ORIGINAL PAPER

G. Raedschelders · C. Debefve · H. Goesaert ·
J. A. Delcour · G. Volckaert · S. Van Campenhout

Molecular identification and chromosomal localization of genes encoding *Triticum aestivum* xylanase inhibitor I-like proteins in cereals

Abstract TAXI (*Triticum aestivum* xylanase inhibitor) proteins are present in wheat flour and are known to inhibit glycosyl hydrolase family 11 endoxylanases, enzymes which are commonly applied in grain processing. Here, we describe the PCR-based molecular identification of genes encoding endoxylanase inhibitors HVXI and SCXI, the TAXI-like proteins from barley (*Hordeum vulgare*) and rye (*Secale cereale*) respectively. The HVXI coding sequence encodes a mature protein of 384 amino acids preceded by a 19 amino acid long signal sequence. SCXI-II/III has an open reading frame encoding a signal peptide of 21 amino acids and a mature protein of 375 amino acids. As for TAXI-I, no introns were detected in the untranslated regions and coding sequences identified. These newly identified sequences allowed us to perform a multiple sequence alignment with TAXI-I and similar proteins. Rice TAXI-type proteins clustered together with the cereal endoxylanase inhibitors. Dicotyledonous proteins with sequence similarity to TAXI-I, including the tomato xyloglucan-specific endoglucanase inhibiting protein, formed a different clade. The TAXI-type proteins may hence be part of a superfamily of proteins all involved in plant responses to biotic or abiotic stress and for which a function as glycosyl hydrolase inhibitors can be suggested. The chromosomal localization of the TAXI-I gene identified on wheat chromosome 3B, of the SCXI-II/III gene identified on rye chromosome 6R, and the presence of a cluster of TAXI-like genes on rice chromosome 1, allowed us to assign the location of TAXI-like genes to the wheat-rye translocation area 3BL/6RL characterized by RFLP markers *XGlb33* and *Xpsr454* and isozyme *Est-5*. In rice, RFLP marker C1310S corresponds to a TAXI-like protein encoding sequence.

G. Raedschelders · C. Debefve · G. Volckaert ·
S. Van Campenhout (✉)
Laboratory of Gene Technology,
Katholieke Universiteit Leuven,
Kasteelpark Arenberg 21, 3001 Leuven, Belgium
e-mail: steven.vancampenhout@agr.kuleuven.ac.be
Tel.: +32-16-329672
Fax: +32-16-321965

H. Goesaert · J. A. Delcour
Laboratory of Food Chemistry,
Katholieke Universiteit Leuven,
Kasteelpark Arenberg 20, 3001 Leuven, Belgium

Introduction

Most cereals contain arabinoxylan as a structural component in their cell walls. Together with cellulose, (1,3;1,4)-β-glucan, arabinogalactanpeptide and other minor constituents, they are referred to as non-starch polysaccharides. Although they are minor constituents of cereal grains (8.5% in rye, 6.6% in wheat and 5.7% in barley) (Henry 1985), they significantly affect cereal processing and/or end use quality (Courtin and Delcour 2002). In particular, the arabinoxylan fraction in wheat and rye is of importance for their breadmaking qualities and for the nutritional properties of food and feed. It is also important for the functionality of barley, as undegraded arabinoxylan, extracted or solubilized from malt, contributes to wort viscosity and lower filtration rates (Viëtor et al. 1993; Leclercq et al. 1999) and is involved in the formation of certain types of beer hazes (Coote and Kirsop 1976).

Endo-(1-4)-β-xylanases (EC 3.2.1.8), also called (endo)xylanases, play a primary role in arabinoxylan hydrolysis, being responsible for hydrolysis of the xylan backbone. Based on amino acid sequence similarities and hydrophobic cluster analysis, these endoxylanases have been grouped into two glycosyl hydrolase classes, families 10 and 11 (Henrissat 1991), which differ in their molecular structures, molecular masses and catalytic properties (Jeffries 1996; Biely et al. 1997). The known cereal endogenous endoxylanases all belong to family 10 and play an important role during germination by degrading the endosperm cell walls (Simpson et al. 2003). Microbial endoxylanases can belong to family 10 as well as 11 and are frequently used to modify the functionality of (arabino)xylan in biotechnological processes such as

breadmaking (Courtin and Delcour 2002), wheat gluten-starch separation (Christophersen et al. 1997) and paper and pulp production (Suurnäkki et al. 1997; Christov et al. 1999). They are also used as components in animal feed preparations (Bedford and Schulze 1998).

The recent discovery of a novel class of cereal proteins that inhibit endoxylanases and thereby affect their functionality in the above cited processes and applications is an important development. In the literature, two structurally different cereal endoxylanase inhibitors are described, namely the TAXI-type (*Triticum aestivum* endoxylanase inhibitor) (Debyser et al. 1997, 1999; Debyser and Delcour 1998; Gebruers et al. 2001, 2004; Goesaert et al. 2001, 2002) and the XIP-type (endoxylanase inhibiting protein) (McLauchlan et al. 1999; Elliott et al. 2002, 2003; Goesaert et al. 2003b) endoxylanase inhibitors. Both types of inhibitors were isolated from wheat (*T. aestivum*) (Debyser and Delcour 1998; McLauchlan et al. 1999; Gebruers et al. 2001, 2002b), durum wheat (*Triticum durum*) (Goesaert et al. 2003a, 2003b), barley (*Hordeum vulgare*) (Goesaert et al. 2001, 2003a, 2003b) and rye (*Secale cereale*) (Goesaert et al. 2002, 2003a, 2003b). In general, all isolated inhibitor fractions were heterogeneous, comprising several isoforms of inhibitors. For the rye TAXI-type inhibitors, small differences in amino acid sequences were found indicating that these isoforms originate from different genes that are yet to be identified (Goesaert et al. 2002).

In wheat, TAXI-type endoxylanase inhibitors with different endoxylanase specificities in their inhibition activities were identified, i.e. TAXI-I and TAXI-II. Both types have similar structures, N-terminal amino acid sequences, pI values (8.8 and at least 9.3, respectively) and occur in two forms, A and B (Gebruers et al. 2001). While form A consists of a single 40 kDa polypeptide chain with at least one intramolecular disulfide bond, form B is made up of two disulfide linked subunits of approximately 30 and 10 kDa. Up to six TAXI-I type inhibitor isoforms and one TAXI-II type inhibitor were isolated from wheat whole meal (Gebruers et al. 2002a). Like wheat, rye contains several iso-inhibitors: four of them, SCXI I, SCXI II, SCXI III and SCXI IV, have 40 kDa and 10 kDa N-terminal sequences that all are highly similar to those of TAXI-I and TAXI-II. Their activity pattern towards family 11 endoxylanases is similar to that of TAXI-I and they have basic pI values (at least 9.0) (Goesaert et al. 2002). In contrast to wheat and rye, barley contains one predominant inhibitor, HVXI (pI value at least 9.0), with N-terminal amino acid sequences of form A and B that are highly similar to those of TAXI and SCXI (Goesaert et al. 2001). All endogenous plant endoxylanases are family 10 endoxylanases, and all tested family 10 endoxylanases are not inhibited by TAXI (Gebruers et al. 2001), HVXI (Goesaert et al. 2001) or SCXI (Goesaert et al. 2002). The specific activities of TAXI (Gebruers et al. 2001), HVXI (Goesaert et al. 2001) and SCXI (Goesaert et al. 2002) towards family 11 endoxylanases suggests a possible role of TAXI-like inhibitors in plant protection against microbial attack; this

places TAXI in line with XIP-1 (Flatman et al. 2002), the polygalacturonase inhibiting proteins (De Lorenzo et al. 2001), pectin lyase inhibitors (Bugbee 1993) and xyloglucan-specific endoglucanase inhibitor (XEGIP) (Qin et al. 2003) which mainly inhibit carbohydrate degrading enzymes produced by microbial pathogens in the early stages of infection.

The TAXI-I gene of *T. aestivum* (Fierens et al. 2003) was the first molecularly identified and characterized TAXI-like gene. It encodes a mature protein of 381 amino acids with a calculated molecular mass of 38.8 kDa, preceded by a signal peptide for secretion of 21 amino acids (Fierens et al. 2003). Sequence similarity was reported with an extracellular dermal glycoprotein (EDGP) of carrot, which has a role in biotic and/or abiotic stress responses (Satoh et al. 1992) and also with functionally non-characterized proteins encoded by multicopy genes in rice and *Arabidopsis thaliana* (Fierens et al. 2003). Here, we report the molecular identification of TAXI-like genes in barley and rye and the chromosomal localization of TAXI-I (Fierens et al. 2003) and these TAXI-like genes.

Materials and methods

Plant materials

For the identification of the TAXI-like genes in barley and rye, the same varieties were used as for the biochemical characterization of the TAXI-like proteins in these cereals by Goesaert et al. (2001, 2002), namely barley (*H. vulgare*) cultivar Hiro and rye (*S. cereale*) cultivar Halo.

For the determination of the chromosomal location of the identified TAXI-I DNA sequence in wheat (Fierens et al. 2003), aneuploid wheat (*T. aestivum,* cv. Chinese Spring) lines developed by Sears (1966) were used, in particular the group 3 nulli-tetrasomic stocks. For the localization of the SCXI DNA sequences, wheat-rye addition lines (Chinese Spring/Imperial) were used (kindly provided by Dr. Ian Dundas, Adelaide University, Australia).

Genomic DNA and RNA isolation

Genomic DNA was isolated from young leaves using the Dneasy Plant Mini kit (Qiagen, Hilden, Germany). Total RNA was isolated from leaf material of one month old embryos and from immature ears (still protected by unfolded upper stem leaves) of *H. vulgare* cultivar Hiro and *S. cereale* cultivar Halo using the Rneasy Plant Mini kit (Qiagen).

Rapid amplification of cDNA ends and PCR

Rapid amplification of cDNA 5′ends (RACE) was performed on isolated *H. vulgare* RNA using the GeneRacer kit (Invitrogen, Carlsbad, USA) with the consensus primer TAXCN31R (5′-TAggCgTTggCgAggAggCA-3′) which allowed us to design the gene-specific primer HVXISTRT (5′-ATggCACgggTgCTCCTC-CTC-3′) (Fig. 1A). For the 3′ RACE reaction on *S. cereale* RNA the 5′/3′ RACE Kit (Roche Diagnostics, Mannheim, Germany) was used with the consensus primer TAXCONF (5′-ACgCgTTCAC-CAAggCCCTg-3′). The gene-specific primer SChalo3R2 (5′-CA-CgCAgTAgATAATCAgAC-3′) was designed based on the sequence information obtained (Fig. 1B). Both consensus primers,
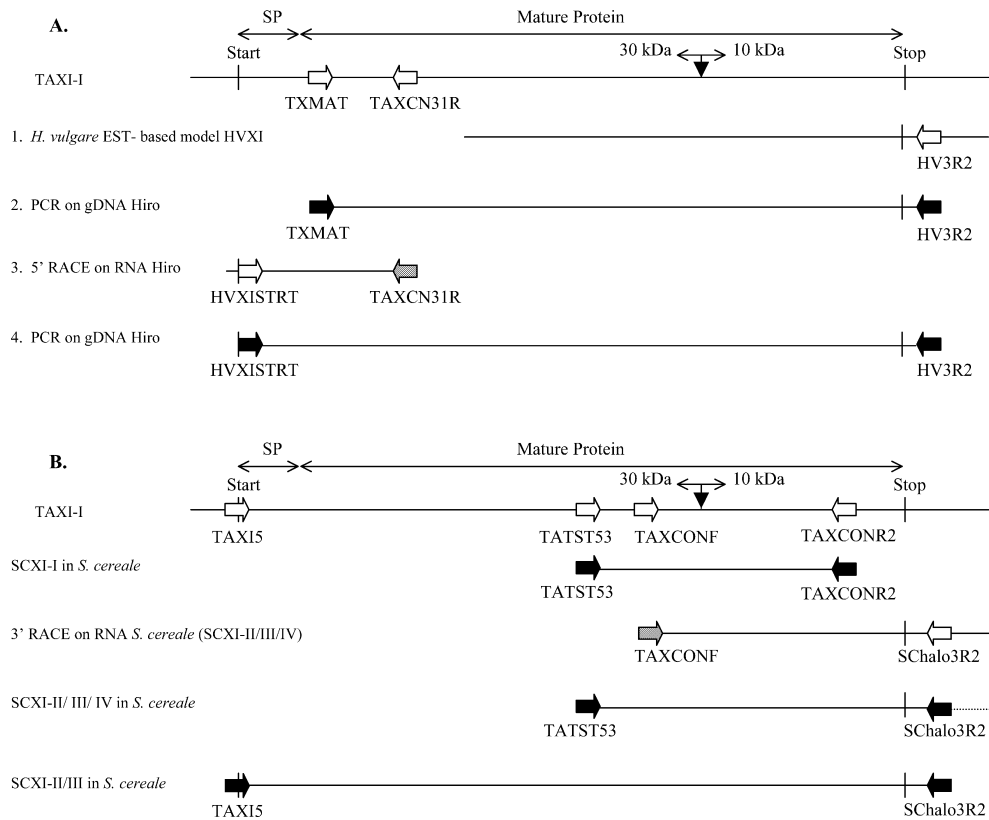
Fig. 1A, B Schematic overview of HVXI and SCXI sequence data and relative positions of primers used for molecular identification. Primers were designed *(white arrows)* and used for PCR *(black)* or RACE *(grey)*. Identified nucleotide sequences are deposited under EMBL accession numbers AJ581529 to AJ581532. **A** HVXI: the EST-based model is constituted by nucleotide sequences with GenBank accession numbers BQ472163, BQ471743, BQ664728, BQ664429, AJ460470, BE602955. **B** SCXI: genomic DNA or total RNA of *Secale cereale* cv. Halo was used as starting material. The identicity between the $3'$ RACE sequence and the SCXI-II/III/IV sequence allowed us to extend the genomic PCR obtained SCXI-II/III/IV sequence *(dotted line)*. SP, signal peptide; ▼, position of internal cleavage site discerning forms A and B

TAXCN31R and TAXCONF, were designed based on the TAXI-I sequence (Fierens et al. 2003).

H. vulgare public database DNA sequence information was used to design the primer HV3R2 (5′-TTCATAAAAgCgggCgC-Tag-3′) (Fig. 1A). The identified TAXI-I sequence (Fierens et al. 2003) was used for the design of PCR primers TAXIUTR (5′-ACAATTCCACgCTCCATCTg-3′), TAXI 5 (5′-CAAgAAAgAT-gCCACCAgTg-3′), TXMAT (5′-gCCTTCCggTgCTCgCTCC-3′), TATST53 (5′-CCgCAATTCACgCAgTCgATg-3′), TAXI 3 (5′-gTAgTggACgAATCCACCTgTC-3′) and TAXCONR2 (5′-CAAC-CCgTAAAgTgCggCAg-3′) (Fig. 1). These oligonucleotides were synthesized by PROLIGO Primers & Probes (Paris, France) and used for (re)amplification using HotStarTaq DNA polymerase (Qiagen). PCR reactions contained PCR buffer, Q-solution, 200 μM of each dNTP (Promega, Leiden, The Netherlands), 1 μM of each primer, 2.5 U HotStarTaq DNA Polymerase and 50 ng genomic or cDNA in a total volume of 30 μl. Each reaction was subjected to a standard PCR program of incubation at 95°C for 15 min, followed by 35 cycles of 1 min at 94°C, 90 s at 58°C and 2 min at 72°C and a final single incubation of 10 min at 72°C, performed on a UNO II (Biometra, Göttingen, Germany) thermocycler.

To check the specificity and yield of the PCR amplifications, 3 μl of the reaction mixture was analyzed on 1% (w/v) agarose gels in TAE electrophoresis buffer (40 mM Tris-HCl pH 7.2, 500 mM sodium acetate, 50 mM EDTA). Following staining with ethidium bromide, DNA was visualized by UV illumination (Sambrook et al. 1989). The size of the fragments was compared with fragments of λ/*Pst*I or the 100 bp DNA ladder (New England Biolabs, Beverly, USA). PCR products were purified using the QIAquick PCR Purification kit (Qiagen) prior to direct sequencing or cloning. Cloning of PCR fragments was done with the TOPO TA Cloning Kit for Sequencing (Invitrogen).

DNA sequence analysis

PCR products or cloned fragments were DNA sequenced using ABI PRISM Big Dye Terminator chemistry (Applied Biosystems, Foster City, USA) in combination with gene-specific and vector-specific primers respectively. Sequencing products were analyzed on a 377 DNA Sequencer (Applied Biosystems). The Sequencher 4.1 package (Gene Codes Corporation, Michigan, USA) was used to correct and align the DNA sequences obtained. These sequences were used to screen public databases using the BLASTn and tBLASTn algorithms (Altschul et al. 1997) for homologous gene fragments.

Amino acid sequences encoded by the identified TAXI-type genes were analyzed using Psort (Nakai and Horton 1999) for signal peptide identification and protein localization, and Protparam (Gill and von Hippel 1989) for calculation of theoretical p*I* and molecular weight. The INTERPRO and PROSITE databases were searched for the presence of existing motifs or patterns in the identified HVXI and SCXI sequences. The ScanProsite Tool (Gattiker et al. 2002) was used to detect potential glycosylation sites. The MEROPS database (http://Merops.sanger.ac.uk) was screened for candidate proteases able to cleave form A into B.

All obtained sequences and their related sequences identified from the databases were compared by pairwise comparisons using

```
TAXI-I      ATCCACTTCCACACCCACAAACAATTCCACGCTCCATCTGTGCAAGAAAG--ATGCCACCAGTGCTCCTCCTCGTCCTGGCCGCCTCGCTCGTGGCGCTGCCGTCGTGCCAAAGCCTTCC   118
SCXI-II/III ----------------------------------------CAAGAAAG--ATGCCACCAGTGCTCCTCCTCGTCCTGGCCGCCTCGCTCGTGGCGCTGCCGTCGTGCCGAAGCCTGCC   76
HVXI        ----------------------------------CAGTTCACGCAGTCCATGGCACGGGTGCTCCTCCTCGCCCTGGCCGCCACGCTCGCCGCGCGCAGGCGTCGTCCAAGGCCCTTCC   83
                                                  * * *  ********* * ** ******** ********* ********* ******  **** * ****** *   *** **

TAXI-I      GGTGCTCGCTCCGGTCACCAAGGACCCCGCCACCTCCCTCTACACAATCCCCTTCCACGACGGCGCCAGCCTCGTCCTCGACGTCGCCGGCCCTCTCGTCTGGTCCACGTGCGATGGCGG   238
SCXI-II/III GGTGCAAGCTCCGGTCACCAAGGACCCCGCCACCTCCCTCTACACCATCCCCTTCCACGACGGCGCCAGCCTCGTCCTCGACGCCGCCGGCCCGTCGTCTGGTCCACCTGCGAGGCCGG   196
HVXI        GGTGCTCGCTCCGGTCACCAAGGACGCCGCCACCTCCCTCTACACAATCCCCTTCCACGACGGCGCCAACCTCGTCCTCGACGTCGCCGGCCCGCTCGTCTGGTCCACGTGCGACGGCGG   203
            *****  ****************** **************** ** *************************** **********      ** ** *  ***** ********* ** ****

TAXI-I      CCAGC------CGCCCGCGGAGATCCCGTGCAGCAGCCCCACCTGCCTCCTCGCCAACGCCTACCCCGCCCCGGGCTGCCCCGCTCCCAGCTGCGGCAGCGATAAGCACGACAAACCGTG   352
SCXI-II/III CCAGC------CGCCGGCGGGGATCCCGTGCGGCAGCCCCACCTGCCTCCTCGCCAACGCCTACCCCGCCCCGGGCTGCCCCGCGCCCACCTGCGGCAGCGA---------CAAGCCCTG   301
HVXI        CCAGCGGCCGCGCGGCGGCTGAGATCACCTGCAGCAGCCCCACCTGCCTCCTCGCCAACGCCTACCCCGCCCCGGGCTGCCCCGCGCCCAGCTGCGGCAGCGACAGGCACGACAAGCCCTG   323
            *****        **** ** ** * *** ***** *********************************************** *** **** *********       *** ** **

TAXI-I      CACGGCGTACCCGTACAACCCGGTCAGCGGCGGCGTGCGCCGCAGGGAGCCTCTCCCACACGAGATTCGTGGCCAACACCACCGACGGGAGCAAGCCGGTGAGCAAGGTGAACGTCGGGGT   472
SCXI-II/III CACGGCGTTCCCATCCAACCCGGTCACCGGCGCCGTGCGCCGCCGGGAGCCTCTTCCACACCAGCTTCGTGGCCAACATCACCGACGGGAGCAAGGAGGCGAAGGTGAAAGTCGGGGT   421
HVXI        CACGGCGTACCCGTCCAACCCGGTCACCGGCGCGTGCGCCGCCGGGAGCCTCTTCCGCGCGAGGCTCGTGGCCAACATCACCGACGGGAACCGCCCGGTGAGCGCCGGTGACCGTCGGGGT   443
            ******** *** *  ********** *****     *** * ****** *  * ** ****    * ** ****     * ********* *** ***** ***** *****  *******

TAXI-I      CCTGGCGGCGTGCGCGCCGAGCAAGCTCCTGGCGTCGCTGCCCCGGGGCTCCACGGGCGTGGCCGGGCTCGCGAACTCCGGCTTGGCGCTGCCGGCGCAGGTGGCATCGCGCGCAGAAGGT   592
SCXI-II/III CCTGGCGGCGTGCGCGCCGAGCAAGCTCCTGGCGTCGCTGCCCCGCGGCTCCACGGGCGTGGCCGGGCTCGCGAACTCCGGCTCGCGGCTGCCGGCGCAGGTGGCGTCCGCGCAGAAGGT   541
HVXI        CCTTGCGGCGTGCGCGCGACGAAGCTGCTCGCGTCGCTCTTCCCGCGGCTCCAGGGGCGTGGCCGGGCTCGCGGGCTGCCGCGCGGCGCGGCGCAGGTGGCGTCGCGCGCAGAAGGT   563
            *** ************* * ****** ** *  ****** * *** ** ***** ** *********************  *** **  * * ***** ******** **  *** ********

TAXI-I      CGCCAACAGGTTCCTCCTCTGCCTCCCCACCGGCGGCCCTGGCGTGGCCATATTTGGCGGCGGCCCGGTCCCGTGGCCGCAATTCACGCAGTCGATGCCTTACACGCCGCTCGTCACCAA   712
SCXI-II/III CGCCAACAGGTTCTTCCTCTGCCTCCCCACCGGCGGCGCTGGCGTGGCCATCTTCGGCGGCGGCCCGCTCCCGTGGCCGCAGTTCACGCAGTCGATGCCGTACACGCCGCTCGTCACCAA   661
HVXI        CTCCCATCGGTTCCTCCTCTGCCTCCCCACGGGCGGCGCCGGCGTGGCCATCCTCGGCGGCGGCCCGCTCCCGTGGCCGCAGTTCACGCAGTCCATGGCCTACACCCCGCTCGTCGCCAA   683
            * * *  ***** **************** ***** *  *********  *  *** *********** ********** ********* *** ** *  **** ** ***** ***  ***

TAXI-I      GGGCGGCAGCCCCGCGCACTACATCTCGGCCAGGTCCATTGTAGTGGGGGACACCCGCGTCCCCGTACCGGAGGGCGCGCTCGCCACCGGCGGCGTGATGCTCAGCACGAGGCTACCCTA   832
SCXI-II/III GGGCGGCAGCCCCGCGCACTACATCTCGCTCAAGTCCATCAAAGTGGACAACACCCGCGTCCCCGTATCGGAGG---------CCACCGGCGGCGTGATGCTCAGCACGAGGCTGCCCTA   772
HVXI        GCAGGGCAGCCCCGCGCACTACGTCTCGGGCACGTCCATCAGGGTGGAGGACACCCGCGTGCCCGTCCCCGGACCGCGCGCTCGCCACCGGCGGCGTGATGCTCAGCACGAGGCTGCCCTA   803
            *  * ********************  *****   *  ******    ****  ***********  *****  *  ****      ** ************************** ******

TAXI-I      CGTCTTGCTCCGCCCCGACGTGTACCGCCCGTTGATGGACGCGTTCACCAAGGCCCTGGCGGCGGCAGCATGCCAACGAGCGCCCGTGGCGCGCGCAGTGGAGGCTGTGGCGCCGTTCGG   952
SCXI-II/III CGCCCTGCTCCGCCGCGACGTGTACCGCCCGTTGGTGGACGCGTTCACCAAGGCCCTGGCGGCGCAGCCAGCCAACGGAGCGCCCGTGGCGCGCGCAGTGCAGCCCGTGGCGCCGTTCGG   892
HVXI        CGTCCTGCTCCGCCGCGACGTGTACCGCCCGTTCGTCGACGCGTTCGCCAAGGCCCTGGCGGCGCAGCATGCCAACGGAGCGCTCGCGCGGCGCGCGGCGTGAACCCCGTGGCGCCGTTCGG   923
            **  * ******** ***************** *  ********* ************************ **      *  *** ***** *   *******   ** * ***********

TAXI-I      GGTGCTACGACACGAAGACGCTGGGCAACAACCTCGGCGGGTACGCGGTGCCCAACGTCCAGCTGGGGCTCAGTGGCGGCAGTGACTGGACGATGACCGGGAAGAACTCGATGGTGGA   1072
SCXI-II/III GGTGTGCTACGACACCAAGACGCTGGGCAACAACCTCGGCGGGTACGCGGTGCCCAACGTCCTGCTGGCGCTCGACGGCGGGGGTGAGTGGGCCATGACCGGGAAGAACTCCATGGTGGA   1012
HVXI        GCTGTGCTACGACGCCAAGACGCTGGGCAACAACCTCGGCGGGTACTCGGTGCCCAACGTGGTGCTGGCGCTCGACGGCGGGGGTGAATGGGCGATGACCGGGAAGAACTCGATGGTGGA   1043
            * ************  ***************************** ************    ***** ***  ********     *** **** ********************* *******

TAXI-I      CGTCAAGCAAGGGACGGCGTGCGTTGCGTTCGTGGAGATGAAGGGAGTGGCGGCCGGCGACGGCAGGGCGCCGGCGGTGATCCTCGGAGGGGCCCAGATGGAGGACTTCGTGCTCGACTT   1192
SCXI-II/III CGTCAAGCCGGGGACGGCGTGCGTCGCGTTCGTGGAGATGAAGGGGGTGGAGGCCGGCGACGGCAGGGCGCCGGCGGTGATCCTCGGAGGGGCCCAGATGGAGGACTTCGTGCTCGACTT   1132
HVXI        CGTCAAGCCGGGGACGGCGTGCGTTGCGTTCGTGGGAGAT---------GGAGGCCGGCGACGGCGGGGCGCCGGCGGTGATCCTCGGAGGGGCCCAGATGGAGGACTTCGTGCTCGACTT   1154
            ********  ************** ***** ****   **          ** ************ ***** *************************************************

TAXI-I      CGACATGGAGAAGAAGCGGCTCGGGTTTAGCAGGCTGCCGCACTTTACGGGTTGCGGCGGC--CTGTAATAATAA--ATCTGTTTAACGACAGGTGGATTCGTCCACTAC---   1298
SCXI-II/III CGACATGGAGAAGAAGCGGCTCGGGTTTACCAGGCTGCCGCACTTTACGGGCTGCGGCAGCGCCTAGAGCTTTCACTGTCTGTATAATAATAAGTCTGATTATCTACTGCGTG   1245
HVXI        CGACATGGAGAAGAAGCGGCTCGGGTTTATCAGGCTGCCGCACTTCACGGGTTGCGGCCAAC--TTGAACTTTAGAAAAAACGCTAGCGCCCGGTTTTATGAA----------   1255
            ******************************* ***************** ***** *****    *   *  *    * *       **      *  **
```

Fig. 2 Alignment at the nucleotide level of TAXI-I (cv. Estica) (Fierens et al. 2003), HVXI (cv. Hiro) and SCXI-II/III (cv. Halo) coding sequences, including the 5′UTR and 3′UTR regions (*italic*) identified. Start and stop codons are *underlined*. Signal sequence is indicated in *bold*. The *arrow* indicates the cleavage site discerning forms A and B

the EMBOSS Align tool (EMBL-EBI) or by multiple sequence alignments using ClustalW (EMBL-EBI).

Results

Identification of the HVXI gene in *Hordeum vulgare*

The complete TAXI-I sequence (Fierens et al. 2003) was used to screen *H. vulgare* expressed sequence tag (EST) sequences in public databases with BLASTn and tBLASTn. This yielded six matching barley nucleotide sequences (with GenBank accession numbers BQ472163, BQ471743, BQ664728, BQ664429, AJ460470 and BE602955). Assembly and manual editing resulted in a contiguous draft sequence of 872 bp including the translation stop codon. Based on this draft sequence, an *H. vulgare* specific primer (HV3R2) was designed (Fig. 1A). PCR was performed on genomic DNA of the *H. vulgare* cultivar Hiro using HV3R2 in combination with the TAXI-I primer TXMAT and amplified a 1,178 bp PCR product corresponding to 99.5% of the mature HVXI protein lacking the N-terminal coding part and the 5′ untranslated region (UTR) (Fig. 1A). To amplify the missing 5′region, 5′RACE was performed on total RNA isolated from leaf material of *H. vulgare* cultivar Hiro using the internal consensus primer TAX-CN31R (Fig. 1A) which also allowed the determination of the missing 5′ sequence. Another primer, HVXISTRT, was designed, which in combination with HV3R2, allowed amplification of the complete gene coding sequence from barley (cultivar Hiro) genomic DNA. Sequencing of this PCR fragment showed that the HVXI coding sequence is 1,209 bp long (Fig. 2). Comparison of the genomic DNA fragment with the cDNA sequence and the available EST data revealed that the HVXI gene does not contain introns. Aligning the deduced amino acid sequence of HVXI with that of TAXI revealed that the open reading frame encodes a signal peptide of 19 amino acids followed by a mature protein of 384 amino acids containing the experimentally determined peptides (Fig. 3), i.e. N-terminal sequences of the 40 kDa (form A) and the dimeric 30 kDa plus 10 kDa (form B) native HVXI protein (Goesaert et al. 2001). The suggested signal

Fig. 3 Alignment at the amino acid level of the identified TAXI-I, HVXI and SCXI-II/III coding sequences (Clustal W, EBI; using default parameters) matching experimentally derived peptides (*underlined*) (Gebruers et al. 2001; Goesaert et al. 2001, 2002). The signal sequence is indicated in *bold*. Numbering starts from the first amino acid of the mature protein. The *boxed* amino acids show the conserved Asn predicted to be a glycosylation site. The *arrow* indicates the cleavage site between the 30 and 10 kDa fragments. All cysteine residues (*shaded*) are conserved (*, identical residues; :, conserved residues; ., semi-conserved residues)

```
TAXI  MPPVLLLVLAASLVALPSCQSLPVLAPVTKDPATSLYTIPFHDGASLVLDVAGPLVWSTC  39
SCXI  MPPVLLLVLAASLVALPSCRSLPVQAPVTKDPATSLYTIPFHDGASLVLDAAGPLVWSTC  39
HVXI  MARVLLLALAATLAAQASSKALPVLAPVTKDAATSLYTIPFHDGANLVLDVAGPLVWSTC  41
      *. ****.***:*.* .*.::*** ******.************.****.*********

TAXI  DGGQ--PPAEIPCSSPTCLLANAYPAPGCPAPSCGSDKHDKPCTAYPYNPVSGACAAGSL  97
SCXI  EAGQ--PPAGIPCGSPTCLLANAYPAPGCPAPTCGSD---KPCTAFPSNPVTGACAAGSL  94
HVXI  DGGQRPPPAEITCSSPTCLLANAYPAPGCPAPSCGSDRHDKPCTAYPSNPVTGACAAGSL  101
      :.**   *** *.*.*****************:****   *****:* ***:********

TAXI  SHTRFVANTTDGSKPVSKVNVGVLAACAPSKLLASLPRGSTGVAGLANSGLALPAQVASA  151
SCXI  FHTSFVANTTDGTKPVSEVKVGVLAACAPSKLLASLPRGSTGVAGLANSGLALPAQVASA  154
HVXI  FRARLVANITDGNRPVSAVTVGVLAACAPTKLLASLPRGSTGVAGLAGSGLALPAQVASA  161
      ::  :*** ***.:*** *.*********:****************.************

TAXI  QKVANRFLLCLPTGGPGVAIFGGGPVPWPQFTQSMPYTPLVTKGGSPAHYISARSIVVGD  217
SCXI  QKVANRFFLCLPTGGAGVAIFGGGPLPWPQFTQSMPYTPLVTKGGSPAHYISLKSIKVDN  214
HVXI  QKVSHRFLLCLPTGGAGVAILGGGPLPWPQFTQSMAYTPLVAKQGSPAHYVSGTSIRVED  221
      ***::**:*******.****:****:*********.******:* ******:*  ** * :

TAXI  TRVPVPEGALATGGVMLSTRLPYVLLRPDVYRPLMDAFTKALAAQHANGAPVARAVEAVA  277
SCXI  TRVPVSE---ATGGVMLSTRLPYALLRRDVYRPLVDAFTKALAAQPANGAPVARAVQPVA  271
HVXI  TRVPVPDRALATGGVMLSTRLPYVLLRRDVYRPFVDAFAKALAAQHANGALAARGVNPVA  281
      *****.:    *************.*** *****::***:****** **** .**.*:.**

TAXI  PFGVCYDTKTLGNNLGGYAVPNVQLGLDGGSDWTMTGKNSMVDVKQGTACVAFVEMKGVA  337
SCXI  PFGVCYDTKTLGNNLGGYAVPNVLLALDGGGEWAMTGKNSMVDVKPGTACVAFVEMKGVE  331
HVXI  PFGLCYDAKTLGNNLGGYSVPNVVLALDGGGEWAMTGKNSMVDVKPGTACVAFVEME---  338
      ***:***:**********:****  *.****.:*:********** ********* :

TAXI  AGDGRAPAVILGGAQMEDFVLDFDMEKKRLGFSRLPHFTGCGGL--  381
SCXI  AGDGRAPAVILGGAQMEDFVLDFDMEKKRLGFTRLPHFTGCGSA--  375
HVXI  AGDGGAPAVILGGAQMEDFVLDFDMEKKRLGFIRLPHFTGCGNLNF  384
      **** ************************** ********* .
```

peptide cleavage site, situated after Ser-(−1) whereby Lys-1 and the following amino acids correspond to the experimentally determined N-terminus of mature HVXI, was confirmed using the PSORT program (Nakai and Horton 1999). PSORT analysis of protein localization indicates secretion of HVXI outside the plant cell. Protparam (Gill and von Hippel 1989) was used to calculate a theoretical molecular mass of 39.4 kDa (as compared to the 40 kDa experimentally determined) and a p*I* of 8.43 (slightly less basic then the >9.0 value determined for the natural HVXI protein) (Goesaert et al. 2001). ScanProsite analysis (Gattiker et al. 2002) indicated the presence of a potential glycosylation at Asn-109.

Identification of SCXI genes in *Secale cereale*

Because of the high sequence similarity of the N-terminal amino acid sequence of TAXI-I with those of the SCXIs isolated from *S. cereale* cultivar Halo (Goesaert et al. 2002), we screened the public databases for TAXI-I homologous rye EST sequences. As no *S. cereale* EST data could be retrieved we tried to amplify PCR fragments from the genomic DNA of *S. cereale* cultivar Halo using existing TAXI-I primers. Primer set TATST53-TAXCONR2 was the only primer combination tested that amplified a clear product from rye (Fig. 1B). Its deduced amino acid sequence corresponded well with the experimentally derived sequence of the SCXI-I 10 kDa polypeptide (21 residues, Goesaert et al. 2002). A 3′ RACE reaction on total RNA isolated from *S. cereale* cultivar Halo immature ears was performed using the TAXI-I

specific primer TAXCONF (Fig. 1B). The resulting PCR fragment was cloned into a TOPO-TA vector and the insert of different clones was used to design a new SCXI-specific primer in the stop codon region, i.e. SChalo3R2 (Fig. 1B). SChalo3R2 was combined with the TAXI-I forward primers TATST53 and TAXI5 (Fig. 1B). The TATST53-SChalo3R2 primer combination amplified a fragment, whose sequence corresponded perfectly with the overlapping obtained 3′ RACE sequences and matched the 21 residue N-terminal amino acid sequence of the 10 kDa fragment of SCXI II, III and IV as determined by Goesaert et al. (2002). The TAXI5-SChalo3R2 primer combination enabled us to amplify the complete coding sequence of a SCXI gene, with a deduced amino acid sequence matching the experimentally derived N-terminal amino acid sequences of the 40 kDa or 30 kDa subunit and 10 kDa subunit polypeptides of SCXI-II/III (Fig. 3). Sequencing of the TAXI5-SChalo3R2 PCR fragment showed that the SCXI-II/III coding sequence identified was 1,188 bp long (Fig. 2). Comparison of the genomic DNA fragment with the cDNA sequences obtained revealed that the SCXI gene does not contain introns. Aligning the deduced amino acid sequence of SCXI II/III with that of TAXI and HVXI revealed that the open reading frame encodes a signal peptide of 21 amino acids followed by a mature protein of 375 amino acids. The PSORT program (Nakai and Horton 1999) predicted that the identified SCXI gene encodes a protein that is secreted outside the plant cell. The predicted cleavage site is situated after Ser-(−1), whereby Leu-1 and the following amino acids correspond to the experimentally determined N-terminus of mature SCXI II/III (Fig. 3). Protpram (Gill and von Hippel 1989) was used to calculate a theoretical molecular mass of 38.1 kDa (compare with the approximately 40 kDa as experimentally determined) and pI of 8.24 for the mature protein, which is slightly less basic then the >9.0 value determined for the natural SCXI protein (Goesaert et al. 2002). ScanProsite analysis (Gattiker et al. 2002) indicates the presence of potential glycosylation at Asn-102.

Chromosomal localization of the TAXI-I and SCXI genes

Previous BLASTp searches of public DNA databases (Fierens et al. 2003) revealed a cluster of eight TAXI-like genes on rice chromosome 1 (GenBank AP003269), one of the first sequenced chromosomes of this model plant (Sasaki et al. 2002). The products of these genes, with protein sequence similarities to TAXI-I ranging from 59% to 72% (Fierens et al. 2003), were annotated in the DNA database as putative extracellular dermal glycoproteins due to their similarity with EDGP from carrot (Satoh et al. 1992). As the linkage map for the entire rice chromosome 1 is conserved relative to the linkage map of chromosome 3 of *Triticum* (Van Deynze et al. 1995), we assigned the identified TAXI-I gene sequence (Fierens et al. 2003) to a specific wheat chromosome via aneuploid analyses on group 3 chromosome nulli-tetrasomic lines. More specif-
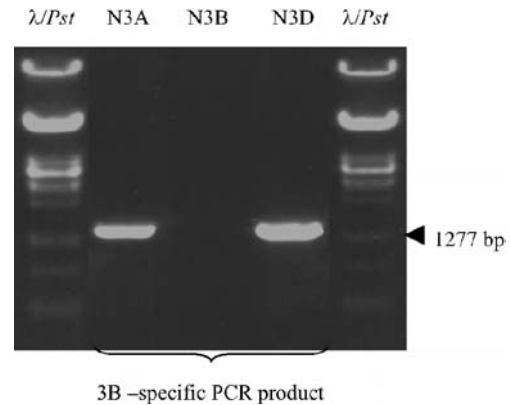

Fig. 4 Aneuploid analysis of genomic DNA isolated from chromosome 3 nulli-tetrasomic wheat (cv. Chinese Spring). HotStarTaq PCR using the TAXI-I specific primer combination TAXIUTR5 (located in the 5′UTR)—TATST33 (located at the stop codon) maps the TAXI-I gene fragment on wheat chromosome 3B by specific amplification of a 1,277 bp product in N3AT3D, N3DT3B and no amplification in N3BT3A. The size marker used is λ/*Pst*I

ically, PCR analysis and sequencing with the TAXI-I specific primer combination TAXIUTR5-TAXI3 (amplifying the complete TAXI-I coding sequence) located the identified TAXI-I sequence on chromosome 3B (Fig. 4).

PCR amplification and sequencing using the SCXI-specific primer combinations TATST53-SChaloR2 (partly amplifying SCXI II/III/IV) and TAX5-SChaloR2 (amplifying SCXI II/III), showed the presence of the identified sequences in the wheat-rye (cv. Chinese Spring/Imperial) addition line CS+6R only (results not shown).

Discussion

Identification of TAXI-like genes

Cereals such as wheat, durum wheat, rye and barley contain endoxylanase inhibiting proteins (Debyser et al. 1997; Debyser and Delcour 1998; Rouau and Surget 1998). The endoxylanase inhibitor TAXI isolated from wheat has homologues in other cereals (Goesaert et al. 2003a), like barley (HVXI; Goesaert et al. 2001) and rye (SCXI; Goesaert et al. 2002). All isolated inhibitor fractions from these cereals were heterogeneous to some extent, comprising several isoforms of the inhibitors (Goesaert et al. 2003a). The molecular identification of the TAXI-I protein as a member of a new class of plant proteins and the availability of TAXI-I sequence-similar EST data from non-wheat cereals such as barley, have enabled us to completely identify the first HVXI and SCXI genes in barley and rye.

The molecular genetic identification of these proteins confirms that HVXI and SCXI are members of the same family of plant proteins as TAXI. Pairwise comparison (Align tool, using the Smith-Waterman algorithm with default parameters) of the identified DNA sequence of HVXI to TAXI-I and SCXI-II/III, showed identities of

Table 1 Overview of characteristics of the different identified TAXI-type inhibitors. *CS* Coding sequence, *SP* signal peptide, *MP* mature protein. Prediction percentage for secretion outside the cell was determined using PSORT (Nakai and Horton 1999)

	Length CS (bp)	% GC	Length SP (aa)	Pred. % outside	Length MP (aa)	Calc. p*I*	Calc. MW (kDa)
TAXI-I	1,206	70	21	82	381	8.24	38.8
HVXI	1,209	73	19	45	384	8.43	39.4
SCXI-II/III	1,188	72	21	82	375	8.24	38.1

88.7% and 87.9% respectively. The SCXI-II/III DNA sequence we identified is 91.8% identical to TAXI-I. Table1 shows an overview of some characteristics of the different identified TAXI-type inhibitors. Although the calculated p*I* is slightly lower than that estimated by isoelectric focusing of native protein samples, the calculated molecular weight of the deduced amino acid sequences corresponds to the experimentally derived data (Goesaert et al. 2001, 2002). All coding sequences have a high GC content. The length of the coding sequence varies and is due to small indels in the mature protein encoding sequence. PSORT analysis of the signal peptide suggests a role of the protein outside the plant cell, for example in plant protection.

As in wheat, TAXI-type endoxylanase inhibitors in barley and rye all occur in two molecular forms, A and B (Gebruers et al. 2001; Goesaert et al. 2001, 2002). The deduced amino acid sequences of the gene sequences identified here agree with the results of the amino acid sequencing of the N-terminal ends of form B of HVXI and SCXI II/III. The cleavage site between the two subunits of form B is located between an Asn and Gly and is conserved among the TAXI-like inhibitors in wheat, barley and rye (Fig. 3). No peptidases able to cleave this bond could be found in the MEROPS database. The TAXI-I sequence contains 12 cysteines, all of which are conserved between the identified HVXI and SCXI sequences (Fig. 3).

Alignment of the amino acid sequences deduced from the identified coding sequences, using ClustalW (EMBL-EBI) (Fig. 3), shows the (semi)conserved residues as well as non-conserved regions of TAXI-like protein. The latter are probably not involved in the specific inhibitory activity towards xylanases. Conserved residues are present over the whole length of the protein, in the 30 kDa polypeptide as well as in the 10 kDa polypeptide. Except for the AATAA consensus signal for polyadenylation in the 3′ untranslated region, identified 5′ and 3′ UTR nucleotide sequences are not well conserved. A database (INTERPRO, PROSITE) search for the presence of catalogued motives or patterns in the identified HVXI and SCXI sequences gives no result.

The isolation of SCXI proteins with different N-terminal amino acid sequences (Goesaert et al. 2002) indicates the presence of different SCXI encoding genes (isoforms) in the *S. cereale* genome. Using specific primer combinations, we identified a complete open reading frame, encoding the SCXI-II/III protein, and fragments of the SCXI-I protein and SCXI II/III/IV protein. The SCXI-I fragment is 96.5% identical to SCXI-II/II and 96.8% to SCXI-II/III/IV, the SCXI-II/III fragment is 98.1% identical to SCXI-II/III/IV. Variation at the nucleotide level is reflected in minor variation at the amino acid level. When we compared these results to the data collected after aligning the identified TAXI, HVXI and SCXI coding sequences, we observed a closer relationship among the SCXI-like fragments.

The HVXI and SCXI gene sequences we identified were used in a search of homologous sequences in the public amino acid and DNA databases using BLAST. As for TAXI-I, amino acid sequence similarity was found with proteins encoded by multicopy genes in rice and *Arabidopsis thaliana*, xyloglucan-specific endoglucanase (XEGIP) from tomato (Qin et al. 2003), extracellular dermal glycoprotein (EDGP) of carrot (Satoh and Fujii 1988; Satoh et al. 1992), basic globulin 7S protein (Bg7S) from soybean (Blackgrove and Gillespie 1975; Blackgrove et al. 1980) and conglutin *γ* (C*γ*) from lupin (Elleman 1977). Cladogram analysis and pairwise comparison of these proteins (Fig. 5) showed a clear clustering of monocotyledonous (wheat, barley, rye and rice) and dicotyledonous (thale cress, tomato, carrot, soybean) proteins. The highly identical endoxylanase inhibiting proteins TAXI-I, HVXI and SCXI-II/III cluster together and show high identity percentages (of about 55%) to one of the rice TAXI-type proteins (BAB89708). Other TAXI-type rice proteins are clustered in diverse, less similar, clades (Fig. 5). The endoxylanase inhibiting activity of these rice proteins remains to be determined and at the moment can not be predicted from sequence comparisons, as conserved amino acid residues involved in the TAXI-endoxylanase interaction have not yet been identified. It should be remarked that Goesaert et al. (2003, 2003) were not able to detect endoxylanase inhibition activity in rice wholemeal extracts and that no TAXI-like proteins could be purified from them using affinity chromatography with immobilized endoxylanase, indicating the absence of TAXI-type endoxylanase inhibitors (Goesaert et al. 2003a). The similarity of the TAXI-type endoxylanase inhibiting proteins with the xyloglucan-specific endoglucanase inhibitor suggests that these proteins belong to a superfamily of glycosyl hydrolase inhibiting proteins in plants. The specific activity towards microbial endoxylanases of proteins found in monocots (TAXI, HVXI and SCXI; Goesaert et al. 2003) and towards endoglucanases for a protein found in dicotyls (XEGIP; Qin et al. 2003) may reflect the relative abundance and importance of these polysaccha-
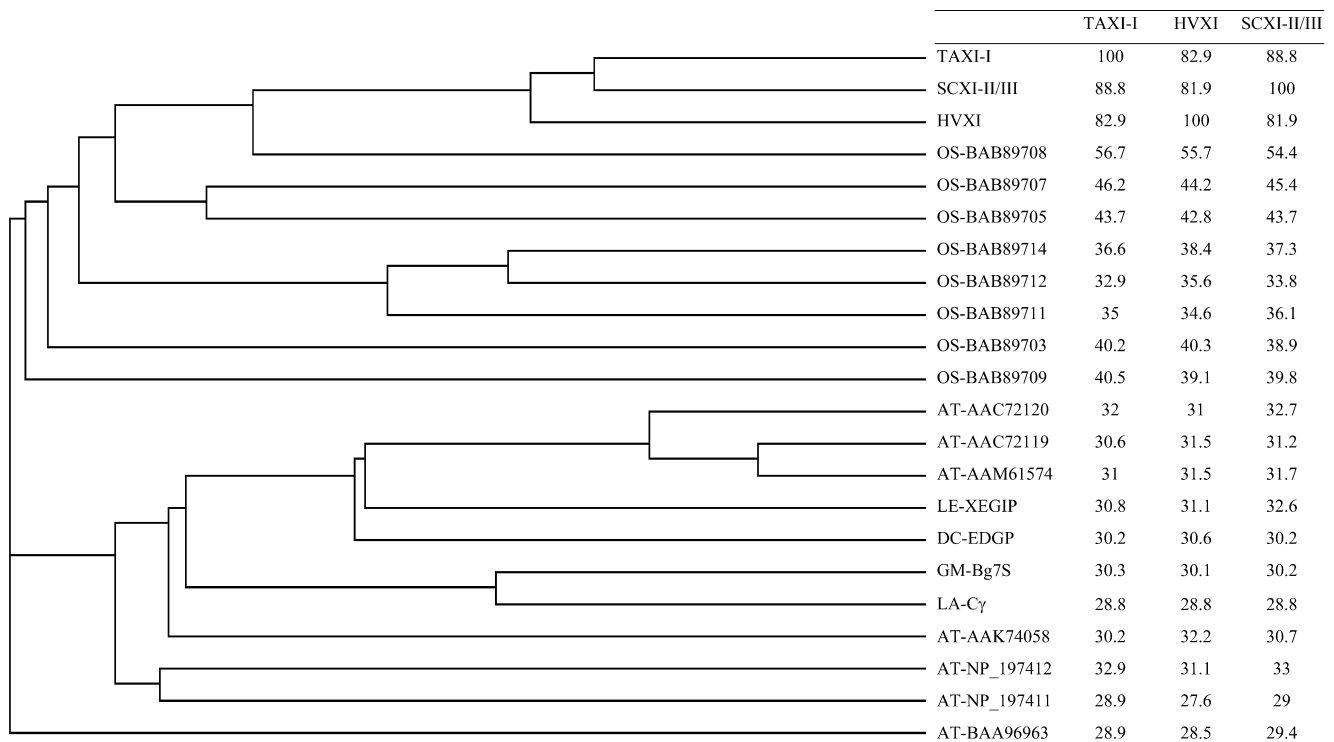
	TAXI-I	HVXI	SCXI-II/III
TAXI-I	100	82.9	88.8
SCXI-II/III	88.8	81.9	100
HVXI	82.9	100	81.9
OS-BAB89708	56.7	55.7	54.4
OS-BAB89707	46.2	44.2	45.4
OS-BAB89705	43.7	42.8	43.7
OS-BAB89714	36.6	38.4	37.3
OS-BAB89712	32.9	35.6	33.8
OS-BAB89711	35	34.6	36.1
OS-BAB89703	40.2	40.3	38.9
OS-BAB89709	40.5	39.1	39.8
AT-AAC72120	32	31	32.7
AT-AAC72119	30.6	31.5	31.2
AT-AAM61574	31	31.5	31.7
LE-XEGIP	30.8	31.1	32.6
DC-EDGP	30.2	30.6	30.2
GM-Bg7S	30.3	30.1	30.2
LA-Cγ	28.8	28.8	28.8
AT-AAK74058	30.2	32.2	30.7
AT-NP_197412	32.9	31.1	33
AT-NP_197411	28.9	27.6	29
AT-BAA96963	28.9	28.5	29.4

Fig. 5 Cladogram (derived from Clustal W alignment, using default parameters) of proteins showing sequence similarity to TAXI-I, HVXI and SCXI-II/III. The table shows percentage amino acid identity with TAXI-I, HVXI and SCXI-II/III, calculated after pairwise comparison using the Align tool (Smith-Waterman algorithm). OS-X, *Oryza sativa* GenBank accession; AT-X, *Arabidopsis thaliana* GenBank accession; LE-XEGIP, *Lycopersicon esculentem* xyloglucan-specific endoglucanase inhibiting protein (AAN87262); DC-EDGP, *Daucus carotus* extracellular dermal glycoprotein (BAA03413); GM-Bg7S, *Glycine max* basic globulin 7S protein (P13917); LA-Cγ, *Lupinus albus* conglutin γ (CAC16394)

rides; (arabino)xylan as a cell wall component is more prevalent in monocots than in dicotyls, where the presence of (xylo)glucan is more important.

Alignment of TAXI-I, HVXI, SCXI, the eight rice TAXI-like proteins, tomato XEGIP, carrot EDGP, soybean Bg7s and lupin Cγ revealed that several amino acids are generally conserved, most notably ten of the 12 TAXI-like protein cysteines, suggesting common structural features in the 3D structure of these proteins. The conserved Asn-Gly cleavage site of TAXI, HVXI and SCXI discussed above is not conserved in the rice, *Arabidopsis* and leguminous proteins. However, peptide bond cleavage behind an Asn-residue at the C-terminal part of large leguminous proteins is well documented (Muntz 1996). For example, the soybean 11S globulin family member glycinin is cleaved at the asparagine residue at the C- terminal part by a cysteine proteinase to yield a mature glycinin (Shimada et al. 1994).

Specific activity towards cell wall degrading microbial endoxylanase indicates a role for TAXI-like endoxylanase inhibitors in plant protection. Detection of the highest TAXI-related inhibition activity in outer kernel tissues (Gebruers et al. 2002b), suggests most of the TAXI related inhibition activity occurs in the outer wheat kernel tissues (bran and shorts) as if to protect the plant.

Observed sequence similarities between the TAXI-like proteins and the characterized EDGP, XEGIP, Bs7s and Cγ proteins suggest a similar function. However, no endoxylanase inhibitory activity of these proteins has been reported. EDGP is a 57 kDa glycoprotein that has been proposed to be involved in pathogen resistance due to its localization in dermal tissues and expression in response to wounding (Satoh et al. 1992). XEGIP is a 51 kDa protein, suggested to be a plant defense protein, functioning as an inhibitor of cell wall degrading microbial glycosyl hydrolase family 12 glycanases. The latter have a 3D structure similar to family 11 endoxylanases, sharing the β-jelly-roll fold. Moreover, a general protective role of XEGIP against biotic or abiotic stresses can be assumed, based on its wide expression in tomato vegetative tissues and high expression in cultured cells (Qin et al. 2003). Leguminous seed proteins Bs7s and Cγ have been considered as storage proteins, but more recent research indicates an additional physiological role as enhanced expression is observed in the response to heat-shock (Duranti et al. 1994; Kagawa and Hirano 1987) and both proteins have the capacity to bind specific proteins or peptides (Komatsu and Hirano 1991; Duranti et al. 1995).

Chromosomal location of TAXI-like genes

PCR on nulli-tetrasomic lines located the identified TAXI-I sequence on wheat chromosome 3B of cv. Chinese Spring. Comparison of this sequence with the reported TAXI-I sequence from cv. Estica (Fierens et al. 2003) revealed no cultivar-specific nucleotide variation. The amino acid sequence of the 6R PCR fragment in the Chinese Spring-Imperial addition is identical to that of the identified Halo sequence, although the nucleotide sequence showed very small variation. The locus of the identified SCXI coding sequence was determined on rye chromosome 6R. In the centromeric regions, gene orders are highly conserved between the homologous group 3 chromosomes of the three wheat genomes and the rye genome (Devos et al. 1992). However, the presence of a 3RL/6RL translocation was reported by Devos and Gale (1993) and Devos et al. (1993). This enables us to delineate the chromosomal location of the TAXI-type genes to the 3RL/6RL translocation area characterized by isozyme *Est-5* and RFLP markers *XGlb33*, *Xpsr1205*, *Xpsr1203* and *Xpsr454* (Devos and Gale 1993; Devos et al. 1993).

Sequencing of the rice chromosome 1 (Sasaki et al. 2002) allowed us to screen the available DNA data in this model plant for the presence of TAXI-like genes. A 50 kb cluster of eight TAXI-like protein encoding sequences (GenBank protein_id: BAB89703, BAB89705, BAB89707, BAB89708, BAB89709, BAB89711, BAB89712 and BAB89714) is situated on rice PAC clone P0504E02 (GenBank AP003269), starting at approximately 40.47 Mb and ending at 40.52 Mb, located in the distal end of chromosome 1. This cluster encompasses a sequence encoding a putative disease resistance protein I2 (protein_id: BAB89710) of the nucleotide binding site-leucine rich repeat type. Isozyme Est-5 marks the distal end of chromosome 1 (Van Deynze et al. 1995). Furthermore a high-density rice map was constructed where the cDNA RFLP marker C1310S maps to the PAC clone P0504E02 (Harushima et al. 1998; Kurata et al. 1994). BLASTn (NCBI) showed that the C1310S cDNA clone sequences (GenBank C98239, D15808) are identical to a TAXI-like protein (BAB89707) encoding sequence. The C1310S TAXI-type cDNA is derived from rice callus tissue. As callus tissue grows in response to wounding, expression of the TAXI-like gene may be induced as a stress response to wounding. The XEGIP and EDGP proteins, structurally similar to TAXI-like proteins, are also secreted into the medium of callus-derived cell suspension cultures and are suggested to play a role in the plant protection against biotic or abiotic stresses (Satoh et al. 1992; Qin et al. 2003).

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Bedford MR, Schulze H (1998) Exogenous enzymes for pigs and poultry. Nutr Res Rev 11:91–114

Biely P, Vršanská M, Tenkanen M, Kluepfel D (1997) Endo-β-1,4-xylanase families: differences in catalytic properties. J Biotechnol 57:151–166

Blackgrove RJ, Gillespie JM (1975) Isolation, purification and characterization of the seed globulins of *Lupinus angustifolius*. Aust J Plant Physiol 2:13–27

Blackgrove RJ, Gillespie JM, Lilley GG, Woods EF (1980) Physicochemical studies of conglutin gamma, a storage globulin from seeds of *Lupinus angustifolius*. Aust J Plant Physiol 7:1–13

Bugbee WM (1993) A pectin lyase inhibitor protein from cell walls of sugar-beet. Phytopathology 83:63–68

Christophersen C, Andersen E, Jakobsen TS, Bagsvaerd PW (1997) Xylanases in wheat separation. Starch 49:5–12

Christov LP, Szakacs G, Balakrishnan H (1999) Production, partial characterization and use of fungal cellulase-free xylanases in pulp-bleaching. Process Biochem 34:511–517

Coote N, Kirsop BH (1976) A haze consisting largely of pentosans. J Inst Brew 82:34

Courtin CM, Delcour JA (2002) Arabinoxylans and endoxylanases in wheat flour bread-making. J Cereal Sci 35:225–243

Debyser W, Delcour JA (1998) Inhibitors of xylanolytic and β-glucanolytic enzymes. European Patent Application WO98/49278

Debyser W, Derdelinckx G, Delcour JA (1997) Arabinoxylan solubilisation and inhibition of the barley malt xylanolytic system by wheat during brewing with wheat wholemeal adjunct: evidence for a new class of enzyme inhibitors. J Am Soc Brew Chem 55:153–156

Debyser W, Peumans WJ, Van Damme EJM, Delcour JA (1999) TAXI, a new class of enzyme inhibitors affecting bread volume. J Cereal Sci 30:39–43

De Lorenzo G, D'Ovidio R, Cervone F (2001) The role of polygalacturonase-inhibiting proteins (PGIPs) in defense against pathogenic fungi. Annu Rev Phytopathol 39:313–35

Devos KM, Gale MD (1993) RFLP report: extended genetic maps of the homologous group 3 chromosomes of wheat, rye and barley. Theor Appl Genet 85:649–652

Devos KM, Atkinson MD, Chinoy CN, Liu CJ, Gale MD (1992) RFLP-based genetic map of the homologous group 3 chromosomes of wheat and rye. Theor Appl Genet 83:931–939

Devos KM, Atkinson MD, Chinoy CN, Francis HA, Harcourt RL, Koebner RMD, Liu CJ, Masojc P, Xie DX, Gale MD (1993) Chromosomal rearrangments in the rye genome relative to that of wheat. Theor Appl Genet 85:673–680

Duranti M, Scarafoni A, Gius C, Negri A, Faoro F (1994) Heat-induced synthesis and tunicamycin-sensitive secretion of the putative storage glycoprotein conglutin γ from mature lupin seeds. Eur J Biochem 222:387–393

Duranti M, Gius C, Scarafoni A (1995) Lectin-like activity of lupin seed conglutin γ, a glycoprotein previously referred to as a storage protein. J Exp Bot 46:725–728

Elleman TC (1977) Amino acid sequence of the smaller subunit of conglutin gamma, a storage globulin of *Lupinus angustifolius*. Aust J Biol Sci 30:33–45

Elliot GO, Hughes RK, Juge N, Kroon PA, Williamson G (2002) Functional identification of the cDNA coding for a wheat endo-1,4-β-D-xylanase inhibitor. FEBS Lett 519:66–70

Elliot GO, Mclauchlan WR, Williamson G, Kroon PA (2003) A wheat xylanase inhibitor protein (XIP-I) accumulates in the grain and has homologues in other cereals. J Cereal Sci 37:187–194

Fierens K, Brijs K, Courtin CM, Gebruers K, Goesaert H, Raedschelders G, Robben J, Van Campenhout S, Volckaert

G, Delcour JA (2003) Molecular identification of wheat endoxylanase inhibitor TAXI-I, member of a new class of plant proteins. FEBS Lett 540:259–263

Flatman R, McLauchlan WR, Juge N, Furniss CS, Berrin JG, Hughes RK, Manzanares P, Ladbury JE, O'Brien R, Williamson G (2002) Interactions defining the specificity between fungal xylanases and the wheat proteinaceous inhibitor, XIP-1. Biochem J 365:773–781

Gattiker A, Gasteiger E, Bairoch A (2002) ScanProsite: a reference implementation of a PROSITE scanning tool. Appl Bioinform 1:107–108

Gebruers K, Debyser W, Goesaert H, Proost P, Van Damme J, Delcour JA (2001) *Triticum aestivum* L. endoxylanase inhibitor (TAXI) consists of two inhibitors, TAXI I and TAXI II, with different specificities. Biochem J 353:239–244

Gebruers K, Goesaert H, Brijs K, Courtin CM, Delcour JA (2002a) Purification of TAXI-like endoxylanase inhibitors from wheat (*Triticum aestivum* L.) whole meal reveals a family of isoforms. J Enzyme Inhib Med Chem 17:61–8

Gebruers K, Courtin CM, Goesaert H, Van Campenhout S, Delcour JA (2002b) Endoxylanase inhibition activity in different European wheat cultivars and milling fractions. Cereal Chem 795:613–616

Gebruers K, Brijs K, Courtin CM, Fierens K, Goesaert H, Rabijns A, Raedschelders G, Robben J, Sansen S, Sørensen JF, Van Campenhout S, Delcour JA (2004) Properties of TAXI-type endoxylanase inhibitors. Biochim Biophys Acta 1696:213–221

Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 182:319–26

Goesaert H, Debyser W, Gebruers K, Proost P, Van Damme J, Delcour JA (2001) Purification and partial characterization of an endoxylanase inhibitor from barley. Cereal Chem 78:453–457

Goesaert H, Gebuers K, Courtin CM, Proost P, Van Damme J, Delcour JA (2002) A family of TAXI-like endoxylanase inhibitors in rye. J Cereal Sci 36:177–185

Goesaert H, Gebruers K, Brijs K, Courtin CM, Delcour JA (2003a) TAXI-type endoxylanase inhibitors in different cereals. J Agric Food Chem 51:3770–3775

Goesaert H, Gebruers K, Brijs K, Courtin CM, Delcour JA (2003b) XIP-type endoxylanase inhibitors in different cereals. J Cereal Sci 38:317–324

Harushima Y, Yano M, Shomura A, Sato M, Shimano T, Kuboki Y, Yamamoto T, Lin SY, Antonio BA, Parco A, Kajiya H, Huang N, Yamamoto K, Nagamura Y, Kurata N, Khush GS, Sasaki T (1998) A high-density rice genetic linkage map with 2,275 markers using a single F_2 population. Genetics 148:479–494

Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J 280:309–16

Henry RJ (1985) A comparison of the non-starch carbohydrates in cereal grains. J Sci Food Agric 36:1243–1253

Jeffries TW (1996) Biochemistry and genetics of microbial xylanases. Curr Opin Biotechnol 7:337–42

Kagawa H, Hirano H (1987) Soybean basic 7S globulin represents a protein widely distributed in legume species. FEBS Lett 226:145–149

Komatsu S, Hirano H (1991) Plant basic 7S globulin-like proteins have insulin and insulin-like growth factor binding activity. FEBS Lett 294:210–212

Kurata N, Nagamura Y, Yamamoto K, Harushima Y, Sue N, Wu J, Antonio BA, Shomura A, Shimizu T, Lin SY, Inoue T, Fukuda A, Shimano T, Kuboki Y, Toyama T, Miyamoto Y, Kirihara T, Hayasaka K, Miyao A, Monna L, Zhong HS, Tamura Y, Wang ZX, Momma T, Umehara Y, Yano M, Sasaki T, Minobe Y (1994) A 300 kilobase interval genetic map of rice including 883 expressed sequences. Nat Genet 8:365–372

Leclercq C, Dervilly G, Saulnier L, Dallies N, Zimmermann D, Roue C (1999) Barley and malt pentosans: structure and functionalities in the brewing industry. In: Proceedings of the 27th congress of the European Brewery Convention, Eur Brew Con, Zoeterwoude, pp 429–436

Mclauchlan WR, Garcia-Conesa MT, Williamson G, Roza M, Ravestein P, Maat J (1999) A novel class of protein from wheat which inhibits xylanases. Biochem J 338:441–446

Muntz K (1996) Proteases and proteolytic cleavage of storage proteins in developing and germinating dicotyledonous seeds. J Exp Bot 47:605–622

Nakai K, Horton P (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci 24:34–36

Qin Q, Bergmann CW, Rose JK, Saladie M, Kolli VS, Albersheim P, Darvill AG, York WS (2003) Characterization of a tomato protein that inhibits a xyloglucan-specific endoglucanase. Plant J 34:327–338

Rouau X, Surget A (1998) Evidence for the presence of a pentosanase inhibitor in wheat flour. J Cereal Sci 28:63–70

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, New York

Sasaki T, Matsumoto T, Yamamoto K, et al. (2002) The genome sequence and structure of rice chromosome 1. Nature 420:312–316

Satoh S, Fujii T (1988) Purification of GP57, and auxin-regulated extracellular glycoprotein of carrots, and its immunocytochemical localization in dermal tissues. Planta 175:364–373

Satoh S, Sturm A, Fujii T, Chrispeels MJ (1992) cDNA cloning of an extracellular dermal glycoprotein of carrot and its expression in response to wounding. Planta 188:432–438

Sears ER (1966) Nullisomic-tetrasomic combinations in hexaploid wheat. In: Riley R, Lewis KR (eds) Chromosome manipulation and plant genetics. Oliver and Boyd, Edinburgh, pp 389–407

Shimada T, Hiraiwa N, Nishimura M, Hara-Nishimura I (1994) Vacuolar processing enzyme of soybean that converts proproteins to the corresponding mature forms. Plant Cell Physiol 35:713–718

Simpson DJ, Fincher GB, Huang AHC, Cameron-Mills V (2003) Structure and function of cereal and related higher plant (1-4)-β-xylan endohydrolases. J Cereal Sci 37:111–127

Suurnäkki A, Tenkanen M, Buchert J, Viikari L (1997) Hemicellulases in the bleaching of chemical pulps. Adv Biochem Eng Biotechnol 57:261–287

Van Deynze AE, Nelson JC, Yglesias ES, Harrington SE, Braga DP, McCouch SR, Sorrells ME (1995) Comparative mapping in grasses. Wheat relationships. Mol Gen Genet 248:744–754

Viëtor RJ, Voragen AGJ, Angelino SAGF (1993) Composition of non-starch polysaccharides in wort and spent grains from brewing trials with malt from a good malting quality barley and a feed barley. J Inst Brew 99:243–248