The final step in the biosynthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor* is the transformation of the labile cyanohydrin into a stable storage form by *O*-glucosylation of (*S*)-*p*-hydroxymandelonitrile at the cyanohydrin function. The UDP-*p*-hydroxymandelonitrile-*O*-glucosyltransferase was isolated from etiolated seedlings of *S. bicolor* employing Reactive Yellow 3 chromatography with UDP-glucose elution as the critical step. Amino acid sequencing allowed the cloning of a full-length cDNA encoding the glucosyltransferase. Among the few characterized glucosyltransferases, the deduced translation product showed highest overall identity to *Zea mays* flavonoid-*O*-glucosyltransferase (*Bs-Mc-2* allele). The substrate specificity of the enzyme was established using isolated recombinant protein. Compared with endogenous *p*-hydroxymandelonitrile, mandelonitrile, benzyl alcohol, and benzoic acid were utilized at maximum rates of 78, 13, and 4%, respectively. Surprisingly, the monoterpenoid geraniol was glucosylated at a maximum rate of 11% compared with *p*-hydroxymandelonitrile. The picture that is emerging regarding plant glucosyltransferase substrate specificity is one of limited but extended plasticity toward metabolites of related structure. This in turn ensures that a relatively high, but finite, number of glucosyltransferases can give rise to the large number of glucosides found in plants.

Cyanogenic glucosides are amino acid-derived secondary plant metabolites (1). The ability to synthesize these glucosides is common across many plant genera, including several plant species that are important food crops (2). The hydrolysis of cyanogenic glucosides, aided by the presence of β-glucosidases and α-hydroxynitrilases, results in the release of cyanide, which has well known toxicological properties (3). It has therefore been suggested that cyanogenic glucosides and their breakdown products may play a role in plant defense (2, 4–6), although increased levels of cyanogenic glucosides in barley (7), which has well known toxicological properties (3). It has therefore been suggested that cyanogenic glucosides and their breakdown products may play a role in plant defense (2, 4–6), since increased levels of cyanogenic glucosides in barley (7), breakdown products may play a role in plant defense (2, 4–6), which has well known toxicological properties (3). It has therefore been suggested that cyanogenic glucosides and their breakdown products may play a role in plant defense (2, 4–6).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank®/EBI Data Bank with accession number(s) AF199453. recipient of an Australian Postgraduate award.

**To whom correspondence should be addressed. Tel.: 45-35283352; Fax: 45-35283333; E-mail: blm@kvl.dk.

This paper is available on line at http://www.jbc.org
Novel xenobiotic substances are also glucosylated by many plant species (28). Plants therefore have a large capability to glucosylate a wide range of different chemical structures. However, the number of glucosyltransferases present in plants and the range of substrate specificities exhibited by these are largely unknown. Whereas numerous glucosyltransferases have been partially purified, few have been characterized in an isolates where biological processes are not duplicated. Such enzymes are essential for the demonstration that around 65 genes may be present in the complete genome assuming an equal distribution. Altogether, over 100 glucosyltransferases have been identified in crops, although the difficulty associated with the separation of glucosyltransferases with similar chromatographic properties has confused the picture (34). The multiplicity of glucosyltransferases can be assessed from the Arabidopsis thaliana genomic sequencing project. Eighteen putative secondary plant metabolism glucosyltransferase-encoding genes have been identified in the 28% of the genome sequenced to date, indicating that around 65 genes may be present in the complete genome assuming an equal distribution. Altogether, over 100 different, putative, secondary plant metabolism glucosyltransferase-encoding cDNAs are available in international databases. Only in a few cases has the protein encoded by these gene sequences been verified (35–39), and only in two instances have authors (34, 40) attempted to investigate the complete specificity of the recombinant protein. In order to gain further insight into the biology of secondary plant metabolism glucosyltransferases, it is necessary to verify functionally the identity of the proteins encoded by these cDNAs by (a) heterologous expression and (b) conducting assays with aglycones of diverse chemical structures.

In the present study we report the isolation, cloning, functional heterologous expression, and characterization of the substrate specificity of a p-hydroxymandelonitrile-glucosyltransferase from S. bicolor. The isolation of this protein and the corresponding cDNA completes the biosynthetic pathway of cyanogetic glucoside biosynthesis.

**EXPERIMENTAL PROCEDURES**

**Biochemicals and Reagents**—All biochemicals were of analytical or higher grade. Substrates and authentic glucosides were obtained from Sigma and Extrasynthese (Genay, France). Dye reagents were obtained from Amersham Pharmacia Biotech and Sigma.

**Plant Materials**—S. bicolor seeds were obtained from Pacific Seeds, Australia (cultivar MR31).

**General Methods**—Protein preparations were concentrated using a Speed Vac Concentrator (Savant) prior to electrophoresis. SDS-PAGE was performed using high-Tris linear 8–25% SDS-polyacrylamide gradient gels (Mini-Protean II, Bio-Rad) (41), and polypeptides were visualized by staining with Coomassie Brilliant Blue R-250. DNA sequencing reactions were carried out using the Thermo Sequenase fluorescent-labeled primer cycle sequencing kit (7-deaza dGTP) (Amersham Pharmacia Biotech) and analyzed using an ALFExpress DNA Sequencer (Amersham Pharmacia Biotech). Sequence computer analysis was performed using programs in the Genetic Computer Group (Madison, WI) sequence analysis package and NCBI BLAST (42).

**Enzyme Assays**—General reaction mixtures (total volume 20 μl) included 100 mM Tris-HCl (pH 7.9), 1–5 μM [3H]UDP-glucose (11.0 GBq/mmol-1, Amersham Pharmacia Biotech), 0–200 μM UDP-glucose, 20 mM aglycone dissolved in water, 25 mM γ-glucuronolactone (β-glucosidase inhibitor), and 0.5–10 μl of protein preparation. At the end of the incubation period (10 min, 30 °C), 2 μl of 10% acetic acid were added to terminate the reaction. BSA (1 mg/ml) was included in assays for the assessment of sbHMNGT yield throughout the purification procedure. Qualitative analyses of recombinant sbHMNGT were performed as outlined above except for incubation period (20 min) and concentrations of reagents as follows: 1.25 mM aglycone (dissolved in ethanol except for flavonoids that were dissolved in ethylene glycol monoether), 1.25 μM [3H]UDP-glucose, 12.5 μM UDP-glucose, 100 ng of recombinant sbHMNGT, 4 μg of BSA. Quantitative analyses were performed as outlined for qualitative analysis except for incubation period (4 min) and concentrations of reagents as follows: 1, 5, or 10 mM aglycone, 5 μM [3H]UDP-glucose, 0.2 mM UDP-glucose, 200 ng of recombinant sbHMNGT, 24 μg of BSA. Reaction mixtures for analysis by NMR spectroscopy (total volume 0.5–1 ml) included 2 mM p-hydroxymandelonitrile or 6.5 mM geraniol, 3 mM UDP-glucose, 2.5 μM recombinant sbHMNGT, 0.5 mM of BSA. After incubation (2 h), glucosides were extracted with ethyl acetate and lyophilized in speed-vac prior to NMR analysis. For TLC, the reaction mixture was applied to Silica Gel 60 F254 plates (Merek), dried, and developed in a solvent containing ethyl acetate:aceton: dichloromethane:methanol:H₂O (40:30:12:10:8, v/v) for 1 h. Plates were dried (1 h, room temperature) and exposed to storage phosphorimaging screens (Molecular Dynamics, Sunnyvale, CA) prior to scanning on a Storm 860 PhosphorImager (Molecular Dynamics). For LSC, reaction mixtures were extracted with 400 μl of ethyl acetate to separate glucosides from unincorporated [3H]UDP-glucose. Two ml of Eoscin A (National Diagnostics, NJ) were added to 250 μl of each ethyl acetate extract and analyzed using a liquid scintillation counter.

Mandelonitrile was used as substrate to assay fractions generated by liquid chromatography.

**Purification**—All procedures were carried out at 4 °C except where indicated. S. bicolor seeds (1 kg) were imbibed in water overnight at room temperature and grown at 30 °C in darkness for 2 days (22). Seedling shoots were harvested and extracted in 2 volumes of ice-cold extraction buffer (250 mM sucrose; 100 mM Tris-HCl (pH 7.5); 50 mM NaCl; 2 mM EDTA; 5% (w/v) of polyvinylpolypyrolidone; 200 μM phenylmethylsulfonil fluoride; 6 mM DTT) using mortar and pestle. The extract was filtered through a nylon mesh prior to centrifugation (20,000 × g, 20 min). The supernatant fraction was subjected to differential ammonium sulfate fractionation (35–70%). The pellet was resus-

1 Annotated sequences deposited in GenBank™ December 15, 1998.

2 The abbreviations used are: PAGE, polyacrylamide gel electrophoresis; BSA, bovine serum albumin; DTT, dithiothreitol; LSC, liquid scintillation counting; sbHMNGT, p-hydroxymandelonitrile-O-glucosyltransferase; sbHMNGT, cDNA encoding sbHMNGT; PCR, polymerase chain reaction; bp, base pair.
pended in buffer A (20 mM Tris-HCl (pH 7.5); 5 mM DTT) and desalted using a Sephadex G-25 (Amersham Pharmacia Biotech) or Bio-Gel P-6 (Bio-Rad) column (2.5 x 20 cm, flow rate: 20 ml/min) equilibrated in buffer A. The first UV-absorbing peak was collected and applied to a Q-Sepharose (Amersham Pharmacia Biotech) column (2.6 x 23 cm, flow rate: 30 ml/min) equilibrated in buffer B. The Q-Sepharose column was washed with buffer B until the base line had stabilized, and proteins were eluted with a linear gradient from 50 to 400 mM NaCl in buffer B (800 ml total). Fractions (10 ml) were collected and 3–5 ml assayed for mandelonitrile and/or p-hydroxymandelonitrile glucosyltransferase activity by LSC. To reduce the salt concentration, combined active fractions (50 mg of protein, 20 ml) were diluted 5-fold in buffer B and concentrated 20-fold using an Amicon YM30 or YM10 membrane prior to storage at –80 °C.

The remainder of the purification was carried out at room temperature. One-quarter of the concentrated material from the Q-Sepharose step (~10–15 mg protein, 5 ml) was applied to a column (1 x 10 cm, flow rate: 10–15 ml/hr) containing Reactive Yellow 3 cross-linked onto 4% beaded agarose (Lot 63H0562) (Sigma) equilibrated in buffer B. The column was washed with buffer B until the base line had stabilized. Proteins were eluted with 10 ml of 2x UDP-glucose in buffer B. Active fractions containing essentially pure sbHMNGT were combined and stored at –80 °C with or without addition of 1 mg/ml BSA.

**Peptide Generation and Sequencing—** sbHMNGT (~5 μg) was subjected to a sequencing using a recombinant protein sequencing kit. For peptide digestion of sbHMNGT (~100 μg), the recombinant protein was precipitated by addition of trichloroacetic acid (10% w/v final concentration), resuspended in 50 μl of 50 mM Tris-HCl (pH 8.0), 5 mM DTT, and 6.4 mM NaCl, incubated (60 °C, 50 min), cooled to room temperature, and diluted with 3 volumes of 30 mM Tris (pH 7.7) and 1.25 mM EDTA. Endoproteasomes Lys-C (Promega, Madison, WI) was added (protease:substrate ratio 1:25 (w/w)) and the reaction allowed to proceed for 24 h at 37 °C. Peptides were purified with Beckman System Gold high performance liquid chromatography equipment fitted with a Vydac 20TPS2 C8 column (2.1 x 250 mm, flow rate: 0.2 ml/min). Peptides were applied in buffer C (0.1% trifluoroacetic acid) and eluted with a linear gradient from 0 to 80% acetonitrile in buffer C. Peptides were collected manually and sequenced as described above.

**PCR Amplification, Cloning, and Library Screening—** First round PCR amplification reactions (total volume: 40 μl) were carried out using 2 units of Taq DNA polymerase (Amersham Pharmacia Biotech), 4 μl of 10× Taq DNA polymerase buffer, 5% (v/v) dimethyl sulfoxide, 1 μl of dNTPs (10 mM), 80 pmol each of primers C2EF (5′-CTTCAGCTTCACTATAG-3′) and T7 (5′-ATATCCAGACTCATAATT-3′), and ~10 ng of genomic DNA template. The PCR amplicon DNA template was prepared from a unidirectional pcDNA1 (Invitrogen, Carlsbad, CA) plasmid library made from 1 to 2 cm high etiolated S. bicolor seedlings (16). Thermal cycling parameters were 95 °C, 5 min, 3 times (95 °C for 5 s, 42 °C for 30 s, 72 °C for 30 s), 32 times (95 °C for 5 s, 50 °C for 30 s, 72 °C for 30 s), and a final 72 °C for 5 min. Second round PCR amplifications were carried out as above, except using primers C2DF (5′-GAAGACTCGGACCCAT-3′) and T7, and 1 μl of first round reaction as DNA template. Thermal cycling parameters were 95 °C, 5 min, 32 times (95 °C for 5 s, 55 °C for 30 s, 72 °C for 30 s), and a final 72 °C for 5 min. The PCR reaction mixtures were subjected to gel electrophoresis using a 1.5% agarose gel, and an ~600 bp band was excised and cleaned using a QiaEX II gel extraction kit after 1.5% agarose gel electrophoresis and resuspended in 0.5 ml of water (data not shown). The use of a Waring blender resulted in less than 50% of the activity as compared with extraction with mortar and pestle. sbHMNGT activity was largely unaffected by both freezing at –80 °C and the addition of glycerol. The addition of relatively high concentrations of DTT were required to retain activity. Thus, lowering the concentration of DTT from 5 to 2 mM resulted in a 10-fold decrease in activity after storage at 4 °C for 2 days. This pronounced effect of DTT was primarily found in crude preparations, whereas partially purified ion-exchange preparations were less responsive to the concentration of reducing agents in contrast to previous results (12).

**RESULTS**

**Yield and Stability of sbHMNGT—** In preparation for sbHMNGT purification, sorghum seeds were germinated in darkness for 1.5–5 days, and extracts made from seedlings were tested for sbHMNGT activity. Under the conditions of growth, a 2-day germination period proved optimal with regard to total sbHMNGT activity, protein concentration, and extract volume (data not shown). The use of a Waring blender resulted in less than 50% of the activity as compared with extraction with mortar and pestle. sbHMNGT activity was largely unaffected by both freezing at –80 °C and the addition of glycerol. The addition of relatively high concentrations of DTT were required to retain activity. Thus, lowering the concentration of DTT from 5 to 2 mM resulted in a 10-fold decrease in activity after storage at 4 °C for 2 days. This pronounced effect of DTT was primarily found in crude preparations, whereas partially purified ion-exchange preparations were less responsive to the concentration of reducing agents in contrast to previous results (12).

**Isolation of sbHMNGT—** Mandelonitrile was employed as a substrate for the assay of sbHMNGT activity throughout purification, although the endogenous substrate of sbHMNGT is p-hydroxymandelonitrile. Previously, mandelonitrile had been shown to be an equally good substrate (12). Furthermore, the absence of a hydroxyl group at the benzene ring corelated the possibility of p-glucosloydemandelopamine synthesis, which would be indistinguishable from dhurrin when the convenient assay based on LSC was employed. Etiochrome synthesis, which would be indistinguishable from dhurrin when the convenient assay based on LSC was employed. Etiochrome synthesis, which would be indistinguishable from dhurrin when the convenient assay based on LSC was employed. Etiochrome synthesis, which would be indistinguishable from dhurrin when the convenient assay based on LSC was employed.
UDP-glucose:p-Hydroxymandelonitrile-O-Glucosyltransferase

N-terminal sequencing of approximately 100 pmol of isolated sbHMNGT yielded phenylthiohydantoin-derivatives at a level 10 times lower than expected (sequence 1, Fig. 3A). The low response suggested either partial blockage of the N terminus or the presence of a co-migrating and fully blocked contaminant. The latter possibility was ruled out because protein digestion by endoproteinase Lys C yielded peptides (sequences 2–5, Fig. 3A) with sequences all contained in a single cDNA (see below). Only sequence 4 exhibited high similarity to other known and putative glucosyltransferases (Fig. 3A).

Cloning of Full-length sbHMNGT—Degenerate oligonucleotides derived from peptide sequence 4 and a plasmid T7 primer were employed in nested PCR reactions using a unidirectional S. bicolor seedling plasmid library as template. An ~600-bp PCR fragment representing the C-terminal portion of sbHMNGT was isolated and shown to encode peptide sequences 4 and 5.

The partial PCR fragment was then employed to screen the library. Approximately 50,000 clones were screened, resulting in seven positive isolates, of which 4 were full-length. Preliminary sequencing indicated that they all represented an identical gene. One clone, sbHMNGT, was chosen for further study. The nucleic acid sequence of the sbHMNGT-encoding cDNA has been deposited in GenBank™ with the accession number XYYYYY.

The deduced translation product comprises 492 amino acid residues and has a predicted molecular mass of 52.9 kDa (Fig. 3A) and a theoretical pI of 5.3. Searching the PROSITE sequence motif data base revealed no extended stretches of identical sequence, except for a UDP-glucosyltransferase signature sequence between residues 368 and 411. Known and putative plant secondary metabolism glucosyltransferases in general exhibit a very low degree of overall similarity, with the exception of the C-terminal part which contains a postulated UDP-glucose binding motif (Fig. 3B). Based on this observation the C-terminal region is thought to encode the UDP-glucose binding domain (21, 44, 45, 57), whereas the N-terminal end of the protein may be responsible for binding the divergent and structurally dissimilar substrate aglycons. A comparison between sbHMNGT and a range of known and putative plant glucosyltransferases (listed in Fig. 3B) revealed that sbHMNGT shares highest overall identity (41.6%) and similarity (51.5%) with a partial putative glucosyltransferase-encoding cDNA from Pisum sativum (Table I). Among the few well characterized glucosyltransferases, sbHMNGT shared highest overall identity (36.7%) and similarity (41.5%) with Zea mays flavonoid-glucosyltransferase.

Functional Expression in E. coli and Characterization of Substrate Specificity—Active soluble recombinant sbHMNGT was synthesised in E. coli as described under “Experimental Procedures.” The enzyme was isolated using the same procedure for the native protein (Fig. 2B), implicating that the recombinant and native protein shared those physical features that were necessary for this highly selective purification system.

The qualitative substrate specificity of recombinant sbHMNGT was compared with the complement of glucosyltransferase activities present in crude desalted extracts of etiolated S. bicolor seedlings (Table II). Fifteen of the 22 substrates tested were glucosylated by the crude extract, whereas only 6 of these substrates were accepted by the recombinant enzyme. Hydroquinone (1,4-benzenediol) and p-hydroxybenzaldehyde were not glucosylated by sbHMNGT although both compounds reported to serve as substrates when non-homogenous preparations were used (12). In addition to mandelonitrile and p-hydroxymandelonitrile, benzyl alcohol, benzoic acid, and the
monoterpenoid geraniol were also utilized as acceptor substrates by sbHMNGT. NMR spectroscopy confirmed the sbHMNGT-mediated in vitro synthesis of geraniol glucoside and dhurrin (p-hydroxy-(S)-mandelonitrile-β-D-glucopyranoside) (data not shown). The multiple substrate specificity of sbHMNGT prompted analyses of the relative efficiency of util-

FIG. 3. The similarity between known and putative plant secondary metabolite glucosyltransferase amino acid sequences. A, multiple amino acid sequence alignment showing the similarity between sbHMNGT and other plant glucosyltransferases. Amino acid sequences (numbered 1–5) determined directly from purified sbHMNGT, and peptides thereof, are underlined. The alignment was created using the PILEUP program (Genetic Computer Group, Madison, WI). Abbreviated sequence names are as follows: zmUFGT (Z. mays flavonoid-glucosyltransferase, Bz-Mc-2 allele, EMBL accession number X13501); vvUFGT (V. vinifera anthocyanin-glucosyltransferase, GenBank™ accession number AF000371); psGT (P. sativum partial putative UDP-glucuronosyltransferase, GenBank™ accession number AF034743); meGT (M. esculenta putative glycosyltransferase, EMBL accession number X77464); zmIAAGT (Z. mays indole/acetic acid-glucosyltransferase, GenBank™ accession number L34847).

B, relative similarity between 17 different known and putative glucosyltransferases prepared using the PLOTSIMILARITY program (Genetic Computer Group, Madison, WI). Amino acid sequences were aligned using PILEUP, and values of relative average similarity between all sequences at each position of the alignment is indicated by the continuous trace. The average similarity across the entire alignment is indicated by the dashed line. Additional sequences not listed in Table II include Cassava (M. esculenta putative glycosyltransferases, EMBL accession numbers X77459, X77459, X77460, X77461, X77462, and X77463), Gentiana triflora (Swiss-Prot Q96493), and Nicotiana tabacum (GenBank™ accession numbers AB000623 and U32644).
expression vector (pSP19 g10L) carrying no insert showed no
Such quantitative measurements demonstrate a greater pref-
partly specific for the nitrile group and that the stereochem-
ble II). The other three non-cyanogenic substrates were only
been made toward understanding the enzymology and molec-
possible to estimate $V_{\text{max}}$ values for the different substrates. Such quantitative measurements demonstrate a greater preference for $p$-hydroxymandelonitrile, the endogenous substrate, compared with mandelonitrile. This is in contrast to previous results obtained using nonhomogeneous preparations (12) (Table II). The other three non-cyanogenic substrates were only utilized at less than a fifth of the maximal rate observed with the cyanohydrins. However, the acceptance of benzyl alcohol and benzoic acid as substrates indicates that sbHMNGT is only partially specific for the nitrile group and that the stereochemistry and/or interactive chemistry of the additional groups present on the hydroxyl-bearing carbon also influence sbHMNGT acceptance. Acetone cyanohydrin, the non- aromatic precursor of the cyanogenic glucoside linamarin present in cassava (M. esculenta), was not glucosylated by sbHMNGT. This suggested that sbHMNGT is exclusively specific for the presence of a benzyl group. The acceptance of geraniol at rates comparable to benzyl alcohol was therefore surprising (Table II). Control reactions with crude extracts of E. coli transformed with the expression vector (pSP19 g10L) carrying no insert showed no evidence of geraniol conjugation (data not shown).

**DISCUSSION**

Thousands of glycosylated secondary plant metabolites exist and yet only recently have a handful of enzymes responsible for glycosidic transfer been purified and characterized in some detail. Glycosylation is the terminal step in the biosynthesis of cyanogenic glucosides, an important class of metabolites present in several major food crops (2). Significant progress has been made toward understanding the enzymology and molecular biology of the synthesis of dhurrin, the cyanogenic glucos-
side present in *S. bicolour*. Although two cytochrome P450 en-
yzmes responsible for the conversion of tyrosine into $p$-hydroxymandelonitrile have been isolated and their cDNAs cloned, the enzyme crucial to the stable accumulation of dhurrin, namely sbHMNGT, has like many other glucosyltransferases proven very difficult to identify at the molecular level (12, 18–21). The difficulty associated with the isolation of plant secondary metabolite glucosyltransferases in general has been attributed to the apparent lability (46) and low concentra-
tion of these proteins (47). *S. bicolour* was therefore chosen as a model plant since young etiolated seedlings are capable of de novo synthesis of dhurrin up to a level corresponding to 30% of the dry weight matter (22). Furthermore, the two enzymes in the biosynthetic pathway that precede sbHMNGT, CYP79A1 (13) and CYP71E1 (14) (Fig. 1), had earlier been isolated from the same source. The critical step in sbHMNGT isolation was Reactive Yellow 3 dye chromatography in association with UDP-glucose elution, which resulted in more than a 100-fold purification in a single step. Recently, Reactive Yellow 3 with UDP-glucose elution was also employed in the purification of UDP-glucose:betanidin-6-O-glycosyltransferase with similar results as in the present study (48). The use of pseudo-affinity dye chromatography together with substrate-specific elution is now emerging as a highly selective key step in plant glucosyltransferase purification (47–51). This significantly shortens the purification procedure and minimizes the time-dependent loss of activity seen in earlier glucosyltransferase purification protocols (29, 46, 52–56). The importance of optimized buffer conditions was highlighted by the strong effect of DTT concentra-
tion on sbHMNGT stability in crude extracts. The differential effect of DTT on the maintenance of a reduced environment at different stages throughout purification may be explained by a higher concentration of low molecular weight radicals and oxidative enzymes in the crude plant extracts. The mechanism by which sbHMNGT activity is affected remains unknown; however, it is interesting to note that the addition of DTT to purified *Z. mays* indole/acetate acid-glycosyltransferase inhibited the formation of inactive multimers (49). The strong inhibitory effect of $p$-hydroxybenzaldehyde on sbHMNGT activity was surprising, given that it is a degradation product of the acceptor substrate, $p$-hydroxymandelonitrile. The addition of BSA to purified sbHMNGT significantly enhanced the apparent activity. Whether this effect is non-selectively chemical or due to allosteric regulation is not known. $p$-Hydroxybenzaldehyde would only be present in high concentration at the site of synthesis, if $p$-hydroxymandelonitrile was released into the cytosol prior to conjugation by sbHMNGT. The consequence of such an *in vitro* effect is further argument for an intimate association between sbHMNGT and CYP71E1 (24), which would minimize any formation of tyrosine-derived $p$-hydroxy-
benzaldehyde. However, this effect may not be relevant *in vivo* since (a) local concentrations of $p$-hydroxymandelonitrile may be much less than *in vitro* (1–5 mM), and (b) the presence of other proteins may diminish the inhibition in a similar manner to BSA.

The sequences of glucosyltransferase-encoding cDNAs ex-
hibit only moderate positional identities, and these are largely confined to discrete blocks specifying the C-terminal regions of the encoded enzymes. This block of relatively well conserved sequence in glucosyltransferases most likely represents a common UDP-glucose binding domain (21, 39, 57). At present it is not possible to predict glucosyltransferase acceptor substrate specificity from amino acid sequence, as no determinant resi-
dues or regions of residues have been established. A larger set of functionally verified glucosyltransferase-encoding cDNAs are required to further our understanding on this matter,
TABLE II
Substrate specificity of recombinant sbHMNGT compared with desalted crude etiolated Sorghum seedling extract

Glucosyltransferase-activity was determined by TLC using $[^{14}C]$UDP-glucose as detailed under “Experimental Procedures.” Filled boxes indicate that a radiolabeled product was visualized after incubation with the respective aglycone substrate. Empty boxes indicate that no radiolabeled substrate could be detected under the experimental conditions employed. Figures in brackets indicate relative $V_{max}$ for each respective aglyca with calculated standard deviations. The estimated $V_{max}$ value for $p$-hydroxymandelonitrile was 1500 mol of product/mol of sbHMNGT/s. n. d., not determined.

| SUBSTRATES                  | Crude Sorghum extract | Recombinant sbHMNGT |
|-----------------------------|-----------------------|---------------------|
| cyanohydrins                |                       |                     |
| 1) mandelonitrile           | ■                     | (77.8 ± 8.6 %) ■    |
| 2) $p$-hydroxymandelonitrile| ■                     | (100 ± 7.2%) ■      |
| 3) acetone cyanohydrin      | □                     | □                   |
| benzyl derivatives          |                       |                     |
| 4) hydroquinone             | ■                     | □                   |
| 5) benzyl alcohol           | ■                     | (13.1 ± 2.1%) ■     |
| 6) benzoic acid             | ■                     | (4.2 ± 0.8%) ■      |
| 7) $p$-hydroxybenzoic acid  | ■                     | (n. d.) ■           |
| 8) $p$-hydroxybenzaldehyde  | ■                     | □                   |
| 9) gentisic acid            | □                     | □                   |
| 10) caffeic acid            | □                     | □                   |
| 11) 2-hydroxy cinnamic acid | ■                     | □                   |
| 12) resveratrol (stilbene)  | ■                     | □                   |
| 13) salicylic acid          | □                     | □                   |
| 14) $p$-hydroxymandelic acid| □                     | □                   |
| flavonoids                  |                       |                     |
| 15) quercetin (flavonol)    | □                     | □                   |
| 16) cyanidin (anthocyanidin)| □                     | □                   |
| 17) biochanin A (isoflavone)| □                     | □                   |
| 18) naringenin (flavanone)  | □                     | □                   |
| 19) apigenin (flavone)      | □                     | □                   |
| others                      |                       |                     |
| 20) indole acetic acid      | (plant hormone)       | □                   |
| (plant hormone)             |                       |                     |
| 21) geraniol (monoterpenoid) | ■                    | (11.0 ± 0.5%) ■     |
| 22) tomatidine (alkaloid)   | ■                     | □                   |
which may not be confirmed until the three-dimensional structure of a secondary plant metabolism glucosyltransferase has been presented. sbHMNGT shares highest degree of overall identity with a putative glucosyltransferase deduced from an unknown clone from *P. sativum*, a *Z. mays* flavonoid-glucosyltransferase (43), and a *Z. mays* indole/acetate acid-glucosyltransferase (39) (Table I). This illustrates that the large number of putative glucosyltransferase-encoding cDNAs tabulated in the nucleic acid data banks, which have been labeled as flavonoid or indole/acetate acid-glucosyltransferase homologues, may be expected to glucosylate a range of other substrates in vivo and thus to be functionally mislabeled at present. The strong sequence and structural similarity exhibited between the glucosinolate-degrading myrosinases and a cyanogenic β-glucosidase from *Trifolium repens* (58), and the presence of CYP79 homologues in glucosinolate-producing plants (59), suggest that there is a strong evolutionary link between cyanogenic glucosides and glucosinolates. It was therefore surprising to find that the deduced sequence of a thiohydroximate-glucosyltransferase from *Brassica napus* showed only moderate overall identity to sbHMNGT. Similarly, none of the putative glucosyltransferase-encoding cDNAs isolated from cassava (*M. esculenta*) (21) shared any strong identity with sbHMNGT. However, sbHMNGT exhibited no activity toward acetone cyanohydrin, the main cyanohydrin aglycone present in cassava. Given that the monocotyledonous maize flavonoid-glucosyltransferase and sbHMNGT both utilize aglycones with benzyl groups, this may simply indicate a stronger co-evolution between sbHMNGT and flavonoid glucosyltransferases than between cyanohydrin-glucosyltransferases of different species. Alternatively, none of the clones isolated by Hughes et al. (21) encode for an acetone cyanohydrin:glucosyltransferase.

Investigations into the quantitative and qualitative substrate specificity of recombinant sbHMNGT showed a strong preference for the cyanohydrin present in *S. bicolor*. Similarly, when recombinant *V. vinifera* anthocyanidin-glucosyltransferase was assayed against a wide range of different aglycones, it was found to be strictly specific for flavonols and anthocyanidins only, with a strong preference for the latter (34). Both sbHMNGT and *V. vinifera* anthocyanidin-glucosyltransferase are involved in the final stages of predominant secondary metabolite biosynthetic pathways. Their presence is necessary for the highly tissue- and development-specific accumulation of their respective glucosides (12, 22–23). A possible scenario then is that the sole in vivo function of these enzymes is related to the glucosylation of unique and single secondary metabolites. It is, nevertheless, likely that enzymes present at the end of biosynthetic pathways have a broader substrate specificity than those preceding upstream, if there is to be any flexibility with respect to the evolution of novel secondary metabolite biosynthesis and xenobiotic catabolism. This is illustrated by the finding that CYP71E1 and sbHMNGT also accept phenylalanine-derived oximes and cyanohydrins (mandelonitrile), recently, whereas the first enzyme of the pathway, CYP79A1, is exclusive for tyrosine (Fig. 1) (60).

The stereochemistry of the cyanohydrin function is very important for substrate recognition. In sorghum the enzyme is stereospecific for the *S*-enantiomer of p-hydroxymandelonitrile (12). The wrong stereochemistry at the chiral carbon atom carrying the cyanohydrin function prevents acceptance of the nitrile group in the active site. On the other hand, the presence of a nitrile group is not necessarily required for substrate recognition by sbHMNGT, as demonstrated by the glucosylation of benzyl alcohol and benzoc acid, although at significantly lower rates compared with mandelonitrile. The above results indicate that sbHMNGT has high specificity for substrates that are closely similar to mandelonitrile, given that aglycones with only slight differences in chemical structure, such as hydroquinone, gentisic acid, and acetone cyanohydrin, do not serve as acceptor substrates. It was, therefore, surprising to find that sbHMNGT also conjugated the monoterpenoid geraniol, with equal efficiency compared with benzyl alcohol. To date there are no reports of the isolation of a monoterpenoid glucosyltransferase, despite the obvious importance of this enzyme class in relation to the aroma of processed fruits and vegetables (61). Initial tertiary structural modeling indicates that this unexpected and broadened specificity may be explained by the similarity of geraniol to benzyl alcohol in particular configurations.

The extensive characterization of sbHMNGT, *V. vinifera* anthocyanidin-glucosyltransferase (34), and a tobacco phenylpropanoid-glucosyltransferase (40), now allows us to address the question of whether glucosylation of the multitude of secondary plant metabolites results from the action of a relatively small number of highly promiscuous enzymes with broad substrate specificity or, at the other extreme, a large number of glucosyltransferases with a tight substrate specificity. The picture that is emerging, at least in vitro, is an intermediate situation. The in-depth characterization of the three glucosyltransferases reveals that a finite number of glucosyltransferases with some, but not very extended, plasticity toward structurally similar secondary metabolites exist. The effective substrate specificity can be further tightened in vivo, through the generation of only a specific set of aglycones. For example, if p-hydroxymandelonitrile and not geraniol is formed in etiolated seedlings of *S. bicolor*, then sbHMNGT exhibits tight specificity for the former metabolite. Alternatively, if a multitude of secondary metabolites, all of which can act as substrates for a particular glucosyltransferase, are present simultaneously, then it is possible that glucosyltransferase promiscuity is exhibited in vivo. The consequences of modulating glucosyltransferase activity, in planta, then becomes an intriguing one, and it remains to be seen whether the metabolism of a single class of metabolites can be influenced in vivo through the modulation of a specific glucosyltransferase expression. The tools and experimental systems available in the area of cyanogenic glucoside metabolism now allows us to address this important question. Hence, the availability of cDNAs encoding for all three enzymes of the cyanogenic glucoside biosynthetic pathway will now enable the preparation of transgenic plants, of acyanogenic cultivars, in which the capability to synthesize dhurrin has been introduced. This, together with antibodies directed toward these enzymes, will now permit in-depth studies of the biological role of this important class of secondary plant metabolites to take place.

Acknowledgments—We thank David Tattersall, Dean Naylor, Anna Stines, Chris Ford, John Strickart Nielsen, Mette Dahl Andersen, Ute Wittstock, Rachel Alice Kahn, Barbara Halkier, and Anna Haldrup for invaluable discussions; Neil Shirley and Jelle Lanstein for amino acid sequence analysis; Carl Erik Olsen for mass spectrometric analysis; and Søren Bak for providing plasmid libraries.

REFERENCES

1. Møller, B. L., and Seigler, D. S. (1998) in *Plant Amino Acids: Biochemistry and Bio/Technology* (Bijay K. Singh, ed.) pp. 563–609, Marcel Dekker, Inc., New York.
2. Jones, D. A. (1998) *Phytochemistry* 47, 155–162.
3. Solomonson, L. P. (1981) in *Cyanide in Biology* (Vennesland, B., Conn, E. E., Knowles, C. J., Westley, J., and Wissing, F., eds) pp. 11–28, Academic Press, NY.
4. Nahrestdt, A. (1985) *Plant Syst. Ecol.* 150, 35–47.
5. Hruska, A. J. (1988) *J. Chem. Ecol.* 14, 2213–2221.
6. Patton, C. A., Ranney, T. G., Burton, J. D., and Walgenbach, J. F. (1997) *J. Am. Soc. Hortic. Sci.* 122, 668–672.
7. Pourmohseni, H., and Benthal, W.-D. (1991) *Angew. Bot.* 65, 341–350.
8. Lüdtke, M., and Hahn, H. (1995) *Biochem. Z.* 324, 433–442.
9. Lieberri, R., Biehl, B., Ginsemann, A., and Junqueira, N. T. V. (1988) *Plant
