Available online at www.sciencedirect.com
SCIENCE DIRECT®
PERGAMON
Phytochemistry 64 (2003) 1069–1076
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Cloning and regiospecificity studies of two flavonoid glucosyltransferases from *Allium cepa*

Catherine M. Kramer[a], Rogerio T.N. Prata[a], Michael G. Willits[a], Vincenzo De Luca[b], John C. Steffens[a], Gerson Graser[a,*]

[a]*Syngenta Biotechnology Incorporated, 3054 Cornwallis Road, Research Triangle Park, NC 27709, USA*
[b]*Biology Department, Brock University, St. Catherines, Ontario, Canada L2S 3A1*

Received 8 May 2003; received in revised form 22 July 2003

Abstract

Two UDP-glucose-dependent flavonoid glucosyltransferases (EC 2.4.1.-) isolated from the epidermal layer of yellow onion (*Allium cepa*) were functionally expressed in *Escherichia coli* and their substrate specificity investigated. The two enzymes exhibited different substrate- and regio-specificity profiles. *A. cepa* UGT73G1 used a wide range of different flavonoid substrates including flavonoids not naturally occurring in onion. Regiospecificity was indicated for hydroxyl-groups of the C-3, C-7 and C-4′ positions of the flavan backbone structure to yield flavonoid mono- and diglucosides. In contrast, *A. cepa* UGT73J1 showed activity only with the flavonoid mono-glucoside isoquercitrin and the isoflavone aglycone genistein, with regiospecificity for the C-7 position. The regiospecificity for both enzymes included positions that are glucosylated in flavonoids of onion bulbs, indicating their involvement in flavonoid biosynthesis in *A. cepa*.

Keywords: *Allium cepa*; Alliaceae; Onion; Flavonoids; Quercetin; Glucosyltransferases; Substrate specificity; Regiospecificity

1. Introduction

Glycosylation is a widespread modification reaction in plant metabolism that is often involved in secondary metabolite biosynthesis (Vogt and Jones, 2000). The transfer of activated monosaccharide units (e.g. UDP-glucose) to an acceptor molecule is catalyzed by glycosyltransferases and can lead to changes in the activity of the acceptor molecule and its subcellular localization (Campbell et al., 1997; Lim et al., 2002). Glycosylation may have a role in increased water solubility, changes in polarity and increased molecular stability (Vogt and Jones, 2000). One of the most widely studied classes of plant glycosides is the large and heterogenic group of polyphenols, with an overwhelming number of native polyphenolic glycosides identified (Harborne and Williams, 2000).

Flavonoids are an important group of polyphenolic natural products ubiquitous in vascular plants, including many vegetables and fruits like onions, broccoli and apples. Flavonoids exhibit a wide range of biological activities that include antioxidant, estrogenic and vasodilatory properties. A diet high in flavonoids has been linked to a multitude of beneficial effects on human health, such as an inverse association with stroke incidence and coronary heart disease (Hertog et al., 1995; Keli et al., 1996). It is also reported that consumption of onion (*Allium cepa*) and other *Allium* vegetables high in flavonoids reduces the risk of stomach cancer (Dorant et al., 1996).

Most of the known flavonoid structures (ca. 6000) accumulate as glycosides in plants. A wide variety of simple monoglycosides, diglycosides, and complex polyglycosides have been identified (Harborne and Baxter, 1999). The glycosylation pattern of flavonoids has a strong influence on bioavailability (Aziz et al., 1998). Quercetin (**1**, Fig. 1) is a major flavonol found in many different foods, where it occurs predominantly as

* Corresponding author. Tel.: +1-919-765-5066; fax: +1-919-541-8585.
E-mail address: gerson.graser@syngenta.com (G. Graser).

0031-9422/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0031-9422(03)00507-7

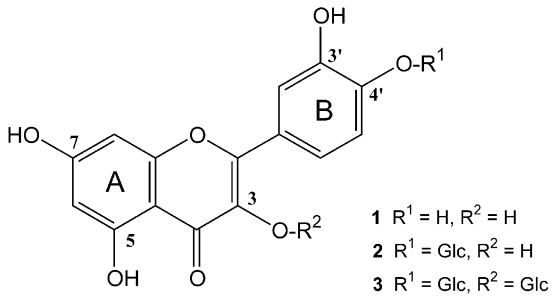

Fig. 1. Quercetin and the glucosylated quercetin derivatives found in yellow onion (*Allium cepa*) epidermal layers: **1** quercetin; **2** quercetin 4′-*O*-glucoside; **3** quercetin 3,4′-*O*-diglucoside.

the diglycoside rutin (quercetin 3-*O*-rutinoside) (Hollman et al., 1997), a substitution pattern that reduces its bioavailability in humans (Hollman et al., 1999). Increasing evidence shows that quercetin monoglucosides are probably absorbed in the small intestine via an active transport mechanism (Walgren et al., 2000a,b), while diglycosides (e.g. rutin) can only be absorbed after cleavage of the sugar molecules by microflora in the colon (Scalbert and Williamson, 2000). Therefore conversion of quercetin glycosides into glucosides can be a promising strategy to enhance bioavailability of quercetin from foods (Olthoff et al., 2000).

Different onion varieties accumulate between 300 and 900 mg/kg fresh weight of quercetin glycosides composed predominantly of quercetin 3,4′-*O*-diglucoside and quercetin 4′-*O*-glucoside (Hertog et al., 1992; Rhodes and Price, 1996) (**2**, **3**, Fig. 1). These two glucosides account for more than 80% of the total flavonoids found in onion. The remaining flavonoid fraction may consist of up to 17 different glucosides including quercetin 3,7-*O*-diglucoside, quercetin 3-*O*-glucoside, quercetin 4′-*O*-glucoside and isorhamnetin 4′-*O*-glucoside (Rhodes and Price, 1996; Price and Rhodes, 1997). The quercetin *O*-glucosides in onion appear to be highly bioavailable, since they are absorbed more rapidly and accumulate at levels up to 3-fold higher in the human bloodstream, compared to rutin (Hollman et al., 1995, 1997).

UDP-glucosyltransferases (UGTs) from onion could be used to increase the bioavailability of quercetin **1** by genetic engineering of other food crops or to manufacture bioavailable flavonoid glucosides through biofermentation. The present study describes a degenerate PCR strategy based on sequence homology to UGTs in GenBank™ to clone and functionally characterize two flavonoid UGTs from onion. Each cloned gene was expressed in bacteria, the protein products were purified and their flavonoid substrate specificities investigated. This is the first report describing the characterization of onion flavonoid UGTs after heterologous over-expression in a bacterial system.

2. Results and discussion

2.1. Sequence

cDNA fragments obtained from onion by degenerate PCR were sequenced and identified with BLASTx (NCBI) as putative UGTs based on sequence homology with known enzymes. Sequences of at least eight different UGT cDNAs were obtained by RACE PCR, and four full-length clones were designated as *UGT73H1* (AY62060) with a length of 487 aa, *UGT73H2* (AY62061), also 487 aa, *UGT73G1* (AY62062), 483 aa, and *UGT73J1* (AY62063), 470 aa in length (Fig. 2). *UGT73H1* and *UGT73H2* had 96% nucleotide and 94% amino acid identity. *UGT73G1* and *UGT73J1* were very different from each other with only 37% amino acid identity and also different from the *UGT73H1* and *UGT73H2* sequences with 36 and 35% amino acid identities, respectively. All four cDNAs contained some sequences identical to the plant secondary product UGT consensus sequence (PSPG box): **WAPQV**EVLAHPAVGCFVT**HCGWNS**TLESISAGVPMVAWPFFADQ. According to Vogt and Jones (2000) overall sequence identity of the plant UGT superfamily is low (~10%), but the highlighted amino acids of the PSPG box are common to 95% of the β-group UGTs analyzed, with underlined amino acids being completely conserved. The results reported here suggest that these onion UGTs belong to the plant β-group UGTs. The underlined amino acids are conserved in all four cDNAs (Fig. 2). The highlighted regions are conserved in both *UGT73H1* and *UGT73H2*, but differ by one amino acid in *UGT73G1* (V→M) and two amino acids in *UGT73J1* (A→V; V→T). *UGT73H1* and *UGT73H2* were 96% identical at the nucleotide level to a consensus sequence of seven published UGT sequences obtained from GenBank.

All four full-length cDNAs were expressed in *Escherichia coli* and were extracted in the soluble fraction. Accumulation of target protein was low but could be detected by SDS-PAGE and Western analysis using 6×His monoclonal antibodies (Fig. 3). Proteins from each clone migrated on SDS-PAGE in the range of 52–54 kDa, a molecular weight typical for glucosyltransferases involved in secondary plant metabolism (Vogt and Jones, 2000).

2.2. Substrate specificity studies

UGT activity was assayed according to standard protocols (Hirotani et al., 2000) with the addition of 0.1% Tween20, which was found to stabilize UGT activity (data not shown). Flavonoid reaction products were characterized by high performance liquid chromatography (HPLC), UV Spectra and in the case of major reaction products, by Liquid Chromatography–Mass

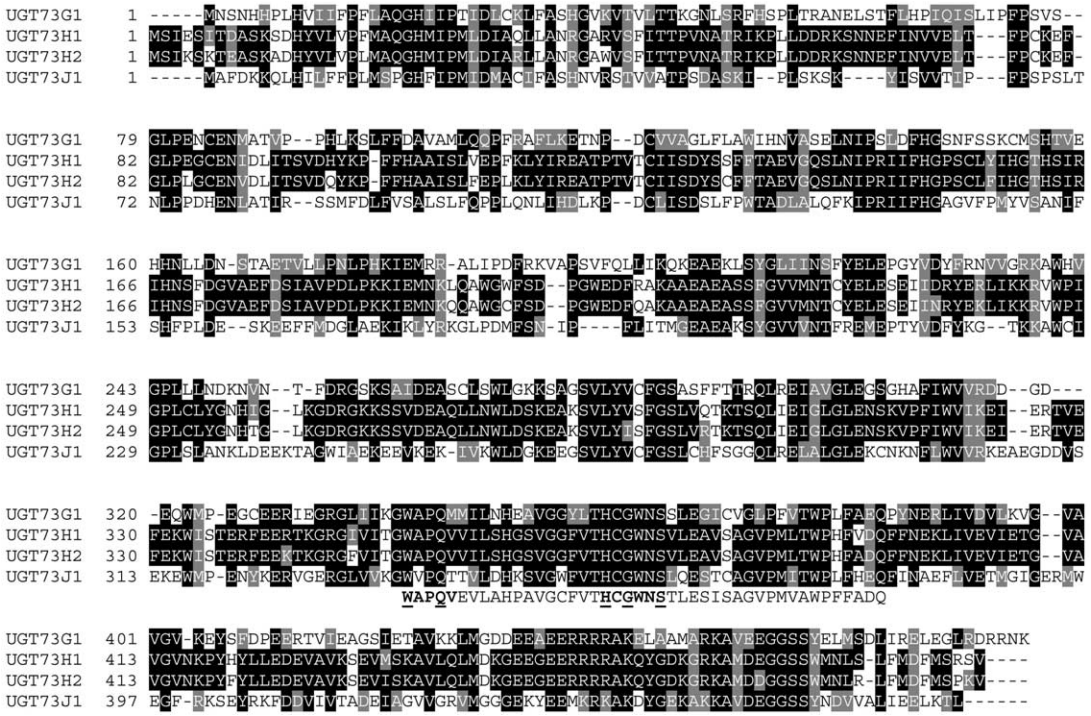

Fig. 2. Alignment of onion UGT amino acid sequences. The UGT sequences were compared using Clustal W, and the boxshading was performed with BOXSHADE 3.21. Black background represents identities, gray background represents similarities, and dashes represent gaps. The PSPG Box is designated below the alignment, with highlighted amino acids being common to 95% of the β-group UGTs analyzed and underlined amino acids being completely conserved.

Spectrometry (LC–MS). The apparent regiospecificity of each UGT was investigated by testing 18 hydroxylated flavonoid and flavonoid-related substrates (summarized in Table 1).

2.3. UGT73G1 produces mono and diglucosides with various flavonoid substrates

UGT73G1 exhibited low substrate specificity, since flavones, flavonoid glucosides, flavanones, isoflavones and chalcones were all glucosylated. Several of the substrates were glucosylated at two positions, resulting in the formation of not only monoglucosides but also diglucosides. Based on the OH-groups available on the flavan backbone structure (Fig. 1), UGT73G1 is most likely active on positions C-3, C-7 and C-4′, since all flavonoids with free OH-groups at these positions were glucosylated (Table 1). Flavonols such as **9**, **10** and **11** with only one of these positions available, were each converted to a single monoglucoside, while substrates offering two of the positions, like **5**, **6**, **7** and **8**, yielded two different monoglucosides and one diglucoside. The only exception was found for the flavanones **16** and **17,** with OH-groups in C-7 and C-4′, and formation of only one monoglucoside product. This result suggests that the double bond between C-2 and C-3 in the flavan backbone structure may be essential for diglucoside formation, because flavonol **7**, offering both positions

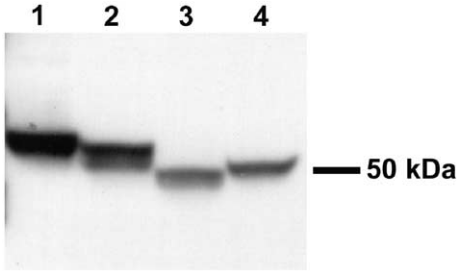

Fig. 3. Accumulation of recombinant UGT73H1 (lane 1), UGT73H2 (lane 2), UGT73G1 (lane 3), and UGT73J1 (lane 4) in *E. coli* BL21(DE3)pLysS. 20 μg HIS-purified protein analyzed by Western blot using 6xHis mAB and chemiluminescent detection.

C-7 and C-4′, was glucosylated with the formation of two monoglucosides and one diglucoside.

The isoflavones **18** and **19** represent a class of flavonoids typically restricted to legumes, with a migration of the B-ring to the C-3 position (Fig. 1). The isoflavone aglycone **18** was converted to a monoglucoside with high efficiency (71%), while the isoflavone **19,** glucosylated already at the C-7 position, was not a substrate for UGT73G1. HPLC retention time and UV spectrum of the glucoside product formed from **18** was identical to isoflavone **19**, indicating that UGT73G1 is active on the C-7 position of **18** resulting in the conversion of **18** to **19**. Similarly, strong activity with the isoflavone

Table 1
Substrate specificity of flavonoid *O*-glucosyltransferases UGT73G1 and UGT73J1 isolated from epidermal layers of *Allium cepa*[a]

Substrates	OH-groups available	UGT73G1	UGT73J1
Flavones			
1 Quercetin	3,5,7,3′,4′	**1MG** (M^- 463, 16%), 2MG (M^- 463), **1DG** (M^- 625), 1DG (M^- 625)	–
4 Isoquercitrin	5,7,3′,4′ (glc in 3)	**1DG** (M^- 625, 11%) 1DG (M^- 625)	**1DG** (M^- 625, 23%)
5 Rhamnetin	3,5,3′,4′ (–OCH_3 in 7)	**1MG** (M^- 477, 100%[b]) 1MG (M^- 477) **1DG** (M^- 638)	–
6 Kaempheride	3,5,7 (–OCH_3 in 4′)	**1MG** (M^- 461, 51%), 1MG, **1DG** (M^- 623)	–
7 7,8,4′-OH-flavone	7,8,4′	**1MG** (32%), 1MG, **1DG**	–
8 3,7-OH-flavone	3,7	**1MG** (6%), 1MG, 1DG	–
9 6,3′4′-OH-flavone	6,3′,4′	**1MG** (M^- 431, 37%)	
10 6,7-OH-flavone	6,7	**1MG** (M^- 415, 91%)	–
11 5,7-OH-flavone	5,7	**1MG** (25%)	–
12 Eupatorin	5,3′ (–OCH_3 in 6,7,4′)	–	–
13 5,6-OH-7-MeO-flavone	5,6 (–OCH_3 in 7)	–	–
14 6,3′-OH-flavone	6,3′	–	–
Flavanones			
15 Hesperetin	5,7, 3′ (–OCH_3 in 4′)	**1MG** (M^- 463, 32%)	–
16 Naringenin	5,7,4′	**1MG** (M^- 433, 34%)	–
17 Liquiritigenin	7,4′	**1 MG** (5%)	–
Isoflavones			
18 Genistein	5,7,4′	**1MG** (M^- 463, 71%)	**1MG** (M^- 463, 100%[c])
19 Genistin	5,4′ (glc in 7)	–	–
Chalcones			
20 Isoliquiritigenin		**1 MG** (M^- 417, 16%)	–

[a] Products were identified by HPLC retention times, UV spectra and for most of the products also by MS. Main products in bold, MG = monoglucoside, DG = diglucoside, glc = glucose. Activity calculations are for the formation of the first product to appear in the enzyme assay in the linear range of product formation.
[b] 100% activity UGT73G1 = 25 pkat/mg protein.
[c] 100% activity UGT73J1 = 2.6 pkat/mg protein.

genistein **18** was recently reported for a flavonoid UGT isolated from *Arabidopsis thaliana* (Willits et al., 2003). Substrate **20**, a chalcone which is a flavonoid precursor, was also glucosylated with the formation of one monoglucoside, emphasizing the low substrate specificity of UGT73G1. Chalcone glycosides have been identified in other plants and are known to be converted to their corresponding flavanone glycosides (Iwashina and Kitajima, 2000), but there is no evidence for accumulation of a chalcone monoglucoside (e.g. naringenin chalcone) or of the corresponding flavanone monoglucoside (e.g., **16**) in onion.

No tri- or tetraglucosides were produced from any of the offered substrates, strongly suggesting that UGT73G1 is only active with aglycones and monoglucosides and not with diglucosides.

The low substrate and regiospecificity of UGT73G1 is consistent with other heterologously-expressed UGT enzymes. For example, two UGTs (betanidin 5-UGT and betanidin 6-UGT) cloned from cell-suspension cultures of *Dorotheanthus bellidiformis* were active with flavonoid structures having OH-groups in either the C-4′ and C-7 or C-3 and C-7 positions, respectively (Vogt et al., 1997).

The major native flavonoid aglycone of onion, substrate **1,** was converted to three monoglucosides, with retention times of 5.98, 6.32 and 7.14 min, and two diglucosides, with retention times of 3.54 and 5.11 min (Fig. 4A). LC–MS analysis for all monoglucosides and diglucoside fractions showed the characteristic M^--ions of 463 and 625, respectively. Based on the retention times and UV profiles, the monoglucoside associated with the 5.98 min peak could be identified as quercetin

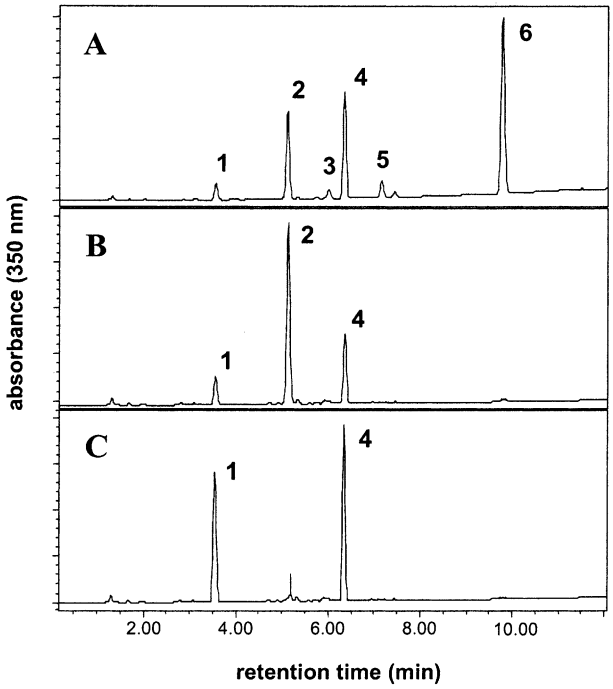

Fig. 4. Analysis of UGT73G1 (A and B) and UGT73J1 (C) activity by HPLC of flavonoid glucosides. Standard enzyme assay conditions with quercetin (A) and isoquercitrin (B and C) as substrates. UV detection, 350 nm: 1 = quercetin 3,7-*O*-diglucoside (RR_t 3.54 min), 2 = quercetin 3,4′-*O*-diglucoside (RR_t 5.11 min), 3 = quercetin 7-*O*-glucoside (RR_t 5.98 min), 4 = quercetin 3-*O*-glucoside (**4**, RR_t 6.32 min), 5 = quercetin 4′-*O*-glucoside (RR_t 7.14 min), 6 = quercetin (**1**, RR_t 9.76 min).

7-*O*-glucoside (no hypsochromic shift of the quercetin-characteristic absorption band I to 372 nm); the 6.32 min monoglucoside product as quercetin 3-*O*-glucoside (**4**) with a hypsochromic shift to 354 nm (Table 1); and the 7.14 min monoglucoside product as quercetin 4′-*O*-glucoside (**2**) with a hypsochromic shift to 365 nm (Markham, 1982; Vogt et al., 1997). Monoglucoside **4** was the first and most abundant product to appear in the enzyme assays followed by the major diglucoside (retention time 5.11 min, Fig. 4A). The other two monoglucosides and the other diglucoside were detected only in small amounts. Using **4** as a substrate for UGT73G1 also yielded two diglucosides with the same retention time as those produced in reactions with substrate **1** (Fig. 4B). This result strongly indicates that **4** is the major intermediate for the formation of the two diglucosides produced from substrate **1**. Based on the identified regiospecificity for UGT73G1, the retention times and the UV profiles, the major diglucoside (RR_t 5.11 min, Fig. 4A), with a hypsochromic shift of the quercetin-characteristic absorption band I from 372 nm to 345 nm, was identified as **3**, and the minor diglucoside, with a hypsochromic shift to 354 nm, as quercetin 3,7-*O*-diglucoside (Markham, 1982). The question of whether **4** is the only precursor for the two diglucosides, or if quercetin 7- and 4′-*O*-glucoside can also act as substrates, was not further investigated. But because of the activity of UGT73G1 with **7**, it is likely that glucosylation in the C-3 position is not essential for diglucoside formation.

2.4. UGT73J1 accepts very narrow range of flavonoid substrates

UGT73J1 showed much narrower substrate specificity than UGT73G1, since only quercetin 3-*O*-glucoside **4** and the isoflavone aglycone genistein **18** were accepted as substrates (Table 1). The diglucoside product from **4** showed the same HPLC retention time, UV spectrum and mass as the 3.55 min diglucoside peak detected in the UGT73G1 assay (Fig. 4C) and suggested that glucosylation also occurred at the C-7 position of **4**. As shown for UGT73G1, the product formed from **18** could be identified as **19** (HPLC retention time 6.33 min, M^- 463), supporting a regiospecificity of UGT73J1 for the C-7 position.

UGT73H1 and UGT73H2 had no activity with any of the flavonoid substrates tested, suggesting that these enzymes are not involved in flavonoid biosynthesis in onion or are not produced as active proteins in our bacterial system. These results were surprising considering the high degree of sequence identity with previously characterized and published UGTs from other plants.

Previous reports on flavonoid UGT activities detected in onion crude extracts showed alternative substrate specificities compared to the data presented here. A partially purified enzyme from *A. cepa* bulbs showed activity for the C-4′ position of quercetin **1** and, with less efficiency and specificity, for several other flavonol aglycones (Latchinian-Sadek and Ibrahim, 1991; Vogt et al., 1997). This enzyme showed no activity with isoflavones. Tsushida and Suzuki (1996) studied the substrate specificity for a UGT from crude extracts of *A. cepa* and reported activity for substrate **1** but not for quercetin glucosides such as **4.**

Characterization of cloned UGTs over-expressed in heterologous systems can yield different results from UGT activities identified in crude or partially purified enzyme extracts (Vogt and Jones, 2000). UGTs involved in secondary metabolism pathways are known to be unstable enzymes that occur only in minute amounts in plants (Vogt and Jones, 2000). Based on the number of UGT-like cDNA sequences identified from *A. cepa* (eight putative flavonoid UGT sequences identified by BLASTx), it is most likely that a whole set of UGT enzymes could be involved in the flavonoid pathway, thus complicating the characterization of such enzyme systems. However, both enzymes described here could be identified as flavonoid UGTs and the substrate and regiospecificity of the enzymes indicate involvement in in vivo biosynthesis of flavonoid glucosides in onion bulbs. Onion contains mainly flavonoid monoglucosides and diglucosides which are glucosylated predominantly in the C-3 and C-4′ position, but glucosylation in the C-

7 position is also known (Price and Rhodes, 1997). UGT73G1 was found to be active on all three of these positions with the production not only of monoglucosides but also diglucosides. The main products of UGT73G1 were identified as quercetin 3-*O*-glucoside and quercetin 3,4′-*O*-diglucoside, the latter shown to be the dominant flavonoid glucoside in onion (Hertog et al., 1992; Rhodes and Price, 1996). UGT73J1 was only found to be active on **4** and produced quercetin 3,7-*O*-diglucoside, the same product formed by UGT73G1 after glucosylation of either substrate **4** or substrate **1**.

3. Conclusion

UGT73G1 was active with a large number of different flavonoids and flavonoid glucosides, including artificial substrates and substrates that do not naturally occur in onion. These data suggest that the substrate versatility of UGT73G1 could be even broader than reported here, because the investigated flavonoids were limited to 18 structures. Another interesting feature of UGT73G1 was the production of five different products from substrate **1** (e.g., three monoglucosides and two diglucosides), illustrating the high affinity this enzyme has for this substrate and the probable role of UGT73G1 in the production of the glucosylated flavonoids found at high levels in *A. cepa*.

Information on substrate specificity of plant derived UGTs is still limited, but recent studies indicate that most plant UGTs are not absolutely substrate specific (Vogt and Jones, 2000; Messner et al., 2003). These characteristics make UGTs interesting targets for biotechnology applications because chemical glucosylation is difficult and expensive (Arend et al., 2001). Because of its distinctive low substrate specificity, UGT73G1 represents an interesting enzyme candidate for engineering flavonoid diversity and was also found to be active in a recently reported in vivo culture technique (data not shown; Willits et al., submitted). Further investigations may show the potential for using the onion flavonoid UGTs reported here as tools for increasing the bioavailability of polyphenols.

Although two other full length putative flavonoid UGTs were cloned from onion, only UGT73G1 and UGT73J1 were determined to be active on the flavonoids tested. The function of the other genes expressed in onion has yet to be elucidated.

4. Experimental

4.1. Cloning of full-length A. cepa cDNAs

The outermost layer of yellow onion was frozen in liquid nitrogen, ground with a mortar and pestle to a fine powder and total RNA extracted as previously described (Lagrimini et al., 1987). The poly $(A)^+$ RNA fraction was isolated using the PolyATtract® mRNA Isolation System (Promega, Madison, WI). First strand cDNA was prepared with the SMART™ RACE cDNA Amplification Kit (BD Biosciences Clontech, Palo Alto, CA) and the provided primers: SMART II Oligonucleotide 5′-AAGCAGTGGTAACAACGCAGAGTA CGCGGG-3′; 3′-RACE cDNA Synthesis Primer 5′-AAGCAGTGGTAACAACGCAGAGTAC$(T)_{30}N_{-1}$ N-3′ and 5′-RACE cDNA Synthesis Primer 5′-$(T)_{25}N_{-1}$N-3′ (N = A, C, G, or T, N_{-1} = A, G or C).

Degenerate primers were designed by sequence homology comparisons to known UGT sequences published in GenBank (accession numbers Y18871, AL021961, U32644, X85138, ACMP12, ALMP28). Degenerate primer sequences were as follows: GT1 5′-CCHKTBATGGCTCAWGGYCACATGATWCC - 3′; GT7 5′-GAAGADTGGYTRCCWGAAGGRTTYGA RGARAGAA-3′; GT4 5′-AYWGGCCATGTYACCA TTGGHASMCCTSC-3′; GT2 5′-TYAAWYCAKTGT AAGAWGAYCCWCCTTCTTCMAYAGC-3′.

PCR was performed with the Advantage 2® PCR Kit (Clontech, Palo Alto, CA) which includes Advantage 2 Polymerase mix containing TITANIUM™ *Taq*DNA Polymerase, TaqStart™ Antibody, and a minor amount of proofreading polymerase.

PCR products were gel purified with QIAquick Gel Extraction Kit (Qiagen Inc., Valencia, CA), cloned into pCR®II-TOPO® (Invitrogen, Carlsbad, CA) and both strands sequenced using dye terminator chemistry on ABI3700 DNA sequencers (Applied Biosystems, Foster City, CA). Sequences were edited and assembled into contigs using Sequencher 4.0.5 (GeneCodes Corp., Ann Arbor, MI). Based on the sequence data, gene-specific primers were designed, and 5′- and 3′-cDNA fragments were produced using the SMART™ RACE cDNA Amplification Kit. PCR products were purified, cloned and sequenced as described above. Full-length cDNAs were amplified using primers specific for the 5′- and 3′-ends of the coding regions.

UGT73H1: GTA1-1 5′-CATATGAGCATAGAAA GTATAACTGACGCCTCAAAGTCTGAT-3′; GTA3-2 5′-CTCGAGTCACACACTGCGAGACATAAAA TCCATGA-3′; UGT73H2: GTA3-7 5′-CATATG AGCA TAAAAAGTAAAACTGAGGCCTCAAAG-3′; GTA3-3 5′-CTCGAGTCACACCTTGGGAGACATAAAGT CCAT-3′; UGT73G1: GTB-3 5′-CATATGAATTCT AATCATCATCCTCTCCATGTCATC-3′; GTB-46 5′-CATATGTCATTTATTTCTTCTATCTCGCAA-3′; UGT73J1: GTC-8 5′-CATATGGCGTTCGACAAAAA ACAACTGCAC-3′; GTC-9 5′-CTCGAGCTACAGGG TCTTCAGTTCCTC-3′.

To facilitate cloning into an *E. coli* expression vector, an *Nde*I restriction enzyme site was incorporated at the

beginning of the coding regions, and either an *Xho*I or *Nde*I site at the end (underlined sequences). Full length amino acid sequences were compared using Clustal W, and boxshading was performed with BOXSHADE 3.21.

4.2. Expression of recombinant A. cepa glucosyltransferases in E. coli

DNAs were digested with *Nde*I or *Nde*I and *Xho*I and cloned into pET19b (Novagen, Madison, WI) with a HIS Tag at the N-terminus. *E. coli* strain BL21(DE3)-pLysS (Novagen) was transformed with the resulting UGT expression vectors. Cells were precultured in 3 ml LB medium (Sambrook et al., 1989) containing ampicillin and chloramphenicol on a 37 °C shaker overnight. Fresh LB medium (50 ml) was inoculated with 1.5 ml overnight culture and grown for 7–8 h on an 18 °C shaker. Cultures were induced with 1 mM isopropyl-1-thio-β-D-galactosidase (IPTG) and grown overnight on an 18 °C shaker. Subsequent protein purification was carried out at 4 °C except where noted otherwise. Cells from each 50 ml culture were pelleted by centrifugation for 10 min at 10,000 *g* and resuspended in 12.5 ml 20 mM Tris (pH 8). The cells were pelleted again, the supernatant discarded and the cell pellets frozen at −80 °C overnight. The frozen pellets were then resuspended in 3 ml BugBuster™ reagent (Novagen), 6 μl Benzonase® nuclease (Novagen) and 120 μl of a 25X stock of Complete™ Protease inhibitor cocktail (Boehringer Mannheim, Germany). Suspensions were incubated for 20 min at room temperature on a shaker. Insoluble cell debris was pelleted by centrifugation at 16,000 *g* for 20 min at 4 °C. Soluble fractions were concentrated in Centricon® Centrifugal Filter Devices (Millipore, Billerica MA) and quantitated with Quanti-Pro BCA Assay Kit (Sigma). Soluble protein (150 μg) was diluted in Tris–Glycine, SDS-PAGE (denaturing, non-reducing) sample buffer, boiled and then separated by electrophoresis on an 8% TRIS–glycine gel (Novex®). Proteins were transferred to PVDF membrane and analyzed by Western blot using a 6xHis mAB directly conjugated to HRP (BD Biosciences Clontech), diluted 1:10,000, and SuperSignal® West Pico Chemiluminescent Substrate (Pierce) for detection.

4.3. Enzyme purification

LB medium (300 ml) with ampicillin was inoculated with 3 ml overnight cultures of *E. coli* transformed with UGT expression vectors. Cultures were grown for 3 h at room temperature, induced with 1 mM IPTG and grown for another 3 h at room temperature. Cells were pelleted by centrifugation for 10 min at 10,000 *g*, resuspended in 10 ml cold 20 mM Tris (pH 8), pelleted again and stored overnight at −80 °C. Each pellet was resuspended in 5 ml BugBuster™ reagent and 5 μl Benzonase® nuclease and incubated at room temperature with gentle shaking for 12 min. Insoluble material was pelleted by centrifugation at 12,000 *g* for 10 min at 4 °C. Supernatants (5 ml each) were desalted with Sephadex™ PD-10 columns (Amersham Pharmacia, Uppsala, Sweden) and concentrated 3-fold in Microcon YM-10 centrifugation devices. Tween20 was added to a final concentration of 1.0% and protein extracts were immediately used for enzyme assays or stored at −80 °C.

4.4. Enzyme assays for substrate specificity studies

UGT enzyme assays contained 10 mM KH_2PO_4 (pH 6.8), 5 mM $MgCl_2$, 0.1% Tween20, 160 μM flavonoid substrate and 500 μM UDP-glucose (Boehringer Mannheim) in a total volume of 110 μl. Flavonoid substrates were purchased from Indofine (Somerville, NJ) and dissolved in DMSO. After addition of 50 μl (10–25 μg protein) of enzyme solution, the reaction mixture was incubated at 37 °C. The reaction was terminated after 15 or 30 min, in the linear range of product formation, by addition of 10 μl 6 N HCl and 80 μl methanol and centrifuged for 10 min at 12,000 *g*. The supernatant was immediately used for HPLC analysis or stored at 4 °C.

4.5. Flavonoid glucoside HPLC analysis

Samples were analyzed with a Waters HPLC 2690 Alliance separations module (Waters, Milford, MA) and a 996 diode-array detector. A 4.6×150 mm Xterra RP_{18} (5 μm) column protected by a C_{18} Nova-pack guard column kept at 40 °C was used for all analysis. HPLC grade acetonitrile (Fisher Scientific, Fair Lawn, NJ), 0.01% trifluoroacetic acid (Pierce, Rockford, IL) and water from a Milli-Q Water System (Millipore) were used as mobile phases. Flavonoid glucosides were separated using a linear gradient from 10 to 85% acetonitrile with a flow rate of 1.5 ml min^{-1} for 20 min.

4.6. Mass spectrometry

LC–MS analyses of the flavonoid glucosides was carried out on an Agilent HP1100 HPLC connected to a Finnigan LCQ Classic Mass Spectrometer equipped with an electrospray ionization (ESI) source. The procedure employed a Phenomenex Synergi 4 μ MAX RP 150×4.6 mm column, operated at a flow rate of 0.4 ml min^{-1} at 25 °C. Flavonoid glucosides were eluted from the HPLC using a linear gradient of 20–40% B in 22 min (solvent A: 0.1% sodium formate, 0.1% formic acid, methanol; solvent B: 0.1% sodium formate, 0.1% formic acid, water, w/v/v). Negative ion mass spectra were acquired from 200 to 2000 a.m.u. at a scan rate of one scan/s. The ESI capillary voltage was −43.6 V and the spray needle voltage 0.76 kV. Helium was used as

damping gas, nitrogen as sheath gas (60 psi) and the capillary temperature was held at 250 °C.

Acknowledgements

We wish to thank Jim Pomes for LC–MS analysis.

References

Arend, J., Warzecha, H., Hefner, T., Stockigt, J., 2001. Utilizing genetically engineered bacteria to produce plant specific glucosides. Biotechnol. Bioeng. 76, 126–131.

Aziz, A.A., Edwards, C.A., Lean, M.E.J., Crozier, A., 1998. Absorption and excretion of conjugated flavonols, including quercetin-4′-*O*-β-glucoside and isorhamnetin-4′-*O*-β-glucosides by human volunteers after the consumption of onions. Free Radic. Res. 29, 257–269.

Campbell, J.A., Davies, G.J., Bulone, V., Henrissat, B., 1997. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. Biochem. J. 326, 929–939.

Dorant, E., van den Brandt, P.A., Goldbohm, R.A., Sturmans, F., 1996. Consumption of onions and a reduced risk of stomach carcinoma. Gastroenterology 110, 12–20.

Harborne, J.B., Baxter, H., 1999. Handbook of Natural Flavonoids. 2 Vols. John Wiley & Sons, Chichester, New York.

Harborne, J.B., Williams, C.A., 2000. Advances in flavonoid research since 1992. Phytochemistry 55, 481–504.

Hertog, M.G.L., Hollman, P.C.H., Katan, M.B., 1992. Content of potentially anticarcinogenic flavonoids of 28 vegetables and 9 fruits commonly consumed in the Netherlands. J. Agric. Food Chem. 40, 2379–2383.

Hertog, M.G.L., Kromhout, D., Aravanis, C., Blackburn, H., Buzina, R., Fidanza, F., Giampaoli, S., Jansen, A., Menotti, A., Nedeljkovic, S., Pekkarinen, M., Simic, B.S., Toshima, H., Feskens, E.J.M., Hollman, P.C.H., Katan, M.B., 1995. Flavonoid intake and long-term risk of coronary heart disease and cancer in the seven countries study. Arch. Intern. Med. 155, 381–386.

Hirotani, M., Kuroda, R., Suzuki, H., Yoshikawa, T., 2000. Cloning and expression of UDP-glucose: flavonoid 7-*O*-glucosyltransferase from hairy root cultures of *Scutellaria baicalensis*. Planta 210, 1006–1013.

Hollman, P.C.H., de Vries, J.H.M., van Leeuwen, S.D., Mengelers, M.J.B., Katan, M.B., 1995. Absorption of dietary quercetin glycosides and quercetin in healthy ileostomy volunteers. Am. J. Clin. Nutr. 62, 1276–1282.

Hollman, P.C.H., van Trijp, J.M.P., Buysman, M.N.C.P., Gaag, M.S.v.d., Mengelers, M.J.B., de Vries, J.H.M., Katan, M.B., 1997. Relative bioavailability of the antioxidants flavonoid quercetin from various foods in man. FEBS Lett. 418, 152–156.

Hollman, P.C.H., Bijsman, M.N.C.P., van Gameren, Y., Cnossen, E.P.J., de Vries, J.H.M., Katan, M.B., 1999. The sugar moiety is a major determinant of the absorption of dietary flavonoid glycosides in man. Free Radic. Res. 31, 569–573.

Iwashina, T., Kitajima, J., 2000. Chalcone and flavonol glycosides from *Asarum canadense* (Aristochiaceae). Phytochemistry 55, 971–974.

Keli, S.O., Hertog, M.G.L., Feskens, E.J.M., Kromhout, D., 1996. Flavonoids, antioxidant vitamins and risk of stroke. The Zutphen study. Arch. Intern. Med. 156, 637–642.

Lagrimini, L.M., Burkhart, W., Moyer, M., Rothstein, S., 1987. Molecular cloning of complementary DNA encoding lignin-forming peroxidase from tobacco: molecular analysis and tissue-specific expression. Proc. Natl. Acad. Sci. U.S.A. 84, 7542–7546.

Latchinian-Sadek, L., Ibrahim, R.K., 1991. Partial purification and some properties of a ring β-O-glucosyltransferase from onion bulbs. Phytochemistry 30, 1767–1771.

Lim, E.-K., Doucet, C.J., Li, Y., Elias, L., Worrall, D., Spencer, S.P., Ross, J., Bowles, D.J., 2002. The activity of Arabidopsis glycosyltransferases toward salicylic acid, 4-hydroxybenzoic acid, and other benzoates. J. Biol. Chem. 277, 586–592.

Markham, K.R., 1982. Techniques of Flavonoid Identification. Academic Press, London.

Messner, B., Thulke, O., Schaffner, A.R., 2003. Arabidopsis glucosyltransferases with activities towards both endogenous and xenobiotic substrates. Planta 217, 138–146.

Olthof, M.R., Hollman, P.C.H., Vree, T.B., Katan, M.B., 2000. Bioavailability of quercetin-3-glucoside and quercetin-4′-glucoside do not differ in humans. J. Nutr. 130, 1200–1203.

Price, K.R., Rhodes, M.J.C., 1997. Analysis of the major flavonol glycosides present in four varieties of onion (*Allium cepa*) and changes in composition resulting from autolysis. J. Sci. Food Agric. 74, 331–339.

Rhodes, M.J.C., Price, K.R., 1996. Analytical problems in the study of flavonoid compounds in onions. Food Chem. 57, 113–117.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York.

Scalbert, A., Williamson, G., 2000. Dietary intake and bioavailability of polyphenols. J. Nutr. 130, 2073S–2085S.

Tsushida, T., Suzuki, M., 1996. Content of flavonol glucosides and some properties of enzymes metabolizing the glucosides in onion. Nippon Shokuhin Kagaku Kaishi 43, 642–649.

Vogt, T., Zimmermann, E., Grimm, R., Meyer, M., Strack, D., 1997. Are the characteristics of betanidin glucosyltransferase from cell-suspension cultures of *Dorotheanthus bellidiformis* indicative of their phylogenetic relationship with flavonoid glucosyltransferases? Planta 203, 349–361.

Vogt, T., Jones, P., 2000. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. Trends Plant Sci. 5, 380–386.

Walgren, R.A., Karnaky, K.J., Lindenmayer, G.E., Walle, T., 2000a. Efflux of dietary flavonoid quercetin 4′-β-glucoside across human intestinal caco-2 cell monolayers by apical multidrug resistance-associated protein-2. J. Pharmacol. Exp. Ther. 294, 830–836.

Walgren, R.A., Lin, J.-T., Kinne, R.K.-H., Walle, T., 2000b. Cellular uptake of dietary flavonoid quercetin 4′-β-glucoside by sodium-dependent glucose transporter SGLT1. J. Pharmacol. Exp. Ther. 294, 837–843.

Willits, M.G., Giovanni, M., Prata, R.T.N., Kramer, C.M., De Luca, V., Steffens, J.C., Graser, G., 2003. Biofermentation of modified flavonoids: an example of in vivo diversification of plant secondary metabolites. Phytochemistry (submitted for publication).