	<i>A. strigosa</i> (A)	96.5	83.0	85.8	92.3
AATI-3.2 ^m	KU958665	143	118	13,438/7.4	<i>A. magna</i> (AC)	83.7	100.0	97.2	85.8
AATI-3.4 ^m	KU958664	143	118	13,422/7.4	<i>A. magna</i> (AC)	83.7	99.3	96.5	85.1
AATI-3.2 ^{mr}	KU958667	143	118	13,438/7.4	<i>A. maroccana</i> (AC)	83.7	100.0	97.2	85.8
AATI-3.3 ^{mr}	KU958666	144	118	13,619/7.4	<i>A. maroccana</i> (AC)	89.4	90.2	92.3	95.8
AATI-3.2 ^{mu}	KU958662	143	118	13,455/7.4	<i>A. murphy</i> (AC)	83.0	99.3	96.5	85.1
AATI-3.4 ^{mu}	KU958663	143	118	13,455/7.4	<i>A. murphy</i> (AC)	83.0	98.6	95.8	85.1

GenBank codes, number of amino acids, computed molecular weights (Mw), and isoelectric points (pI) are reported as well

^a Mature protein

Fig. 4 **a** 478 bp-long cDNA amplicon harbouring the sequence coding for the AATI-2 protein (in **bold**). The 5' and 3' targets of the primer pair 'A2' used for PCR amplification (see Table 1) are underlined; **b** 542 bp-long amplicon obtained by combining cDNA fragments amplified by 3' and 5' RACE using primers '3RA3' and '5RA3', respectively (see Table 1). The 432 bp-long DNA sequence coding for AATI-3 is in **bold**

a
GTCATCGAACCGGGTCTAGCTAGAACAATGAATAATATGGCGTCCAACCTCCAGTAGAAATCTCTCTCTCTGCCACC GTTCTGC
TCTATGTCTTCGCCCTGCCGCCGCACACGAGTTACTGCTTCCAGGGATGGTGGGTAGCCCGAGCGATCTGATCCCAAGCTG
CCGTGACTACGTGGAACAAACGACCTGCGGCTTCGATACCAGAGGGCCGCAATTCATGGCCAAGGATGACTGTGCAAGCAAGC
TTGCAGCAATCCGAGCACTGCCGGTGGCAGGCGTTGCGCTGCTTATAGGCGGAGGTCCCGTCTGATCAGATCGGCGGCC
TCATCGACATGGCTGGATGCCCCAGGGAGCCGCAGAGGATTTCCGAGAACTACGTCACGCGGGGCAATGCAACTTGGCC
ACCATTCAACAACGCTCCGTACTGTCTCGCTATGGATGA GTCTCAGTGGCACTAGAGA⁴⁷⁸

b
CACCCAGCAGTGAAGTGCACGAAAGGAACTACACCAACGAATCAACCTAGCTAGCATAACA **TGGCGTCCAACCTCACTGCCGT**
TTTCTCTGTTGGCCGCCCTCTGGTCTCTGCTTTGCGCGTGCCGGCGCCACCGGTGATTAAGTCTACCACTCGATGGGTCTTCCG
AGCCGTCCGCTCGATGGCTGCCGTGATACGTGGCACAACAACTTGGCGCACCGTATCTCGGGGCGCCGTCCGCCCAATC
GAAAAGTTGATGTACCAAGTTGCTGCTGGAATTTTACAGATTCGGCAGCACTGCCGGTCCAGGCGCTGCGCTATTGATGGGG
TCTGATCCAGAAACGAGTGGCTCAAGAACTGCCAGGATGCCCATTTGAGCCGAGAGGACTTTGCCAGAACTATCCCAACA
CCTAGGCAGTGCACCTGATTACA GAATACAACACAAGATACGTGCTGGAAATGGACAAGTTCCTCGTAGTACAGATCGCTCATG
AATAAGCATGATGCATCCATGGATGTGTG⁵⁴²

of either 478 or 481 bp in single plants of cv. Donata. When cloned in *E. coli*, these fragments were found to contain the genes coding for two protein pairs, i.e., AATI-2.1 plus AATI-2.3 and AATI-2.2 plus AATI-2.2* (Fig. 3). In particular, AATI-2.1 and AATI-2.3 differed from each other in a single L to M¹⁰⁴ substitution and exhibited an odd number of 11 C residues. Assuming a cleaved peptide of 25 amino acids, the mature form of this protein pair contains 120 amino acids, 5–6 amino acids less than the mature AATI-1 proteins. Mature AATI-2.2 and AATI-2.2* revealed the same primary structure with 120 amino acids, the only difference between these proteins being located in the cleaved peptide at position 26 (Fig. 3). The protein pairs AATI-2.1/AATI-2.3 and AATI-2.2/AATI-2.2* in their mature forms showed 91.7–92.5 % similarity. By contrast, computed sequence similarity between mature AATI-1 and AATI-2 proteins was ≤ 42.4 %.

Based on the polypeptides QCCLEFSQIR and CREYVA obtained by mass spectrometric analysis of the AATI-3 spot of oat cv. Donata, primers "3RA3" and "5RA3" (Table 1) were designed to carry out, respectively, 3' and 5' RACE amplification of total RNA extracted from immature seeds of cv. Donata. The resulting cDNA fragments were then combined to form a 542-bp-long amplicon containing a coding DNA sequence of 432 bp (Fig. 4b). When amplified with primer pair "A3" positioned at the ends of this amplicon (Table 1; Fig. 4b), single plants of cvs Donata and Primula produced PCR fragments, 542 or 545 bp in size, which were found to contain genes coding for four proteins named AATI-3.1, AATI-3.2, AATI-3.2*, and AATI-3.3 (Fig. 3). Proteins AATI-3.2 and AATI-3.2* share the same 25-mer signal peptide and contain 118 amino acids in their mature forms. Sequence similarity between mature AATI-3.2 and AATI-3.2* is as high as 97.2 %, with only four amino acids difference in their primary structure. Proteins AATI-3.1 and AATI-3.3 show a short signal peptide of 23 amino acids and a mature structure of 120–121 amino acids. Sequence similarity between mature AATI-3.1 and

AATI-3.3 is 93.7 %. By contrast, these proteins exhibit ≤ 87.9 % similarity with AATI-3.2 and AATI-3.2*. Finally, sequence similarity of mature AATI-3 proteins with their AATI-1 and AATI-2 counterparts is ≤ 45.0 %. However, mature AATI-1, AATI-2, and AATI-3 proteins show three domains with sequence similarity ≥ 79.2 % (Fig. 3, underlined sequences).

AATI proteins in diploid and tetraploid *Avena* species

Upon A-PAGE fractionation, reduced and alkylated proteins extracted from the starch granules of diploid A-genome *A. canariensis*, *A. damascena*, *A. longiglumis*, and *A. strigosa*, diploid C-genome *A. clauda* and *A. ventricosa*, tetraploid AC-genome *A. insularis*, *A. magna*, *A. maroccana*, and *A. murphy*, and tetraploid AB-genome *A. barbata* appeared as two distinct bands corresponding to vromindolines (VIN) and Avena α -amylase/trypsin-inhibitor (AATI) proteins (Fig. 1a, lanes 2 and 3 and data not shown). The A-PAGE patterns of these wild species were indistinguishable from those of cultivated hexaploid oats (Fig. 1a, lane 1). After fractionation by two-dimensional A-PAGE x SDS-PAGE, the AATI band of diploid and tetraploid oat species split into three spots (AATI-1, AATI-2, and AATI-3), with the only exception of *A. clauda*, which lacked the AATI-3 spot (data not shown). The genes coding for AATI-1, AATI-2, and AATI-3 in wild oat species were investigated by PCR amplification of genomic DNA from single plants using the primer pairs 'A1', 'A2', and 'A3' designed for hexaploid oats (Table 1), and the amplicons were then cloned in *E. coli*. Each of the A- and C-genome diploid species analyzed here produced a single clone coding for AATI-1, AATI-2 or AATI-3. By contrast, the AB- and AC-genome species released two clones coding for AATI-1, one clone coding for AATI-2 and up to three clones coding for AATI-3 (Table 2). Alignments of the amino acid sequences of AATI-1, AATI-2, and AATI-3 proteins in their native forms reveal high similarities among

the analyzed diploid, tetraploid, and hexaploid oat species. In particular, AATI-1.1^s in A-genome *A. strigosa* (allele *Aati-A^s1a*, GenBank KX014751) and AATI-1.1^b in AB-genome *A. barbata* (allele *Aati-A^b1a*, GenBank KX014755) exhibit 100 % similarity with AATI-1.1 in the hexaploid cvs Donata and Hinoat. Therefore, this latter protein is likely to be encoded by the A-genome and thus provisionally assigned to allele *Aati-A1a* (GenBank KX014747). In addition, AATI-1.1 in cvs Donata and Hinoat shows 98.0–99.3 % similarity with its counterparts in *A. longiglumis* (allele *Aati-A^l1a*, GenBank KX014752), *A. canariensis* (allele *Aati-A^c1a*, GenBank KX014753), and *A. damascena* (allele *Aati-A^d1a*, GenBank KX014754). Protein AATI-2.1^l in the A-genome species *A. longiglumis* (allele *Aati-A^l2a*, GenBank KU958681) is identical to protein AATI-2.1 in *A. sativa* suggesting that AATI-2.1 is encoded by an A-genome chromosome carrying the allele *Aati-A2a* (GenBank KU958672). Furthermore, AATI-3.1 shows a sequence similarity as high as 96.5 % with its counterpart in *A. strigosa* (allele *Aati-A^s3a*, GenBank KU958668). Finally, proteins with 90.2–100 % similarity with AATI-1.1, AATI-1.3, AATI-2.1, AATI-2.3, or AATI-3.2 in cvs Donata and Primula were found to occur in AC-genome species *A. insularis*, *A. magna*, and *A. maroccana* (Table 2).

AATI deposition on starch granules in developing kernels

The starch granules extracted from the immature kernels of oat cvs Donata and Primula collected at 7 DPA revealed the first traces of AATI proteins, as determined by A-PAGE fractionation (Fig. 5, lane 1). Deposition of AATI proteins on the starch granules reached its maximum at 35 DPA and slightly decreased at 42 DPA (Fig. 5, lane 6). AATI and vromindolines (VIN) followed the same pattern of accumulation except that the first traces of VIN proteins appeared at 14 DPA (Fig. 5, lane 1), one week later with respect to AATI proteins.

Discussion

Alpha-amylases (α -1,4-glucan-4-glucanohydrolases) constitute a family of endo-amylases that catalyze the hydrolysis of α -D-(1 $\rightarrow$ 4)-glucan linkages in starch components and other carbohydrates, thus playing a key role in carbohydrate metabolism of microorganisms, plants and animals (Franco et al. 2002). Seeds of cereal species, such as wheat, barley, rye, triticale, sorghum, rice, and finger millet, contain proteinaceous α -amylase inhibitors of 120–160 amino acid residues, which occur as monomeric, homodimeric, or heterotetrameric molecules differentially active

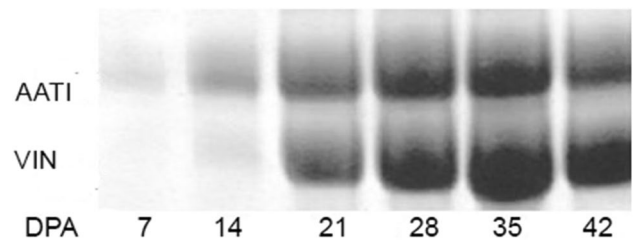


Fig. 5 A-PAGE fractionation of starch-bound proteins extracted from kernels collected at 7–42 days post anthesis (DPA) in oat cv. Donata. AATI Avena- α -amylase-trypsin inhibitors, VIN vromindolines

against α -amylases from *Coleoptera* and *Lepidoptera* phytophagous species (Gutierrez et al. 1990, 1993; Franco et al. 2002). In nucleotide and amino acid sequence comparisons, the 37 AATI proteins identified here in diploid, tetraploid, or hexaploid oat species show a high level of homology compared with α -amylase inhibitors from wheat and barley. At each ploidy level, oat α -amylase inhibitors were found to belong to three families named AATI-1, AATI-2, and AATI-3 with contrasting primary structures, molecular weights, and isoelectric points (pI) (Table 2). The remarkable structural differences amongst the AATI-1, AATI-2, and AATI-3 inhibitors likely affect their specificity and activity against α -amylases of different phytophagous insects, as demonstrated for other cereal α -amylase inhibitors (Feng et al. 1996; Franco et al. 2000, 2002). Amongst the cereal α -amylase inhibitors deposited in the GenBank, the AATI proteins exhibit the highest values of sequence similarity with the components of the wheat tetrameric α -amylase inhibitors (WTAI), which are referred to as CM proteins because of their solubility in chloroform/methanol mixtures (Garcia-Olmedo et al. 1982). Each WTAI, about 60 kDa in size, contains one copy of either CM1 or CM2 protein, one copy of either CM16 or CM17 protein, and two copies of CM3 protein. In sequence comparisons, AATI-1 proteins have the highest level of similarity (55.3 %) with CM16 inhibitor, whereas AATI-2 proteins show approximately 60 % similarity with CM1 and CM2 inhibitors (Table 3). By contrast, similarity values as low as 23.2–26.1 % were found between AATI proteins and wheat monomeric WMAI-1 (synonymous 0.28) and homodimeric WDAI-2 (syn. 0.19) members of the inhibitor family.

In diploid A-genome oat species, a sole DNA sequence was found to code for AATI-1, AATI-2, or AATI-3 proteins (Table 2), whereas in tetraploid AC-genome species, there were two different sequences coding for AATI-1, two for AATI-3, and one for AATI-2. In *A. sativa*, four sequences for each of the three AATI family were identified, suggesting that AATI inhibitors are encoded by a multigene family dispersed over the A, C, and D genomes. The coding regions of AATI-1.1 and AATI-1.1* in the single seeds of

Table 3 Average similarity (%) of native AATI proteins with 0.28 protein (WMAI, monomeric wheat α -amylase inhibitor), 0.19 protein (WDAI, dimeric wheat α -amylase inhibitor), and wheat chloroform/methanol (CM) proteins

	0.28	0.19	CM1	CM2	CM3	CM16	CM17
AATI-1	23.2	23.6	47.5	49.6	53.0	55.3	51.1
AATI-2	26.1	26.1	59.6	60.3	54.5	55.6	52.4
AATI-3	23.4	25.0	55.3	58.2	54.6	51.1	46.3

cv. Hinoat show a single bp difference that results in P to S amino acid substitution at the C-terminal end of the mature proteins (Fig. 3). Similarly, in cv. Donata, genes coding for AATI-2.2 and AATI-2.2* differ from each other for the presence/absence of a single codon in the signal peptide region. These observations suggest that a gene duplication may have occurred at the loci coding for AATI-1.1/AATI-1.1* and AATI-2.2/AATI-2.2* protein pairs in the analyzed oat cultivars. Dispersion of genes coding for WTAI subunits CM1, CM2, CM3, CM16, and CM17 on homoeologous B- and D-genome chromosomes also occurs in polyploid wheat species (Carbonero and Garcia-Olmedo 1999).

AATI-1.1 and AATI-2.1 proteins in *A. sativa* have identical counterparts in *A. strigosa* and *A. longiglumis*, respectively, suggesting that these species may be close relatives and descendants of the ancestral diploid species that participated in the formation of hexaploid oats. This conclusion is in conflict with the very recent finding based on the next-generation DNA sequencing that *A. longiglumis*, but not *A. strigosa*, belongs to the genetic cluster of the closest living relatives of the ancestral donors of A and D genomes to hexaploid oats (Chew et al. 2016). Nevertheless, Chen and Armstrong (1994) and Linares et al. (1998), proposed *A. strigosa* as a likely donor of the A-genome to the ancestral tetraploid and, subsequently, to hexaploid oats. In addition, the comparison of the *Vin-2* genes coding for vomindolines, another class of oat starch-bound proteins, suggested *A. longiglumis* as the likely donor of the A-genome to hexaploid oats (Alfieri et al. 2014). According to Linares et al. (1998), *A. sativa* inherited two A-genomes from diploid progenitors, and one of them evolved into the D genome.

In common wheat, α -amylase inhibitors represent up to 4 % of total proteins (Schuppan and Zevallos 2015) and contribute positively to the biological value of endosperm proteins because of their more balanced amino acid composition compared with major storage proteins, i.e., prolamins. A-PAGE fractionation of the starch granules bound proteins in common wheat and barley reveals relatively large amounts of tryptophan-rich indolines with no traces of α -amylase inhibitors (Corona et al. 2001; Galassi et al. 2012), suggesting that in these cereal species, deposition of α -amylase inhibitors occurs, mainly, in the endosperm matrix. In this context, it is worth noting that also in einkorn wheat *Triticum monococcum*, a component of the

2S protein family, named Einkorn Trypsin Inhibitor (ETI), was found to accumulate on the starch granules (Taddei et al. 2009), and it shows about 41 % sequence similarity with AATI-2.2. On the other hand, diploid, tetraploid, and hexaploid oats synthesize AATI proteins as precursors that undergo specific cleavage of a leader peptide of 25–28 amino acid residues and then accumulate on the starch granules, together with more abundant amounts of vomindoline (VIN) proteins. Densitometric scanning of A-PAGE gels revealed that the amount of AATI associated with 12 mg of starch granules in oat cultivars Donata and Primula was roughly threefold lower than that of vomindolines and about twofold higher than that of puroindolines in soft common wheat cvs Bolero and Bramante (data not shown). Due to the relatively high proportion of the limiting amino acids lysine and threonine in the AATI-1 and AATI-3 proteins, the important accumulation of these inhibitors could contribute to the well-known superior amino acid composition of oats with respect to rice, rye, sorghum, maize, and wheat. Furthermore, AATI and vomindolines could interact in determining the extra-soft texture of oat kernels by reducing the adhesion between the starch granule surface and the surrounding endosperm matrix (Greenblatt et al. 1995; Gazza et al. 2015).

Amongst the wheat α -amylase inhibitors, WMAI 0.28, WDAI 0.19, and N-glycosylated CM16 have been identified as the major allergens associated with the baker's asthma, an allergic disease caused by the inhalation of wheat flour (Amano et al. 1998). In barley, rye, and rice, some members of the α -amylase inhibitor family are also major allergens (Urisu et al. 1991; García-Casado et al. 1995). In addition, CM3 and WDAI 0.19 inhibitors and likely all wheat dimeric and tetrameric α -amylase inhibitors along with some homologous variants in other cereal species have been shown to act as potent activators of the human innate immune system (Junker et al. 2012; Schuppan and Zevallos 2015). In particular, these inhibitors may stimulate monocytes, macrophages, and dendritic cells to produce pro-inflammatory cytokines, such as tumor necrosis factor (TNF) and interleukins IL-8 and IL-12, and trigger immune reactions in patients with celiac disease, non-celiac wheat sensitivity, irritable bowel syndrome, or non-intestinal immune disorders. The current European regulations include oats in the list of gluten-free ingredients, and this is expected to improve their nutritional

value due to the unique composition of oat kernels with respect to proteins, lipids, vitamins, minerals, and fiber. However, several studies have shown significant variation in immunoreactivity amongst the oat cultivars analyzed (Ballabio et al. 2011; Maglio et al. 2011), casting serious doubt on safety of oat when consumed by celiac patients (Haboubi et al. 2006; Comino et al. 2015). The relatively high similarity between oat AATI proteins and wheat CM proteins indicates the need to investigate whether AATI inhibitors are able to elicit immune responses, with broad implications for the onset of allergy or intestinal inflammatory diseases.

Acknowledgements The authors would like to dedicate this article to their beloved colleague and friend Norberto, in memory of his passion for the genetics of cereals.

Compliance with ethical standards

Funding This study was partly funded by the VROTECH project (grant number DM 19177-7303-08) by Ministry of Agriculture, food, and forestry (Mipaaf, Italy) to Norberto E. Pogna.

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

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Alfieri M, Gazza L, Pogna NE, Redaelli R (2014) Gene sequences of vomindolines in *Avena* species. *Gen Resour Crop Evol* 61:1481–1490
- Amano M, Ogawa H, Kojima K, Kamidaira T, Suetsugu S, Yoshihama M, Satoh T, Samejima T, Matsumoto I (1998) Identification of the major allergens in wheat flour responsible for Baker's asthma. *Biochem J* 330:1229–1234
- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1994) Genomic extraction using the modified CTAB method. *Current protocols in molecular biology*. Wiley, New York
- Ballabio C, Uberti F, Manfredelli S, Vacca E, Boggini G, Redaelli R, Catassi C, Lionetti E, Peñas E, Restani P (2011) Molecular characterisation of 36 oat varieties and in vitro assessment of their suitability for coeliacs' diet. *J Cereal Sci* 54:110–115
- Carbonero P, Garcia-Olmedo F (1999) A multigene family of trypsin/ α -amylase inhibitors from cereals. In: Shewry PR, Casey R (eds) *Seed Proteins*. Kluwer Academic Publishers, Dordrecht, pp 617–633
- Chen Q, Armstrong K (1994) Genomic in situ hybridization in *Avena sativa*. *Genome* 37:607–612
- Chew P, Meade K, Hayes A, Harjes C, Bao Y, Beattie AD, Puddephat I, Gusmini G, Tanksley SD (2016) A study on the genetic relationships of *Avena* taxa and the origins of hexaploid oat. *Theor Appl Genet*. doi:10.1007/s00122-016-2712-4
- Comino I, Moreno ML, Sousa C (2015) Role of oats in celiac disease. *World J Gastroenterol* 21:11825–11831
- Corona V, Gazza L, Boggini G, Pogna NE (2001) Variation in friabilin composition as determined by A-PAGE fractionation and PCR amplification, and its relationship to grain hardness in bread wheat. *J Cereal Sci* 34:243–250
- Cunsolo V, Muccilli V, Saletti R, Foti S (2012) Mass spectrometry in the proteome analysis of mature cereal kernels. *Mass Spectrom Rev* 31:448–465
- Feng GH, Richardson M, Chen MS, Kramer KJ, Morgan TD, Reece GR (1996) Alpha-amylase inhibitors from wheat: a sequences and patterns of inhibition of insect and human alpha-amylases. *Insect Biochem Mol Biol* 26:419–426
- Franco OL, Rigden DJ, Melo FR, Grossi-de-Sà MF (2000) Activity of wheat alpha-amylase inhibitors towards bruchid alpha-amylases and structural explanation of observed specificities. *Eur J Biochem* 267:1466–1473
- Franco OL, Rigden DJ, Melo FR, Grossi-de-Sà MF (2002) Plant alpha-amylase inhibitors and their interaction with insect alpha-amylases: structure, function and potential for crop protection. *Eur J Biochem* 269:397–412
- Fric P, Gabrovská D, Nevořal J (2011) Celiac disease, gluten-free diet and oats. *Nutr Rev* 69:107–115
- Galassi E, Gazzelloni G, Taddei F, Muccilli V, Gazza L, Pogna N (2012) Kernel texture and hordoinolines patterns in barley (*Hordeum vulgare*). *Mol Breed* 30:1551–1562
- García-Casado G, Armentia A, Sánchez-Monge R, Sánchez LM, López-Otín C, Salcedo G (1995) A major Baker's asthma allergen from rye flour is considerably more active than its barley counterpart. *FEBS Lett* 364:36–40
- García-Olmedo F, Carbonero P, Jones BL (1982) Chromosomal location of genes that control wheat endosperm proteins. *Adv Cereal Sci Technol* 5:1–47
- Gazza L, Conti S, Taddei F, Pogna NE (2006) Molecular characterization of puroindolines and their encoding genes in *Aegilops ventricosa*. *Mol Breed* 17:191–200
- Gazza L, Taddei F, Conti S, Gazzelloni G, Muccilli V, Janni M, D'Ovidio R, Alfieri M, Redaelli R, Pogna NE (2015) Biochemical and molecular characterization of *Avena* indolines and their role in kernel texture. *Mol Genet Genom* 290:39–54
- Greenblatt GA, Bettge AD, Morris CF (1995) The relationship among endosperm texture, friabilin occurrence, and bound polar lipids on wheat starch. *Cereal Chem* 72:172–176
- Gutierrez C, Sanchez-Monge R, Gomez L, Ruiz-Tapiador M, Castañera P, Salcedo G (1990) α -amylase activities of agricultural insect pests are specifically affected by different inhibitor preparations from wheat and barley endosperms. *Plant Sci* 72:37–44
- Gutierrez C, Garcia-Casado G, Sánchez-Monge R, Gomez L, Castañera P, Salcedo G (1993) Three inhibitor types from wheat endosperm are differentially active against α -amylases of Lepidoptera pests. *Entomol Exp Appl* 66:47–52
- Haboubi NY, Taylor S, Jones S (2006) Coeliac disease and oats: a systematic review. *Postgrad Med J* 82:672–678
- Junker Y, Zeissig S, Kim S-J, Barisani D, Wieser H, Leffler DA, Zevallos V, Libermann TA, Dillon S, Freitag TL, Kelly CP, Schuppan D (2012) Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. *JEM* 209:2395–2408
- Knowlton KF, Allen MS, Erickson PS (1996) Lasalocid and particle size of corn grain for dairy cows in early lactation. 1. Effect on performance, serum metabolite, and nutrient digestibility. *J Dairy Sci* 79:557–564
- La Vieille S, Pulido OM, Abbott M, Koerner TB, Godefroy S (2016) Celiac disease and gluten-free oats: a Canadian position based on a literature review. *Can J Gastroenterol Hepatol*. doi:10.1155/2016/1870305
- Linares C, Ferrer E, Fominaya A (1998) Discrimination of the closely related A and D genomes of the hexaploid oat *Avena sativa* L. *Proc Nat Acad Sci USA* 95:12450–12455
- Maglio M, Mazzarella G, Barone MV, Gianfrani C, Pogna N, Gazza L, Stefanile R, Camarca A, Colicchio B, Nanayakkara M, Miele

- E, Iaquinto G, Giardullo N, Maurano F, Santoro P, Troncone R, Auricchio S (2011) Immunogenicity of two oats varieties, in relation to their safety for celiac patients. *Scand J Gastroenterol* 46:1194–1205
- Meydani M (2009) Potential health benefits of avenanthramides of oats. *Nutr Rev* 67:731–735
- Peterson DM (1992) Composition and nutritional characteristics of oat grain and products. In: Marshall HG, Sorrels ME (eds) *Oat science and technology*. Crop Science Society of America, Madison, pp 265–292
- Peterson DM (2001) Oat antioxidants. *J Cereal Sci* 33:115–129
- Schuppan D, Zevallos V (2015) Wheat amylase trypsin inhibitors as nutritional activators of innate immunity. *Dig Dis* 33:260–263
- Singh R, De S, Belkheir A (2013) *Avena sativa* (Oat), a potential nutraceutical and therapeutic agent: an overview. *Crit Rev Food Sci Nutr* 53:126–144
- Taddei F, Gazza L, Conti S, Muccilli V, Foti S, Pogna NE (2009) Starch-bound 2S proteins and kernel texture in einkorn, *Triticum monococcum* ssp *monococcum*. *Theor Appl Genet* 119:1205–1212
- Tsilo TJ, Hareland GA, Chao S, Anderson JA (2011) Genetic mapping and QTL analysis of flour color and milling yield related traits using recombinant inbred lines in hard red spring wheat. *Crop Sci* 51:237–246
- Urisu A, Yamada K, Masuda S, Komada H, Wada E, Kondo Y, Horiba F, Tsuruta M, Yasaki T, Yamada M et al (1991) 16-kilodalton rice protein is one of the major allergens in rice grain extract and responsible for cross-allergenicity between cereal grains in the *Poaceae* family. *Int Arch Allergy Appl Immunol* 96:244–252
- Wani SA, Shah TR, Bazaria B, Nayik GA, Gull A, Muzaffar K, Kumar P (2014) Oats as a functional food: a review. *Univ J Pharm* 3:14–20
- Wondra KJ, Hancock JD, Behnke KC, Hines RH, Stark CR (1995) Effects of particle size and pelleting on growth performance, nutrient digestibility, and stomach morphology in finishing pigs. *J Anim Sci* 73:757–763