An α-L-arabinofuranosidase and a β-β-d-xylosidase, designated ARA-I and XYL, respectively, have been purified about 1,000-fold from extracts of 5-day-old barley (Hordeum vulgare L.) seedlings using ammonium sulfate fractional precipitation, ion exchange chromatography, chromatofocusing, and size-exclusion chromatography. The ARA-I has an apparent molecular mass of 67 kDa and an isoelectric point of 5.5, and its catalytic efficiency during hydrolysis of 4′-nitrophenyl α-L-arabinofuranoside is only slightly higher than during hydrolysis of 4′-nitrophenyl β-β-d-xyloside. Thus, the enzyme is actually a bifunctional α-L-arabinofuranosidase/β-β-d-xylosidase. In contrast, the XYL enzyme, which also has an apparent molecular mass of 67 kDa and an isoelectric point of 6.7, preferentially hydrolyzes 4′-nitrophenyl β-β-d-xyloside, with a catalytic efficiency ~30-fold higher than with 4′-nitrophenyl α-L-arabinofuranoside. The enzymes hydrolyze wheat flour arabinoxylan slowly but rapidly hydrolyze oligosaccharide products released from this polysaccharide by (1→4)-β-β-d-xylopyranosyl endohydrolase. Both enzymes hydrolyze (1→4)-β-β-d-xylomaltose, and ARA-I can also degrade (1→5)-α-L-arabinofuranohexose. ARA-I and XYL cDNAs encode mature proteins of 748 amino acid residues which have calculated molecular masses of 79.2 and 80.5 kDa, respectively. Both are family 3 glycoside hydrolases. The discrepancies between the apparent molecular masses obtained for the purified enzymes and those predicted from the cDNAs are attributable to COOH-terminal processing, through which about 130 amino acid residues are removed from the primary translation product. The genes encoding the ARA-I and XYL have been mapped to chromosomes 2H and 6H, respectively. ARA-I and XYL transcripts are most abundant in young roots, young leaves, and developing grain, whereas XYL mRNA is detected in most barley tissues.

Heteroxylans are major constituents of cell walls in the Poaceae, which include many commercially important cereals and pasture grasses. In the endosperm of barley grains and in elongating coleoptiles, these polysaccharides may comprise 20–70% by weight of the walls (1) and consist of a backbone of (1→4)-β-β-linked β-d-xylopyranosyl residues substituted predominantly with α-L-arabinofuranosyl residues. The α-L-arabinofuranosyl residues can be linked to O-3, O-2, or both O-3 and O-2 of xylopyranosyl residues of the (1→4)-β-β-d-xylose backbone, and other substituents or short side chains are also detected in low abundance (2, 3). The α-L-arabinofuranosyl residues can be esterified with hydroxycinnamic acids, in particular ferulic acid, which may form cross-bridges between adjacent arabinoxylan chains, or with lignin, by oxidative dimerization (4).

As observed with many wall components in higher plants, the arabinoxylans of primary cell walls can be restructured during normal growth and development. For example, newly synthesized arabinoxylans in maize coleoptiles are deposited in the walls in a highly substituted form, but arabinoxylan residues are removed later, and this leads to significant changes in the physicochemical properties of the polysaccharides and hence in the walls themselves (5, 6). Removal of α-L-arabinofuranosyl residues is also observed when wall arabinoxylans are degraded (7, 8). The presence of α-L-arabinofuranosidasises in germinated barley grain or in isolated aleurone layers has been taken as evidence that these enzymes perform this function during the mobilization of the starchy endosperm after cereal grain germination (9, 10), although this activity has also been attributed to a separate group of enzymes, known as arabinoxylans α-L-arabinofuranohydrolases (11, 12).

Most α-L-arabinofuranosidasises are so designated because they can hydrolyze the synthetic aryl glycoside 4′-nitrophenyl α-L-arabinofuranoside (4NPA),1 and although they are presumed to be responsible for changes in arabinoxylans during wall modification or degradation, this class of enzymes may be subdivided into several quite distinct groups. Thus, α-L-arabinofuranosidasises have been classified in glycoside hydrolase families 3, 43, 51, 54, and 62 (13; afmb.cnrs-mrs.fr), and members of each family exhibit characteristic substrate specificities, action patterns, and reaction mechanisms (14–19), and three-dimensional structures (20). Although most of the characterized enzymes are from saprophytic or rumen microorganisms, several plant α-L-arabinofuranosidasises have also been identi-

1 The abbreviations used are: 4NPA, 4′-nitrophenyl α-L-arabinofuranoside; ARA-I, α-L-arabinofuranosidase; HCA, hydrophobic cluster analysis; HPLC, high performance liquid chromatography; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; 4NPX, 4′-nitrophenyl β-β-d-xylopyranoside; RACE, rapid amplification of cDNA ends; RFLP, restriction fragment length polymorphism; XYL, β-β-d-xylosidase.
family 51 arabinoxylan arabinoarabinohydrolases, which remove α-L-arabinofuranosyl residues from polymeric arabinoxylans, have been purified from germinated barley grain, and their primary structures have been defined (11, 12). There are other reports of the purification or partial purification of higher plant α-L-arabinofuranosidas, but in most cases no amino acid sequence information is available, and it is therefore not possible to classify the enzymes accurately, to draw conclusions about their reaction mechanisms, or to identify their true substrates and biological functions.

Here we have purified a bifunctional family 3 α-L-arabinofuranosidase/β-D-xylanase (ARA-I) from young barley seedlings, defined its kinetic and enzymic properties, and determined its complete amino acid sequence from corresponding cDNAs. The enzyme is unable to hydrolyze arabinoxylans at a significant rate but could play an important role in the complete depolymerization of arabinoxylans through its ability to hydrolyze oligosaccharides released from the polysaccharide by (1→4)-β-D-xylan endohydrolases. In parallel, a family 3 β-D-xylanase (XYL) was purified and characterized.

EXPERIMENTAL PROCEDURES

Materials—DEAE-cellulose (DE52) was from Whatman. CM-Sepharose CL-6B, Polybuffer Exchanger 94, Polybuffer 96, Polybuffer 74, and molecular mass marker proteins were from Amersham Biosciences. Bio-Gel P-100 was from Bio-Rad Laboratories. Phenylmethylsulfonyl fluoride, bovine serum albumin, 4NPA, 4-nitrophenyl-β-D-xylopyranoside (4NPX), all other aryl glycosides, laminarin, and arabinogalactan were from Sigma. Commissar Brilliant Blue reagent was from Pierce, and wheat arabinoxylan, (1→5)-α-L-arabinofuranohexaose, (1→4)-β-D-xylopyranopentaose, and barley (1→3,1→4)-β-D-glucan were from Megazyme (Bray, Ireland). Larchwood (1→4)-β-D-xylan and the Thermomyces endoxylanase were kindly provided by Dr. Peter Biely (Institute of Chemistry, Slovak Academy of Sciences, Bratislava, Slovak Republic).

Enzyme Purification—Barley (Hordeum vulgare L. cv. Clipper) (3 kg dry weight) was surface sterilized in 0.2% (w/v) AgNO3, washed successively with sterile water, 0.5 M NaCl, and sterile water, and steeped for 5 days in the dark at any stage during this period. The germinated barley material was maintained at 4°C for 10 days in the dark with regular application of fresh antibiotic solution for 5 days in the dark at 22°C. Assays were performed at 37°C, except for amino acid sequence analysis, as follows. First, the enzyme was assayed spectrophotometrically using 4NPA and 4NPX, with a final concentration of 0.1 M sodium acetate buffer, pH 5.0, containing 10 mM EDTA, 10 mM sodium azide, 3 mM 2-mercaptoethanol, and 3 mM phenylmethylsulfonyl fluoride. Ammonium sulfate fractional precipitation was performed as described previously (21), and the enzyme purification procedures were as shown in Scheme 1.

Enzyme Assays—Activities of α-L-arabinofuranosidase and β-D-xylanase were determined spectrophotometrically using 4NPA and 4NPX, respectively. Assays were performed at 37°C in 50 mM sodium acetate buffer, pH 5.0, containing 4 mM sodium azide and 0.04% (w/v) substrate. Reactions were terminated by the addition of 2 volumes of saturated sodium tetraborate solution. One unit of activity is defined as the amount of enzyme releasing 1 μmol of 4-nitrophenol/min, as measured by absorbance at 410 nm.

Protein Determination and PAGE—Protein contents of pooled fractions and purified proteins were measured using the Coomassie Brilliant Blue reagent. Purity of column fractions and purified proteins was assessed by SDS-PAGE (22) on 12.5% (w/v) polyacrylamide gels and stained with Coomassie Brilliant Blue R-250.

Substrate Specificities—Aryl glycosides were used as substrates with purified proteins in standard assays at a final concentration of 2.5 mM. Rates of hydrolysis of polymeric substrates, at a final concentration of 0.2% (w/v), were determined from the increase in reducing sugars (23, 24). Oligosaccharide and monosaccharide products liberated during enzymic hydrolyses were analyzed by thin layer chromatography on silica gel plates (Merck). Plates were developed in ethyl acetate:acetic acid:water (3:2:1 v/v), and sugars were detected using the orcinol reagent (21).

Kinetic Analyses—Kinetic parameters of purified enzymes were determined against 4NPA and 4NPX in a concentration range of 0.25–7.0 mM. Assays were performed in triplicate, in 50 mM sodium acetate buffer, pH 4.7, containing 4 mM sodium azide and 160 μg ml−1 bovine serum albumin. S.E. values for assays were less than 5%. Kinetic data were processed by a proportional weighted fit using a nonlinear regression analysis program based on Michaelis-Menten enzyme kinetics (25).

Amino Acid Sequence Analysis—NH2-terminal sequence analyses of proteins and peptides generated by CNBr, trypsin, or Lys-C and purified by reversed phase HPLC were performed on a Hewlett-Packard G1005A automated protein sequencer, using Edman degradation chemistry. Phenylthiohydantoin-derivatives were identified by reversed phase HPLC using a Hewlett-Packard HPLC 1090 system to identify the COOH terminus of ARA-I, the purified enzyme was subjected to extensive amino acid sequence analysis, as follows. First, ARA-I was reduced with 5 mM dithiothreitol, 6 mM iodoacetamide, and 50 mM Tris-HCl buffer, pH 8.5, at 65°C for 30 min and alkylated with 20 mM iodoacetamide for 20 min at 20°C. Dithiothreitol was added to the alkylated solution to remove excess iodoacetamide. Second, the reduced, alkylated protein was fragmented with trypsin or Lys-C. Prior to trypsin digestion, buffer exchange into 100 mM Tris-HCl buffer, pH 7.0, containing 4 mM urea and 2 mM dithiothreitol, was performed using a Nanosep Centricron (PALL Life Sciences, Ann Arbor, MI) with a 3-kDa cutoff. After diluting the ARA-I four times with water, trypsin (Promega) was added for 16 h at 20°C. For Lys-C digestion, the buffer exchange was into 25 mM Tris-HCl buffer, pH 7.0, containing 4 mM urea. Digestion with endoproteinase Lys-C (Promega) was performed for 16 h at 20°C. The reactions were stopped by adjusting the pH to 3 with 2% (v/v) trifluoroacetic acid.

The Lys-C and tryptic peptides were separated on a Vydac C18 protein column (250 × 2.1 mm, 5 μm; Hesperia, CA). The fluent were (A) (0.05% (v/v) trifluoroacetic acid and (B) 0.04% (v/v) trifluoroacetic acid in 70% (v/v) acetonitrile. The flow rate was 0.2 ml/min, and the gradient was composed of 2% (v/v) B for 90 min and 70–100% (v/v) B for an additional 10 min. The column temperature was 40°C, and protein was detected by absorbance at 214 and 280 nm. Where necessary to separate poorly resolved peptides, selected fractions were rechromatographed using a shallower gradient. Some peptides were fragmented further with 100 μM CNBr in 0.1% HCl for 16 h at 20°C.
**Mass Spectrometry**—The ARA-I enzyme was desalted by reversed phase HPLC on a 290 × 21 mm POROS 20R1 column (Applied Biosystems) with 1% (v/v) formic acid and a 5–100% (v/v) acetonitrile gradient at 0.5 ml/min at 40 °C over 15 min. For MALDI-TOF mass spectrometry, desalted ARA-1 and XYL (0.5 μl) mixed with 1 μl of matrix solution (1% w/v protein in 50% v/v acetonitrile, 1% v/v formic acid) were spotted onto a target plate and analyzed in a Voyager-DE STR mass spectrometer (Applied Biosystems). Peptide fragments corresponding to each gene were amplified by 30 cycles of PCR consisting of 94°C, 40 s; 55°C, 40 s; 72°C, 30 s. Amplified products were detected by gel electrophoresis, and DNA bands were observed under ultraviolet light. For each set of primers, amplified DNA was excised from the agarose gel and purified for subsequent DNA sequence analysis. DNA sequences in every case, exactly matched the DNA sequence of the respective cDNA clone.

**Reverse Transcription-PCR**—DNA samples were prepared from 3 μg of total RNA from developing grains (6 days postanthesis), leaf (5 days after germination), coleoptile (4 days), rootlets (5 days), aleurone (3 days), and scutella (3 days). Reverse transcriptase reactions were primed with (dT)₇ primer. PCRs contained 1 μl of cDNA and 0.1 μg of each gene-specific primer. DNA fragments corresponding to each gene were amplified by 30 cycles of PCR consisting of 94°C, 40 s; 55°C, 40 s; 72°C, 30 s. Amplified products were detected by gel electrophoresis, and DNA bands were observed under ultraviolet light. For each set of primers, amplified DNA was excised from the agarose gel and purified for subsequent DNA sequence analysis. DNA sequences in every case, exactly matched the DNA sequence of the respective cDNA clone.

**Genetic Mapping of ARA-I and XYL Genes**—Barley mapping populations and parental lines were screened at high stringency with probes corresponding to each of the XYL and ARA-I cDNAs, as described (30). The DNA probe for ARA-I was the 2,400-bp 3′-end RACE fragment, and for XYL was the 1,800-bp 3′-end RACE fragment. Filters were hybridized with radiolabeled DNA probes at 65°C and washed in 0.1× SSC, 0.1% (v/v) SDS at 65°C to remove nonspecifically bound probe DNA. Chromosomal locations for the ARA-I and XYL genes were allocated by correlation with genetic markers using the Mapmaker and JoinMap software (31, 32).

**Hydrophobic Cluster Analysis and Molecular Modeling**—Hydrophobic cluster analysis (HCA) was performed using DrawHCA (33); www.lmc.pj.jussieu.fr/~callevau/hca_method.html). The three-dimensional molecular models of ARA-I and XYL were built using the three-dimensional structure of barley β-D-glucan glucohydrolase (www.rcsb.org/pdb; PDB accession code 1ERQ; 34) as a template, and Modeller, version 6.1 (35). The models were checked manually on O (36) and evaluated using PROCHECK (37) and PROSAII (38). The three-dimensional structure of α-L-arabinofuranosidase was built and minimized in SYBYL 6.62 (39), and molecular surfaces of β-D-xylopyranosidase (PDB accession code 1B3V) and α-L-arabinofuranosidase were computed in GRASP (40), using a probe size of 1.4 Å and all atoms having fixed radii.

### RESULTS

**Purification of Barley ARA-I and XYL**—Preliminary experiments showed that α-L-arabinofuranosidase and β-D-xylosidase activities in young barley seedlings reached a peak 4–5 days after germination. From 3 kg of germinated barley grain, an extract of soluble proteins from 5-day-old barley seedlings yielded 45 and 46 units of 4NPA- and 4NPX-hydrolyzing activities, respectively (Table I). The steps used for the purification of the ARA-I and XYL isoenzymes are summarized in Scheme 1.

Two peaks of α-L-arabinofuranosidase activity were resolved on the DEAE-cellulose column at −90 mM NaCl and 140 mM NaCl (Fig. 1A) and were designated ARA-I and ARA-II, respectively. Attempts to purify ARA-II completely were unsuccessful, and this isoenzyme will not be described here. The ARA-I fractions contained both β-D-xylosidase activity at −15% of their α-L-arabinofuranosidase activity, throughout the purification of ARA-I (Fig. 1B) and in the final purified enzyme preparation. It appeared therefore that although ARA-I had a preference for 4NPA, it could also hydrolyze 4NPX. The final ARA-I preparation was purified 1,080-fold and represented about 2% of the initial activity (Table I). It should be noted that the true purification factor of ARA-I was likely to be significantly higher than 1,080-fold because of the ARA-II present in the initial tissue extracts. The XYL enzyme was separated from ARA-I during the initial DEAE-cellulose chromatography step (Scheme 1) and, after resolution from other proteins on CM-cellulose (Fig. 1C), was ultimately purified 960-fold (Table I).

In view of the apparent activity of the ARA-I on both 4NPA and 4NPX during the purification process, particular care was taken to evaluate the purities of the final enzyme preparations. The final ARA-I preparation appeared as a single protein of molecular mass 67 kDa after SDS-gel electrophoresis (Fig. 2A) and had an isoelectric point of 5.5. Although the NH₂-terminal sequence could not be obtained with ARA-I, presumably be...
cause the N\textsubscript{2}-terminal residue was blocked to Edman degradation, the sequences of 10 tryptic peptides from the enzyme were determined. For the final XYL preparation, which had a molecular mass of 67 kDa (Fig. 2B) and an isoelectric point of 6.7, a minor protein band greater than 100 kDa could be seen but was not resolved by additional chromatography steps (data not shown). An N\textsubscript{2}-terminal sequence of 38 amino acid residues was obtained for XYL, together with the sequences of six tryptic peptides. For both ARA-I and XYL, amino acid sequence comparisons with protein sequences in the data bases revealed that the enzymes were members of the family 3 group of glycoside hydrolases (13).

**Kinetic Analyses**—The kinetic parameters for ARA-I and XYL using 4NPA and 4NPX, respectively, are shown in Table II. The $K_m$ values for ARA-I against aryl glycosides revealed that this enzyme had a relatively low $K_m$ for 4NPX compared with 4NPA, although the catalytic rate was higher on 4NPA (Table II).

Comparison of kinetic parameters for XYL against 4NPA and 4NPX demonstrated a definite preference for 4NPX (Table II). XYL had a relatively low $K_m$ value for 4NPX (1.7 ± 0.04 mM) compared with the $K_m$ value for 4NPA (24.8 ± 0.04 mM), the catalytic rate constant for 4NPX was twice that for 4NPA, and the catalytic efficiency factor for 4NPX was about 30 times that measured for 4NPA (Table II).

**Substrate Specificities**—The pH optimum for both ARA-I and XYL was 4.7 (data not shown). The preferred aryl glycoside substrates for ARA-I and XYL were 4NPA and 4NPX, respectively. ARA-I could also hydrolyze 4NPX, 4NP-D-galactopyranoside, and 4NP-L-arabinopyranoside, with 20, 16, and 11% of the specific activity observed for 4NPA, respectively (Table III). The preference of XYL for 4NPX was more pronounced, with specific activities for other aryl glycosides no greater than 3% of that for 4NPX (Table III).

Of the polysaccharides examined, ARA-I and XYL hydrolyzed only arabinoxylan, but hydrolytic rates were very low, and activity could be detected only after prolonged incubation of the enzymes with this substrate (Fig. 3A). When ARA-I was incubated with the arabinoxylan, small amounts of both L-arabinose and D-xylose were released, but XYL released only D-xylose (Fig. 3A). Neither enzyme hydrolyzed larch arabinoxylan, (1→3,1→4)-D-glucan, laminarin, or CM-xylan (data not shown).

In the presence of (1→4)-D-xylan endohydrolase, hydrol-
tion of the (1 \rightarrow 4)-\beta-D-xylosylpyrano-
ose to \(\beta\)-xylose within 1 h. ARA-I was less efficient in
hydrolyzing xylopyranopentaose and after 1 h, had degraded
this substrate to approximately the same degree as XYL after
endohydrolysis of the arabinoxylan greatly enhanced the re-

cysis rates increased dramatically, and both ARA-I and XYL
released large amounts of \(\beta\)-xylose; ARA-I also released some
4-NP-\(\beta\)-D-glucopyranoside and 4-NP-\(\beta\)-L-arabinopyranoside
were designed to correspond with ARA-I tryptic amino acid
sequences, and PCR was performed using all combinations of
forward and reverse primers with reverse transcribed, 3-day-
old barley seedling mRNA. The nucleotide sequence of a 500-bp
PCR product corresponded exactly with the amino acid se-
quences of various tryptic peptides from purified ARA-I, and
the PCR product was therefore used to probe a cDNA library
from 24–48 h gibberellic acid\(_3\)-induced aleurone layers. Of
200,000 cDNA clones screened, a single 1,900-bp ARA-I cDNA
was identified. The missing 5'-end fragment of the cDNA
was subsequently isolated during additional screening of the cDNA
library, and 3'-RACE was used to generate the missing 3'-end
fragment. The fragments were used to assemble the full nucle-
otide sequence of the ARA-I cDNA, and a similar strategy was
used to assemble a near full-length cDNA encoding the XYL
enzyme (data not shown). In both cases a strong bias in codon
usage was evident for the two genes, with G or C residues found
in the wobble base position of about 95% of ARA-I and XYL
codons in the region encoding the mature enzymes (data not
shown).

Primary Structures of the Enzymes—The complete amino
acid sequences of the barley ARA-I and XYL enzymes were
deduced from the nucleotide sequences of the corresponding
cDNAs (Fig. 4). For the XYL enzyme, 38 amino acids were
sequenced from the NH\(_2\) terminus of the purified enzyme. This
showed that the NH\(_2\)-terminal residue of the mature enzyme
was the Ala residue of the ADPPF sequence indicated in Fig. 4.

Isolation of cDNAs—Degenerate oligonucleotide primers
were designed to correspond with ARA-I tryptic amino acid
sequences, and PCR was performed using all combinations of

|  |  |  |
|---|---|---|
| ARA-I | 4NP-P | 4NP-PX |
| | 6.5 | 45.6 |
| k\(_{cat}\) | 4.8 | 3.6 |
| U\(^{-1}\) | 3.57 \times 10\(^{3}\) | 3.6 \times 10\(^{3}\) |

Relative activities of ARA-I and XYL against aryl glycosides

| Aryl glycoside | ARA-I (\%) | XYL (\%) |
|---|---|---|
| 4NP-\(\alpha\)-L-arabinofuranosidase | 100 | 3 |
| 4NP-\(\alpha\)-L-arabinopyranosidase | 11 | 2 |
| 4NP-\(\beta\)-L-xylopyranosidase | 20 | 100 |
| 4NP-\(\beta\)-D-glucopyranosidase | 0 | 1 |
| 4NP-\(\beta\)-D-galactopyranosidase | 16 | 3 |

No activity was detected on 4NP-\(\beta\)-D-mannopyranoside, 4NP-\(\alpha\)-D-
xylopyranosides, 4NP-\(\alpha\)-D- and 4NP-\(\beta\)-D-mannopyranosides, 4NP-\(\alpha\)-D-
glucopyranoside or 4NP-\(\alpha\)-L-galactopyranoside.

Activity expressed as percent of specific activity against 4NP-P.
Activity expressed as percent of specific activity against 4NP-PX.
Fig. 3. Substrate specificities of ARA-I and XYL. A, thin layer chromatography of hydrolytic products released by ARA-I and XYL (first and second lanes, respectively) from wheat flour arabinoxylan. The third lane shows products released from the arabinoxylan by the Thermomyces endoxylanase. Products of the endoxylanase mixed with ARA-I or XYL are shown in the fourth and fifth lanes, respectively. Standards were L-arabinofuranose, D-xylopyranose, and oligoxylopyranosides xylobiose-xylohexaose (sixth, seventh, and eighth lanes, respectively). B, products released when ARA-I and XYL were incubated with xylopyranopentaose after 0.5, 10, and 60 min. C, products released when ARA-I and XYL were incubated with (1→5)-α-L-linked arabinofuranohexaose for 10 and 60 min.

Amino acid differences were observed, but these were attributed to differences in the varieties used to isolate the enzyme and the cDNA library. The NH$_2$-terminal sequence of ARA-I (AEAQAQAPVF) was predicted using the SigCleave program and corresponded to the experimentally determined sequence for the NH$_2$ terminus of the XYL enzyme (Fig. 4). In both cases the −3, −1 rule of von Heijne (42) was satisfied, but there was no obvious reason why the NH$_2$-terminal residue of ARA-I would be blocked to Edman degradation (43).

If it is accepted that the NH$_2$-terminal residue of ARA-I is as indicated in Fig. 4, the ARA-I and XYL cDNAs both encode mature polypeptides of 748 amino acids, and these show 51% positional identity for the two enzymes. Signal peptides of 29 residues (Fig. 4) were detected for each enzyme, and these have characteristics typical of those from other eukaryotic signal peptides that direct nascent polypeptides to the endoplasmic reticulum (44). The molecular mass calculated from the deduced amino acid sequence was 79,184 Da for ARA-I, which had a calculated isoelectric point of 5.7. The corresponding values calculated for XYL were 80,500 Da and 6.5 for XYL, respectively. Although the calculated isoelectric point values correspond well with the values of 5.5 and 6.7 for ARA-I and XYL, respectively, as determined from the purified enzymes, the molecular masses deduced from the cDNA sequences are considerably higher than the values of 67 kDa obtained for the purified enzymes on SDS gels run under reducing conditions (Fig. 2). This would suggest that processing of the primary translation product might have occurred and that a peptide fragment of up to about 12 kDa might have been removed during enzyme maturation. However, the NH$_2$-terminal end of XYL has certainly not been processed, and the presence of a Lys-C peptide starting at residue 29 of ARA-I suggests that a fragment of 12 kDa has not been cleaved from the NH$_2$ terminus of this enzyme either (Fig. 4). Internal processing of the enzymes would also be possible, but the 5 min 2-mercaptoethanol included here in the gel loading buffer during electrophoresis and during enzyme purification would dissociate individual peptide chains linked by disulfide bonds (45), and we could find no evidence for internal processing of the two enzymes.

In view of the discrepancy between observed and predicted apparent molecular mass values, ARA-I and XYL were examined by MALDI-TOF mass spectrometry. Broad peaks of 69.4 kDa and 68.2 kDa, with widths of 3–5 kDa, were obtained for ARA-I and XYL, respectively (data not shown). The amino acid sequence of ARA-I was investigated in more detail. Using a combination of MALDI-TOF and electrospray ionization mass spectrometry of tryptic and Lys-C peptides, coupled with Edman sequence analysis of the peptides, most of the NH$_2$-terminal and central peptides predicted from the ARA-I cDNA sequence could be identified (Fig. 4). However, the COOH-terminal region of the enzyme predicted by the cDNA sequence was not detected in any of the enzymic digests, and it was concluded on this basis and on the basis of domain prediction (46) that the actual COOH terminus of the mature ARA-I enzyme is in the vicinity of Met-614 (Fig. 4).

Catalytic Amino Acid Residues and Active Sites—Based on multiple sequence alignments (47) and HCA (33), the putative catalytic nucleophiles are predicted to be Asp-275 for ARA-I and Asp-268 for XYL (Fig. 4). These residues are absolutely conserved in family 3 glycoside hydrolases (19). Prediction of the catalytic acid/base residues of the enzymes is somewhat more complicated. Multiple sequence alignments and HCA clearly identify two candidate amino acid residues for this role (data not shown). Hrmova et al. (48) suggested that the catalytic acid/base for β-D-xylosidase-like members of the family 3 group would correspond to Glu-479 for ARA-I and Glu-472 for XYL. However, at this stage we believe that the catalytic residues could just as easily be Glu-481 for ARA-I and Glu-474 for XYL. Based on three-dimensional structure of the barley β-D-glucan glucohydrolase is used as a template (35), indicate that the Glu-481 residue is more appropriately positioned with respect to the known catalytic acid/base residue Glu-491 of the β-D-glucan glucohydrolase (data not shown). In these models the catalytic acid/base and nucleophile amino acid residues are about 6.5 Å apart. However, it must be pointed out that the sequence identities of the template and target enzymes are ~30% and that this is considered in the “twilight zone” of reliability of the molecular modeling programs (49).

Other features of family 3 glycoside hydrolases which can be observed in the ARA-I and XYL sequences include the conserved WGR and KH motifs, beginning at residues Trp-147 and Lys-192 for ARA-I, and Trp-139 and Lys-185 for XYL. These motifs are probably involved in substrate binding (34, 48). In addition, sequences similar to the conserved COOH-terminal antiparallel
loop of family 3 enzymes (34) were present in the region starting
at about residue 559 in ARA-I and 562 in XYL (Fig. 4).

**Expression Analysis of ARA-I and XYL Genes**—

Transcript levels of ARA-I and XYL in various barley tissues, including
developing grain, aleurone, scutellum, rootlets, coleoptiles, and
leaves, both etiolated and light-grown, were assessed by North-
ern hybridization analyses, but hybridization signals were very
low. Gene-specific primers were therefore synthesized to amplify
specifically short DNA fragments, using reversed transcription
PCR, from reverse transcribed total RNA from each plant tissue
sample. Sense oligonucleotide primers corresponded to cDNA
sequences within the coding region of respective cDNAs, and
antisense primers were designed for sequences within the unique
3'-untranslated regions of respective cDNAs.

ARA-I transcripts were detected in developing grains and in
the vegetative tissues of rootlets, coleoptiles, and leaves (Fig.
5). ARA-I does not appear to be transcribed in aleurone or
scutellum tissue 3 days after germination. XYL transcripts
were found in each of the barley tissues and at relatively high
levels (Fig. 5).

**Genetic Mapping of ARA-I and XYL Genes**—Single dominant
bands were evident when Southern hybridization analyses of
parental lines and mapping population DNA samples were
probed with ARA-I and XYL cDNAs (data not shown). This
suggested that single genes encoding these enzymes are pres-
ent on the barley genome, although it should be noted that the
hybridizations were performed at high stringency to avoid

---

**Fig. 4.** Alignment of ARA-I and XYL amino acid sequences. Identical residues in both sequences are shown in red. The pairwise alignment was prepared using ClustalX (41). The ARA-I and XYL cDNAs encode mature polypeptides of 748 amino acid residues with signal peptides of 29 residues each (shading). The amino acid residues are numbered from the NH2-terminal amino acid residues (arrowheads) of the mature enzymes. An arrowhead is also used to indicate the likely COOH terminus of ARA-I, and asterisks indicate potential N-glycosylation sites (marked as CHO). Arrows indicate the putative catalytic nucleophiles (Asp-275 for ARA-I and Asp-268 for XYL) and putative catalytic acid/bases (Glu-481 for ARA-I and Glu-474 for XYL). Blue overlines above the sequence of ARA-I indicate the amino acid sequences that were confirmed by either NH2-terminal or peptide sequencing with MALDI-TOF or electrospray ionization quadrupole mass spectrometry analyses, after proteolytic cleavage by Lys-C, trypsin, or CNBr.

**Fig. 5.** Reverse transcription-PCR analysis for ARA-I and XYL.
DNA was amplified from total RNA preparations of developing barley
grain (DG), aleurone (Al), scutellum (Sc), root (R), coleoptile (C), etio-
lated leaf (L_e), and light-grown leaf (L_l), using ARA-I (A) and XYL (B)
gen-specific primers. M indicates molecular markers, and C is a con-
trol, in which no cDNA was present.

---

cross-hybridization between the ARA-I and XYL probes, which
are about 50% identical, and related genes might therefore
have gone undetected. Indeed, preliminary amino acid se-
quence analysis of a protein band enriched in ARA-II revealed differences with ARA-I sequences. Despite the fact that ARA-II was not completely purified, it could be concluded from the sequence comparisons that there are at least two genes encoding α-L-arabinofuranosidases in barley (data not shown).

Restriction fragment length polymorphisms (RFLPs) for the DNA probes were rare, with only one RFLP for ARA-I found in DNA digested with HindIII from the parents Chebec and Harrington and for XYL RFLPs were present only for Clipper and Sahara genomic DNA digested with EcoRI or DraI. The ARA-I gene is located on the long arm of barley chromosome 2H, between the molecular markers ABC165 and BCD512, and XYL is found near the centromere of barley chromosome 6H, between markers Bmag9 and BCD269 (Fig. 6).

**DISCUSSION**

An α-L-arabinofuranosidase and a β-D-xylosidase, both of which are family 3 glycoside hydrolases (13), were purified ~1,000-fold (Table I) from extracts of germinated barley grain as outlined in Scheme 1. The purified enzymes have been designated ARA-I and XYL, respectively. A second, less abundant, α-L-arabinofuranosidase isoenzyme, designated ARA-II, was detected in the extracts (Fig. 1A) but was not purified to homogeneity. The ARA-I and XYL enzymes have apparent molecular masses of 67 kDa (Fig. 2) and isoelectric points of 5.5 and 6.7, respectively. Examination of their substrate specificities and kinetic properties indicated that α-L-arabinofuranosidase ARA-I can also hydrolyze 4NP-β-D-xylopyranoside, 4NP-β-D-galactopyranoside, and 4NP-α-L-arabinopyranoside at significant rates, whereas the XYL enzyme has a more restricted, or “tighter,” specificity for β-D-xylosides (Table III). Thus, the XYL enzyme hydrolyzes not only 4NPX but also, with a 30-fold lower catalytic efficiency, 4NPA. The catalytic efficiency factor for ARA-I was of the same order of magnitude for 4NPA and 4NPX but slightly higher for 4NPA (Table II). For ease of expression we have referred to the enzyme here as an α-L-arabinofuranosidase, but because ARA-I can hydrolyze both substrates efficiently, we acknowledge that it should probably be referred to as a bifunctional α-L-arabinofuranosidase/β-D-xylopyranosidase and that both activities might be important for its biological function in planta. Certain family 43 (50, 51), 54 (52–54), and 62 (16) α-L-arabinofuranosidases and β-D-xylopyranosidases show similar flexibility in their substrate specificities.

Using amino acid sequences generated from the purified...
barley ARA-I and XYL enzymes, several cDNAs were isolated, and near full-length cDNA sequences were subsequently assembled (Fig. 4). Deduced amino acid sequences indicated that both enzymes have a typical endoplasmic reticulum-targeting signal peptide (Fig. 4) that presumably directs secretion from cells in which they are synthesized. This is a significant observation, given recent indications that the (1 → 4)-β-D-xylan endohydrolase involved in arabinoxylan depolymerization in germinated barley grain is not located in the endomembrane secretory compartment of aleurone layers but is found instead in the cytosol and is likely to be released from aleurone layers only after programmed cell death (56, 57). In isolated aleurone layers, α-l-arabinofuranosidases and β-D-xylosidases are secreted and can be detected in the surrounding medium much earlier than the (1 → 4)-β-D-xylan endohydrolases (10). Thus, the secretion from aleurone layers of endohydrolases, α-l-arabinofuranosidases, and β-D-xylosidases involved in arabinoxylan degradation is clearly not coincident.

Although the NH₂ terminus of ARA-I could not be defined with certainty, the cDNAs encode primary translation products of 748 amino acid residues (Fig. 4). The calculated molecular masses of the enzymes, based on these deduced amino acid sequences, are about 80 kDa. This value is much higher than the apparent molecular mass values of 67 kDa observed during SDS-gel electrophoresis of the purified enzymes (Fig. 2) and represents a much longer polypeptide than other plant members of the family 3 group of glycoside hydrolases (19). Mass spectrometry was therefore used to examine further the molecular masses of the two enzymes and in both cases confirmed that the enzymes were ∼ 67 kDa in size. Particular attention was paid to ARA-I, for which the analysis of proteolytic peptides accounted for all regions of the enzyme except the COOH terminus predicted from the cDNA sequence (Fig. 4). The amino acid sequence data also suggested that the COOH termini of the enzymes were heterogeneous and that a single COOH-terminal residue could therefore not be identified. At this stage the weight of evidence suggests that the COOH termini of both ARA-I and XYL are close to the Met-614/Tyr-606 residue of ARA-I/XYL, respectively (Fig. 4). Thus, more than 130 amino acid residues appear to have been removed from the COOH terminus during post-translational processing of the enzymes. These values may be compared with the 605 residues found in the family 3 barley β-D-glucan glucohydrolase (58). No biological rationale for COOH-terminal processing of the barley ARA-I and XYL enzymes can be provided at this stage. In the case of the barley (1 → 4)-β-D-xylan endohydrolase, both NH₂- and COOH-terminal processing of the primary translation product occurs (56, 57).

Comparison of the amino acid sequences of the mature enzymes with other members of the family 3 glycoside hydrolases suggested that COOH-terminal processing does not occur in all members of the family. Although β-D-xylosidases from other higher plants are similar in size to those purified here from barley,7 the β-D-xylosidases from Aspergillus niger (59), Aspergillus oryzae (60), and Erwinia chrysanthemi (61) are much larger (∼ 85 kDa) than the barley enzymes and correspond in size to those predicted from cDNA sequences.

The relatively relaxed substrate specificities observed here for the family 3 ARA-I and XYL enzymes from barley can be rationalized in terms of their predicted three-dimensional structures. The three-dimensional structure of a family 3 β-D-glucan glucohydrolase from barley has been solved (62), and although it is the only family 3 crystal structure available, it has been used to model three-dimensional structures of other family 3 enzymes (19). Molecular modeling suggests that the barley ARA-I and XYL enzymes have overall structures similar to that of the β-D-glucan glucohydrolase from barley, although the three-dimensional conformation of the 130-amino acid residues COOH-terminal region of ARA-I and XYL, which is not present in the β-D-glucan glucohydrolase group, cannot be modeled (data not shown).

The barley β-D-glucan glucohydrolase has a broad specificity for different linkage types in unsubstituted oligomeric and polymeric β-D-glucan substrates (48, 63), probably because only two glucosyl residues of the substrate enter the active site pocket and because the glucosyl residue bound at subsite +1 is located between two tryptophan residues that allow some positional flexibility (48). The remainder of the substrate projects away from the enzyme surface, and activity is therefore relatively independent of substrate shape and hence of linkage type (34). The barley ARA-I and XYL enzymes examined here also exhibit some flexibility in substrate specificity. Both 4NP and 4NFX can fit in their catalytic sites. To provide a structural rationale for this observation, the three-dimensional structure of β-D-xylopyranose was taken from the Protein Data Bank and the three-dimensional structure of α-l-arabinofuranose was built. When the two structures were superimposed, a similar stereochemistry was observed about C-1, C-2, and C-3 in both pentoses, and their overall hydrodynamic volumes were also similar (Fig. 7). It is therefore not surprising that the active site of ARA-I can accommodate both substrates.

In addition to the aryl glycosides, both enzymes hydrolyze linear oligosaccharides, but neither hydrolyzes substituted polysaccharides. Similarly, the fact that no arabinose is removed from oligoarabinooligosaccharides released by (1 → 4)-β-D-xylan endohydrolase action (Fig. 3) suggests that neither enzyme is able to hydrolyze substituted oligomeric substrates completely. Only unsubstituted oligosaccharides or oligoarabinopyranosides with two to three unsubstituted xylosyl residues at their non-reducing ends would be expected to fit into a substrate-binding pocket of the shape found in other family 3 enzymes (19, 48), and only xylose would be released.

Family 3 glycoside hydrolases from higher plants can be

---

8 S. Khan, GenBank accession number AC009243, unpublished data.
grouped into two major clades, based on amino acid sequence alignments (48). One group contains the broad specificity β-D-glucan glucohydrolases, and the other contains β-D-xylosidases and α-L-arabinofuranosidases. As expected, the ARA-I and XYL enzymes characterized here fall into the second group (data not shown). Although the catalytic amino acid residues, corresponding to Glu-481 and Asp-275 for ARA-I and Glu-474 and Asp-268 for XYL (Fig. 4), are conserved in higher plant family 3 glycoside hydrolases, Hrmova et al. (48) provided a structural explanation for the differences in substrate specificity of the two groups. Thus, the conserved amino acid residue Asp-95 in the β-D-glucan glucohydrolase group that binds the C6-OH of the glycosyl residue bound at subsite $\text{-1}$ is not found in the α-L-arabinofuranosidase/β-D-xylosidase group. Clearly, the pentoses L-arabinofuranose and D-xylose have no C6-OH group, and the α-L-arabinofuranosidase/β-D-xylosidases have a Glu residue in the position corresponding to Asp-95 of the β-D-glucan glucohydrolase group.

When the phylogeny of the α-L-arabinofuranosidase/β-D-xylosidase group of family 3 enzymes is examined in more detail (Fig. 8), the higher plant representatives are clearly separated from the fungal representatives. There is one bacterial sequence of Thermotoga neapolitana in this group (Fig. 8). In most cases the true substrate specificities of enzymes encoded by the genes shown in Fig. 8 have not been investigated, and the β-D-xylosidase assignment of identity is based on similarities with a small number of partially characterized enzymes. The dual α-L-arabinofuranosidase/β-D-xylosidase specificity of the barley ARA-I has not been reported for other members of family 3 (13). It is noteworthy that the barley ARA-I is some distance from XYL in the phylogenetic tree, and this may eventually provide clues for more detailed classification of closely related enzymes in this family.

To provide some insight into the likely biological functions of the barley ARA-I and XYL enzymes, expression patterns of the genes were investigated, together with the action of the enzymes on well defined oligomeric and polymeric substrates. Reverse transcription-PCR showed the presence of XYL mRNA in all tissues examined. However, ARA-I mRNA appeared to be absent, or in very low abundance, in the aleurone layer and scutellum of germinated grain (Fig. 5). This is somewhat surprising, given that α-L-arabinofuranosidase activity, measured by activity on 4NPA, has been widely reported in the media surrounding isolated barley aleurone layers (9, 10, 65). At the substrate specificity level, XYL was able to hydrolyze (1→3)-β-D-xylopyranopentaose to xylose relatively quickly but exhibited no activity against (1→5)-α-L-arabinofuranohexaose (Fig. 3, B and C). In contrast, ARA-I hydrolyzed (1→5)-α-L-arabinofuranohexaose to arabinose and (1→4)-β-D-xylopyranopentaose to xylose, albeit at slow rates (Fig. 3, B and C).

Neither enzyme hydrolyzed arabinoxylan at a significant rate, but both ARA-I and XYL rapidly released xylose from

**Fig. 8.** Unrooted radial phylogenetic tree of plant family 3 β-D-xylosidases. Amino acid sequences of α-L-arabinofuranosidase/β-D-xylosidase and β-D-xylosidase were aligned with ClustalW (64). Branch lengths are drawn to scale. The GenBank accession numbers of the sequences are shown.
oligoarabinoxylans by the action of (1 \rightarrow 4)-\beta-D-xylanase.

The low levels of arabinose in these hydrolysates suggests that neither enzyme can bypass substituted xylosyl residues in oligoarabinoxylans. In further hydrolysis of oligosaccharides released from arabinoxylans by endohydrolases in germinated barley grain. The enzymes could also play an important role during cell wall turnover in elongating coleoptiles and in other tissues during normal growth and development.

Acknowledgments—We thank Dr. Peilin Xu for providing the aleurone cDNA library, Dr. Ken Chalmers for assistance with the genetic mapping, Dr. Neil Shirley for help with the protein analyses, and Dr. Yoji Hayasaka and Kris Ferguson for elements of the mass spectrometric analyses. Professor Bruce Stone, Dr. Hugues Driguez, Dr. Ross De Gori, and Dr. Andrew Harvey provided valuable advice in various aspects of the project.

REFERENCES

1. Fincher, G. B. (1992) in Barley: Genetics, Molecular Biology and Bio/Technology (Shewry, P. R., ed) pp. 433–437, CAB Publishing, Wallingford, U. K.
2. Viter, R. J., Kormelink, F. J. M., Angelino, S. A. G. F., and Voragen, A. G. J. (1994) Carbohydr. Polym. 24, 113–118
3. Carpita, N. C. (1996) Annu. Rev. Plant Physiol. Plant Mol. Biol. 47, 445–476
4. Iyama, K., Lam, T. T. T., and Stone, B. A. (1993) Plant Physiol. 104, 315–320
5. Andrewartha, K., Phillips, D. R., and Stone, B. A. (1997) Carbohydr. Res. 77, 191–204
6. Gibart, P., and Carpita, N. C. (1991) Plant Physiol. 97, 551–561
7. Wilke, K. C. B. (1979) Adv. Carbohydr. Chem. Biochem. 36, 215–264
8. Heyn, A. N. J. (1986) Plant Sci. 47, 75–82
9. Taiz, L., and Hornigan, W. A. (1976) Plant Physiol. 56, 380–386
10. Banki, M., Li, C. D., Langridge, P., and Fincher, G. B. (1997) Mol. Gen. Genet. 253, 599–608
11. Ferré, H., Broberg, A., Duus, J. Ø, and Thomsen, K. K. (2000) Eur. J. Biochem. 156, 387–393
12. Lee, B. C., Burton, R. A., Hrurnva, M., and Fincher, G. B. (2001) Biochem. J. 356, 181–189
13. Henrissat, B. (1998) Biochem. Soc. Trans. 26, 153–156
14. Pitsin, S. M., Voragen, A. G. J., and Eldertman, G. (1996) FEBS Lett. 398, 7–11
15. McKie, V. A., Black, G. W., Millward-Sadler, S. J., Hazelwood, G. P., Laurie, L. D., and Gilbert, H. J. (1997) Biochem. J. 323, 547–553
16. Vincent, P., Shareck, F., Dupont, C., Morosoli, R., and Kluepfel, D. (1997) Biochem. J. 322, 845–852
17. Gielkens, M. M. C., Visscher, J., and de Graaf, L. H. (1997) Curr. Genet. 31, 22–29
18. Matsuo, N., Kaneko, S., Kano, A., Kobayashi, H., and Kumakabe, I. (2000) Biochem. J. 346, 9–15
19. Harvey, A. J., Hrurnva, M., De Gori, R., Varghese, J. N., and Fincher, G. B. (2000) Proteins 41, 257–269
20. Zverlov, V. V., Liebl, W., Bachelein, M., and Schwarz, W. H. (1998) FEBS Microbiol. Lett. 164, 337–343
21. Hrurnva, M., and Fincher, G. B. (1993) Biochem. J. 289, 453–461
22. Laemmli, U. K. (1970) Nature 227, 680–685
23. Nelson, N. (1944) J. Biol. Chem. 153, 375–380
24. Somogyi, M. (1927) J. Biol. Chem. 69, 19–25
25. Leatherbarr, R. J. (1998) GraFit Version 4.0, Erithacus Software, Horley, Surrey, U. K.
26. Frohman, M. A., Dush, M. K., and Martin, G. R. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 8998–9002
27. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
28. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
29. Deverieux, J., Haeberli, P., and Smithies, O. (1984) Nucleic Acids Res. 12, 387–395