Correspondence

Identification of four *Arabidopsis* genes encoding hydroxycinnamate glucosyltransferases

Carsten Milkowski, Alfred Baumert, Dieter Strack*

First published online 23 November 2000

Members of the Brassicaceae accumulate sinapate esters as major phenylpropanoid secondary metabolites, which are fluorescent UV-protective compounds, as shown with *Arabidopsis thaliana* [1]. A pivotal enzyme in sinapate ester biosynthesis is UDP-glucose:sinapate glucosyltransferase (SGT, EC 2.4.1.120) which catalyzes the transfer of glucose from UDP-glucose to sinapate and some other hydroxycinnamates (HCAs), including 4-coumarate, caffeate and ferulate [2]. 1-*O*-Sinapoylglucose is the immediate acyl donor in reactions leading to sinapine (sinapoylcholine) in seeds [3] and sinapoylmalate plants [4,5].

The first gene in sinapate ester biosynthesis has been cloned from *Arabidopsis* [6] and encodes 1-*O*-sinapoylglucose:malate sinapoyltransferase (SMT, EC 2.3.1.92). The second gene involved in this metabolism has very recently been cloned from rape (*Brassica napus*) [7]. It encodes UDP-glucose:SGT. Here we report the cloning and identification of four *Arabidopsis* glucosyltransferase (GT) genes encoding cinnamate and HCA GTs. These genes have previously been putatively assumed to encode indole-3-acetate (IAA) GTs ([8]; Y. Nakamura, NCBI protein database accession number BAB00006). We unambiguously demonstrate, however, that the heterologously expressed *Arabidopsis* enzymes do not catalyze the synthesis of the IAA glucose ester, but rather specifically cinnamate and HCA glucose esters. This is a prominent example of misidentification of new genes. It shows again that the assignment of genes without functional expression carries a high risk of error.

As the corresponding nucleotide sequences retrieved from the databases reveal no introns, the *Arabidopsis* genes encoding the proteins BAB00006, D71419, E71419 and F71419 were amplified by polymerase chain reaction (PCR) with genomic *Arabidopsis* DNA as template. Genomic DNA was isolated according to Brandstädter et al. [9] from rosette leaves of 5-week-old *Arabidopsis* plants (*Arabidopsis thaliana* Heynh. ecotype Columbia) grown in the green house. The following specific primers were used: BAB00006: 5′-ATG GAG CTA GAA TCT TCT CC-3′ (forward) and 5′-TTA AAA GCT TTT GAT TGA TCC-3′ (reverse); E71419: 5′-ATG GAC CCG TCT CGT CAT ACT CAT G-3′ (forward) and 5′-CTA GTG TTC TCC GTT GTC TTC TCT CG-3′ (reverse); F71419: 5′-ATG GAG ATG GAA TCG TCG TTA CCT C-3′ (forward) and 5′-TTA CAC GAC ATT ATT AAT GTT TGT C-3′ (reverse); D71419: 5′-ATG GTG TTC GAA ACT TGT CCA TCT CC-3′ (forward) and 5′-CTA GTA TCC ATT ATC TTT AGT CTT CG-3′ (reverse). The PCR products were inserted into the *Sma*I-linearized vector pQE-32 (Qiagen). Sense and antisense constructs were used to transform the *Escherichia coli* strain M15 (pREP4).

E. coli M15 (pREP4) cells harboring the *Arabidopsis* genes on the pQE-32 vector were grown to logarithmic phase. Expression of the recombinant *Arabidopsis* proteins was then induced by incubating the cells for another 20 h at 30°C in the presence of 1 mM isopropylthiogalactoside (IPTG).

The enzyme assays were performed as described [7]. In short: IPTG-induced *E. coli* cells harboring the *Arabidopsis* GT genes were collected, suspended in lysis buffer (0.1 M Mes buffer, pH 6.0), disrupted by sonification and centrifuged. Aliquots of the supernatants were incubated for 30 min at 30°C with 4 mM UDP-glucose and various acceptor molecules (cinnamate, HCAs, benzoate, hydroxybenzoates or IAA, 2 mM each) in the presence of 0.1 M Mes buffer (pH 6.0). The reactions were terminated by adding trifluoroacetic acid. Enzymatic products were identified by liquid chromatographic-mass spectrometry (LC-MS) and photodiode array-high performance liquid chromatography (PDA-HPLC) analyses. Quantification was achieved by external standardization with cinnamate and HCAs purchased from Sigma. Further details on PDA-HPLC and LC-MS are described elsewhere [7].

In a recent paper [7] we presented results on cloning and heterologous expression of a rape cDNA encoding UDP-glucose:sinapate GT (SGT1). The rape SGT1 sequence showed 64, 63 and 61% sequence identities to putative *Arabidopsis* IAA GTs [7]. In addition, a recent data base search showed the highest identity of rape SGT1 (81% on amino acid level) to a new entry assigned also to an *Arabidopsis* IAA GT-like protein (Y. Nakamura, NCBI protein database accession number BAB00006). These results prompted us to clone and functionally express the four *Arabidopsis* genes in *E. coli*, expecting to find rape SGT1-like enzymes.

We amplified the four respective genes by PCR with genomic DNA isolated from rosette leaves of 5-week-old *Arabidopsis* plants. These genes were introduced into *E. coli* cells as sense and antisense constructs and their expression induced by the addition of IPTG. Crude *E. coli* protein extracts were assayed for GT activities. Testing potential acceptors, it was found that none of the recombinant enzymes was active towards IAA (Table 1). Thus, the earlier putative assignments of these genes are incorrect. Instead, the recombinant enzymes catalyzed specifically the formation of cinnamate glucose ester (1-*O*-cinnamoylglucose) and HCA glucose esters (1-*O*-4-coumaroyl-, -caffeoyl-, -feruloyl- and -sinapoylglucose) with markedly different acceptor specificities. Possible glucose esterification of benzoate, 4-hydroxybenzoate and salicylate was not observed.

Identification of the products, cinnamoyl- and HCA-glucose esters, was achieved by chromatographic comparison (PDA-HPLC) with standard compounds available from a previous study [7] and by LC-MS. The latter gave the expected molecular mass for cinnamoylglucose of 310, for 4-coumaroylglucose of 326, for caffeoylglucose of 342; for feruloylglucose of 356, and for sinapoylglucose of 386. Neither heat-denatured proteins (5 min at 95°C) nor that from *E. coli*

Table 1
Acceptor specificity of the recombinant *Arabidopsis* GTs in 0.1 M Mes buffer (pH 6.0)

Substrates[a]	Relative activities (%)			
	AtSGT1 (BAB00006)[b]	AtHCAGT1 (F71419)[b]	AtHCAGT2 (D71419)[b]	AtHCAGT3 (E71419)[b]
Sinapate	100[c]	100[d]	57	32
Ferulate	7	77	77	94
Cinnamate	–[g]	24	24	76
4-Coumarate	–	66	100[e]	100[f]
Caffeate	–	23	72	45
Benzoate	–	–	–	–
4-Hydroxybenzoate	–	–	–	–
Salicylate	–	–	–	–
Indole-3-acetate (IAA)	–	–	–	–

[a]2 mM at 4 mM UDP-glucose.
[b]NCBI protein database accession numbers.
[c]35.5 μkat kg^{-1} protein.
[d]17.9 μkat kg^{-1} protein.
[e]173 μkat kg^{-1} protein.
[f]87.6 μkat kg^{-1} protein.
[g]–, no product detected by the HPLC system applied.

harboring cDNA antisense constructs showed any GT activities.

One protein (accession number BAB00006 with 81% identity and 87% similarity to rape SGT1) exhibited unexpectedly high specificity towards sinapate. Only ferulate was also esterified, but with a relative activity of only 7% compared to sinapate. Cinnamate, 4-coumarate and caffeate were not accepted at all by this protein (based on the HPLC system applied). In contrast, the other enzymes showed broad acceptor specificities. F71419 (63% identity and 76% similarity to rape SGT1) catalyzed the esterification of sinapate (100%), ferulate (77%), 4-coumarate (66%), cinnamate (24%) and caffeate (23%). E71419 (64% identity and 76% similarity to rape SGT1) and D71419 (61% identity and 74% similarity to rape SGT1) converted cinnamate and all the HCAs tested, however, with highest activities towards 4-coumarate (Table 1). In conclusion, BAB00006 can be classified as UDP-glucose:SGT (AtSGT), E71419, F7141 and D71419 as UDP-glucose:HCA GTs (AtHCAGTs). We propose that the four *Arabidopsis* genes and the previously identified rape SGT1 gene [7] as well as a limonoid GT from *Citrus unshiu* [10] belong to a distinct subgroup of GTs catalyzing the formation of 1-*O*-acylglucosides (β-acetal esters).

We thank Jürgen Schmidt (IPB) for MS analysis. We also thank Toni Kutchan (IPB) for critical reading of the manuscript. This work is part of the research project 'NAPUS 2000 – Gesunde Lebensmittel aus transgener Rapssaat', supported by the Bundesministerium für Bildung und Forschung.

References

[1] Landry, L.G., Chapple, C.C.S. and Last, R. (1995) Plant Physiol. 109, 1159–1166.
[2] Strack, D. (1980) Z. Naturforsch. 35c, 204–208.
[3] Strack, D., Knogge, W. and Dahlbender, G. (1983) Z. Naturforsch. 38c, 21–27.
[4] Tkotz, N. and Strack, D. (1980) Z. Naturforsch. 35c, 835–837.
[5] Mock, H.-P., Vogt, T. and Strack, D. (1992) Z. Naturforsch. 47c, 680–682.
[6] Lehfeldt, C., Shirley, A.M., Meyer, K., Ruegger, M.O., Cusumano, J.C., Viitanen, P.V., Strack, D. and Chapple, C. (2000) Plant Cell 12, 1295–1306.
[7] Milkowski, C., Baumert, A. and Strack, D. (2000) Planta 211, 883–886.
[8] Bevan, M. et al. (1998) Nature 391, 485–488.
[9] Brandstädter, J., Roßbach, C. and Theres, K. (1994) Planta 192, 69–74.
[10] Kita, M., Hirata, Y., Moriguchi, T., Endo-Inagaki, T., Matsumoto, R., Hasegawa, S., Suhayda, C.G. and Omura, M. (2000) FEBS Lett. 469, 173–178.

*Corresponding author. Fax: (49)-345-5582 106.
E-mail: dstrack@ipb-halle.de

Abteilung Sekundärstoffwechsel, Institut für Pflanzenbiochemie, Weinberg 3, 06120 Halle (Saale), Germany

PII: S0014-5793(00)02270-5