frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 10 May 2022
doi: 10.3389/fpls.2022.899740

Hololectin Interdomain Linker Determines Asparaginyl Endopeptidase-Mediated Maturation of Antifungal Hevein-Like Peptides in Oats

*Shining Loo†, Stephanie V. Tay†, Antony Kam, Warren Lee and James P. Tam**

School of Biological Sciences, Nanyang Technological University, Singapore, Singapore

Heveins and hevein-containing (hev-) lectins play important roles in stress and pathogenic responses in plants but cause health concerns in humans. Hev-hololectins contain multiple modular hev-peptide domains and are abundantly present in cereals and pseudocereals. However, it is unclear why some cereal hev-hololectins are presented as different forms of proteolytically processed proteoforms. Here we show the precursor architectures of hev-hololectins lead to different processing mechanisms to give either hololectins or hevein-like peptides. We used mass spectrometry and datamining to screen hev-peptides from common cereals, and identified from the oat plant *Avena sativa* nine novel hevein-like peptides, avenatide aV1–aV9. Bioinformatic analysis revealed that asparaginyl endopeptidase (AEP) can be responsible for the maturation of the highly homologous avenatides from five oat hev-hololectin precursors, each containing four tandemly repeating, hev-like avenatide domains connected by AEP-susceptible linkers with 13–16 residues in length. Further analysis of cereal hev-hololectins showed that the linker lengths provide a distinguishing feature between their cleavable and non-cleavable precursors, with the cleavables having considerably longer linkers (>13 amino acids) than the non-cleavables (<6 amino acids). A detailed study of avenatide aV1 revealed that it contains eight cysteine residues which form a structurally compact, metabolic-resistant cystine-knotted framework with a well-defined chitin-binding site. Antimicrobial assays showed that avenatide aV1 is anti-fungal and inhibits the growth of phyto-pathogenic fungi. Together, our findings of cleavable and non-cleavable hololectins found in cereals expand our knowledge to their biosynthesis and provide insights for hololectin-related health concerns in human.

Keywords: hololectin, hevein, oats, biosynthesis, celiac diseases, anti-fungal, asparaginyl endopeptidase

OPEN ACCESS

Edited by:
Pan Liao,
Purdue University, United States

Reviewed by:
Pitter F. Huesgen,
Julich Research Center, Helmholtz Association of German Research Centres (HZ), Germany
Fuai Sun,
University of California, Davis, United States

***Correspondence:**
James P. Tam
JPTam@ntu.edu.sg

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Metabolism and Chemodiversity, a section of the journal Frontiers in Plant Science

Received: *19 March 2022*
Accepted: *20 April 2022*
Published: *10 May 2022*

Citation:
Loo S, Tay SV, Kam A, Lee W and Tam JP (2022) Hololectin Interdomain Linker Determines Asparaginyl Endopeptidase-Mediated Maturation of Antifungal Hevein-Like Peptides in Oats. Front. Plant Sci. 13:899740. doi: 10.3389/fpls.2022.899740

INTRODUCTION

Asparaginyl endopeptidases (AEPs), also known as legumains and vacuolar protein endopeptidases (VPEs), are cysteine proteases that cleave the carboxyl-terminal side of Asx (Asp/Asn) (Csoma and Polgár, 1984; Kembhavi et al., 1993; Yamada et al., 2020). Functional AEPs are widely distributed in plants, mammals, protozoan parasites, trematodes like *Schistosoma mansoni*, and

insects like ticks, but not in bacteria (Dall and Brandstetter, 2016). In plants, AEPs play important roles in different plant organs and different stages of plant development and death. They are involved in the processing of peptides, proteins and their precursors like seed storage proteins, for growth and development as well as regulating programmed cell death and environmental stress responses (Vorster et al., 2019). In animals, AEPs are pivotal in the endosome/lysosomal degradation system and are implicated in antigen processing (Manoury et al., 1998, 2002).

Asparaginyl endopeptidases (C13 family) share the cysteine and histidine residues in the active site with other cysteine proteases such as papains (C1 family) and caspases (C14 family), but have little sequence similarity to them (Rawlings et al., 2010). Functionally, AEPs can mediate the biosynthesis of peptides and proteins through selective proteolysis of exposed Asx residues in the mature domains. Recent reports include the maturation of cysteine-rich peptides with protease inhibitory and insecticidal activities, such as roseltides from *Hibiscus sabdariffa* and jasmintides from *Jasminum sambac* (Loo et al., 2016; Kumari et al., 2018; Kam et al., 2019a,b). Unlike other cysteine proteases which solely break peptide bonds, certain AEPs can reverse their enzymatic direction to act as ligases to form peptide bonds (Nguyen et al., 2014, 2015; Harris et al., 2015; Hemu et al., 2019; Du et al., 2020; Chen et al., 2021; Liew et al., 2021). This highly unusual ligating function assisted the AEP-mediated biosynthesis of cyclic peptides through head-to-tail cyclization, forming cyclotides, sunflower seed trypsin inhibitors, and orbitides from seed storage proteins (Mylne et al., 2011; Fisher et al., 2018; Franke et al., 2018). The dual function of AEPs further enable them to act as a splicing enzyme, through a sequence of post-translational cut-and-join events, in the maturation of the cyclic trypsin inhibitor MCoTI-I/II from *Momordica cochinchinese* of the squash family, and circular-permutated plant lectins (Carrington et al., 1985; Bowles et al., 1986; Mylne et al., 2012; Du et al., 2020; Liew et al., 2021; Nonis et al., 2021).

Plant lectins are a superfamily of carbohydrate-binding proteins that serve as defense mechanisms against other plants and fungi (Peumans and Van Damme, 1995; Van Damme et al., 2008; Tsaneva and Van Damme, 2020). They are classified based on the number of carbohydrate-binding domains present in their mature sequences (Peumans and Van Damme, 1995; Van Damme et al., 2008). An example is the chitin-binding domain which interacts with chitin, a common naturally-occurring polysaccharide found in the exoskeleton of insects and the cell wall of fungi (Lenardon et al., 2010). Merolectins, chimerolectins, and hololectins represent three main types of plant lectins containing chitin-binding domains (Peumans et al., 2001; Porto et al., 2012).

Merolectins, the simplest and smallest form of lectins, have a single carbohydrate-binding domain of 29-46 amino acid residues (Peumans et al., 2001). A representative example of merolectin is hevein (hev), a cysteine-rich peptide (CRP) which is derived from the rubber tree (*Hevea brasiliensis*) and was the first reported chitin-binding peptide (Archer, 1960; Rodriguez-Romero et al., 1991; Van Parijs et al., 1991; Andersen et al., 1993; Gidrol et al., 1994). Chimerolectins, such as the *Urtica dioica* agglutinin (UDA), are chimeras which have singly- or tandemly-arrayed hev-peptide domains with a protein cargo such as chitinases (Does et al., 1999; Peumans et al., 2001). In contrast, hololectins have tandem repeats of hev-peptide domains (2–7 repeats), but lack a protein cargo (Peumans et al., 2001).

Heveins, hev-like peptides, and hev-peptide domains are characterized by a conserved cysteine motif, CXnCXnCCXnCXnC which possesses a tandemly connecting CC motif at CysIII and CysIV, and a cystine-knot disulfide connectivity (Rodriguez-Romero et al., 1991; Andersen et al., 1993). The chitin-binding site of heveins consists of a SXφGφ motif in intercystine loop 3, and in loop 4, a GXXXXφ motif (X represents any amino acid and φ represents aromatic acid residues, Phe, Tyr or Trp) (Tam et al., 2015). The conserved chitin-binding site (accession no. PS00026) recognizes and binds to planar chitin monomers (Kini et al., 2015, 2017).

Common cereals such as wheat, rye, barley, oats, corn, and rice contain a high concentration of hololectins (De Punder and Pruimboom, 2013). Representative examples include wheat germ agglutinin (WGA) from wheat, oryza sativa agglutinin (OSA) from rice, and barley hololectins (Wright et al., 1985; Lerner and Raikhel, 1989; Smith and Raikhel, 1989; Zhang et al., 2000). WGA, a 200-amino-acid protein, is present at ∼0.5 g/kg in wheat germ (Peumans and Van Damme, 1996) and contains four interconnected hev-peptide domains (Wright et al., 1985; Smith and Raikhel, 1989). The multivalent hev-peptide domains in WGA and other hololectins play important roles in their cell agglutination activities (Mishra et al., 2019). In addition, the avidity contributed by the repeating hev-peptide domains in WGA that binds strongly not only to chitin found in fungi, but also sialic acid found in the gastrointestinal tract, could trigger celiac diseases (Shaw et al., 1991; De Punder and Pruimboom, 2013). Thus, the presence of hololectins in edible cereal might be a health concern (De Punder and Pruimboom, 2013).

In general, cereal hololectins are processed and released as a protein with multiple hev-peptide domains from their respective precursors without further bioprocessing by an endopeptidase during their maturation (Peumans et al., 2001). Recently, our laboratory identified a cleavable hololectin from *Chenopodium quinoa*, a common edible pseudocereal (Loo et al., 2021b). The quinoa hololectin precursor which contains two hev-peptide domains is processed by a cathepsin-like endopeptidase to release hev-like chenotides, which are anti-fungal (Loo et al., 2021b). Thus, unlike most hololectin precursors which are resistant to proteolytic processing during their maturation, the quinoa hololectin precursors are cleaved to give two identical hev-like peptides. Apart from chenotides, the only other known example of cleavable-hololectins is Sm-Amp-1 from *Stellaria media* (Slavokhotova et al., 2017). However, the molecular basis underpinning the difference between cleavable and non-cleavable hololectins in their biosynthesis remains undetermined.

Here, we report the identification of a novel family of anti-fungal hev-like peptides from oats (*Avena sativa*) termed avenatides aV1–aV9 which are derived from hololectin precursors. Unlike majority of cereal hololectins which

are presented as proteoforms with multiple hev-domains, oat hololectins are presented as proteolytically processed proteoforms with a single hev-domain. A structural feature distinguishing these two families can be found in their precursor architecture and interdomain linkers that contribute to their different mechanisms of biosynthesis.

MATERIALS AND METHODS

Plant Material

All plant materials were purchased from local grocery stores, including *Avena sativa, Briza maxima, Cajanus cajan, Coix lacryma, Elymus canadensis, Glycine max, Phaseolus vulgaris, Hordeum jubatum, Hordeum vulgare, Secale cereale, Sorghum bicolor, Triticum aestivum, Vigna umbellate, Vigna radiata, Vigna unguiculata,* and *Vigna angularis*. Authentications were done by Mr. Paul Leong from the Singapore Botany Center based on macroscopic and microscopic analyses. Voucher samples were deposited at the Nanyang Technological University Herbarium, School of Biological Sciences, Singapore.

Isolation and Purification of Avenatide aV1

Small scale screening of oats and other cereals and legumes were performed by vortexing 0.1 g oats with 1 mL water for 1 h. The crude extract was centrifuged at 9,500 rpm for 10 min and the resulting supernatant was subjected to a C18 Zip-tip and eluted with 80% ACN.

For large scale extraction, 2 kg of oats were homogenized in 20 L of water for 3 h. The crude extracts were centrifuged at 9,500 rpm for 20 min at 4°C. The supernatant was filtered before loading on a flash column packed with 500 g C18 powder (Grace, MD, United States) in a Büchner funnel. Elution was performed using increasing concentrations of ethanol (20–80%). Eluents containing avenatide aV1 were pooled and purified using multiple rounds of SCX- and RP-HPLC in which fractions from SCX-HPLC containing avenatide aV1 were pooled and further purified by RP-HPLC. MALDI-TOF MS was used to identify the presence and assess the purity of avenatide aV1 in the eluted fractions.

Sequence Determination

The primary sequences of avenatides were determined by LC-MS/MS sequencing as described previously (Loo et al., 2016). The RP-HPLC-enriched avenatide samples were re-dissolved in 20 mM DTT at 37°C for 1 h followed by *S*-alkylation with 200 mM iodoacetamide at 37°C for 1 h. The mixture was desalted with a C18 Zip-tip and subjected to analysis on a Dionex UltiMate 3000 UHPLC system equipped with an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific Inc., Bremen, Germany). The composition of mobile phase A and B were 0.1% FA in deionized water and 0.1% FA in 90% acetonitrile with 10% deionized water, respectively. Mass spectra were acquired with LTQ Tune Plus software (Thermo Fisher Scientific, Bremen, Germany) using a positive mode with alternating Full FT-MS as previously described (Kumari et al., 2018). The data analysis were performed using PEAKS studio (version 7.519, Bioinformatics Solutions, Waterloo, ON, Canada) with a precursor ion tolerance of 10 ppm and fragment ion tolerance of 0.05 Da. Carbamidomethylation at Cys was set as a fixed modification. Deamidation of Asp and Glu, oxidation of Met, acetylation at Lys and N-term were set as variable modification. Peptide sequencing was performed using PEAKS DB protein identification which integrates database search of in-house eight-cysteine hev-peptide library with *de novo* sequencing using the following filtering parameters: peptide hit threshold (-10logP) was set as 30.0 and *de novo* score (%) threshold was set as 15. The data generated in this study are publicly available via the ProteomeXchange consortium through the partner repository JPOST (Okuda et al., 2017). JPOST accession: PXD033161.

Nuclear Magnetic Resonance Structural Study

All nuclear magnetic resonance (NMR) experiments were conducted on a BRUKER Avance 800 NMR spectrometer with a cryogenic probe at 25°C. The concentration of each peptide was around 1 mM, in a solution containing 5% D_2O and 95% H_2O (pH 3.5). For 1H, 1H-2D TOCSY and NOESY, the mixing times were 80 and 200 ms, respectively. The spectrum width was 12 ppm for both dimensions. The NMR spectra were processed using NMRPipe software (Delaglio et al., 1995). All data analysis were performed using Sparky software based on results of the 2D NOESY and TOCSY experiment. The proton chemical shift assignments for each amino acid residue were achieved by 2D TOCSY and NOESY while the proton-proton distances restraints were obtained from 2D NOESY based on the intensities of NOE cross-peaks. The solution structures of avenatide aV1 were calculated using CNSsolve 1.3 software. Proton-proton distance restraints and hydrogen bonds were employed in a standard simulated annealing protocol. The distance restraints were divided into three classes based on NOE cross-peak intensities: strong, $1.8 < d < 2.9$ Å, medium, $1.8 < d < 3.5$ Å and weak, 1.8 A $< d < 5$ Å. Eight hydrogen bonds were used in the structure calculation. A total of 100 structures were calculated and the 10 lowest energy structures were chosen for data statistics and presentation. The structure was verified using the PROCHECK program (Laskowski et al., 1996) and presented using Chimera version 1.6.2 (Huang et al., 1996; Pettersen et al., 2004). Accession code(s): PDB ID 6M5C.

Chitin-Binding Assay

A chitin-binding assay was performed as described previously (Loo et al., 2021b). Briefly, *S*-alkylated and purified avenatide aV1 were mixed with chitin beads (80 μL) (New England BioLabs, United Kingdom) in chitin binding buffer and incubated at 25°C for 30 min. After incubation, the mixture was washed with chitin binding buffer to remove unbound compounds. Elution of bound peptide was performed with 1 M acetic acid. The supernatant and eluent were analyzed using RP-UPLC and MALDI-TOF MS to assess binding and elution.

Peptide Stability Assay

Purified avenatides aV1 and *S*-alkylated aV1 were incubated under the stated conditions and recommended buffer solution. At each time interval, aliquots of samples were taken and RP-UPLC was performed in triplicate.

Exoproteolytic Enzyme Stability Assay

Purified avenatide aV1 (200 μM) were added to 50 mM Tris-HCl, 100 mM sodium chloride with 100 nM carboxypeptidase A or 20 mM tricine and 0.05% bovine serum albumin (pH 8.0) with 20 U/mL aminopeptidase I. The mixture was incubated in a 37°C water bath for 4 h. At each time-point (0 and 6 h), 20 μL of the treated sample was aliquoted and quenched with 5 μL 1 M hydrochloric acid. RP-UPLC was performed to determine the amounts of avenatide aV1 present before and after treatment.

Anti-fungal Assay

Four phyto-pathogenic fungal strains from the China Center of Industrial Culture Collection (Beijing, China) were used to examine the anti-fungal activity of avenatide aV1: *Alternaria alternata* (CICC 2465), *Curvularia lunata* (CICC 40301), *Fusarium oxysporum* (CICC 2532), and *Rhizoctonia solani* (CICC 40259). Fungal strains were grown on potato dextrose agar plates at 25°C.

The half maximal inhibitory concentration levels (IC_{50}) of avenatide aV1 were determined using a microbroth dilution assay (Loo et al., 2021b). Fungal spores were harvested from a 4-day old, actively growing fungal plate and suspended in half-strength potato dextrose broth. In the 96-well microplate, 1×10^5 cells/mL of spore suspension was mixed with peptides at varying concentrations and incubated at 25°C for 24 h. The cells were then fixed with 100% methanol for 15 min. Staining

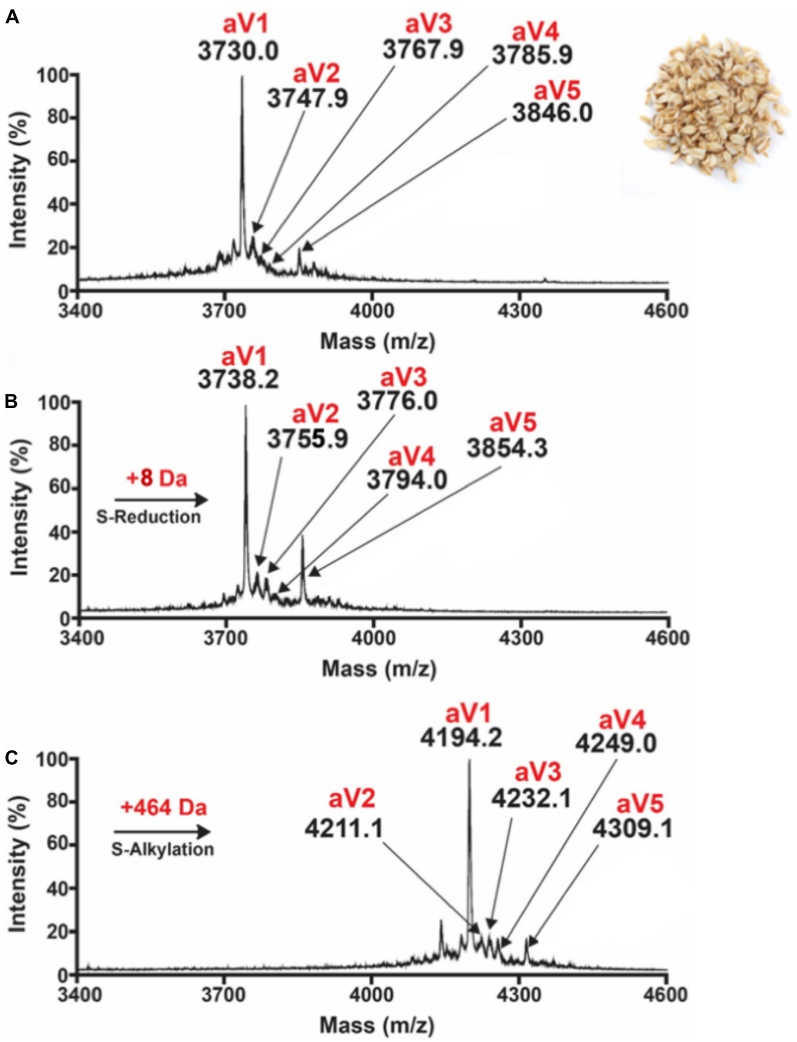

FIGURE 1 | MALDI-TOF MS profile of **(A)** aqueous oat extracts. Clusters of peaks between 2,000 and 4,000 Da indicate the presence of putative cysteine-rich peptides. Avenatides aV1–aV5 are labeled at corresponding peaks. **(B)** MALDI-TOF MS profile of aqueous oat extracts after *S*-reduction by DTT, and **(C)** *S*-alkylation by iodoacetamide (IAM) to give their corresponding linear forms and a gain of 58 Da for each *S*-alkylated Cys. Based on the increase of 464 Da, each avenatide is calculated to contain 8 cysteine residues.

was done for 45 min with crystal violet dye. MilliQ water was used to remove excess dye. Elution was performed using 1:1 (v/v) ethanol/0.1 N HCl. Absorbance was measured at 570 nm.

Data-Mining and Bioinformatics Analysis

Genes encoding avenatides, hololectin OAT_OCH1-5 (accession number: GO581539.1, GO581912.1, GO582252.1, GO583188.1, GO585827.1) were obtained from the NCBI GenBank and translated using the ExPaSy translation tool. Signal peptide cleavage sites were identified using SignalP 4.0 (Petersen et al., 2011).

Data-mining was performed to collect datasets for hev-peptide domain precursor sequences using a combination of motif searches and BLAST algorithms. For motif searches, the "Search Sequence Database" function in MOTIF Search[1] was used to interrogate Genbank, UniProt and RefSeq databases using a common hev-peptide domain input; C-{C}n-C-{C}n-C-C-{C}n-C-{C}n-C-{C}n-C, where "{C}" represents any amino acid except cysteine. For data-mining using BLAST, avenatide precursor sequences were used to perform BLAST on NCBI

[1]https://www.genome.jp/tools/motif/MOTIF2.html

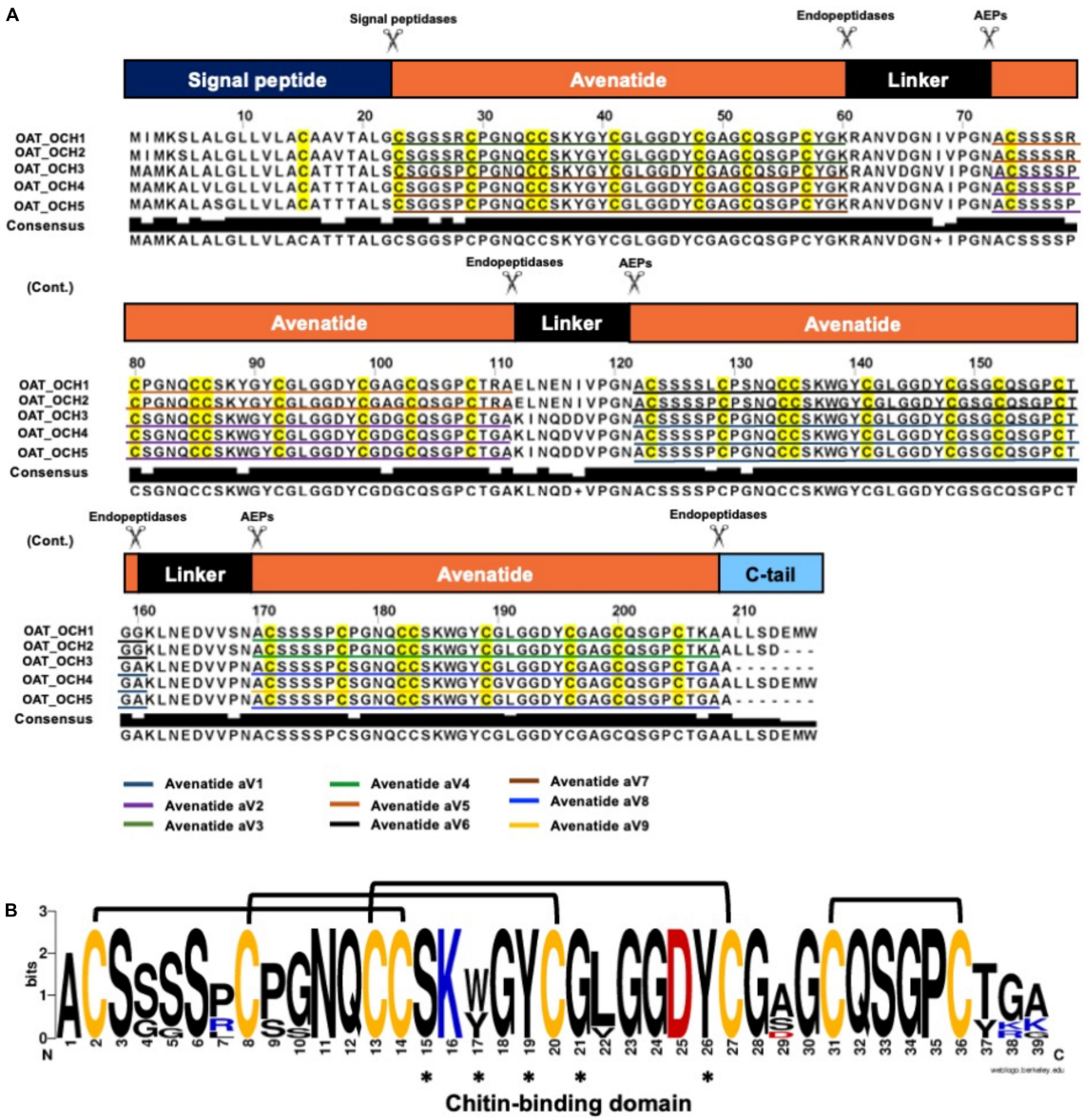

FIGURE 2 | Biosynthetic precursors and sequence comparison of avenatides. **(A)** Alignment of avenatide precursor sequences. Avenatide precursors contain an N-terminal signal peptide, four tandem repeating mature hev-peptide domains, and a C-terminal tail. The signal peptide is cleaved by signal peptidase, whereas the C-terminal domain and linkers are likely cleaved by an asparaginyl endopeptidase at the conserved Asn-Ala dipeptide motif to release the mature avenatide domain with Ala as the N-terminal amino acid in avenatides. **(B)** Consensus sequence of avenatides. The disulfide connectivity is based on the NMR-determined structure shown in **Figure 3** and amino acid residues contributing to the conserved chitin-binding sites are indicated by asterisks.

Genbank and OneKP. BLAST searches were confined to the taxa "Viridiplantae" (taxid:33090). Full sequences were retrieved from the databases and converted to.fasta formats. Then, sequences were re-interrogated using the "Search Motif Library" function with the Pfam motif library selected in MOTIF Search[2] to identify sequences with chitin-binding domains. Sequences were then sorted to identify hololectins containing multiple chitin-binding domains. Bioinformatics analysis was performed by identification of hololectin linkers which are defined as the sequences between two cysteinyl residues of two adjacent hev-peptide domains.

RESULTS

Mass Spectrometry Profiling of Hev-Peptides From Oats, Cereals and Legumes

We used a mass spectrometry (MS)-driven approach to profile hev-like peptides in the aqueous extract of oats, and selected cereals and legumes, including *B. maxima*, *C. cajan*, *C. lacryma*, *E. canadensis*, *G. max*, *P. vulgaris*, *H. jubatum*, *H. vulgare*, *S. cereale*, *S. bicolor*, *T. aestivum*, *V. umbellate*, *V. radiata*, *V. unguiculata*, and *V. angularis* (Loo et al., 2016, 2017, 2021a,b; Tam et al., 2018; Kam et al., 2019a,b; **Supplementary Figures 1–19**). Since hev-peptides contain 38–42 amino acids and eight cysteine residues, we focused on peaks around 4 kDa in their MS profiles. Also, because hev-peptides are CRPs, we performed a mass-shift assay to determine their mass increase, and in turn, their cysteine content. After *S*-reduction of cystine residues with dithiothreitol (DTT) and *S*-alkylation of the liberated cysteine residues with iodoacetamide (IAM), each *S*-alkylated cysteine residue results in an increase of 58 Da.

Preliminary profiling by MALDI-TOF-MS on 20 crop samples revealed the absence of prominent peaks around 4 kDa with eight cysteine residues in the aqueous extracts of *B. maxima*, *C. cajan*, *C. lacryma*, *E. canadensis*, *G. max*, *P. vulgaris*, *H. jubatum*, *H. vulgare*, *S. cereale*, *S. bicolor*, *T. aestivum*, *V. umbellate*, *V. radiata*, *V. unguiculata*, and *V. angularis*. In contrast, oats displayed prominent clusters of putative CRPs around 4,000 Da (**Figure 1A** and **Supplementary Figures 1–19**). From these clusters, five avenatides, termed aV1-aV5, with relative monoisotopic molecular weights $[M + H]^+$ of 3,730, 3,747, 3,767, 3,785, 3,846, and 3,858 Da, respectively, were identified and annotated. Mass-shift assay revealed a mass increase of 464 Da, indicating that avenatide aV1–aV5 are putative eight-Cys hev-peptides (**Figures 1B,C**).

Primary Sequence of Avenatides

To investigate the identities of avenatides, we used LC-MS/MS *de novo* sequencing assisted by database search with in-house hev-peptide library for determining the primary sequence of avenatides aV1-aV5 from oat extracts. This combined approach yielded the primary sequence of avenatide aV1 as A**C**SSSSP**C**PGNQ**CC**SKWGY**C**GLGGDY**C**GSG**C**QSGP**C**TGA.

[2]https://www.genome.jp/tools/motif/MOTIF.html

TABLE 1 | Sequence comparison of the mature peptide sequences of avenatides and reported eight-cystine-hevein-like peptides.

Peptide	Species	Amino acid sequence	Mass (Da)[a]	Charge[b]
Loop		I II III IV V VI		
aV1	*A. sativa*	---A--C---SSSSPCPGNQCCSKWGYCGLGGDYCGSG--CQSGPCTGA--	3,729	0
aV2	*A. sativa*	---A--C---SSSSPCSGNQCCSKWGYCGLGGDYCGDG--CQSGPCTGA--	3,747	−1
aV3	*A. sativa*	------C---SGSSRCPGNQCCSKYGYCGLGGDYCGAG--CQSGPCYGK--	3,767	+2
aV4	*A. sativa*	---A--C---SSSSPCPGNQCCSKWGYCGLGGDYCGAG--CQSGPCTKA--	3,784	+1
aV5	*A. sativa*	---A--C---SSSSRCPGNQCCSKYGYCGLGGDYCGAG--CQSGPCTRA--	3,848	+2
aV6	*A. sativa*	---A--C---SSSSPCPGNQCCSKWGYCGLGGDYCGSG--CQSGPCTGA--	3,740	+1
aV7	*A. sativa*	---A--C---SSSSLCPSNQCCSKWGYCGLGSDYCGSG--CQSGPCTGA--	3,805	0
aV8	*A. sativa*	------C---SGSSPCPGNQCCSKYGYCGLGGDYCGAG--CQSGPCYGK--	3,708	+1
aV9	*A. sativa*	---A--C---SSSSPCSGNQCCSKWGYCGVGGDYCGAG--CQSGPCTGA--	3,689	0
Avesin A	*A. sativa*	WSG---C------SPCPGNECCSKYGYCGLGGDYCGAG--CQSGPCYG---	3,680	−1
mO1	*M.oleifera*	----QNCGRQAGNRACANQLCCSQYGFCGSTSEYCSRANGCQSN-CRGG--	4,603	+3
mO2	*M.oleifera*	----QNCGRQAGNRACANGLCCSQYGFCGSTSEYCSRANGCQSN-CRGG--	4,532	+3
gB1	*G. biloba*	---DPTCSKLGD-FKCNPGRCCSKFNYCGSTAAYCGRGN-CIAQ-CPSNVS	4,715	+3
gB2	*G. biloba*	---DPTCSKLGD-FKCNPGRCCSKFNYCGSTAAYCGRGN-CIAQ-CPS---	4,417	+3
Hevein	*H. brasiliensis*	----EQCGRQAGGKLCPNNLCCSQWGWCGSTDEYCSPDHNCQSN-CKD---	4,715	−2

[a]*Mass (Da): represents the experimentally found molecular weight.*
[b]*Charge: represents the total charge of the molecule, and calculated by the sum of positive (lysine, arginine, and histidine residues) and negative (glutamate and aspartate residues) charges.*
Colored in green represents chitin binding domain. Colored in yellow represents cysteine motif.

In addition to avenatides aV1-aV5 (**Supplementary Figures 20, 21**), a database search from NCBI revealed the primary sequences of avenatides aV6–aV9 (**Figure 2**). The primary sequences of avenatides aV1–aV9 have 38–39 residues and are highly homologous (>90% similarity), with an overall charge ranging from -1 to +2. Similar to other hev-peptides and hev-peptide domains, all avenatides are both Cys- and Gly-rich (aV1 contains 9 Gly) and contain an evolutionarily conserved hev-peptide-like cysteine motif ($CX_nCX_nCCX_nCX_nC$) with a tandemly-connecting cysteine at the third and fourth position, and a chitin-binding site characterized by a highly conserved motif SK**X**(Y/W)G**Y** in intercysteine loop 3 (**Table 1**), followed by a GLGGD**Y** motif in loop 4 (the three invariable aromatic amino acids lettered in bold). Examples of other 8-Cys hev-peptides with similar cysteine motif and chitin binding site include ginkgotide gB1 from *Ginkgo biloba* and morintide mO1 from *Moringa olfeia* (Wong et al., 2016; Kini et al., 2017). In addition, the primary sequence of avenatides showed ∼70% similarity to avesin A, the first and only oat-derived hev-peptide domain, which was reported in 2003 (Li and Claeson, 2003). Avesin A is a 37-residue hev-peptide domain that has the primary sequence of WSG**C**SP**C**PGNE**CC**SKYGY**C**GLGGDY**C**GAG**C**QSGP**C**YG, and an overall -1 charge (Li and Claeson, 2003).

Biosynthesis and Precursor Architecture of Avenatides

From the NCBI database, we identified five full-length avenatide precursors, which we termed oat-cleavable hololectins (OCH1-5). **Figure 2A** shows the avenatide-precursors OCH1-5 as a three-domain hololectin precursor consisted of: (1) an N-terminal signal peptide, (2) four tandem-repeating mature domains of chitin-binding-avenatides joined by three linkers, and (3) a C-terminal tail. Sequence comparison showed that the tandem domains of OCH1 and OCH2 are identical and consisted of avenatides aV5, aV6, and aV4, with linkers between each avenatide domain. Similar architectural arrangements were found in OCH3-5 which consisted of avenatides aV7, aV2, aV1, aV8, and aV9. The linkers in avenatide precursors OCH1-5, also known as hinges or connecting peptides between two different avenatide domains, have 13–16 residues and are Asn/Asp-rich (> 20%). The presence of cleavage sites located at the N-terminal Asn-Ala residues of each avenatide domain suggests the involvement of asparaginyl endopeptidases in their bioprocessing to give avenatides aV1-9 with Ala as the N-terminus. Indeed, the predicted and calculated masses from the mass spectrometry of avenatides aV1, aV2, aV4, and aV5, further supported the AEP-mediated cleavage at the Asn-Ala dipeptide site. This finding suggested that the hydrolase activity of AEP could be involved in the release of each avenatide.

Structure of Avenatide aV1 and Biochemical Assay Its Chitin-Binding Activity

Avenatide aV1-aV9 contain a chitin-binding motif similar to other hev-peptides (**Figure 2B**). To confirm the disulfide connectivity of avenantide, we used NMR spectroscopy to determine the solution structure of avenatide aV1 (PDB: 6M5C). All spin-spin systems of avenatide aV1 were identified, and approximately 98% of the proton resonances were unambiguously assigned. The solution structure of avenatide aV1 was determined based on a total of 260 NMR-derived distance restraints and eight hydrogen bonds. The NMR ensemble of the 10 lowest-energy avenatide aV1 structures was determined (**Figure 3A**). The RMSD value of the 10 best structures for residues Ser3-Gly9 and Gly18-Thr37 was 1.23 ± 0.26 Å, and for all heavy atoms was 1.68 ± 0.27 Å (**Supplementary Tables 1, 2**). The structure of avenatide aV1 was well-defined by several medium- and long-range NOEs consisting of two short extended anti-parallel beta-strands (B1: Cys13-Ser15 and B2: Try19-Gly21) (**Figure 3B**). Avenatide aV1 has cystine-knot disulfide connectivity at its N terminus and an additional disulfide-bonded loop at its C terminus (CysI–CysIV, CysII–CysV, CysIII–CysVI,

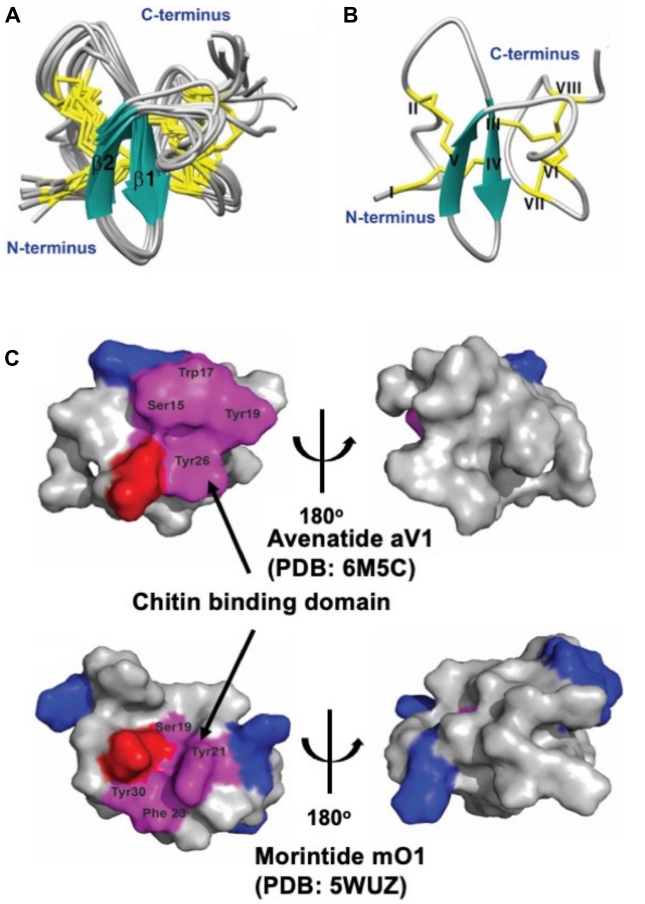

FIGURE 3 | Solution structure of avenatide aV1and comparison of the chitin-binding region with morintide mO1 which belongs to hevein-like domain. **(A)** Superposition of the avenatide aV1 backbone traces from the final 10 ensemble solution structures and restrained energy minimized structure. **(B)** Ribbon representation of avenatide aV1 structure. **(C)** Surface topology comparison of avenatide aV1 (PDB: 6M5C), and morintide mO1 (PDB: 5WUZ). Residues highlighted in purple represent the chitin-binding site. Residues highlighted in blue and red are basic (Arg, His, and Lys) and acidic (Asp, Glu), respectively.

CysVII–CysVIII). The N and C termini of avenatide aV1 are not close in proximity. Surface topography of the chitin binding site of avenatide aV1 (Ser-15, Trp-17, Tyr-19 and Tyr-26) revealed its similarity to morintide mO1, a hev-peptides isolated from *Moringa oleifera* (**Figure 3C**; Kini et al., 2017). Notably, the three conserved aromatic residues, Trp-17, Tyr-19, and Tyr-26 in the avenatide aV1 chitin-binding site, play an essential role in binding to planar chitin monomers (Kini et al., 2015, 2017).

To confirm the chitin-binding activity of avenatides, the representative avenatide aV1 together with the control, the linear *S*-alkylated aV1, were incubated with chitin beads at 25°C for 1 h. Analysis of the elution profiles by C18 reversed phase high-performance liquid chromatography (RP-HPLC) revealed a complete depletion of avenatide aV1 from the incubation solution, indicating its binding to the chitin beads (**Figure 4**). Avenatide aV1 binding activity was confirmed by elution with

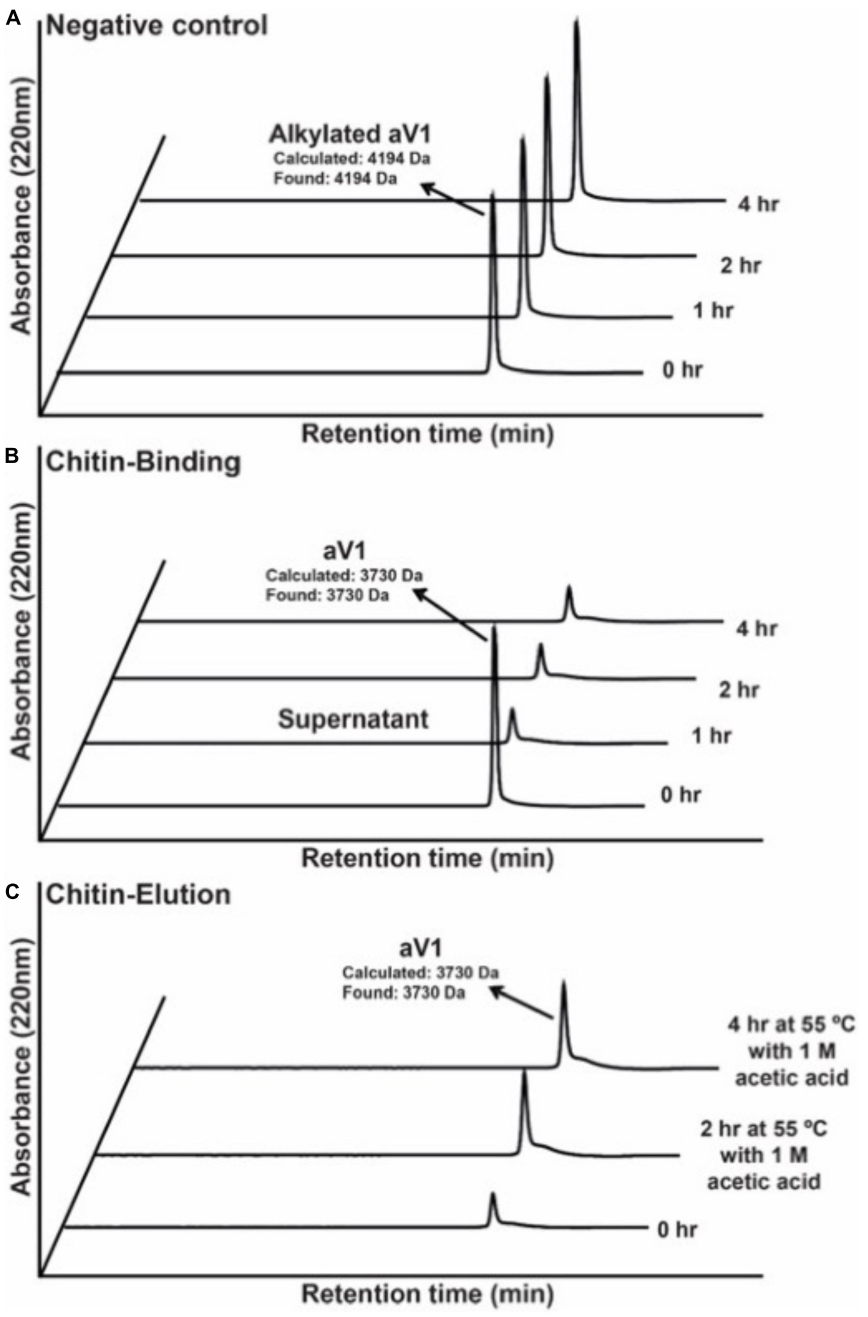

FIGURE 4 | Comparison of chitin-binding activities of **(A)** *S*-alkylated aV1 (iodoacetamido-) and **(B)** avenatide aV1 (iodoacetamido-) using chitin resins. The supernatants were analyzed by RP-HPLC. **(C)** Elution profile of avenatide aV1 from chitin resin using 1M acetic acid at 55°C. The supernatants were analyzed by RP-HPLC.

∼40% 1 M acetic acid at 55°C (**Figure 4**). In contrast, the linear *S*-alkylated aV1 was not retained by chitin beads, suggesting that the 3-dimensional structure of the aV1 chitin-binding site is important for its chitin-binding activities.

Avenatide aV1 Is Hyperstable

Cysteine-rich peptides cross-linked by multiple disulfides are known for their stability against heat, acid, and proteolytic degradation (Kini et al., 2015, 2017; Loo et al., 2016, 2017, 2021a,b; Wong et al., 2016; Tam et al., 2018; Kam et al., 2019a,b). To investigate the involvement of the disulfide scaffold on the stability of avenatides, we performed peptide degradation assays on folded avenatide aV1 and compared with its linearized *S*-alkylated form. Our results showed that avenatide aV1 is highly stable against heat-, acid-, endopeptidase- (represented by trypsin), and exopeptidase- (representing by carboxypeptidase A) mediated degradation. In all conditions, >80% of the peptides were retained after treatment, as monitored by RP-HPLC (**Figure 5**). In contrast, under the same conditions, the linearized *S*-alkylated avenatide aV1 showed substantial reduced stability, indicating the importance of the cystine-knot scaffold in conferring the hyperstability of avenatides.

Avenatide aV1 Is an Anti-fungal Peptide

To investigate the anti-fungal activities of avenatide aV1, we performed microbroth dilution assay using four phyto-pathogenic fungal strains of *A. alternata*, *C. lunata*, *F. oxysporum,* and *R. solani*. A microbroth dilution assay showed that avenatide aV1 has anti-fungal activity against all four fungal strains as evidenced by IC_{50} values of 239, 74, 53, and 77 μM for *A. alternata*, *C. lunata*, *F. oxysporum*, and *R. solani*, respectively (**Figure 6A**). To show that avenatide aV1 inhibits hyphae growth, *F. oxysporum* fungal spores were treated with different concentrations of avenatide aV1. Microscopic analysis revealed that avenatide aV1 stunted hyphae growth (**Figure 6B**).

Length of Interdomain Linkers in Protein Precursors Determine Cleavable- and Non-cleavable Hololectins in Cereals

To understand the difference between endopeptidase-susceptible and -resistant hololectins that give cleavable- and non-cleavable hololectins, we performed a BLAST search of hololectins containing hev-peptide domains using the NCBI and OneKP database (Johnson et al., 2008; Matasci et al., 2014). The search criteria include the presence of an evolutionarily conserved cysteine motif ($CX_nCX_nCCX_nCX_nC$) with a tandemly connecting CC motif at the position of CysIII and CysIV typical of a hev-peptide domain, and a chitin-binding site having a SXXG and GXXXXφ motif at the inter-cysteine loop 3 and 4, respectively (**Figure 2**). We further refined the search to identify putative AEP-susceptible hololectins based on the presence of Asn/Asp residues in their linkers.

A total of 121 hololectin precursor sequences with 280 hololectin linkers were identified from 44 plant species

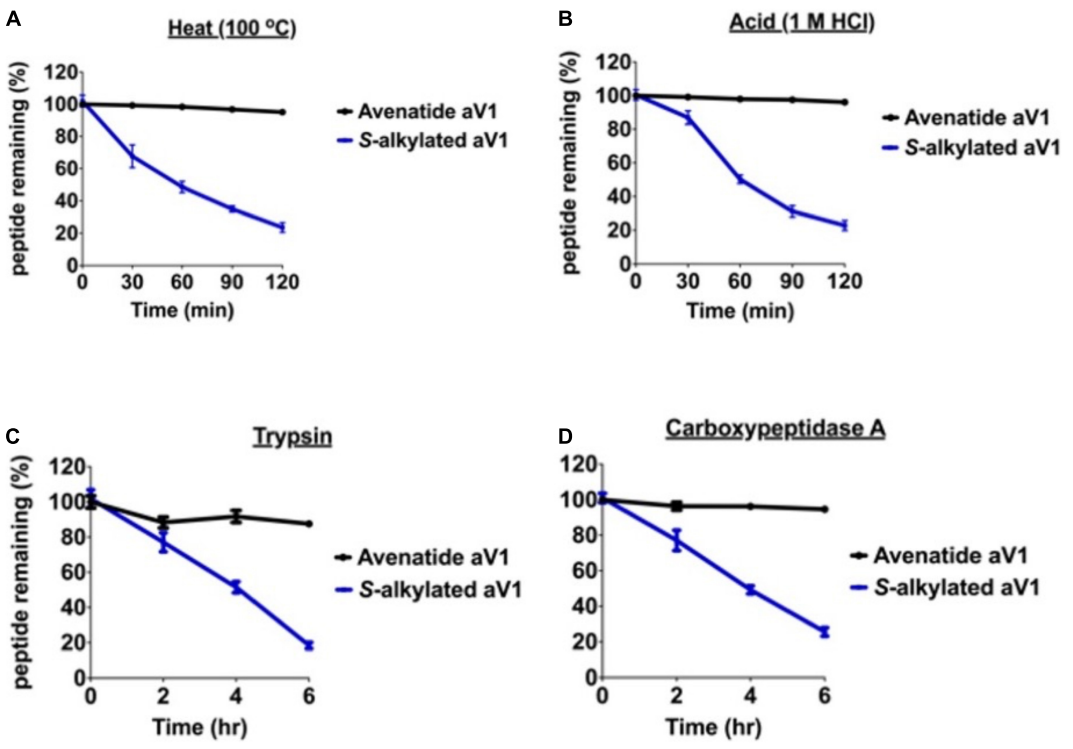

FIGURE 5 | Stability comparison of avenatide aV1 and *S*-alkylated aV1 (iodoacetamido-) under **(A)** heat (95°C), **(B)** acid (1M HCl), **(C)** trypsin, and **(D)** carboxypeptidase A treatment as analyzed by RP-HPLC (n = 3).

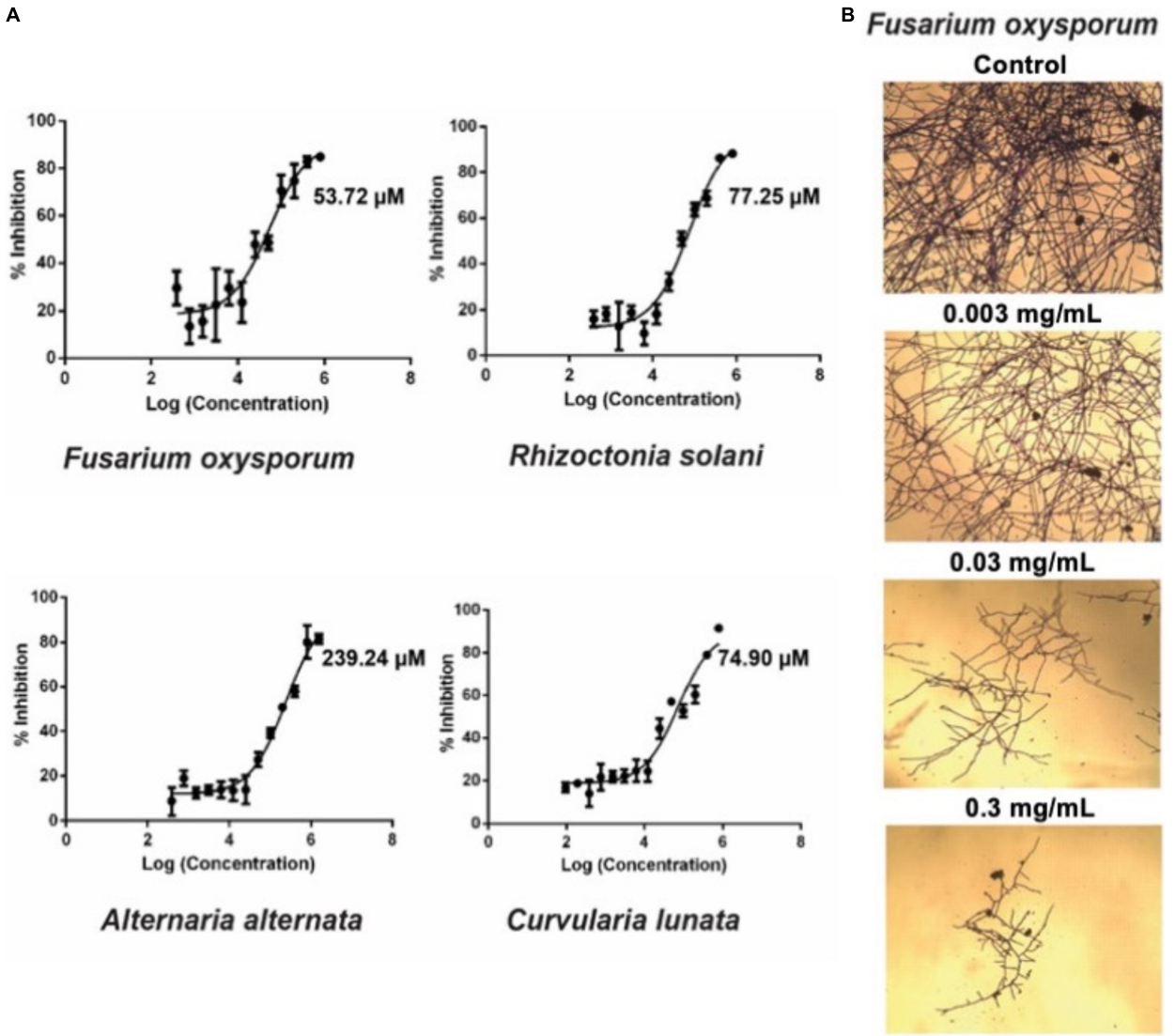

FIGURE 6 | Antifungal assays of avenatide aV1. **(A)** Fungal inhibition of avenatide aV1 against *Fusarium oxysporum*, *Rhizoctonia solani*, *Alternaria alternata*, and *Curvularia lunata*. The IC_{50} was calculated based on the dose-response curve obtained from the micro-broth dilution assay. **(B)** Bright field microscopy of hyphal growth inhibition with avenatide aV1. *F. oxysporum* treated with different concentrations of avenatide aV1. Formation of stunted hyphae ends indicated that avenatide aV1 inhibits hyphal growth at the ends of the fungal mycelia.

(**Supplementary Figure 22** and **Supplementary Tables 3–6**). After refinement, 194 putatively AEP-susceptible Asn/Asp-containing linkers of varying length were identified (**Figure 7**). The Asn/Asp-containing linkers were then sorted based on the number of amino acid residues in their linkers. These Asn/Asp-containing linkers formed two major clusters: long and short linkers. The long-linker cluster, containing 13–18 amino acid residues, are found in 32 hololectin linker members that include oats, with OCH1-5 accounting for 15 out of 32 putatively AEP-cleavable and Asn/Asp-containing linkers (**Figures 7**, **8**). In contrast, the short-linker cluster contains 1–6 amino acid residues from 152 hololectin linker members which include wheat, barley and millet (**Figures 7**, **8**). An interesting observation is that short Asn/Asp-containing linkers, but not the long linkers usually contain positive charge amino acids such as Lys and Arg (**Figure 7B**).

DISCUSSION

Cereals such as wheat, barley, rice, sorghum, and rye are a rich source of hev-hololectins, but not oats. This study provides an explanation to this divergence based on their biosynthesis. In contrast to many cereal-derived hololectins, oats have cleavable hev-hololectin precursors in which the mature domains of their gene products could be processed by an AEP and then released as small subunits in the form of hev-like peptides, avenatide aV1–aV9. A determining factor is found in their interdomain

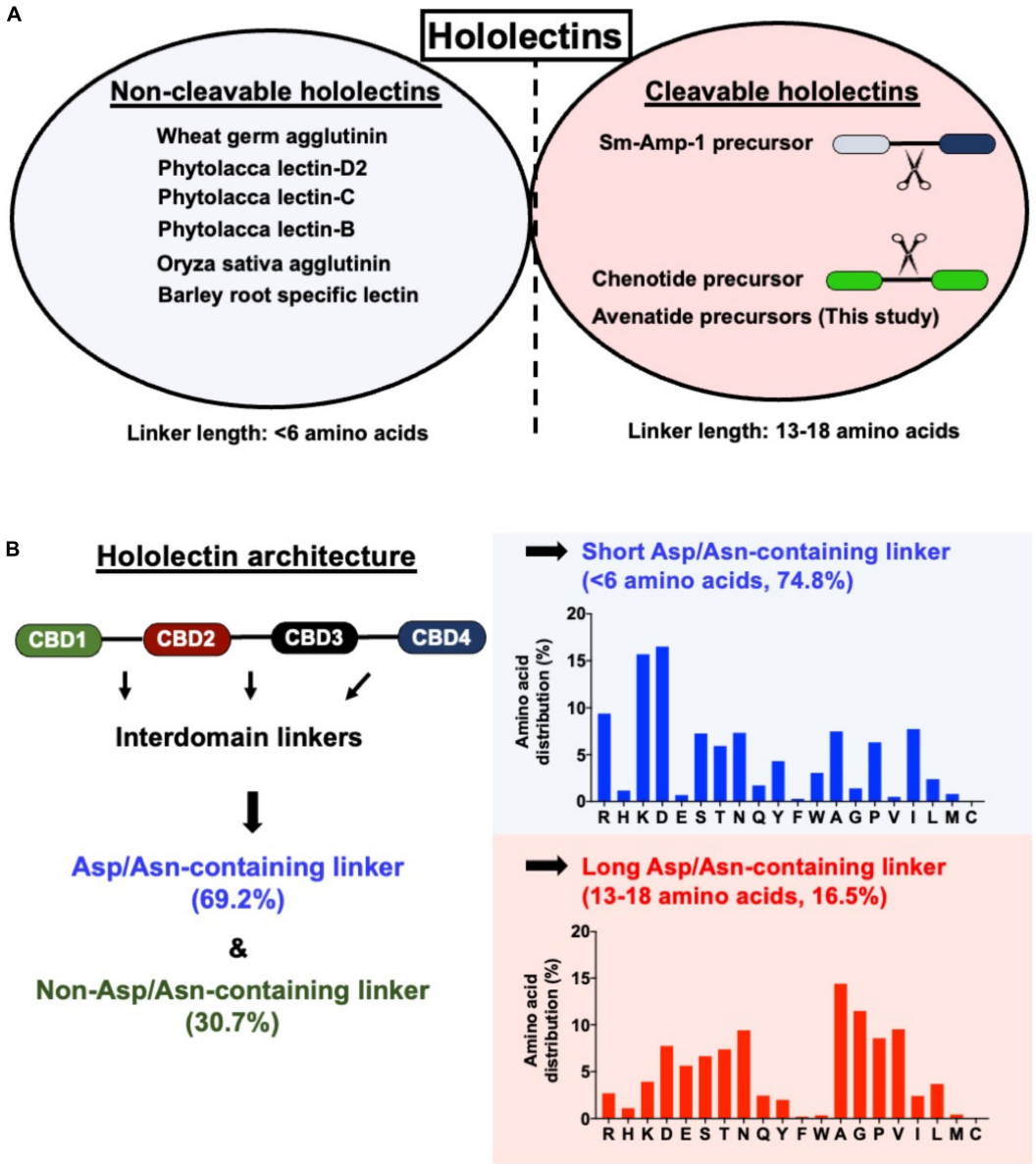

FIGURE 7 | (A) Comparison of linker length of reported cleavable hololectins (13–18 amino acids) and non-cleavable hololectins (4–6 amino acids). **(B)** Database search revealed a total of 121 hololectin precursor sequences with 280 hololectin linkers from NCBI and Onekp database. 69.2% are Asn/Asp-containing hololectin linkers that are susceptible to AEP processing. Among these Asn/Asp-containing hololectin linkers, 74.8% are classified as short linkers (<6 amino acids), and 16.5% are long linkers (13–18 amino acids). Short linkers, but not long linkers are rich in positive charge amino acids (e.g., Arg and Lys).

linker. Oat hev-hololectin precursors, such as OCH1-5, contain long, flexible and putatively AEP-susceptible linkers connecting the tandem-repeating hev-like domains. In contrast, most other cereals containing non-cleavable hololectin precursors have short linkers of 1-6 amino acid residues and which are resistant to bioprocessing. Thus, the interdomain linkers of hololectin precursors provide a key to distinguish cleavable from non-cleavable hololectins.

Avenatide aV1–aV9 with 8 Cys residues are 38-39 amino acids in length. They share a very high sequence similarity to each other (>90%) and a conserved chitin-binding site. Indeed, avenatide aV1 differs only 1 to 3 residues from the other seven members, aV2–aV9. Like other hev-peptides, avenatides are Gly-rich. Together, the 39-residue avenatide aV1 contains a total of 17 Cys and Gly residues, accounting for >43% of its amino acid composition. Like other hev-peptides, avenatides possess an evolutionarily conserved cysteine motif ($CX_nCX_nCCX_nCX_nC$) that is arranged in a disulfide-dense cystine-knot framework and a characteristic chitin-binding site as shown in **Figure 3** (Rodriguez-Romero et al., 1991; Andersen et al., 1993). Also, like other hev-peptides (Tam et al., 2015; Wong et al., 2016, 2017; Kini et al., 2017; Loo et al., 2021b), avenatide aV1 is

FIGURE 8 | Sequence alignment of AEP-resistant hololectin (e.g., WHEAT_AGI1, P10968; WHEAT_AGI2, P02876; WHEAT_AGI3, XP_037472823.1; BARLEY_AGI, P15312; MILLET_K3YJM9, XP_004972860.1) and putatively AEP-cleavable hololectin OAT_OCH1-5 (e.g., GO581539.1, GO581912.1, GO582252.1, GO583188.1, GO585827.1) precursor sequences from common cereals.

antifungal and capable in binding to chitin to inhibit phytopathogenic fungal strains. Thus, we can firmly conclude that avenatides belong to the family of hev-peptides based on their size (38–39 amino acids), abundance of glycine and cysteine residues, the presence of the characteristic cysteine motif, and the three essential aromatic amino acid residues which form the chitin-binding site.

The biosynthesis of hev-peptides can be broadly categorized into two major types based on their precursor architectures. The first type (type A) has a three-domain arrangement: a signal peptide, a single hev-peptide domain, and a C-terminal tail which can be short or long (protein-cargo carrying type). Examples of short hev-precursors include morintide mO1 and ginkgotide gB1, both of which contain a short C-terminal tail (<20 amino acids) (Wong et al., 2016; Kini et al., 2017). A subtype of the single-hev-peptide domain precursor is the cargo-carrying hev-peptide precursors with a long C-terminal cargo. Examples are hevein and EeCBP-1 whose precursors are significantly larger, and the C-terminal peptide is a functional protein (Lee et al., 1991; Van den Bergh K. et al., 2002; Van den Bergh K. P. et al., 2002).

The second architectural type (type B) belongs to the multi-modular hev-family in which their mature domains contain 2–7 hev-peptide domains (**Figure 9**). Type B family contains both precursors that are cleavable and non-cleavable (Loo et al., 2021b). Cleavable hev-hololectins are processed by an endopeptidase to release individual hev-peptide domains as hev-peptides. Examples include chenotides from *C. quinoa*, Sm-Amp-1 from *S. media*, and, as demonstrated in this study, avenatides from oats (Slavokhotova et al., 2017; Loo et al., 2021b). The type B precursor architecture, such as avenatide OCH1-5 is hololectin-like (>200 amino acid residues), comprising a signal peptide, four tandem repeats of highly similar hev-peptide domains connected by linkers, and a short C-terminal tail (Wright et al., 1985; Smith and Raikhel, 1989). This precursor architecture and

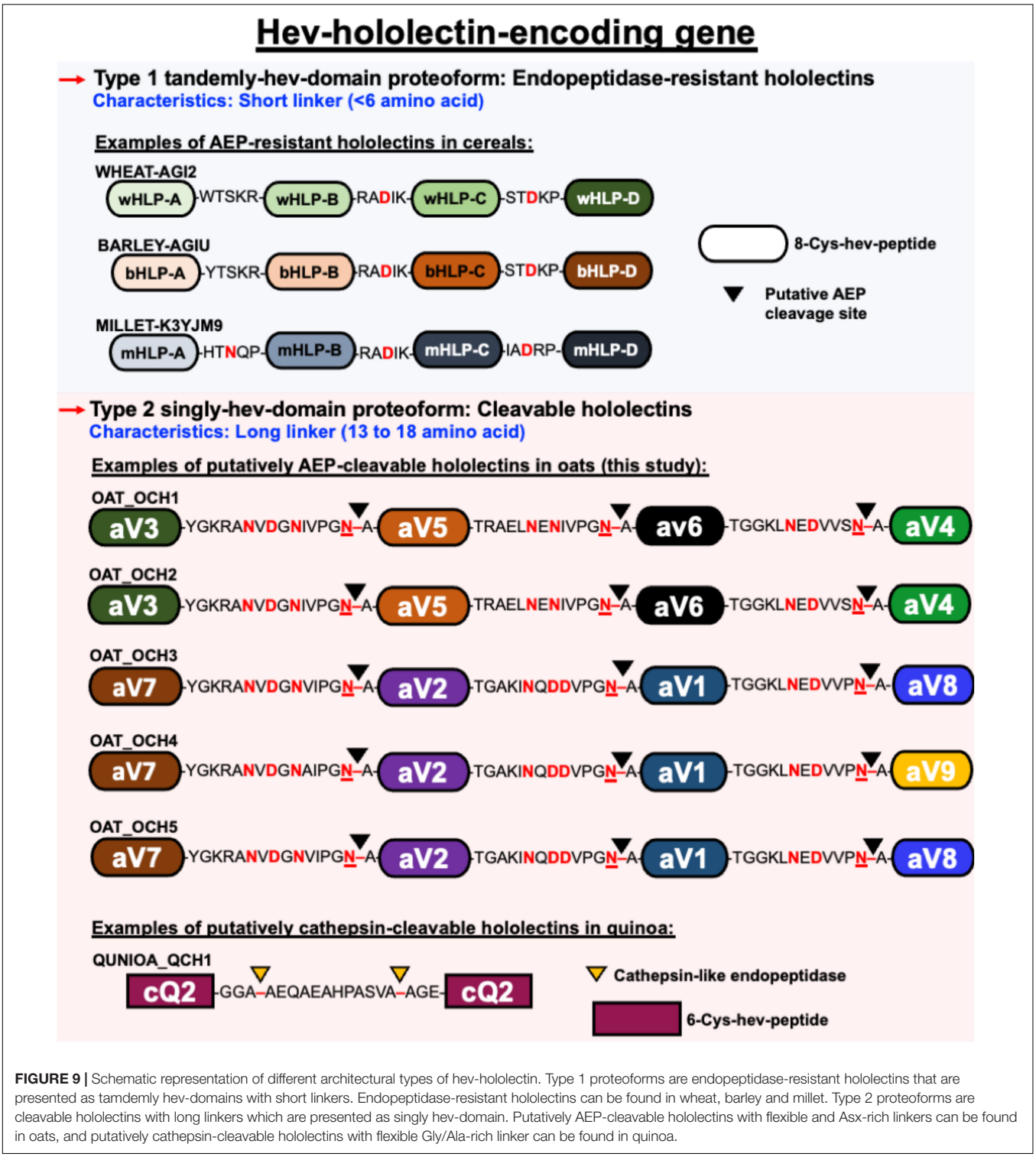

FIGURE 9 | Schematic representation of different architectural types of hev-hololectin. Type 1 proteoforms are endopeptidase-resistant hololectins that are presented as tamdemly hev-domains with short linkers. Endopeptidase-resistant hololectins can be found in wheat, barley and millet. Type 2 proteoforms are cleavable hololectins with long linkers which are presented as singly hev-domain. Putatively AEP-cleavable hololectins with flexible and Asx-rich linkers can be found in oats, and putatively cathepsin-cleavable hololectins with flexible Gly/Ala-rich linker can be found in quinoa.

length (>200 amino acid residues) is comparable to the cereal hololectins like WGA, OSA, and phytolacca lectins, all of which contain four hev-peptide domains (Wright et al., 1985; Smith and Raikhel, 1989; Yamaguchi et al., 1995, 1996; Yamaguchi et al., 1997; Zhang et al., 2000). However, the bioprocessing of oat hev-hololectin precursors differs from hololectins such as WGA and OSA, both of which are non-cleavable hev-hololectins. WGA and OSA are expressed as a single-chain multi-modular hev-protein proteoform.

A prominent feature that distinguishes cleavable from non-cleavable hololectins is the length of their linkers, sequences between two cysteinyl residues of two adjacent hev-peptide

domains (**Figure 9**). Non-cleavable hololectins have short linkers that are generally between 1 and 6 amino acids long. In contrast, cleavable hololectins have long linkers of 13–18 amino acids. The avenatide precursors OCH1-5 have linkers of 13–16 amino acids, similar in length to other cleavable hololectins (e.g., chenotides and Sm-Amp-1). The short linkers in non-cleavable hololectins could be crucial to maintain cooperative inter-domain interactions and to improve overall structural protein stability while preventing cleavage by endopeptidases. Long linkers, on the other hand, are likely to be susceptible to proteolytic cleavage, resulting in releasing individual mature tandem-repeating hev-peptide domains as hev-peptides.

The oat precursors OCH1-5 have long linkers that are rich in Asn and Asp (accounting for ∼23% of the total residues) and are susceptible to proteolytic cleavage by AEPs. Our predicted and calculated mass from mass spectrometry for avenatides aV1, aV2, aV4, and aV5 indicate that their N terminus is Ala and cleavage would occur between the conserved Asn-Ala dipeptide at the N-terminus of avenatide precursors. This finding suggests that the hydrolase activity of AEP could be involved in the release of each avenatide. Hololectins such as WGA, barley hololectin and millet hololectin also have Asn/Asp in their linkers which are short, ranging from 3 to 5 amino acids. Such short linkers could hinder AEP access, and in turn, prevent further processing to render them non-cleavable. Overall, avenatide precursors, with tandem-repeating hev-peptide domains, represent the first example of cleavable hololectins that could be bioprocessed by an AEP. Gene amplification through cleavable hololectin precursors could be an evolutionarily advantageous trait to boost the biosynthetic efficiency of these hev-peptide domains and benefit plant survival and reproduction (Panchy et al., 2016).

As functional foods, oats are less likely to cause celiac disease compared to other members of the cereal family (Rashid et al., 2007). Indeed, cereal lectins such as WGA from wheat and glutens are known to trigger celiac disease because the multivalent carbohydrate-binding properties of WGA promote interactions with sialic acid on surface tissues of the gastrointestinal tracts. Such interactions could result in intestinal inflammation and reduce nutrient absorption (De Punder and Pruimboom, 2013). The absence of hev-hololectins by the gut-friendly oats as a functional food could be a distinguish feature among cereals (Wright et al., 1985; Smith and Raikhel, 1989; Mishra et al., 2019). Furthermore, the combination cleavable precursor architecture and the need for AEP as a processing enzyme could increase the biosynthetic efficiency of hev-peptides in oats to provide an evolutionary advantage for plant survival and reproduction.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

SL, ST, AK, and WL designed, performed, and analyzed the experiments. SL, AK, and JT wrote the manuscript. All authors reviewed the results and approved the final version of the manuscript.

FUNDING

This research was supported in part by the Competitive Research Grant by Nanyang Technological University Internal Funding - Synzyme and Natural Products (SYNC) and the AcRF Tier 3 funding (MOE2016-T3-1-003). SL and AK are recipients of the Mistletoe Research Fellowship.

ACKNOWLEDGMENTS

The authors would like to thank Jing-Song Fan, Huang Jiayi, and Wong Kaho for their assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.899740/full#supplementary-material

REFERENCES

Andersen, N. H., Cao, B., Rodriguez-Romero, A., and Arreguin, B. (1993). Hevein: NMR assignment and assessment of solution-state folding for the agglutinin-toxin motif. *Biochemistry* 32, 1407–1422. doi: 10.1021/bi00057a004

Archer, B. L. (1960). The proteins of *Hevea brasiliensis* Latex. 4. isolation and characterization of crystalline hevein. *Biochem. J.* 75, 236–240. doi: 10.1042/bj0750236

Bowles, D. J., Marcus, S. E., Pappin, D., Findlay, J., Eliopoulos, E., Maycox, P. R., et al. (1986). Posttranslational processing of concanavalin A precursors in jackbean cotyledons. *J. Cell Biol.* 102, 1284–1297. doi: 10.1083/jcb.102.4.1284

Carrington, D., Auffret, A., and Hanke, D. (1985). Polypeptide ligation occurs during post-translational modification of concanavalin A. *Nature* 313, 64–67. doi: 10.1038/313064a0

Chen, Y., Zhang, D., Zhang, X., Wang, Z., Liu, C.-F., and Tam, J. P. (2021). Site-Specific protein modifications by an engineered Asparaginyl Endopeptidase from *Viola canadensis*. *Front. Chem.* 9:768854. doi: 10.3389/fchem.2021.768854

Csoma, C., and Polgár, L. (1984). Proteinase from germinating bean cotyledons. evidence for involvement of a thiol group in catalysis. *Biochem. J.* 222, 769–776. doi: 10.1042/bj2220769

Dall, E., and Brandstetter, H. (2016). Structure and function of legumain in health and disease. *Biochimie* 122, 126–150. doi: 10.1016/j.biochi.2015.09.022

De Punder, K., and Pruimboom, L. (2013). The dietary intake of wheat and other cereal grains and their role in inflammation. *Nutrients* 5, 771–787. doi: 10.3390/nu5030771

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277–293. doi: 10.1007/BF00197809

Does, M. P., Ng, D. K., Dekker, H. L., Peumans, W. J., Houterman, P. M., Van Damme, E. J., et al. (1999). Characterization of *Urtica dioica* agglutinin isolectins and the encoding gene family. *Plant Mol. Biol.* 39, 335–347. doi: 10.1023/a:1006134932290

Du, J., Yap, K., Chan, L. Y., Rehm, F. B., Looi, F. Y., Poth, A. G., et al. (2020). A bifunctional asparaginyl endopeptidase efficiently catalyzes both cleavage and cyclization of cyclic trypsin inhibitors. *Nat. Commun.* 11, 1–11. doi: 10.1038/s41467-020-15418-2

Fisher, M. F., Zhang, J., Taylor, N. L., Howard, M. J., Berkowitz, O., Debowski, A. W., et al. (2018). A family of small, cyclic peptides buried in preproalbumin since the Eocene epoch. *Plant Direct* 2:e00042. doi: 10.1002/pld3.42

Franke, B., Mylne, J., and Rosengren, K. (2018). Buried treasure: biosynthesis, structures and applications of cyclic peptides hidden in seed storage albumins. *Nat. Prod. Rep.* 35, 137–146. doi: 10.1039/c7np00066a

Gidrol, X., Chrestin, H., Tan, H. L., and Kush, A. (1994). Hevein, a lectin-like protein from *Hevea brasiliensis* (rubber tree) is involved in the coagulation of latex. *J. Biol. Chem.* 269, 9278–9283. doi: 10.1016/s0021-9258(17)37104-1

Harris, K. S., Durek, T., Kaas, Q., Poth, A. G., Gilding, E. K., Conlan, B. F., et al. (2015). Efficient backbone cyclization of linear peptides by a recombinant asparaginyl endopeptidase. *Nat. Commun.* 6, 1–10. doi: 10.1038/ncomms10199

Hemu, X., El Sahili, A., Hu, S., Wong, K., Chen, Y., Wong, Y. H., et al. (2019). Structural determinants for peptide-bond formation by asparaginyl ligases. *Proc. Natl. Acad. Sci. U.S.A* 116, 11737–11746. doi: 10.1073/pnas.1818568116

Huang, C. C., Couch, G. S., Pettersen, E. F., and Ferrin, T. E. (1996). "Chimera: an extensible molecular modeling application constructed using standard components," in *Proceedings of the Pacific Symposium on Biocomputing* (Singapore: World Scientific) 724.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and Madden, T. L. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Res.* 36, W5–W9. doi: 10.1093/nar/gkn201

Kam, A., Loo, S., Dutta, B., Sze, S. K., and Tam, J. P. (2019a). Plant-derived mitochondria-targeting cysteine-rich peptide modulates cellular bioenergetics. *J. Biol. Chem.* 294, 4000–4011. doi: 10.1074/jbc.RA118.006693

Kam, A., Loo, S., Fan, J.-S., Sze, S. K., Yang, D., and Tam, J. P. (2019b). Roseltide rT7 is a disulfide-rich, anionic, and cell-penetrating peptide that inhibits proteasomal degradation. *J. Biol. Chem.* 294, 19604–19615. doi: 10.1074/jbc.RA119.010796

Kembhavi, A. A., Buttle, D. J., Knight, C. G., and Barrett, A. J. (1993). The two cysteine endopeptidases of legume seeds: purification and characterization by use of specific fluorometric assays. *Arch. Biochem. Biophys.* 303, 208–213. doi: 10.1006/abbi.1993.1274

Kini, S. G., Nguyen, P. Q., Weissbach, S., Mallagaray, A., Shin, J., Yoon, H. S., et al. (2015). Studies on the chitin binding property of novel cysteine-rich peptides from *Alternanthera sessilis*. *Biochemistry* 54, 6639–6649. doi: 10.1021/acs.biochem.5b00872

Kini, S. G., Wong, K. H., Tan, W. L., Xiao, T., and Tam, J. P. (2017). Morintides: cargo-free chitin-binding peptides from *Moringa oleifera*. *BMC Plant Biol.* 17:68. doi: 10.1186/s12870-017-1014-6

Kumari, G., Wong, K. H., Serra, A., Shin, J., Yoon, H. S., Sze, S. K., et al. (2018). Molecular diversity and function of jasmintides from *Jasminum sambac*. *BMC Plant Biol.* 18:144. doi: 10.1186/s12870-018-1361-y

Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–486. doi: 10.1007/BF00228148

Lee, H., Broekaert, W., Raikhel, N., and Lee, H. (1991). Co-and post-translational processing of the hevein preproprotein of latex of the rubber tree (*Hevea brasiliensis*). *J. Biol. Chem.* 266, 15944–15948. doi: 10.1016/s0021-9258(18)98499-1

Lenardon, M. D., Munro, C. A., and Gow, N. A. (2010). Chitin synthesis and fungal pathogenesis. *Curr. Opin. Microbiol.* 13, 416–423. doi: 10.1016/j.mib.2010.05.002

Lerner, D. R., and Raikhel, N. V. (1989). Cloning and characterization of root-specific barley lectin. *Plant physiol.* 91, 124–129. doi: 10.1104/pp.91.1.124

Li, S.-S., and Claeson, P. (2003). Cys/Gly-rich proteins with a putative single chitin-binding domain from oat (*Avena sativa*) seeds. *Phytochemistry* 63, 249–255. doi: 10.1016/s0031-9422(03)00116-x

Liew, H. T., To, J., Zhang, X., Hemu, X., Chan, N.-Y., Serra, A., et al. (2021). The legumain McPAL1 from *Momordica cochinchinensis* is a highly stable Asx-specific splicing enzyme. *J. Biol. Chem.* 297:101325. doi: 10.1016/j.jbc.2021.101325

Loo, S., Tay, S. V., Kam, A., Tang, F., Fan, J.-S., Yang, D., et al. (2021b). Anti-Fungal hevein-like peptides biosynthesized from quinoa cleavable hololectins. *Molecules* 26:5909. doi: 10.3390/molecules26195909

Loo, S., Kam, A., Li, B. B., Feng, N., Wang, X., and Tam, J. P. (2021a). Discovery of hyperstable noncanonical plant-derived epidermal growth factor receptor agonist and analogs. *J. Med. Chem.* 64, 7746–7759. doi: 10.1021/acs.jmedchem.1c00551

Loo, S., Kam, A., Xiao, T., and Tam, J. P. (2017). Bleogens: cactus-derived anti-candida cysteine-rich peptides with three different precursor arrangements. *Front. Plant Sci.* 8:2162. doi: 10.3389/fpls.2017.02162

Loo, S., Kam, A., Xiao, T., Nguyen, G. K., Liu, C. F., and Tam, J. P. (2016). Identification and characterization of roseltide, a knottin-type neutrophil elastase inhibitor derived from *Hibiscus sabdariffa*. *Sci. Rep.* 6:39401. doi: 10.1038/srep39401

Manoury, B., Hewitt, E. W., Morrice, N., Dando, P. M., Barrett, A. J., and Watts, C. (1998). An asparaginyl endopeptidase processes a microbial antigen for class II MHC presentation. *Nature* 396, 695–699. doi: 10.1038/25379

Manoury, B., Mazzeo, D., Fugger, L., Viner, N., Ponsford, M., Streeter, H., et al. (2002). Destructive processing by asparagine endopeptidase limits presentation of a dominant T cell epitope in MBP. *Nat. Immunol.* 3, 169–174. doi: 10.1038/ni754

Matasci, N., Hung, L.-H., Yan, Z., Carpenter, E. J., Wickett, N. J., Mirarab, S., et al. (2014). Data access for the 1,000 Plants (1KP) project. *Gigascience* 3:17. doi: 10.1186/2047-217X-3-17

Mishra, A., Behura, A., Mawatwal, S., Kumar, A., Naik, L., Mohanty, S. S., et al. (2019). Structure-function and application of plant lectins in disease biology and immunity. *Food Chem. Toxicol.* 134:110827. doi: 10.1016/j.fct.2019.110827

Mylne, J. S., Chan, L. Y., Chanson, A. H., Daly, N. L., Schaefer, H., Bailey, T. L., et al. (2012). Cyclic peptides arising by evolutionary parallelism via asparaginyl-endopeptidase–mediated biosynthesis. *Plant Cell* 24, 2765–2778. doi: 10.1105/tpc.112.099085

Mylne, J. S., Colgrave, M. L., Daly, N. L., Chanson, A. H., Elliott, A. G., McCallum, E. J., et al. (2011). Albumins and their processing machinery are hijacked for cyclic peptides in sunflower. *Nat. Chem. Biol.* 7, 257–259. doi: 10.1038/nchembio.542

Nguyen, G. K., Kam, A., Loo, S., Jansson, A. E., Pan, L. X., and Tam, J. P. (2015). Butelase 1: a versatile ligase for peptide and protein macrocyclization. *J. Am. Chem. Soc.* 137, 15398–15401. doi: 10.1021/jacs.5b11014

Nguyen, G. K., Wang, S., Qiu, Y., Hemu, X., Lian, Y., and Tam, J. P. (2014). Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat. chem. Biol.* 10, 732–738. doi: 10.1038/nchembio.1586

Nonis, S. G., Haywood, J., Schmidberger, J. W., Mackie, E. R., da Costa, T. P., Bond, C. S., et al. (2021). Structural and biochemical analyses of concanavalin A circular permutation by jack bean asparaginyl endopeptidase. *Plant Cell* 33, 2794–2811. doi: 10.1093/plcell/koab130

Okuda, S., Watanabe, Y., Moriya, Y., Kawano, S., Yamamoto, T., Matsumoto, M., et al. (2017). jPOSTrepo: an international standard data repository for proteomes. *Nucleic Acids Res.* 45, D1107–D1111. doi: 10.1093/nar/gkw1080

Panchy, N., Lehti-Shiu, M., and Shiu, S. H. (2016). Evolution of gene duplication in plants. *Plant Physiol.* 171, 2294–2316. doi: 10.1104/pp.16.00523

Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi: 10.1038/nmeth.1701

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. doi: 10.1002/jcc.20084

Peumans, W. J., and Van Damme, E. (1995). Lectins as plant defense proteins. *Plant Physiol.* 109, 347–352. doi: 10.1104/pp.109.2.347

Peumans, W. J., Van Damme, J., Barre, A., and Rougé, P. (2001). Classification of plant lectins in families of structurally and evolutionary related proteins. *Adv. Exp. Med. Biol.* 491, 27–54. doi: 10.1007/978-1-4615-1267-7_3

Peumans, W., and Van Damme, E. (1996). Prevalence, biological activity and genetic manipulation of lectins in foods. *Trends Food Sci. Technol.* 7, 132–138. doi: 10.1016/0924-2244(96)10015-7

Porto, W. F., Souza, V. A., Nolasco, D. O., and Franco, O. L. (2012). In silico identification of novel hevein-like peptide precursors. *Peptides* 38, 127–136. doi: 10.1016/j.peptides.2012.07.025

Rashid, M., Butzner, D., Burrows, V., Zarkadas, M., Case, S., Molloy, M., et al. (2007). Consumption of pure oats by individuals with celiac disease: a position statement by the Canadian Celiac Association. *Can. J. Gastroenterol. Hepatol.* 21, 649–651. doi: 10.1155/2007/340591

Rawlings, N. D., Barrett, A. J., and Bateman, A. (2010). MEROPS: the peptidase database. *Nucleic Acids Res.* 38, D227–D233.

Rodriguez-Romero, A., Ravichandran, K. G., and Soriano-Garcia, M. (1991). Crystal structure of hevein at 2.8 a resolution. *FEBS Lett.* 291, 307–309. doi: 10.1016/0014-5793(91)81308-u

Shaw, L., Yousefi, S., Dennis, J. W., and Schauer, R. (1991). CMP-N-acetylneuraminic acid hydroxylase activity determines the wheat germ agglutinin-binding phenotype in two mutants of the lymphoma cell line MDAY-D2. *Glycoconj. J.* 8, 434–441. doi: 10.1007/BF00731295

Slavokhotova, A. A., Shelenkov, A. A., Korostyleva, T. V., Rogozhin, E. A., Melnikova, N. V., Kudryavtseva, A. V., et al. (2017). Defense peptide repertoire of Stellaria media predicted by high throughput next generation sequencing. *Biochimie* 135, 15–27. doi: 10.1016/j.biochi.2016.12.017

Smith, J. J., and Raikhel, N. V. (1989). Nucleotide sequences of cDNA clones encoding wheat germ agglutinin isolectins A and D. *Plant Mol. Biol.* 13, 601–603. doi: 10.1007/BF00027321

Tam, J. P., Nguyen, G. K., Loo, S., Wang, S., Yang, D., and Kam, A. (2018). Ginsentides: cysteine and glycine-rich peptides from the ginseng family with unusual disulfide connectivity. *Sci. Rep.* 8, 1–15. doi: 10.1038/s41598-018-33894-x

Tam, J. P., Wang, S., Wong, K. H., and Tan, W. L. (2015). Antimicrobial peptides from plants. *Pharmaceuticals* 8, 711–757. doi: 10.3390/ph8040711

Tsaneva, M., and Van Damme, E. J. (2020). 130 years of Plant Lectin Research. *Glycoconj. J.* 37, 533–551. doi: 10.1007/s10719-020-09942-y

Van Damme, E. J., Lannoo, N., and Peumans, W. J. (2008). Plant lectins. *Adv. Bot. Res.* 48, 107–209.

Van den Bergh, K. P., Proost, P., Van Damme, J., Coosemans, J., Van Damme, E. J., and Peumans, W. J. (2002). Five disulfide bridges stabilize a hevein-type antimicrobial peptide from the bark of spindle tree (*Euonymus europaeus* L.). *FEBS lett.* 530, 181–185. doi: 10.1016/s0014-5793(02)03474-9

Van den Bergh, K., Van Damme, E., Peumans, W., and Coosemans, J. (2002). Ee-CBP, a hevein-type antimicrobial peptide from bark of the spindle tree (*Euonymus europaeus* L.). *Meded. Rijksuniv. Gent. Fak. Landbouwkd. Toegep. Biol. Wet.* 67, 327–331.

Van Parijs, J., Broekaert, W. F., Goldstein, I. J., and Peumans, W. J. (1991). Hevein: an antifungal protein from rubber-tree (*Hevea brasiliensis*) latex. *Planta* 183, 258–264. doi: 10.1007/BF00197797

Vorster, B. J., Cullis, C. A., and Kunert, K. J. (2019). Plant vacuolar processing enzymes. *Front. Plant Sci.* 10:479. doi: 10.3389/fpls.2019.00479

Wong, K. H., Tan, W. L., Kini, S. G., Xiao, T., Serra, A., Sze, S. K., et al. (2017). Vaccatides: antifungal glutamine-rich hevein-like peptides from *Vaccaria hispanica*. *Front. Plant Sci.* 8:1100. doi: 10.3389/fpls.2017.01100

Wong, K. H., Tan, W. L., Serra, A., Xiao, T., Sze, S. K., Yang, D., et al. (2016). Ginkgotides: proline-rich hevein-like peptides from gymnosperm *Ginkgo biloba*. *Front. Plant Sci.* 7:1639. doi: 10.3389/fpls.2016.01639

Wright, H. T., Brooks, D. M., and Wright, C. S. (1985). Evolution of the multidomain protein wheat germ agglutinin. *J. Mol. Evol.* 21, 133–138. doi: 10.1007/BF02100087

Yamada, K., Basak, A. K., Goto-Yamada, S., Tarnawska-Glatt, K., and Hara-Nishimura. (2020). Vacuolar processing enzymes in the plant life cycle. *New Phytol.* 226, 21–31. doi: 10.1111/nph.16306

Yamaguchi, K., Mori, A., and Funatsu, G. (1995). The complete amino acid sequence of lectin-C from the roots of pokeweed (*Phytolacca americana*). *Biosci. Biotechnol. Biochem.* 59, 1384–1385. doi: 10.1271/bbb.59.1384

Yamaguchi, K., Mori, A., and Funatsu, G. (1996). Amino acid sequence and some properties of lectin-D from the roots of pokeweed (*Phytolacca americana*). *Biosci. Biotechnol. Biochem.* 60, 1380–1382. doi: 10.1271/bbb.60.1380

Yamaguchi, K., Yurino, N., Kino, M., Ishiguro, M., and Funatsu, G. (1997). The amino acid sequence of mitogenic lectin-B from the roots of pokeweed (*Phytolacca americana*). *Biosci. Biotechnol. Biochem.* 61, 690–698. doi: 10.1271/bbb.61.690

Zhang, W., Peumans, W. J., Barre, A., Astoul, C. H., Rovira, P., Rougé, P., et al. (2000). Isolation and characterization of a jacalin-related mannose-binding lectin from salt-stressed rice (*Oryza sativa*) plants. *Planta* 210, 970–978. doi: 10.1007/s004250050705

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.