Plant Physiol. (1995) 109: 31–39

Temporal and Spatial Expression of Amygdalin Hydrolase and (*R*)-(+)-Mandelonitrile Lyase in Black Cherry Seeds[1]

Liansheng Zheng[2] and Jonathan E. Poulton*

Department of Biological Sciences, The University of Iowa, Iowa City, Iowa 52242

In black cherry (*Prunus serotina* Ehrh.) macerates, the cyanogenic diglucoside (*R*)-amygdalin undergoes stepwise degradation to HCN catalyzed by amygdalin hydrolase (AH), prunasin hydrolase, and (*R*)-(+)-mandelonitrile lyase (MDL). A near full-length AH cDNA clone (pAH1), whose insert encodes the isozyme AH I, has been isolated and sequenced. AH I exhibits several features characteristic of β-glucosidases of the BGA family, including their likely nucleophile center (isoleucine-threonine-glutamic acid-asparagine-glycine) and acid catalyst (asparagine-glutamic acid-proline/isoleucine) motifs. The temporal expression of AH and MDL in ripening fruit was analyzed by northern blotting. Neither mRNA was detectable until approximately 40 days after flowering (DAF), when embryos first became visible to the naked eye. Both mRNAs peaked at approximately 49 DAF before declining to negligible levels when the fruit matured (82 DAF). Taken together with enzyme activity data, these time courses suggest that AH and MDL expression may be under transcriptional control during fruit maturation. In situ hybridization analysis indicated that AH transcripts are restricted to the procambium, whereas MDL transcripts are localized within cotyledonary parenchyma cells. These tissue-specific distributions are consistent with the major locations of AH and MDL protein in mature seeds previously determined by immunocytochemistry (E. Swain, C.P. Li, and J.E. Poulton [1992] Plant Physiol 100: 291–300).

Approximately 3000 species of higher plants, including such agronomically important crops as cassava, sorghum, and rosaceous stone fruits, exhibit cyanogenesis (HCN release). In most cases, this HCN is generated during the degradation of cyanogenic glycosides, a group of β-glycosylated α-hydroxynitriles, by specific β-glycosidases and α-hydroxynitrile lyases. A general feature of cyanogenic species is that tissue disruption or infection is required to initiate the large-scale catabolism of these glycosides to HCN. Consequently, it is generally accepted that undamaged plants avoid premature, and possibly suicidal, cyanogenesis by some critical compartmentation of cyanoglycosides and their catabolic enzymes at either tissue or subcellular levels (Kojima et al., 1979; Poulton, 1988; Pancoro and Hughes, 1992).

Long regarded as highly cyanogenic, the kernels of *Prunus* species (Rosaceae) are a rich source of the cyanogenic diglucoside (*R*)-amygdalin [the β-gentiobioside of (*R*)-mandelonitrile] and its catabolic enzymes. In black cherry (*Prunus serotina*) seed macerates, three glycoproteins cooperate in cyanogenesis (Poulton, 1993). AH cleaves the β(1→6)-glycosidic bond of amygdalin, yielding the monoglucoside (*R*)-prunasin, which is subsequently hydrolyzed to (*R*)-mandelonitrile by PH. The dissociation of mandelonitrile to benzaldehyde and HCN may proceed nonenzymically, but this reaction is greatly accelerated by MDL (EC 4.1.2.10), which constitutes approximately 10% of the soluble protein of black cherry seeds. These enzymes were purified to homogeneity and characterized (Poulton, 1993). Monospecific polyclonal antisera raised against each of the deglycosylated proteins allowed us to gain some insights into the temporal and spatial regulation of cyanogenesis in maturing cherry fruits. The three catabolic enzymes, which first appeared within developing seeds about 6 weeks after flowering (Swain et al., 1992a), were localized at the tissue and subcellular levels by colloidal gold immunocytochemistry. AH and PH are restricted to protein bodies of the procambium, whereas MDL occurs primarily within protein bodies of the cotyledonary parenchyma cells (Swain et al., 1992b). When tissue printing subsequently localized amygdalin to the cotyledonary parenchyma cells, it became clear that premature cyanogenesis was precluded in intact black cherry and plum (*Prunus domestica*) seeds by segregation of AH and amygdalin in different tissues (Poulton and Li, 1994).

To begin exploring at the molecular level how cyanogenesis is temporally and spatially regulated within developing stone fruits, we first constructed a λgt11 cDNA expression library using poly(A)$^+$ RNA isolated from mid-maturation black cherry seeds. Screening this library yielded a full-length MDL cDNA clone, designated pMDL1, and two partial-length putative AH clones (Cheng and Poulton, 1993; Li, 1993). In the present paper, we report the isolation and characterization of a near full-length AH cDNA clone (designated pAH1) whose insert encodes AH I, one of the four known AH isozymes (Li et al., 1992). Furthermore, we describe how pAH1 and pMDL1 have been used to study the temporal and spatial expression of AH and MDL in maturing black cherry fruits.

[1] This work was supported by National Science Foundation grant No. IBN 9218929.

[2] Present address: Monsanto Agrochemical Co., St. Louis, MO 63198.

* Corresponding author; e-mail jepoultn@vaxa.uiowa.edu; fax 1-319-335-3620.

Abbreviations: AH, amygdalin hydrolase; DIG, digoxigenin; MDL, (R)-(+)-mandelonitrile lyase; PH, prunasin hydrolase.

MATERIALS AND METHODS

Plant Material

Developing fruits (29–82 DAF) were collected from a single black cherry (*Prunus serotina* Ehrh.) tree growing locally, immediately frozen in liquid N_2, and stored at −70°C.

RNA Isolation and Analysis

Total RNA was isolated from developing seeds ($n = 40$) essentially as described by Sharrock and Quail (1989), separated by electrophoresis (10 μg/lane) on denaturing 1.2% (w/v) agarose gels containing 1.2 M formaldehyde, and blotted onto nylon membranes (Micron Separations Inc., Westboro, MA) (Sambrook et al., 1989). After the samples were UV cross-linked, prehybridization and hybridization were undertaken at 65°C in 0.25 M Na_2HPO_4, pH 7.4, containing 1 mM EDTA, 1% (w/v) BSA, and 7% (w/v) SDS. The membranes were probed with ^{32}P-labeled pAH1 and pMDL1 inserts generated by random priming (Boehringer Mannheim) and subsequently washed twice (30 min each) at a maximum stringency of 0.1× SSC containing 0.1% (w/v) SDS at 65°C. Autoradiography was performed overnight at −80°C with intensifying screens.

Assay of AH and MDL Protein Levels in Developing Seeds

Immature fruits ($n = 40$), stored at −70°C since harvest, were halved with a razor blade. The developing seeds were rapidly excised and homogenized in a mortar at 4°C with 0.2 g of polyvinylpolypyrrolidone, 1 g of sand, and 15 mL of 0.1 M His-HCl buffer, pH 6.0. After the macerate was centrifuged twice for 25 min at 12,100*g*, an aliquot (2.5 mL) of the final supernatant liquid was chromatographed on a Sephadex G-25 column (8.3 × 1.5 cm) using 20 mM His-HCl buffer, pH 6.0. AH and MDL enzyme activities were assayed in duplicate as previously described (Swain et al., 1992a).

Isolation and Sequencing of AH cDNA Clones

Poly(A)$^+$ RNA was purified from total RNA using the PolyATract mRNA isolation system (Promega). Following the manufacturer's instructions (Stratagene), we constructed a cDNA library (3.2×10^5 plaque-forming units) in λZAPII using poly(A)$^+$ RNA isolated from mid-maturation seeds. Two partial-length putative AH cDNA clones, identified in previous work (Li, 1993), were labeled by random priming (Boehringer Mannheim) and utilized to screen approximately 2.5×10^5 plaque-forming units by standard methods (Sambrook et al., 1989). The longest cDNA insert recognized by both probes was subcloned into pBluescript SK(−) by in vivo excision, yielding the clone pAH1. Its insert (designated AH1) was sequenced completely in both directions by the dideoxy chain-termination method (Sanger et al., 1977). Sequencing analysis was performed using the University of Wisconsin Genetics Computer Group software package (Devereux et al., 1984).

In Situ RNA Localization

AH and MDL mRNAs were localized in paraffin-embedded seed sections by the nonisotopic DIG-labeling system (Boehringer Mannheim) essentially as described by Cox and Goldberg (1988). Sense and antisense DIG-labeled riboprobes were generated from both pAH1 and pMDL1 using T7 and T3 RNA polymerases and reduced to approximately 300 nucleotides in length by alkaline hydrolysis.

Seeds excised from immature fruits (49 DAF) were fixed in Histochoice medium following the manufacturer's instructions (Amresco, Solon, OH) before being embedded in Paraplast wax (Monoject Scientific, St. Louis, MO). Tissue sections (15 μm thick) were mounted on poly-L-Lys-coated glass slides. After deparaffinization with xylene, the sections were hydrated and treated with proteinase K. Hybridization was performed at 42°C for 16 h with the riboprobe (0.3 μg mL^{-1}) in 50 mL of hybridization buffer (50% [v/v] formamide, 5× SSC, 2% blocking reagent [Boehringer Mannheim], 0.1% [w/v] *N*-lauroylsarcosine, and 0.02% [w/v] SDS). The nonspecifically bound riboprobe was removed by three washings (10 min each) with 4× SSC containing 1 mM DTT followed by digestion for 30 min at 37°C with RNase A (20 μg mL^{-1}) in 10 mM Tris-HCl, pH 7.5, containing 0.5 M NaCl and 1 mM EDTA. The sections were finally washed following the method of Cox and Goldberg (1988). Immunological detection was performed using 5-bromo-4-chloro-3-indolyl-phosphate (30 μg mL^{-1}) and nitroblue tetrazolium (60 μg mL^{-1}) as chromogens. After 10 h, color development was terminated by adding 10 mM Tris-HCl, pH 8.0, containing 1 mM EDTA. Where indicated, sections were stained with Fast Green FCF (Sigma). The slides were permanently mounted in Cytoseal 60 (Stephens Scientific, Riverdale, NJ), examined under an Olympus model BH-2 microscope, and photographed using Kodak Ektar 100 film.

RESULTS AND DISCUSSION

Isolation and Characterization of AH and MDL cDNA Clones

β-Glycosidases and α-hydroxynitrile lyases involved in cyanogenesis in rosaceous stone fruits were among the earliest enzymes to be described in the scientific literature (Liebig and Wöhler, 1837; Rosenthaler, 1908). During the past 3 decades, these catabolic enzymes have been highly purified from many species, allowing the characterization of their major kinetic and physical properties (for review, see Poulton, 1993). More recently, their tissue and subcellular localizations have been determined by immunocytochemistry (Swain et al., 1992b; Swain and Poulton, 1994a, 1994b). However, little is known about the molecular biology of cyanogenesis in rosaceous stone fruits; this constitutes the major goal of our current research.

In 1993, we screened a λgt11 expression library constructed from poly(A)$^+$ RNA isolated from mid-maturation black cherry seeds, using polyclonal antibodies monospecific for AH and MDL. Although this initial screen yielded a full-length MDL cDNA clone designated pMDL1 (Cheng and Poulton, 1993), we were able to identify only two

partial-length AH clones (Li, 1993). Their inserts, which were 551 and 294 nucleotides in length, respectively, were 73 and 56% identical with the *Trifolium repens* linamarase (Hughes, 1993). Using such inserts as probes to rescreen this library failed to yield any longer clones, thereby indicating the need for alternative libraries. A new cDNA library was therefore constructed in λZAPII using poly(A)$^+$ RNA isolated from mid-maturation cherries by a method based on that of Sharrock and Quail (1989). Screening this library with the AH partial-length clones yielded 59 putative AH cDNA clones that hybridized to both probes. The longest insert, as revealed by PCR analysis, was subsequently subcloned into pBluescript SK(−) for double-strand sequencing, yielding a cDNA clone designated pAH1.

The nucleotide and derived amino acid sequences of the pAH1 insert are shown in Figure 1. The sequence, which is 1859 nucleotides long, contains a 1656-nucleotide open reading frame (beginning at position 1 and ending at position 1653 before a TGA stop codon) and 203 nucleotides of 3′ noncoding region terminated by a poly(A) tail. An amino acid sequence corresponding to the known N terminus of the isozyme AH I (Li et al., 1992) is present within the open reading frame and is underlined in Figure 1. The identification of this sequence confirms that the pAH1 insert encodes the isozyme AH I. However, since the known N terminus of AH I is encoded by nucleotides 70 to 117, it is probable that the putative primary translation product is processed by removal of an N-terminal signal sequence. The presence of this putative signal peptide is

```
   1  ACG AAG TTG GGC TCT TTG CTC TTA TGT GCG CTT CTC CTC GCT GGC TTT GCA TTG ACA AAT AGC AAA GCT GCG
   1   T   K   L   G   S   L   L   L   C   A   L   L   L   A   G   F   A   L   T   N   S   K   A   A

  73  AAA ACA GAT CCA CCC ATT CAC TGT GCT TCT CTC AAC AGG AGC AGT TTC GAT GCT CTC GAA CCA GGG TTC ATA
  25   K   T   D   P   P   I   H   C   A   S   L   N...R...S   S   F   D   A   L   E   P   G   F   I

 145  TTT GGC ACA GCC TCA GCA GCT TAC CAG TTC GAA GGT GCT GCA AAA GAA GAT GGT AGA GGA CCA AGT ATA TGG
  49   F   G   T   A   S   A   A   Y   Q   F   E   G   A   A   K   E   D   G   R   G   P   S   I   W

 217  GAT ACC TAC ACC CAC AAC CAT TCA GAA AGG ATC AAA GAT GGC AGT AAT GGA GAT GTC GCT GTT GAT CAA TAT
  73   D   T   Y   T   H   N...H...S   E   R   I   K   D   G   S   N   G   D   V   A   V   D   Q   Y

 289  CAC CGA TAT AAG GAA GAT GTG AGG ATT ATG AAG AAA ATG GGG TTT GAT GCT TAT AGG TTT TCT ATC TCG TGG
  97   H   R   Y   K   E   D   V   R   I   M   K   K   M   G   F   D   A   Y   R   F   S   I   S   W

 361  TCC AGA GTC TTG CCA AAT GGA AAG GTA AGT GGG GGC GTG AAT GAG GAT GGA ATC AAA TTT TAC AAC AAT CTC
 121   S   R   V   L   P   N   G   K   V   S   G   G   V   N   E   D   G   I   K   F   Y   N   N   L

 433  ATC AAT GAA ATC CTA CGT AAT GGT CTA AAA CCA TTT GTG ACA ATC TAT CAT TGG GAT CTT CCC CAA GCT TTA
 145   I   N   E   I   L   R   N   G   L   K   P   F   V   T   I   Y   H   W   D   L   P   Q   A   L

 505  GAG GAC GAA TAC GGT GGT TTC TTA AGC CCT AAT ATT GTC GAT CAC TTT AGA GAC TAT GCA AAC CTT TGT TTT
 169   E   D   E   Y   G   G   F   L   S   P   N   I   V   D   H   F   R   D   Y   A   N   L   C   F

 577  AAG AAA TTT GGC GAT CGA GTA AAA CAC TGG ATC ACG TTG AAT GAG CCA TAT ACC TTT AGT AGC AGT GGT TAT
 193   K   K   F   G   D   R   V   K   H   W   I   T   L   N   E   P   Y   T   F   S   S   S   G   Y

 649  GCA TAC GGG GTC CAT GCA CCA GGA CGA TGC TCT GCT TGG CAA AAA CTA AAT TGC ACT GGT GGG AAT TCG GCA
 217   A   Y   G   V   H   A   P   G   R   C   S   A   W   Q   K   L   N...C...T   G   G   N   S   A

 721  ACT GAA CCA TAT TTG GTG ACA CAC CAC CAA CTC CTT GCT CAT GCA GCG GCT GTA AAA TTG TAC AAA GAT GAA
 241   T   E   P   Y   L   V   T   H   H   Q   L   L   A   H   A   A   A   V   K   L   Y   K   D   E

 793  TAT CAG GCA TCT CAA AAT GGC TTG ATA GGA ATA ACA TTG GTG TCA CCT TGG TTT GAG CCT GCT TCG GAG GCA
 265   Y   Q   A   S   Q   N   G   L   I   G   I   T   L   V   S   P   W   F   E   P   A   S   E   A

 865  GAG GAA GAT ATA AAT GCT GCA TTT CGA TCT TTG GAT TTT ATT TTT GGA TGG TTT ATG GAC CCG TTG ACA AAT
 289   E   E   D   I   N   A   A   F   R   S   L   D   F   I   F   G   W   F   M   D   P   L   T   N

 937  GGT AAC TAT CCG CAC CTC ATG CGA TCA ATT GTT GGG GAA CGA TTA CCA AAT TTC ACG GAA GAA CAA TCC AAG
 313   G   N   Y   P   H   L   M   R   S   I   V   G   E   R   L   P   N...F...T   E   E   Q   S   K

1009  TTG CTA AAG GGG TCA TTT GAT TTT ATT GGA CTA AAT TAT TAT ACA ACT AGA TAT GCA AGC AAT GCA CCT AAG
 337   L   L   K   G   S   F   D   F   I   G   L   N   Y   Y   T   T   R   Y   A   S   N   A   P   K

1081  ATT ACT TCT GTA CAT GCA AGC TAC ATA ACA GAT CCT CAA GTT AAT GCT ACA GCT GAG CTT AAG GGG GTC CCC
 361   I   T   S   V   H   A   S   Y   I   T   D   P   Q   V   N...A...T   A   E   L   K   G   V   P

1153  ATT GGT CCA ATG GCT GCT TCA GGC TGG TTA TAT GTT TAT CCC AAA GGA ATT CAC GAT CTT GTA CTT TAC ACA
 385   I   G   P   M   A   A   S   G   W   L   Y   V   Y   P   K   G   I   H   D   L   V   L   Y   T

1225  AAG GAA AAG TAT AAT GAT CCC CTC ATT TAC ATT ACT GAG AAT GGG GTT GAT GAG TTC AAT GAT CCC AAA TTA
 409   K   E   K   Y   N   D   P   L   I   Y [ I   T   E   N   G ] V   D   E   F   N   D   P   K   L

1297  TCA ATG GAG GAA GCC CTC AAA GAT ACC AAT AGA ATT GAC TTT TAT TAT CGT CAC CTT TGT TAC CTT CAA GCA
 433   S   M   E   E   A   L   K   D   T   N   R   I   D   F   Y   Y   R   H   L   C   Y   L   Q   A

1369  GCC ATC AAA AAG GGT TCT AAA GTG AAG GGT TAC TTT GCA TGG TCA TTT CTA GAC AAC TTT GAA TGG GAT GCA
 457   A   I   K   K   G   S   K   V   K   G   Y   F   A   W   S   F   L   D   N   F   E   W   D   A

1441  GGA TAC ACT GTT CGA TTT GGT ATC AAC TAC GTG GAT TAC AAT GAC AAT TTA AAA AGG CAC TCT AAA CTC TCA
 481   G   Y   T   V   R   F   G   I   N   Y   V   D   Y   N   D   N   L   K   R   H   S   K   L   S

1513  ACG TAC TGG TTC ACA AGT TTC CTC AAG AAG TAC GAA AGA AGT ACG AAA GAA ATC CAA ATG TTT GTG GAA AGT
 505   T   Y   W   F   T   S   F   L   K   K   Y   E   R   S   T   K   E   I   Q   M   F   V   E   S

1585  AAA CTA GAA CAT CAA AAG TTT GAA TCC CAA ATG ATG AAT AAA GTA CAA AGC TCT CTA GCA GTC GTC GTG TGA
 529   K   L   E   H   Q   K   F   E   S   Q   M   M   N   K   V   Q   S   S   L   A   V   V   V   *

1657  GTTGTGTTTTCAGTTTTTTTAGTTTGACTTTGTATGAGTCGAAATAAGTTGAAACAACTCATTTTGAGTTTGTTGTATCTCCTTGTGAGAATTT
      AGTTTTAAGTTTGCTTTAAGTTGTTGATGTCTGGTTTATCTTCGTGTGCTGAGTTTATGCAATAAAGGTGTTTCGACCCTTTGTGGTTGTGGTT
      GGAAACATAAAAAAA
```

Figure 1. Nucleotide and derived amino acid sequences of the pAH1 cDNA insert. The singly underlined sequence within the open reading frame corresponds to the known N terminus of the mature AH I protein. Putative *N*-glycosylation sites are underlined by dotted lines. The star indicates the stop codon. The motif Ile-Thr-Glu-Asn-Gly, which includes the predicted active site nucleophile Glu^{421}, is within a box.

consistent with the protein body location of AH in black cherry seeds (Swain et al., 1992b). The processed protein has a predicted molecular mass of 60.3 kD and a pI of 6.76, which correlates well with the observed values for the glycosylated protein of 62 kD (by SDS-PAGE) and 6.6, respectively. Consistent with the known glycoprotein nature of AH I, five putative *N*-glycosylation sites (Asn-X-Ser/Thr) are present in the deduced amino acid sequence, although Asn^{329} may remain unglycosylated because this site has Pro at position +1 of the consensus sequence.

During the past decade, the availability of the primary structures of several hundred glucan hydrolases (EC 3.2.1.x) has allowed these enzymes to be reclassified based on their structural similarities rather than according to their substrate specificities (Henrissat, 1991). Sequence analysis indicates that β-glucosidases fall into two distinct families, designated by Béguin (1990) as families BGA and BGB. The BGA family includes β-glucosidases, phospho-β-glycosidases, thio-β-glucosidases, and β-galactosidases from organisms as diverse as archaebacteria, bacteria, plants, and mammals. The BGB family includes fungal enzymes and the β-glucosidases of rumen bacteria.

The common structural features shown by BGA enzymes suggest a shared mechanism for the enzymatic hydrolysis of β-glycosidic bonds. The BGA enzymes catalyze glycoside hydrolysis by a double-displacement mechanism in which an enzymatic nucleophile attacks the substrate forming a glucosyl-enzyme intermediate (Sinnott, 1990). The latter is then hydrolyzed, releasing the sugar with overall retention of anomeric configuration. Aglycone group departure may be aided by a protonated amino acid side chain with acid catalytic function. Withers et al. (1990) identified the active site nucleophile (Glu^{358}) of the *Agrobacterium faecalis* β-glucosidase by inactivating the enzyme with the mechanism-based inhibitor 2′,4′-dinitrophenyl-2-deoxy-2-fluoro-β-D-glucopyranoside. The catalytic importance of this amino acid residue, which occurs as part of an Ile/Val-Thr-Glu-Asn-Gly motif that is highly conserved in members of the BGA family (but not of the BGB family), was subsequently confirmed by site-directed mutagenesis (Trimbur et al., 1992). The nature of the acid-base catalytic residue remains less certain. Trimbur et al. (1992) suggested the participation of Asp^{374}, which exists as part of an Asp-X-Arg-X-Tyr sequence that is highly conserved in the BGA family glucosidases. More recently, however, the inactivation of cassava linamarase by *N*-bromoacetyl-β-D-glucopyranosylamine identified Glu^{198} as the likely active site carboxylate group with acid catalytic function (Keresztessy et al., 1994b). This residue occurs within a Asn-Glu-Pro/Ile motif that is also highly conserved in BGA family β-glycosidases (with the exception of myrosinases).

Comparison of the deduced amino acid sequence of the mature AH I protein with other known sequences in the data base leads us to conclude that *P. serotina* AH belongs to the BGA family of β-glycosidases for the following reasons: (a) AH I exhibits high homology (37–65% identity) with other members of the BGA family, even including the prokaryotic and mammalian members (Table I). Highest similarity was observed with white clover linamarase (65.5% identity; 78.5% similarity), the white clover noncyanogenic β-glucosidase of unknown physiological function (59.1% identity; 73.1% similarity), and cassava linamarase (51.7% identity; 69.5% similarity). Like AH, these linamarases hydrolyze both cyanogenic glycosides and the chromogenic substrate *p*-nitrophenyl-β-D-glucoside. Somewhat less homology (43–48% identity) was observed with *O*-β-glucosidases from the monocots maize and oats and with thioglucosidases from the Brassicaceae. (b) Contrasting with the magnitude of the foregoing values, the AH I amino acid sequence exhibits only low homology (16–23% identity) with six members of the BGB family (data not shown). (c) The deduced AH I sequence includes the highly conserved Ile-Thr-Glu-Asn-Gly motif (putative active site nucleophilic center) at residues 396 to 400 of the mature protein. (d) The sequence also contains the Asn-Glu-Pro

Table I. *Identity (percentage) of the deduced amino acid sequence of the mature AH I protein (less signal sequence) with several glucosidases of the BGA family, as determined by the FastA program*

Except for AH 1, the enzymes are identified by their GenBank accession numbers: AH I, *P. serotina* (this work); X56733, *T. repens* linamarase (Hughes, 1993); X56734, *T. repens* noncyanogenic β-glucosidase (Hughes, 1993); S35175, *Manihot esculenta* linamarase (Hughes, 1992); X74217, *Zea mays* (Brzobohaty et al., 1993); X78433, *Avena sativa* (Gus-Mayer et al., 1994); X59879, *Sinapis alba* myrosinase (Xue et al., 1992); X60214, *Brassica napus* myrosinase (Falk et al., 1992); L11454, *Arabidopsis thaliana* thioglucosidase (Chadchawan et al., 1993); M96979, *Bacillus circulans* β-glucosidase (Paavilainen et al., 1993); M19033, *Agrobacterium faecalis* (Wakarchuk et al., 1988); M61841, human lactase-phlorizin hydrolase domain 3 (Mantei et al., 1988). Where known, the sequences of mature proteins (lacking signal or transit peptides) were used for analysis.

Enzyme	X56733	X56734	S35175	X74217	X78433	X59879	X60214	L11454	M96979	M19033	M61841
AH I	65.5	59.0	51.7	47.4	43.7	46.9	46.1	48.1	43.4	36.8	40.3
X56733		59.3	50.7	47.4	45.6	48.5	48.4	48.6	42.1	36.9	42.6
X56734			42.6	45.9	45.2	46.0	45.2	46.2	39.7	36.6	38.3
S35175				41.6	41.6	40.5	40.9	41.7	38.4	33.1	37.6
X74217					61.0	40.9	41.0	41.0	39.0	34.3	37.4
X78433						39.5	39.1	39.2	38.2	33.8	40.0
X59879							90.8	70.1	39.1	32.1	39.3
X60214								72.5	38.9	33.6	38.5
L11454									35.5	30.8	36.7
M96979										46.3	39.5
M19033											35.1

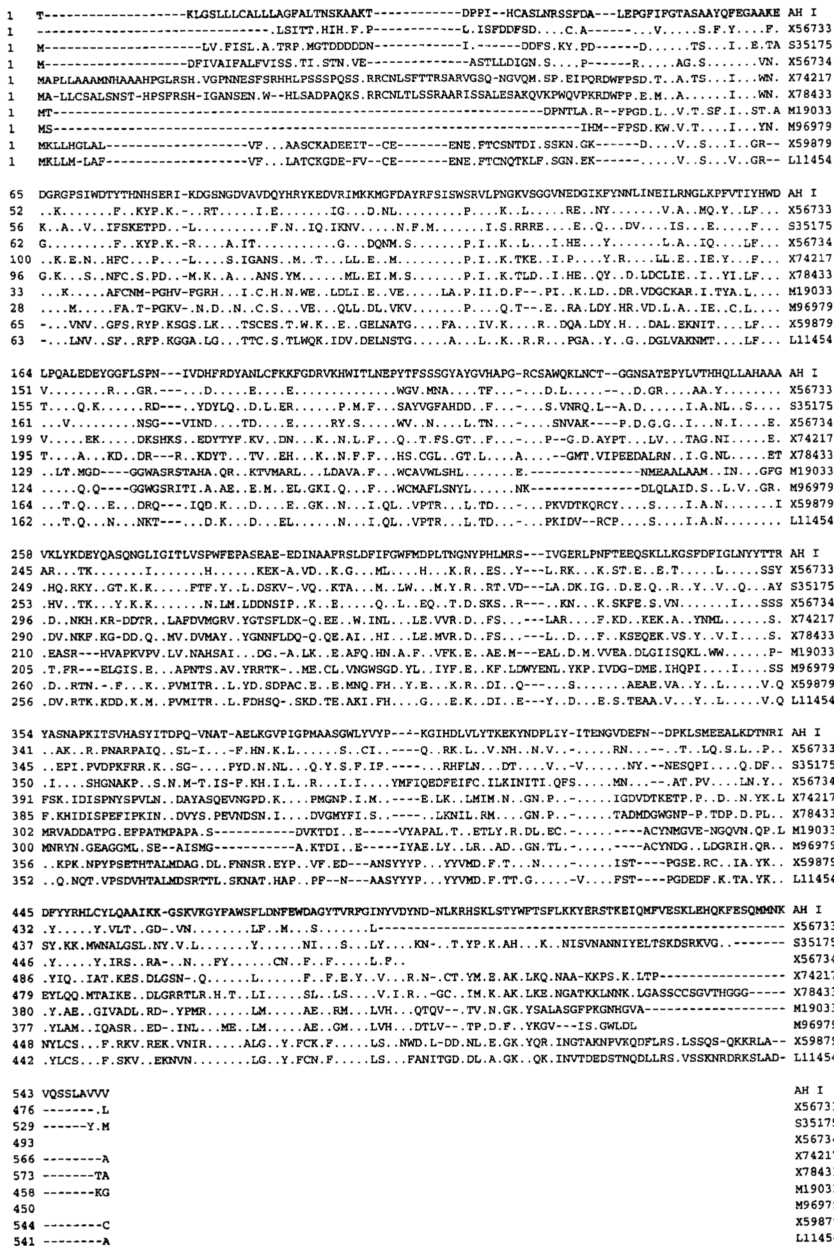

Figure 2. Multiple sequence alignment of the deduced amino acid sequence of AH I with sequences of nine other members of the BGA family of β-glucosidases. Amino acid alignment was performed using the DNASTAR (Madison, WI) Megalign Clustal Program (PAM250 residue weight table). The sources of β-glucosidase sequences (identified by GenBank accession number, except for AH I) used in this comparison are: AH I, *P. serotina* (this work); X56733, *T. repens* linamarase (Hughes, 1993); S35175, *M. esculenta* linamarase (Hughes et al., 1992); X56734, *T. repens* noncyanogenic β-glucosidase (Hughes, 1993); X74217, *Z. mays* (Brzobohaty et al., 1993); X78433, *Avena sativa* (Gus-Mayer et al., 1994); M19033, *Agrobacterium faecalis* (Wakarchuk et al., 1988); M96979, *B. circulans* β-glucosidase (Paavilainen et al., 1993); X59879, *S. alba* myrosinase (Xue et al., 1992); L11454, *Arabidopsis thaliana* myrosinase (Chadchawan et al., 1993). Residues identical with those of AH I are shown by dots, whereas introduced gaps are represented by dashed lines.

motif (putative acid-base catalyst) at residues 183 to 185 of the mature protein. (e) The sequence includes four His residues that are conserved within the four other sequenced plant β-glucosidases (Fig. 2). In this context, the presence of a reactive His residue at the active center of cassava linamarase should be noted (Keresztessy et al., 1994a). (f) Three of the five putative *N*-glycosylation sites within the AH I sequence are shared by white clover linamarase. The high homology between AH I and other BGA family β-glucosidases, not only at the putative active site moieties but also scattered throughout their entire sequences, is clearly illustrated by the multiple sequence alignment in Figure 2.

Developmental Expression of AH and MDL in Maturing Fruits

In previous work (Swain et al., 1992a), biochemical changes related to cyanogenesis were monitored during the maturation of black cherry fruits. It was shown that, concomitant with cotyledon development during phase II, the seeds begin accumulating both amygdalin and the catabolic enzymes AH, PH, and MDL and, from that time onward, are therefore highly cyanogenic when disrupted. In contrast, the pericarp remains acyanogenic throughout the entire ripening process because it lacks the catabolic enzymes.

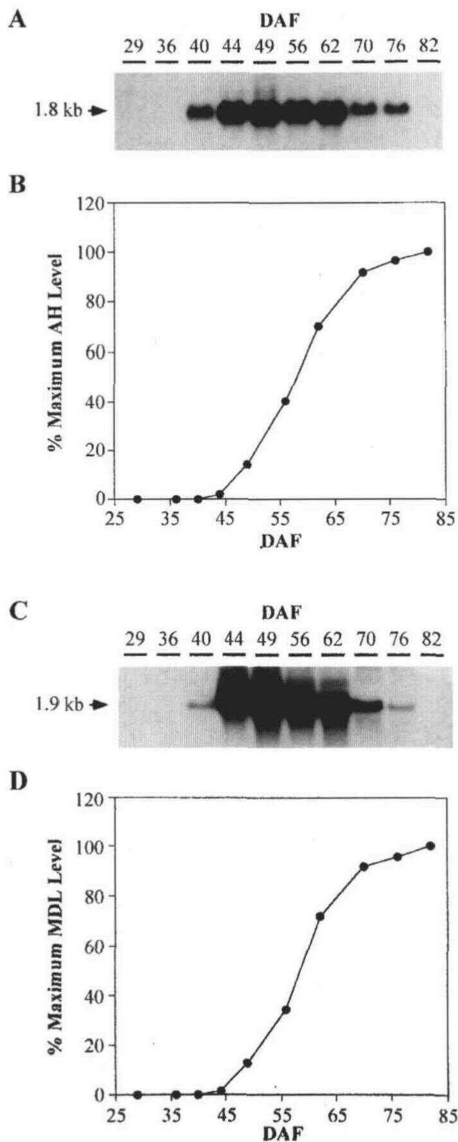

Figure 3. Temporal accumulation of AH1 and MDL1 mRNAs and proteins during fruit maturation in *P. serotina*. A and C, Northern blot analyses. Total RNA (10 μg) isolated from developing seeds at the times indicated (DAF) was fractionated on a denaturing agarose gel and blotted onto nylon membranes. The blots were hybridized with ^{32}P-labeled pAH1 insert (A) or pMDL1 insert (C) under conditions described in "Materials and Methods." B and D, Estimation of AH (B) and MDL (D) protein levels by direct enzyme assay. Enzyme activities are given as a percentage of the maximum level observed for each enzyme (AH, 48.8 μmol min^{-1} seed^{-1}; MDL, 10.8 μmol min^{-1} seed^{-1}). Each data point represents the mean of duplicates, the range of which did not exceed the dimensions of the symbol shown.

In the current study, we have now utilized northern blots to examine the temporal expression of AH1 and MDL1 transcripts in ripening fruits. Total RNA was isolated at approximately weekly intervals from maturing seeds (29–82 DAF) and probed using ^{32}P-labeled pAH1 and pMDL1 cDNA inserts. As Figure 3, A and C, illustrates, neither mRNA was detectable during phase I of fruit ripening. However, when embryos first became visible to the naked eye during early phase II (40 DAF), both transcripts became detectable, increasing to reach a maximum at approximately 49 DAF. Subsequently, transcript levels declined and were undetectable at full fruit maturity (82 DAF). It should be noted that the mRNAs detected by the AH1 and MDL1 probes were approximately 1.8 and 1.9 kb in length, respectively; these values correlate well with the known sizes of the cDNAs for these enzymes (Cheng and Poulton, 1993).

The levels of AH and MDL proteins were also measured during fruit ripening by direct enzyme assay of seed homogenates. Confirming previously published data (Swain et al., 1992a), Figure 3, B and D, illustrates that AH and MDL activities were undetectable until 44 DAF. They then increased rapidly during mid-phase II, essentially reaching a plateau at full fruit maturity (82 DAF). Comparison of their respective temporal patterns of mRNA and protein accumulation during fruit maturation suggests that the expression of AH and MDL may be regulated at the transcriptional level, although run-on transcription studies are required to confirm this tentative conclusion.

In Situ Localization of AH1 and MDL1 mRNAs in Immature Embryos

The spatial expression patterns of AH and MDL mRNAs were analyzed by in situ hybridization using antisense and sense DIG-labeled riboprobes transcribed from pAH1 and pMDL1, respectively. Immature seeds collected 49 DAF were selected for analysis because northern analysis had shown that they exhibit the highest transcript levels (Fig. 3). When seed sections were hybridized with the antisense AH1 riboprobe, intense labeling was observed exclusively within the procambial cells (Fig. 4, A and B). By contrast, control hybridizations using the DIG-labeled sense AH1 riboprobe gave no positive signal (Fig. 4C). The tissue-specific localization of AH1 mRNA within the procambial cells is in accordance with our previous immunocytochemical data showing that AH protein is restricted to the protein bodies of that tissue (Swain et al., 1992b).

MDL, which constitutes approximately 10% of the soluble protein of black cherry seeds, is believed to be multifunctional, serving both as a storage protein and in cyanogenesis (Swain et al., 1992b). In situ hybridization analysis showed that the spatial expression of MDL1 mRNA differs greatly from that of AH1 mRNA, although both transcripts show similar temporal expression patterns in developing seeds (Fig. 3). In contrast to the procambial location of AH1 transcripts, MDL1 mRNA exhibited a spatial expression pattern more characteristic of storage proteins, being restricted to the cotyledonary parenchyma cells (Fig. 4D). With the antisense MDL1 riboprobe, hybridization signals were strongest at the interior of the cotyledon and diminished sharply toward the periphery of that organ. No signal was observed when seed sections were exposed to the sense MDL1 riboprobe (Fig. 4E). The presence of MDL1 mRNA in the cotyledonary storage parenchyma cells correlates well with the known major location of MDL protein in mature seeds (Swain et al., 1992b). Although previous

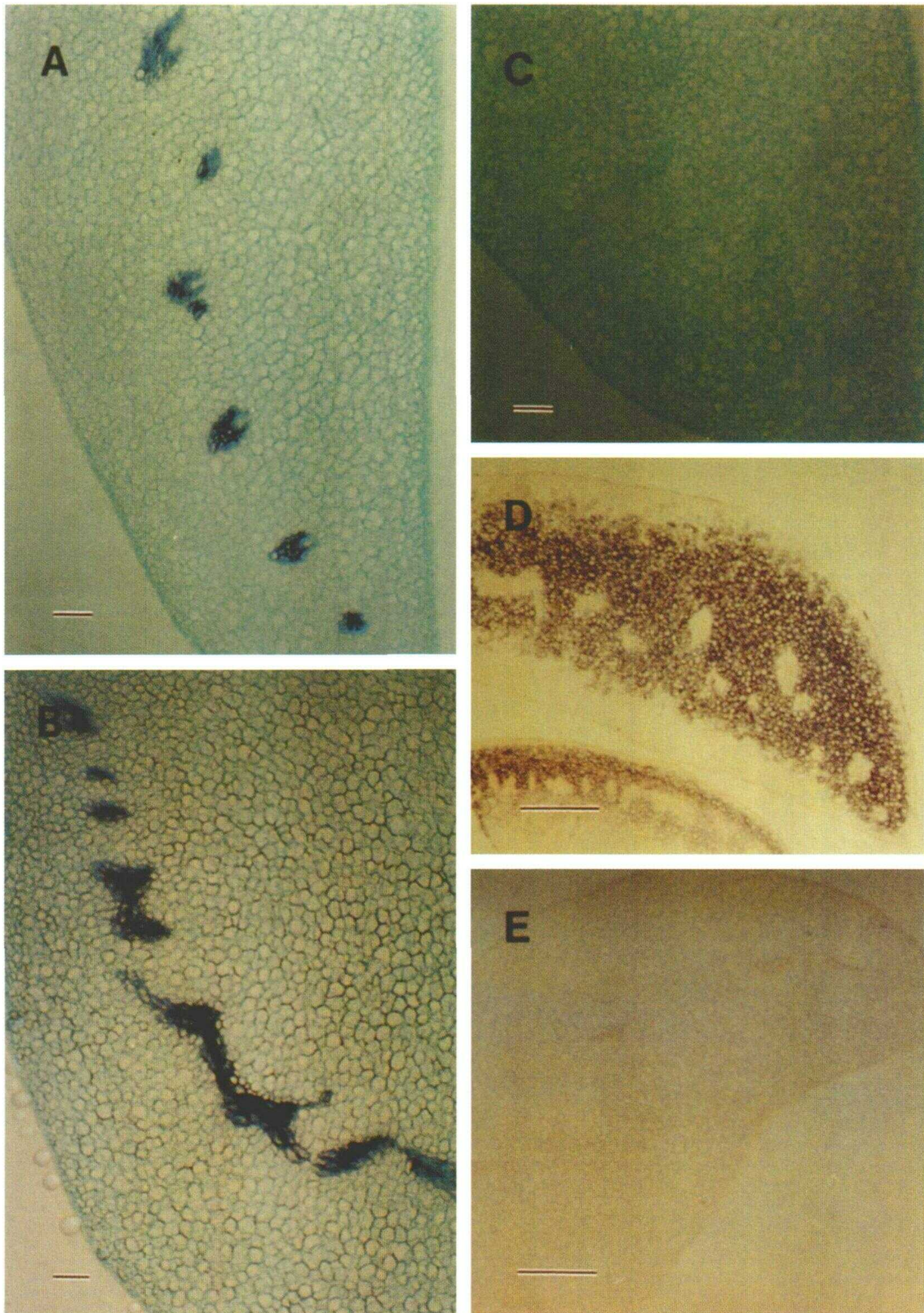

Figure 4. Localization of AH and MDL expression in immature black cherry seeds (49 DAF) by in situ hybridization. Tissue was fixed, embedded in paraffin, sectioned, and hybridized in situ with DIG-labeled sense or antisense transcripts synthesized from pAH1 and pMDL1. The dark blue or purple represents hybridization to target mRNAs, indicating their tissue distributions. A, Transverse section probed with AH1 antisense riboprobe, stained with Fast Green FCF. Bar, 200 μm. B, Longitudinal section probed with AH1 antisense riboprobe, stained with Fast Green FCF. Bar, 200 μm. C, Transverse section probed with AH1 sense riboprobe, stained with Fast Green FCF. Bar, 200 μm. D, Transverse section probed with MDL1 antisense riboprobe. Bar, 500 μm. E, Transverse section probed with MDL1 sense riboprobe. Bar, 500 μm.

immunocytochemical studies had also detected minor amounts of MDL protein in procambial cells, MDL1 mRNA was not detectable in such cells during the present study (Fig. 4, D and E). Assuming that in situ hybridization and immunocytochemistry are equally sensitive to localizing small amounts of their respective target molecules, we offer two explanations that might account for this apparent discrepancy. First, it should be noted that our immunocytochemical study was undertaken using fully mature seeds, whereas the in situ hybridization analysis performed here utilized immature (49 DAF) seeds. It is therefore possible that procambial MDL expression occurs after 49 DAF and was therefore not detected by our in situ analysis. Alternatively, because black cherry has several MDL isozymes (Yemm and Poulton, 1986), it is possible that the procambial MDL protein is encoded by a distinct lyase gene whose mRNA hybridizes poorly, if at all, to the MDL1 probe used here.

In conclusion, the isolation and sequencing of the AH cDNA clone pAH1 constitute important steps toward a better understanding of the molecular biology of cyanogenesis in rosaceous stone fruits. Confirming previous immunocytochemical data, our in situ analysis has again demonstrated the remarkable tissue-specific expression of AH and MDL in rosaceous stone fruits. Seeking to understand the molecular mechanisms that underlie such patterns of expression, we are currently characterizing AH and MDL genomic clones with the goal of identifying those promoter regions that confer tissue specificity in black cherry seeds.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Chun Ping Li (University of Arizona) for isolation of the partial-length AH cDNA clones, Drs. Richard Sjolund and Shelley Plattner for assistance with photographic reproductions, the University of Iowa DNA Facility for DNA sequencing, and Dr. Asim Esen (Virginia Polytechnic Institute) for helpful discussions.

Received March 20, 1995; accepted June 7, 1995.

Copyright Clearance Center: 0032-0889/95/109/0031/09.
The GenBank accession number for the sequence reported in this article is U26025.

LITERATURE CITED

Béguin P (1990) Molecular biology of cellulose degradation. Annu Rev Microbiol **44:** 219–248

Brzobohaty B, Moore I, Kristoffersen P, Bako L, Campos N, Schell J, Palme K (1993) Release of active cytokinin by a β-glucosidase localized to the maize root meristem. Science **262:** 1051–1054

Chadchawan S, Bishop J, Thangstad OP, Bones AM, Mitchell-Olds T, Bradley D (1993) *Arabidopsis* cDNA sequence encoding myrosinase. Plant Physiol **103:** 671–672

Cheng I-P, Poulton JE (1993) Cloning of cDNA of *Prunus serotina* (*R*)-(+)-mandelonitrile lyase and identification of a putative FAD-binding site. Plant Cell Physiol **34:** 1139–1143

Cox KH, Goldberg RB (1988) Analysis of plant gene expression. *In* CH Shaw, ed, Plant Molecular Biology: A Practical Approach. IRL Press, Oxford, UK, pp 1–34

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res **12:** 387–395

Falk A, Xue J, Lenman M, Rask L (1992) Sequence of a cDNA clone encoding the enzyme myrosinase and expression of myrosinase in different tissues of *Brassica napus*. Plant Sci **83:** 181–186

Gus-Mayer S, Brunner H, Schneider-Poetsch HAW, Rudiger W (1994) Avenacosidase from oat: purification, sequence analysis and biochemical characterization of a new member of the BGA family of β-glucosidases. Plant Mol Biol **26:** 909–921

Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J **280:** 309–316

Hughes MA (1993) Molecular genetics of plant cyanogenic β-glucosidases. *In* A Esen, ed, Symposium Series 533. American Chemical Society, Washington, DC, pp 153–169

Hughes MA, Brown K, Pancoro A, Murray BS, Oxtoby E, Hughes J (1992) A molecular and biochemical analysis of the structure of the cyanogenic β-glucosidase (linamarase) from cassava (*Manihot esculenta* Cranz). Arch Biochem Biophys **295:** 273–279

Keresztessy Z, Kiss L, Hughes MA (1994a) Investigation of the active site of the cyanogenic β-D-glucosidase (linamarase) from *Manihot esculenta* Crantz (cassava). I. Evidence for an essential carboxylate and a reactive histidine residue in a single catalytic center. Arch Biochem Biophys **314:** 142–152

Keresztessy Z, Kiss L, Hughes MA (1994b) Investigation of the active site of the cyanogenic β-D-glucosidase (linamarase) from *Manihot esculenta* Crantz (cassava). II. Identification of Glu-198 as an active site carboxylate group with acid catalytic function. Arch Biochem Biophys **315:** 323–330

Kojima M, Poulton JE, Thayer SS, Conn EE (1979) Tissue distributions of dhurrin and of enzymes involved in its metabolism in leaves of *Sorghum bicolor*. Plant Physiol **63:** 1022–1028

Li CP (1993) Cyanogenesis in rosaceous stone fruits: temporal and spatial expression of amygdalin hydrolase in *Prunus serotina* seeds. PhD thesis. University of Iowa, Iowa City

Li CP, Swain E, Poulton JE (1992) *Prunus serotina* amygdalin hydrolase and prunasin hydrolase. Purification, N-terminal sequencing, and antibody production. Plant Physiol **100:** 282–290

Liebig J, Wöhler F (1837) Ueber die Bildung des Bittermandeloels. Liebigs Ann Chem **22:** 1–24

Mantei N, Villa M, Enzler T, Wacker H, Boll W, James P, Hunziker W, Semenza G (1988) Complete primary structure of human and rabbit lactase-phlorizin hydrolase: implications for biosynthesis, membrane anchoring and evolution of the enzyme. EMBO J **7:** 2705–2713

Paavilainen SK, Hellman J, Korpela T (1993) Purification, characterization, gene cloning, and sequencing of a new beta-glucosidase from *Bacillus circulans* subsp. alkalophilus. Appl Environ Microbiol **59:** 927–932

Pancoro A, Hughes MA (1992) *In-situ* localization of cyanogenic β-glucosidase (linamarase) gene expression in leaves of cassava (*Manihot esculenta* Cranz) using non-isotopic riboprobes. Plant J **2:** 821–827

Poulton JE (1988) Localization and catabolism of cyanogenic glycosides. Ciba Found Symp **140:** 67–91

Poulton JE (1993) Enzymology of cyanogenesis in rosaceous stone fruits. *In* A Esen, ed, Symposium Series 533. American Chemical Society, Washington, DC, pp 170–190

Poulton JE, Li CP (1994) Tissue level compartmentation of (*R*)-amygdalin and amygdalin hydrolase prevents large-scale cyanogenesis in undamaged *Prunus* seeds. Plant Physiol **104:** 29–35

Rosenthaler L (1908) Durch Enzyme bewirkte asymmetrische Synthesen. Biochem Z **14:** 238–253

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA **74:** 5463–5467

Sharrock RA, Quail PH (1989) Novel phytochrome sequences in

Arabidopsis thaliana: structure, evolution, and differential expression of a plant regulatory photoreceptor family. Genes Dev **3:** 1745–1757

Sinnott ML (1990) Catalytic mechanisms of enzymic glycosyl transfer. Chem Rev **90:** 1171–1202

Swain E, Li CP, Poulton JE (1992a) Development of the potential for cyanogenesis in maturing black cherry (*Prunus serotina*) fruits. Plant Physiol **98:** 1423–1428

Swain E, Li CP, Poulton JE (1992b) Tissue and subcellular localization of enzymes catabolizing (*R*)-amygdalin in mature *Prunus serotina* seeds. Plant Physiol **100:** 291–300

Swain E, Poulton JE (1994a) Utilization of amygdalin during seedling development of *Prunus serotina*. Plant Physiol **106:** 437–445

Swain E, Poulton JE (1994b) Immunocytochemical localization of prunasin hydrolase and mandelonitrile lyase in stems and leaves of *Prunus serotina*. Plant Physiol **106:** 1285–1291

Trimbur DE, Warren RAJ, Withers SG (1992) Region-directed mutagenesis of residues surrounding the active site nucleophile in β-glucosidase from *Agrobacterium faecalis*. J Biol Chem **267:** 10248–10251

Wakarchuk WW, Greenberg NM, Kilburn DG, Miller RC Jr, Warren RAJ (1988) Structure and transcription analysis of the gene encoding a cellobiase from *Agrobacterium* sp. strain ATCC 21400. J Bacteriol **170:** 301–307

Withers SG, Warren RAJ, Street IP, Rupitz K, Kempton JB, Aebersold R (1990) Unequivocal demonstration of the involvement of a glutamate residue as a nucleophile in the mechanism of a "retaining" glycosidase. J Am Chem Soc **112:** 5887–5889

Xue J, Lenman M, Falk A, Rask L (1992) The glucosinolate-degrading enzyme myrosinase in Brassicaceae is encoded by a gene family. Plant Mol Biol **18:** 387–398

Yemm RS, Poulton JE (1986) Isolation and characterization of multiple forms of mandelonitrile lyase from mature black cherry (*Prunus serotina* Ehrh.) seeds. Arch Biochem Biophys **247:** 440–445