Title

Functional characterization and reclassification of an enzyme previously proposed to be a limonoid UDP-glucosyltransferase

Journal

Journal of the Science of Food and Agriculture, 100(13)

ISSN

0022-5142

Authors

Cui, Youtian
Allmon, Steven D
Siegel, Justin B

Publication Date

2020-10-01

DOI

10.1002/jsfa.10547

Peer reviewed

HHS Public Access

Author manuscript

J Sci Food Agric. Author manuscript; available in PMC 2021 October 01.

Published in final edited form as:

J Sci Food Agric. 2020 October ; 100(13): 4870–4878. doi:10.1002/jsfa.10547.

Functional characterization and reclassification of an enzyme previously proposed to be a limonoid UDP-glucosyltransferase

Youtian Cui[1], **Steven D. Allmon**[2], **Justin B. Siegel**[1,3,*]

[1]Department of Chemistry, University of California Davis, One Shields Avenue, Davis, California 95616, United States

[2]The Coca-Cola Company, Innovation Center @ Apopka, 2651 Orange Ave, Apopka, Florida 32703, United States

[3]Department Biochemistry & Molecular Medicine and the Genome Center, University of California Davis, One Shields Avenue, Davis, California 95616, United States

Abstract

Background: A major problem in the orange industry is “delayed” bitterness, which is caused by limonin, a bitter compound developing from its non-bitter precursor limonoate A-ring lactone (LARL) during and after extraction of orange juice. The glucosidation of LARL by limonoid UDP-glucosyltransferase (LGT) to form non-bitter glycosyl-limonin during orange maturation has been demonstrated as a natural way to debitter by preventing the formation of limonin.

Result: Here, the debittering potential of heterogeneously expressed glucosyltransferase, maltose-binding protein (MBP) fused to *cu*GT from *Citrus unishiu Marc* (MBP-*cu*GT), which was previously regarded as LGT, was evaluated. An LC-MS method was established to determine the concentration of limonin and its derivatives. The protocols to obtain its potential substrates, LARL and limonoate (limonin with both A and D ring open), were also developed. Surprisingly, MBP-*cu*GT did not exhibit any detectable effect on limonin degradation when Navel orange juice was used as the substrate; MBP-*cu*GT was unable to biotransform either LARL or limonoate as purified substrates. However, it was found that MBP-*cu*GT displayed a broad activity spectrum towards flavonoids, confirming the enzyme produced was active under the conditions evaluated *in vitro*.

Conclusion: Our results based on LC-MS demonstrated that *cu*GT functionality was incorrectly identified. Its active substrates, including various flavonoids but not limonoids, highlights the need for further efforts to identify the enzyme responsible for LGT activity to develop biotechnology-based approaches for producing orange juice from varietals that traditionally have a delayed bitterness.

Keywords

Limonin; limonoid UDP-glucosyltransferase; glycosyltransferase; flavonoids; delayed bitterness

[*]**Corresponding Author:** jbsiegel@ucdavis.edu (J.B.S.).

1. Introduction

Citrus juice is an important component in our daily diet and is associated with many health benefits by supplementing the necessary nutrients and bioactive phytochemicals, particularly polyphenols.[1] However, the citrus juice industry deals with bitterness mainly caused by compounds such as bitter limonoids, bitter flavonoids, and some sour organic acids.[2, 3] Among them, limonin, an intensely bitter tetracyclic triterpenoid dilactone (Figure 1), was commonly noted as the main contributor. [4–7] It has been found that 6 ppm of limonin in juice is the threshold, with greater concentrations of limonin reducing the acceptance of products by customers.[8] When product has greater concentrations of limonin, the producers have to apply extra steps to remove limonin or blend the juice with more sugar, which in turn affects the taste and reduces its beneficial effects. Therefore, technologies to prevent delayed bitterness could have a significant commercial impact on the citrus industry.

Limonin and its chemically related, highly oxygenated tetracyclic triterpenoids can be categorized as limonoids.[9] Progress has been made to deepen the understanding of citrus limonoids, including the biosynthetic pathways in plants and the corresponding enzymes catalyzing these biotransformations.[10] It has shown that limonin derives from the non-bitter precursor, limonoate A-ring lactone (LARL), catalyzed by limonoate D-ring lactone hydrolase at a physiologically acidic pH.[11] Such reaction is reversible and can be accelerated by strong acids or bases without the enzyme's involvement. LARL can also accept the glucose transferred from UDP-glucose to produce non-bitter limonin 17-β-D-glucopyranoside (glycosyl-limonin) catalyzed by a limonoid glycosyltransferase (LGT).[12] Such naturally occurring biosynthetic pathways offer an alternative method to degrade limonin by enzymatic reactions (Figure 1).

Recently, debittering orange juice *in vitro* using enzymes has gained attention due to its advantage of environmental sustainability and maintenance of bioactive limonoids. For example, a protein with LGT activity was reported to be purified from pummelo albedo tissue and utilized for debittering experiments on citrus juice.[13] A protein with LGT activity was also purified from the albedo tissues of Navel orange (*Citrus sinensis*) cultivars.[14] Yet, despite the acknowledgement of the physiological function of LGT, the primary sequences of LGTs from pummelo and Navel orange are still unknown. To date, a single study on *cu*GT from *Citrus unishiu* Marc. (Uniprot ID Q9MB73) is the only report that directly demonstrates a specific protein sequence having LGT function.[15] However, follow-up studies confirming the protein sequence identified are limited. In fact, one recent study explored the activity of a related annotated GT enzyme from *Citrus paradisi* (*cp*GT, previously referred to as PGT8, Uniprot ID B2YGX8) with >98% sequence identity to *cu*GT.[16] *cp*GT was only able to be solubly produced in a eukaryotic system. Furthermore, the homolog was not observed to have any LGT activity. With only a single study directly observing LGT activity from a defined glycosyl-transferase family,[15] and a second report of a close homolog not being able to reproduce the original observation,[16] it is critical to independently evaluate the sequence–function relationship of the previously reported *cu*GT being used to define annotations across the entire LGT family.

Therefore, we aimed to further characterize the *cu*GT from *Citrus unishiu* and its ability to catalyze limonoids from Navel orange juice, purified LARL, and limonoate (limonin with both A and D rings open) under various conditions. Consistent with previous findings, *cu*GT can be solubly expressed and purified as a fusion protein of MBP in *E. coli*.[15] Nevertheless, no activity was observed using HPLC-MS even though the recombinant GT was demonstrated to be functional because it was observed to have both activity and specificity in its ability to accept specific compounds from a panel of flavonoids as aglycon. As previous efforts only utilized indirect methods and impure substrates to characterize activity, the efforts here using a direct measurement of function indicate that *cu*GT has been incorrectly labeled and that the protein sequence family for the LGT function observed in citrus albedo and seeds has yet to be identified.

2. Materials and Methods

2.1 Bacterial strains, plasmids, and chemical reagents

The bacterial strain used for cloning was *Escherichia coli* DH5α, the pET29 (+b) plasmids containing the protein encoding genes were expressed in *E. coli* BLR (DE3). All genes were purchased from Twist Biosciences as synthetic genes optimized for *E. coli* codon usage. The sequences of genes encoding *cu*GT and *cu*GT fused with MBP and GST in the present work are listed in Table S1.

All chemical reagents used were analytical grade. Limonin was purchased from ChromaDex and glucosyl-limonin was purchased from LKT Laboratories, Inc. Limonin was prepared in acetonitrile (ACN) at 100 ppm as stock and limonin glucoside was dissolved in MiliQ water at 10 mM. The stocks were stored at −20 °C. The flavonoids, salicylic acid, and sinapic acid were purchased from Sigma Aldrich.

2.2 Protein expression and purification

E. coli was first grown overnight as the starter culture at 37 °C in Terrific Broth medium (1% tryptone, 0.5% yeast extract, 0.5% NaCl,) supplemented with Kanamycin (50 μg/mL final concentration) and $MgSO_4$ (1 mM final concentration). The culture for protein expression was diluted by ~ 50-fold to 500 mL from the starter culture. The cultures were then grown until OD_{600} to ~0.6 at 37 °C, and IPTG was supplemented to final concentration of 0.5 mM for induction at 16 °C for 24 h. At the end of induction, cells were centrifuged (4,700 x *g*., 4 °C, 10 min), supernatant was removed, cells were resuspended in 40 mL lysis buffer (50 mM HEPES, pH 7.0, 300 mM NaCl, 10% glycerol, 1 mM $MgSO_4$, 15 mM imidazole, 1 mM TCEP), and sonicated for 2 min at 4 °C. Lysed cells were centrifuged at 4,700 x *g* at 4 °C for 30 min to remove cell debris. Supernatant was loaded on a gravity flow column with 1 mL of cobalt slurry (Thermo Fisher Scientific, CAT# PI-90091), which was pre-balanced with 30 mL of wash buffer (50 mM HEPES, pH 7.0, 300 mM NaCl, 10% glycerol, 1 mM MgSO4, 15 mM Imidazole, 1 mM TCEP). The cobalt resin was then washed three times with 10 mL wash buffer; proteins were eluted with 0.6 mL of elution buffer (50 mM HEPES, pH 7.0, 300 mM NaCl, 10% glycerol, 1 mM $MgSO_4$, 1 mM TCEP, 200 mM imidazole). Protein samples were immediately buffer exchanged with spin concentrators (Satorius, CAT# VS0112) into storage buffer (50 mM HEPES, pH 7.0, 300 mM NaCl, 10%

glycerol, 1 mM $MgSO_4$, 1 mM TCEP) and stored at 4 °C until activity characterization. Protein concentrations were determined using a Synergy H1 spectrophotometer (Biotek) by measuring absorbance at 280 nm using their calculated extinction coefficients with the ExPASy ProtParam. The protein samples were further analyzed by 12% SDS-PAGE gel.

2.3 Detection and quantification of limonin, limonoid glucoside, and flavonoid glucosides by liquid chromatography mass spectrometry (LC-MS)

Reverse-phase high-performance liquid chromatography and mass spectroscopy (LC-MS) for analysis were carried out using Agilent 1260 series instruments with Poroshell 120 EC-C18 (Agilent, 4.6 × 50 mm, 2.7 μm) column. Mass spectroscopy was carried out using an Agilent 6120 single quadrupole spectrometer with electrospray ionization (ESI) in either positive-ion mode or negative-ion mode. The gas temperature was 350 °C, drying flow was 13.0 L/min, and capillary voltage was 4300 V. Each sample was analyzed in triplicate.

For limonin quantification, the mobile phase consisted of the following gradient: 70% H_2O with 0.1% formic acid as mobile phase A and 30% ACN with 0.1% formic acid as mobile phase B for 8 mins; 10% mobile phase A and 90% mobile phase B from 8 to 19 min; mobile phase A was decreased to 70% with 30% mobile phase B until 25 min. Podophyllotoxin, the internal standard, was prepared in pure ACN with 125 ppm concentration as stock and was diluted 100 times for the reaction mixture. The HPLC flow rate was 0.8 mL/min and the injection volume was 3 μL. The parameter of the mass spectrum was adjusted to positive SIM mode with m/z detector set as 471.2 (limonin) and 397.2 (podophyllotoxin).

For glucosyl-limonin quantification and detection, the HPLC protocol was similar to the protocol used for limonin, except for the composition of mobile phases, in which 0.01% formic acid was contained as additive. MS in ESI- used the SIM mode to detect m/z 649.2 for glucosyl-limonin.

For the quantification of flavonoid glucosides and related chemicals, the same HPLC protocol as limonin detection was performed, except the λ of UV/Vis detector was set to λ= 350 nm for apigenin and genistein, λ=210 nm for salicylic acid and sinapic acid, and λ= 270 nm for the rest of the flavonoids. Scan ESI+ mode employed an m/z of 100 to 800 to confirm flavonoid glucosides and the m/z corresponding to product peaks was used to match the theoretical value to verify the product.

2.4 Preparation of LARL and limonoate

The purification of LARL was performed by solid-phase extraction (SPE), in which the cartridge with ODS-5 packing (Whatman 6803-0507, 500 mg, 6 mL volume) was used. The SPE cartridge was first washed with two cartridge volumes of ACN, followed by five cartridge volumes of milli-Q water to balance the cartridge. Ten milliliters of freshly squeezed Navel orange juice, which was centrifuged for 10 min at ~4,000 x g to get rid of the leftover albedo, peel, and oil, was loaded into the cartridge. The cartridge was then washed with five cartridge volumes of pure milli-Q water to remove the residual orange juice attached to the cartridge. To elute the LARL, 2 mL of Tris buffer at pH 7.0 with 20% ACN was injected into the cartridge and the eluted sample was collected in a Falcon tube.

LARL was quantified by monitoring the developed limonin. Purified LARL was acidified by concentrated sulfuric acid to accelerate the development. The protocol for limonin quantification was executed after the 1-hour acid treatment.

Limonin was transformed to limonoate by treatment with a sodium hydroxide solution at pH 10 for 2 hours. Subsequently, direct infusion mass spectrometry using the negative ion detection mode at a range of 300–800 m/z was used to verify the production of limonoate.

2.5 LGT Activity Assay

Reaction mixtures were composed of x μL aglycone (dependent on the concentration of stock solution, final concentration was 0.5 mM) either in an aqueous solution or DMSO, 1 μL UDP-glucose (final concentration 1 mM), 10 μL of purified GT (~ 1 mg/mL), 50 μM $MnCl_2$, and (100 - x) μL reaction buffer (pH 7.0 / pH 5.5 Tris buffer or pH 4.0 acetate buffer). The reaction was incubated at 37 °C for 24 hours and quenched by 100 μL of ACN. If LARL and limonoate were substrates, the reaction was treated with strong sulfuric acid (pH 2.0) for 1 hour before reaction quenching to convert the unused substrate into limonin. All the reaction mixtures were then subject to LC-MS for further quantification. Each reaction sample was repeated three times. The P-value representing significance in the pairwise comparison of activity levels between reaction samples and control samples was determined using a paired t-test.

2.6 Sequence analysis and homology modeling

Sequences of *cu*GT and *cp*GT were pairwise aligned using MUSCLE.[17] InterProScan was used to predict the existence of Plant Secondary Product Glucosyltransferase (PSPG) consensus sequence. In order to identify templates to be used for homology modeling, sequences of *cu*GT and *cp*GT were searched using HMMER against the PDB database.[18] The sequences greater than 25% similarity were used for sequence alignment with targets. To increase the sampling efficiency of the model, *cu*GT and *cp*GT sequences were trimmed by removing the C-terminus (position 471–511 in protein sequence) portion of the sequence that did not match any template. Three hundred models were generated using the RosettaCM protocol,[19] and the single lowest energy model for each sequence was selected and checked visually to avoid possible mismatches.

3. Results and Discussion

3.1 Expression, purification, and functional characterization of *cu*GT

In order to explore the possibility of debittering orange juice by recombinant enzymes, a gene that has been reported to encode limonoid glucosyltransferase from *C. unishiu* (Uniprot ID Q9MB73; *cu*GT)[15] was synthesized and expressed in *E. coli*. However, no soluble expression was observed, consistent with the report of the annotated *cp*GT.[16] Indeed, the expression of plant glycosyltransferase genes in bacteria, especially in *E. coli*, has been reported in published literature,[20] but very few studies have shown the expression of LGT cloned from citrus plants. While the isolated protein was not solubly produced, it is well established that co-expression tags can result in the production of soluble protein.[21] Therefore, we explored the use of two common co-expression tags, glutathione S-transferase

(GST) and maltose-binding protein (MBP), by creating synthetic genes that encoded these tags as fusion proteins on the N-terminus of *cu*GT. The GST tag was originally reported to enable soluble expression of *cu*GT in *E. coli*.[15] The MBP tag is commonly used in plant glycosyltransferase production and characterization with the purpose of increasing solubility in a bacterial expression system.[22, 23] For example, the flavonoid 7-O-glycosyltransferase CsUGT75L12 identified in *Camellia sinensis* was heterogeneously expressed in *E. coli* fused with MBP while maintaining its flavonoid glycosylation activity.[23] Our results showed that MBP fusion enabled *cu*GT to be readily expressed and purified using standard recombinant protein production techniques; however, no expression of *cu*GT was observed from the GST fusion construct using the same conditions (Figure. 2A, Figure S1). The isolated MBP fused protein has the predicted molecular weight (103 kDa), and it showed a clear band on SDS-PAGE.

To quantitatively characterize the MBP-*cu*GT activity, we developed a method to quantify the differential production of limonin before and after the treatment of LARL solutions with MBP-*cu*GT. This indirect method was required as there are no readily available isolated standards of purified LARL. Briefly, this method measures the difference between the concentration of limonin before and after acidifying, in which any LARL present in the solution spontaneously converts to limonin under acidic conditions. Any increase in limonin observed after acidification is assumed to result from the cyclization of LARL. If LARL is converted to glucosyl-limonin, then the cyclization cannot occur and no difference in limonin after acidification would be observed. To measure the level of limonin, a modified electrospray ionization liquid chromatography-mass spectrometry (ESI LC-MS) protocol was adopted.[24] The concentration of limonin was calibrated by the ratio of its ion intensity to podophyllotoxin. The limit of detection for limonin was < 0.5 ppm and the R^2 of the standard curve was 0.99 (Figure S2 of Supporting Information).

Initially we evaluated if Navel orange juice mixed with MBP-*cu*GT and 1 mM UDP-glucose would result in a decrease of LARL in the juice. A range of pH and temperatures were evaluated, from which LARL and limonin concentrations were measured. We did not observe any conditions in which the incubation of juice with MBP-*cu*GT was able to significantly decrease limonin, as indicated by the P-value ($P > 0.1$) of the pairwise comparison of limonin and LARL levels. This indicated that MBP-*cu*GT might not have sufficient activity to be utilized directly for debittering Navel orange juice.

3.2 Preparation of LARL and limonoate

While no LGT activity was observed with the recombinantly produced MBP-*cu*GT fusion protein under a variety of pHs and temperatures, it is possible that the interference from orange juice, including metals, salts, and other metabolites or proteins present, may inhibit LGT activity.

Therefore, we developed a method to isolate LARL in order to characterize potential LGT activity on the purified substrate in a simplified environment. LARL is naturally occurring in orange juice and can be generated enzymatically from limonin by limonin hydrolase in orange seeds. Previously described methods for the *in vitro* enzymatic generation of LARL depends on the purified limonin hydrolase from orange seeds.[25] However, the tedious

purification procedures of limonin hydrolase and the complicated follow-up purification of product make isolation of LARL difficult. In addition, the heterologous expression of limonin hydrolase and direct treatment of limonin is not possible as there is no experimentally verified gene sequence of limonin hydrolase available. Therefore, we developed an SPE protocol to obtain LARL from citrus juice and control the quality of resultant LARL by selected ion mode mass spectrometry. Following the same method as that used for the determination of LARL in orange juice, acidic conditions facilitated the conversion from LARL into limonin, which enabled us to indirectly evaluate the concentration of LARL. As shown in Figure 3A, >20 ppm LARL was purified in the eluted sample from the column, while <0.5 ppm of limonin contamination (less than 4%) was observed.

In addition to evaluating MBP-*cu*GT activity on LARL, we wanted to explore if there was activity on limonoate (limonin with both the A- and D- rings are hydrolyzed). This was critical as previous cuGT activity was evaluated on a product mixture of hydrolyzed limonoate and not directly on LARL.[13, 15] To generate limonoate, we incubated limonin at pH 10 with subsequent direct infusion into the ESI- MS with single-quadrupole mode monitoring from 200–800 m/z. The peak at m/z 487.2 corresponding to LARL can be easily observed after 5 mins, indicating the fast conversion from limonin to a limonin species with either the A- or D-ring hydrolyzed. Over time, a peak at m/z 505.2 was also observed, which would correspond to the expected mass of limonoate. The peak at m/z 487.2 eventually disappeared while the peaks at m/z 469.1 (limonin) and 505.2 (limonoate) were dominant after 2 hours at pH 10.0 (Figure 3C). The excess of limonin was then removed by centrifugation due to the water insolubility of limonin.

3.3 No significant activities were observed with LARL and limonoate as substrates

The LARL and limonoate prepared by the above protocols could serve as the substrates for the LGT enzyme assay. The treatment of LARL with LGT for 24 h at 37 °C resulted in no significant decrease of converted limonin compared to the control (Figure 3. B). There was no activity observed for MBP-*cu*GT on the naturally isolated LARL. To exclude the possibility that glycosyl-limonin was not stable under the acidic condition in which LARL was converted into limonin, we evaluated the pH stability of glycosyl-limonin at pH 2.0. When 60.0 ppm of glycosyl-limonin was exposed to pH 2.0 for 1 hour, 59.8 ppm of glycosyl-limonin remained, suggesting that glycosyl-limonin was stable under acidic conditions (Figure. S3).

Previous reports of LGT enzymes utilized a synthetic form of limonoids that contained a mixture of limonoate and monolactone. Therefore, we used our similarly prepared synthetic limonoate to evaluate LGT activity. Similar to the result observed for LARL, no significant change was observed in limonin developed from the limonoate with and without MBP-*cu*GT and UDP-glucose incubation (Figure. 3D).

While this is not consistent with the data reported in the *cu*GT characterization effort, it is consistent with the lack of activity observed for *cp*GT, which had 98.8% sequence similarity to *cu*GT.[16] We aligned and analyzed sequences to determine whether the sequences of *cp*GT and *cu*GT likely account for the inactivity towards limonoids. InterProScan was used to

predict the existence of a consensus sequence of Plant Secondary Product Glucosyltransferase (PSPG) at the C-terminus of both proteins.[26] The PSPG motif represents the nucleotide diphosphate sugar binding site and determines the substrate scope and specificity.[27] Nevertheless, both proteins share an identical PSPG motif, with a total of only six amino acid substitutions observed between the protein sequences (Supplementary information S4).

In order to further evaluate the potential influence of these substitutions on the protein structure and function, three dimensional molecular models were generated for *cu*GT and *cp*GT using RosettaCM.[19] The models generated used templates with 25–35% identity and 50–63% similarity with coverage across 91% of the sequence. Based on previous studies evaluating the accuracy of molecular modeling tools for enzymes, these are expected to be high quality and topologically accurate models.[19, 28] The root-mean-square deviation (RMSD) of the *cp*GT and *cu*GT structures is approximately 0.8 Å, indicating the predicted similarity of overall structures and substrate binding sites (Supplementary information S5A). Most importantly, the substitutions are predicted to occur at least 15 Å away from the active site pocket (Supplementary information S5B). While the residues distal to the active site have been reported effect activity,[29] these mutations generally modulate activity less than 10-fold and are not the differentiating factor between activity observed and not observed for an enzyme. Given the overall high sequence and structural similarity, coupled with their predicted active sites being 100% conserved, we expect that these two enzymes are functionally equivalent.

3.4 Functional characterization of LGT

The lack of LGT activity with Navel orange juice, LARL, and limonoate indicated that either MBP-*cu*GT was previously incorrectly identified as LGT, or that the MBP-*cu*GT protein we produced was inactive. Therefore, it was critical to evaluate whether MBP-*cu*GT had any glycosyl transferase activity to rule out the possibility of the produced protein being inactive. Based on the sequence, *cu*GT is in the GT1 glycosyltransferase family, one of the ubiquitous glycosyltransferase families controlling the widespread modification of secondary metabolites in plants. Many previously characterized plant GT1 glycosyltransferases present substrate promiscuity towards flavonoids and are able to biotransform flavonoid aglycones to glycosides in various positions[30–32]. We hypothesized that an active MBP-*cu*GT enzyme would have some level of promiscuity and catalyze the glycosylation of flavonoids. To evaluate this hypothesis, MBP-*cu*GT was screened for enzyme activity with 12 substrates as acceptors, including 10 flavonoids with different hydroxyl group locations and heteroatoms (N-) and two structurally similar compounds (Figure 4A). The substrate library was designed to give a high-level understanding of substrate regio-selectivity and specificity for these enzymes. However, a broader and more comprehensive panel of substrates will be needed to fully understand the scope of molecules these enzymes can work on, providing potential insights into their natural function.

In the screen, 10 out of the 12 chemicals were shown to be the active substrates for MBP-*cu*GT, which was verified by the HPLC-MS (Figure. 4A, Supplementary information Figure S6-S15). The MBP-*cu*GT could function on two atom types: nitrogen (e.g., 6-aminoflavone)

and oxygen (e.g., 6-hydroxyflavone). For flavonoids with one nucleophilic atom available, for example, 7-hydroxyflavone, its product peak appeared before the substrate (Figure 4B) and the glycosylated product, flavonoid-7-O-glycoside, was verified by mass spectrometry.

For substrates with two nucleophilic groups available, two different reaction patterns were observed. When MBP-*cu*GT was reacted with apigenin, two positions were glycosylated, resulting in the formation of a mixture of glucosides (Figure 4C). Based on the observed mass of the peaks, these include both monoglucosides and diglucosides. In contrast, when genistein was evaluated as a substrate, glycosylation only occurred at one site and a single peak with a mass corresponding to a single glucosylation event was observed (Supplementary information Figure. S6). Because the only difference between the two chemicals was the position where the six-membered heterocyclic C-ring was substituted by phenol ring B, MBP-*cu*GT clearly demonstrates some degree of substrate specificity. Given these results, we can conclude that the MBP-*cu*GT is active but does not catalyze the glucosylation of LARL or limonoate, and therefore is incorrectly identified as a LGT.

Our results also highlight that it is still a challenge to predict substrate specificity and recognition mode of GTs by phylogenetic analyses. The sequence identity of *cu*GT with gallate 1-beta-glycosyltransferase from *Quercus robur*[33] and cinnamate beta-D-glucosyltransferase from *Fragaria ananasa*[34] are 72% and 66.5%, respectively. These homologs have been observed to utilize substrates with a carboxylic acid functional group. Therefore, it was unexpected that MBP-*cu*GT was inactive on both salicylic acid and sinapic acid. It is presently unclear what determines MBP-*cu*GT substrate preference and future studies will investigate selectivity on hydroxyl and amino group. Considering the size and potential function of MBP, future experiments will be conducted on how it affects kinetic and catalytic properties. In addition, the closely related *cp*GT was not reported to glycosylate flavonoids. We speculate that the difference in activity was due to the assay time. In the *cp*GT study, a very short reaction time of five minutes was used, which is generally only useful for highly active enzymes functioning on their native substrate. When screening for promiscuous or low-level enzyme activities, a multi-hour time frame is more commonly used, such as the 24-hour assay used in this study.[35, 36]

4. Conclusion

In this study, we developed the protocol to obtain limonin derivates and the corresponding LC-MS method for activity characterization, which were used to explore the activity of MBP-*cu*GT on limonin and its derivates, LARL and limonoate. We observed that the MBP *cu*GT previously annotated as a limonoid UDP-glucosyltransferase is not able to utilize limonoids as substrates. Instead, MBP-*cu*GT is able to catalyze glycosylation of various flavonoids, indicating that it displays promiscuity and its function has been incorrectly identified. Our research highlights the need for further efforts to identify the enzyme responsible for LGT activity in order to develop biotechnology-based approaches for producing orange juice from varietals that traditionally have a delayed bitterness.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Betsy Tse who coordinated the project and provided critical insight and perspective as the project progressed.

Primary funding for this research was provided by Coca-Cola. Additional support was received by J.B.S and Y.C. from the University of California, Davis, the National Science Foundation Award Numbers 1827246, 1805510, 1627539, the National Institute Of Environmental Health Sciences of the National Institutes of Health (NIH) under Award Number P42ES004699, NIH Award Number R01 GM 076324-11, and the Rosetta Commons. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or National Science Foundation.

5. Abbreviations

LARL	limonoate A-ring lactone
LGT	limonoid UDP-glucosyltransferase
MBP	maltose-binding protein
HPLC-MS	liquid chromatography–mass spectrometry
ACN	acetonitrile
TCEP	tris(2-carboxyethyl)phosphine
ESI	electrospray ionization
UDP	uridine diphosphate
SPE	solid-phase extraction
IPTG	Isopropyl β-D-1-thiogalactopyranoside

7. References

1. Jayaprakasha GK, Patil BS, In vitro evaluation of the antioxidant activities in fruit extracts from citron and blood orange. Food Chem 101:410–418 (2007).
2. Paula BMD, Raithore S, Manthey JA et al., Active taste compounds in juice from oranges symptomatic for Huanglongbing (HLB) citrus greening disease. LWT – Food Sci *Technol* 91:518–525 (2018).
3. Raithore SAK, Johannes and Manthey, John A. and Plotto, Anne and Bai, Jinhe and Zhao, Wei and Baldwin, Elizabeth, Mitigation of Off-flavor in Huanglongbing-affected Orange Juice Using Natural Citrus Non-volatile Compounds. J Agric Food Chem 68:1038–1050 (2020). [PubMed: 31814396]
4. Narnoliya LK, Jadaun JS, Biotechnological Avenues for Fruit Juices Debittering. Green Bio-processes 119–149 (2019).
5. Hasegawa S, Limonin bitterness in citrus juices. Flavor Chemistry (1999).
6. Glabasnia A, Hofmann T, On the non-enzymatic liberation of limonin and C 17-epilimonin from limonin-17-β-d-glucopyranoside in orange juice. European Food Research and Technology 228:55–63 (2008).

7. Raithore S, Dea S, McCollum G et al., Development of delayed bitterness and effect of harvest date in stored juice from two complex citrus hybrids. J Sci Food Agric 96:422–429 (2016). [PubMed: 25615579]
8. Guadagni DG, Maier VP, Turnbaugh JG, Effect of some citrus juice constituents on taste thresholds for limonin and naringin bitterness. J Sci Food Agric 24:1277–1288 (1973).
9. Gualdani R, Cavalluzzi MM, Lentini G, Habtemariam S, The Chemistry and Pharmacology of Citrus Limonoids. Molecules 21:1530 (2016).
10. Manners GD, Citrus limonoids: analysis, bioactivity, and biomedical prospects. J Agric Food Chem 55:8285–8294 (2007). [PubMed: 17892257]
11. Maier VP, Hasegawa S, Hera E, Limonin D-ring-lactone hydrolase. A new enzyme from citrus seeds. Phytochemistry 8:405–407 (1969).
12. Roy A, Saraf S, Limonoids: overview of significant bioactive triterpenes distributed in plants kingdom. Biological and Pharmaceutical Bulletin 29:191–201 (2006). [PubMed: 16462017]
13. Karim MR, Hashinaga F, Preparation and properties of immobilized pummelo limonoid glucosyltransferase. Process Biochem 38:809–814 (2002).
14. Hasegawa S, Suhayda CG, Hsu WJ, Robertson GH, Purification of limonoid glucosyltransferase from navel orange albedo tissues. Phytochemistry 46:33–37 (1997).
15. Kita M, Hirata Y, Moriguchi T et al., Molecular cloning and characterization of a novel gene encoding limonoid UDP-glucosyltransferase in Citrus. FEBS Lett 469:173–178 (2000). [PubMed: 10713266]
16. Devaiah SP, Owens DK, Sibhatu MB et al., Identification, Recombinant Expression, and Biochemical Analysis of Putative Secondary Product Glucosyltransferases from *Citrus paradisi*. J Agric Food Chem 64:1957–1969 (2016). [PubMed: 26888166]
17. Edgar RC, MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797 (2004). [PubMed: 15034147]
18. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD, HMMER web server: 2018 update. Nucleic Acids Res 46:W200–W204 (2018). [PubMed: 29905871]
19. Bertolani SJ, Siegel JB, A new benchmark illustrates that integration of geometric constraints inferred from enzyme reaction chemistry can increase enzyme active site modeling accuracy. PLoS One 14:e0214126 (2019). [PubMed: 30947258]
20. Yang M, Fehl C, Lees KV et al., Functional and informatics analysis enables glycosyltransferase activity prediction. Nat Chem Biol 14:1109–1117 (2018). [PubMed: 30420693]
21. Nakatsuka T, Sato K, Takahashi H, Yamamura S, Nishihara M, Cloning and characterization of the UDP-glucose: anthocyanin 5-O-glucosyltransferase gene from blue-flowered gentian. J Exp Bot 59:1241–1252 (2008). [PubMed: 18375606]
22. Yin Q, Shen G, Chang Z, Tang Y, Gao H, Pang Y, Involvement of three putative glucosyltransferases from the UGT72 family in flavonol glucoside/rhamnoside biosynthesis in Lotus japonicus seeds. J Exp Bot 68:597–612 (2017). [PubMed: 28204516]
23. Dai X, Zhuang J, Wu Y et al., Identification of a Flavonoid Glucosyltransferase Involved in 7-OH Site Glycosylation in Tea plants (*Camellia sinensis*). Sci Rep 7:5926 (2017). [PubMed: 28725058]
24. Manners GD, Breksa AP, Schoch TK, Hidalgo MB, Analysis of bitter limonoids in citrus juices by atmospheric pressure chemical ionization and electrospray ionization liquid chromatography-mass spectrometry. J Agric Food Chem 51:3709–3714 (2003). [PubMed: 12797731]
25. Breksa AP, Manners GD, Determination of limonin D-ring lactone hydrolase activity by solid phase extraction with indirect fluorescence detection. J Agric Food Chem 52:3772–3775 (2004). [PubMed: 15186095]
26. Quevillon E, Silventoinen V, Pillai S et al., InterProScan: protein domains identifier. Nucleic Acids Res 33:W116–W120 (2005). [PubMed: 15980438]
27. Osmani SA, Bak S, Moller BL, Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. Phytochemistry 70:325–347 (2009). [PubMed: 19217634]
28. Dill KA, MacCallum JL, The protein-folding problem, 50 years on. science 338:1042–1046 (2012). [PubMed: 23180855]

29. Chen J, Yuan Z, Zhang H et al., Cit1, 2RhaT and two novel CitdGlcT s participate in flavor-related flavonoid metabolism during citrus fruit development. J Exp Bot 70:2759–2771 (2019). [PubMed: 30840066]

30. Pandey RP, Gurung RB, Parajuli P, Koirala N, Sohng JK, Assessing acceptor substrate promiscuity of YjiC-mediated glycosylation toward flavonoids. Carbohydr Res 393:26–31 (2014). [PubMed: 24893262]

31. Song C, Gu L, Liu J et al., Functional characterization and substrate promiscuity of UGT71 glycosyltransferases from strawberry (*Fragaria × ananassa*). Plant Cell Physiol 56:2478–2493 (2015). [PubMed: 26454881]

32. Williams GJ, Zhang C, Thorson JS, Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. Nat Chem Biol 3: 657–62 (2007). [PubMed: 17828251]

33. Mittasch J, Böttcher C, Frolova N, Bönn M, Milkowski C, Identification of UGT84A13 as a candidate enzyme for the first committed step of gallotannin biosynthesis in pedunculate oak (*Quercus robur*). Phytochemistry 99:44–51 (2014). [PubMed: 24412325]

34. Lunkenbein S, Bellido M, Aharoni A et al., Cinnamate metabolism in ripening fruit. Characterization of a UDP-glucose:cinnamate glucosyltransferase from strawberry. Plant Physiol 140:1047–1058 (2006). [PubMed: 16443693]

35. Irmisch S, Jo S, Roach CR et al., Discovery of UDP-glycosyltransferases and BAHD-acyltransferases involved in the biosynthesis of the antidiabetic plant metabolite montbretin A. The Plant Cell 30:1864–1886 (2018). [PubMed: 29967287]

36. Gantt RW, Peltier-Pain P, Singh S, Zhou M, Thorson JS, Broadening the scope of glycosyltransferase-catalyzed sugar nucleotide synthesis. Proc Natl Acad Sci U S A 110:7648–7648 (2013). [PubMed: 23610417]

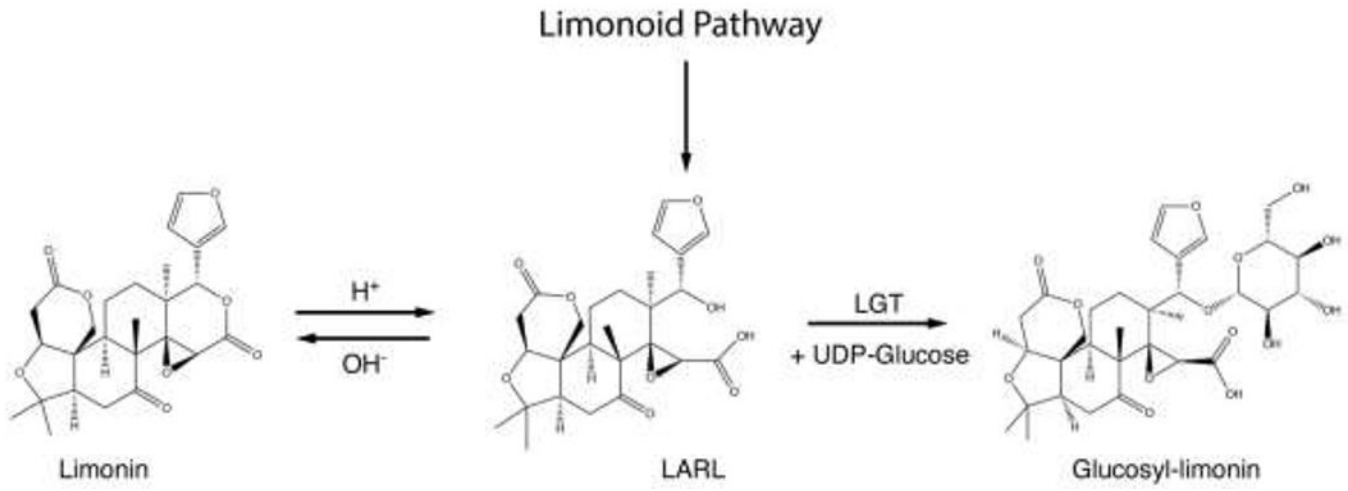

Figure 1.

Biosynthetic pathway of limonin and its related enzymatic reaction. LARL, the precursor of limonin, could be catalyzed by LGT to form non-bitter glucosyl-limonin, which serves as the naturally debittering approach.

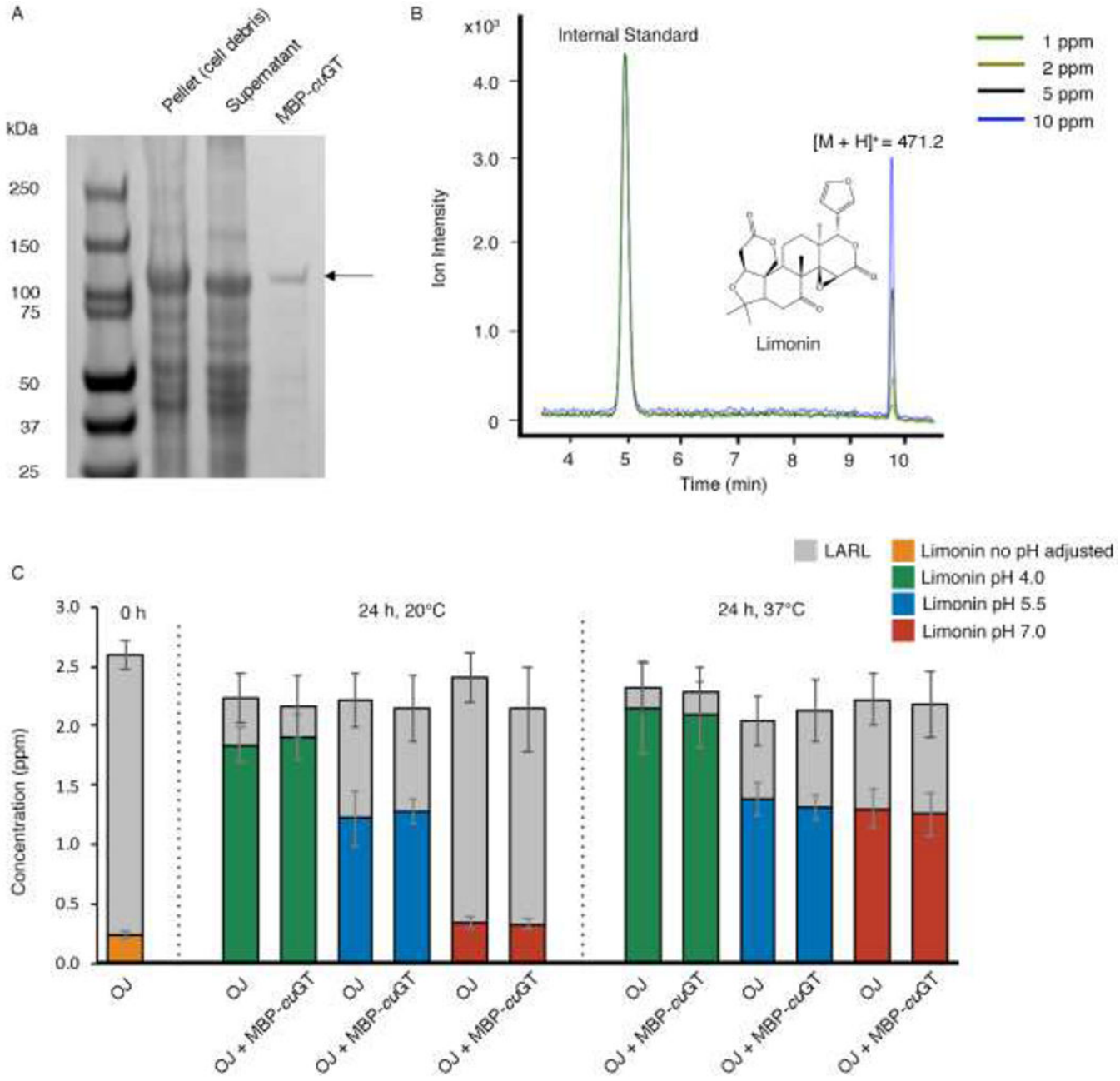

Figure 2.

Expression and purification of *cu*GT with MBP tag (MBP-*cu*GT), and evaluation of its activity towards Navel orange juices. (A) LGT with co-expression of MBP tag at N-terminus was solubly expressed in *E. coli* and purified by IMAC as shown by 12 % SDS-PAGE gel (B) The quantification of limonin by reverse phase HPLC ESI+ LC-MS spectrum. Selected ion mode (SIM) in ESI+ at m/z 471.2 corresponding to $[M + H]^+$ of limonin was executed and various concentrations of limonin were sampled to test the sensitivity and accuracy of this method. The standard curve of limonin was drawn by the ion intensity calibrated with internal standard Podophyllotoxin at m/z 397.2 (Figure S2 of Supporting Information). (C) LGT activity assay with Navel orange juice as substrate in different conditions. Purified

MBP-*cu*GT was blended with 1mM UDP-glucose and diluted Navel orange juice in which the pH was adjusted to 4.0, 5.5 and 7.0 and temperature held at either 20 °C and 37 °C for 24 hours. Half of the sample was directly analyzed for limonin concentration, and the second half acidified and after a 24 h incubation was evaluated for limonin. The increase in limonin observed in acidified samples was assumed to be a result of the cyclization of LARL into limonin. No significant decrease in levels of limonin and LARL was observed between MBP-*cu*GT-treated and control orange juice, as indicated by p-value > 0.1 calculated by the t-test. Error bars indicate the standard deviation of the triplicate samples

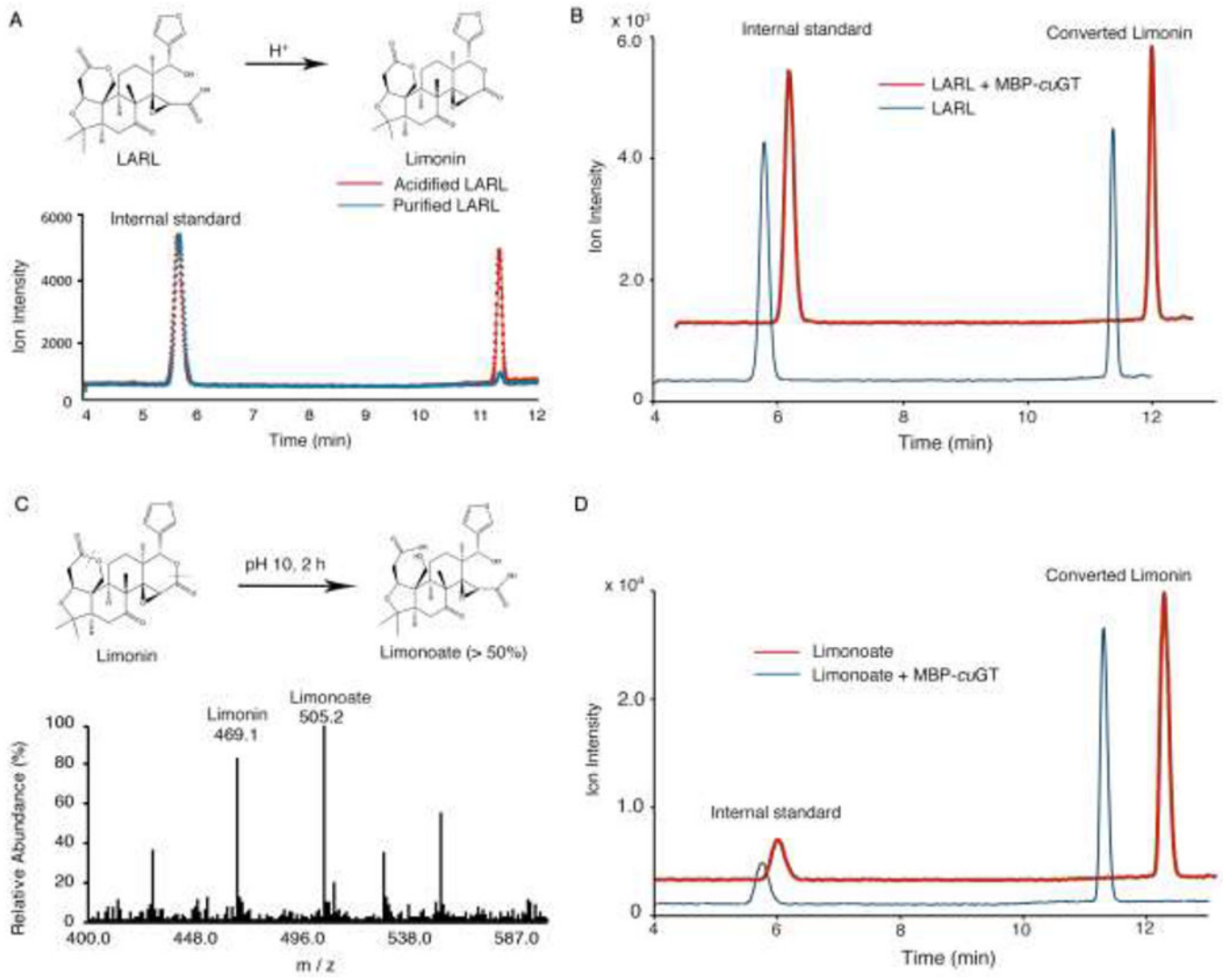

Figure 3.

Preparation of LARL and limonoate as substrates and corresponding MBP-*cu*GT activity assays. (A) LARL could be purified by SPE (See materials and methods) and converted into limonin sequentially by strong acid, enabling its quantification by the same HPLC-MS method for limonin. Purified LARL only contained a small amount of residual limonin as contamination (< 5%). (B) HPLC-MS trace of purified LARL treated with 1mM UDP-glucose and MBP-*cu*GT (1 mg/mL) and the LARL control. No decrease in the observed limonin after incubation 37 °C, 24 hours with UDP-glucose and MBP-*cu*GT is observed, indicating that no LGT activity was observed. (C) Preparation of limonoate by strong base treatment. Because limononate does not occur naturally in citrus juice, limonoate (506.5 Da) can be synthesized by incubating limonin (470.5 Da) under basic conditions for an extended period of time. (D) HPLC-MS trace of limonin converted from limonoate by strong acid, which is used to quantify concentrations of limonoate. Limonoate (>10 ppm) was mixed with UDP-glucose (1 mM) and MBP-*cu*GT (1 mg/mL) under the assay condition 37 °C for

24 hours in tris buffer. As observed, no significant change in the level of limonoate is observed after incubation with MBP-cuGT, indicating the lack of LGT enzymatic activity.

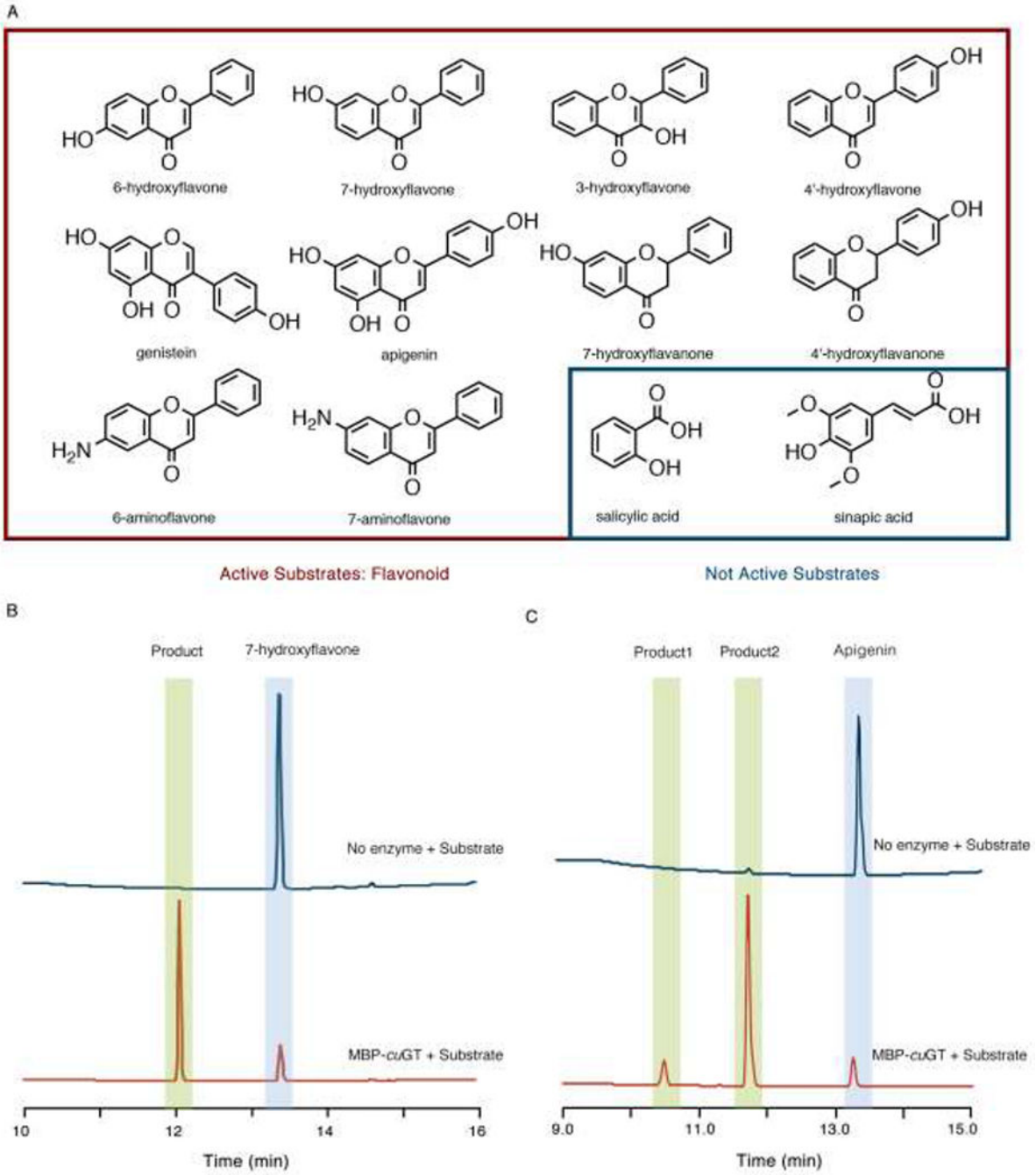

Figure 4.

Activity profile of MBP-*cu*GT towards flavonoids and related chemicals. The substrates were mixed with MBP-*cu*GT (1 mg/mL) and UDP-glucose (1 mM) at 37 °C for 24 hours. The qualitative activity was obtained by comparing the spectrums between the reaction samples and the corresponding controls without enzyme added. (A) The flavonoids accepted by LGT were enclosed by the red box while not active substrates were in the blue box. (B) HPLC spectrum of biotransformation of 7-hydroxyflavone by MBP-*cu*GT as representative of active flavonoid substrates. The product of flavone-7-o-glucoside was then verified by

MS. (C) HPLC spectrum of biotransformation of apigenin by MBP-*cu*GT. Apigenin possessing three hydroxyl functional groups and two of them were glycosylated, resulting in two different products.