Pergamon

Bioorganic & Medicinal Chemistry 10 (2002) 1731–1741

BIOORGANIC & MEDICINAL CHEMISTRY

Arbutin Synthase, a Novel Member of the NRD1β Glycosyltransferase Family, is a Unique Multifunctional Enzyme Converting Various Natural Products and Xenobiotics†

Tobias Hefner,[a] Joachim Arend,[a] Heribert Warzecha,[b] Karsten Siems[c] and Joachim Stöckigt[a,*]

[a]*Johannes Gutenberg-University Mainz, Institute of Pharmacy, Department of Pharmaceutical Biology, Staudinger Weg 5, D-55099 Mainz, Germany*

[b]*Boyce Thompson Institute for Plant Research, Cornell University, Tower Road, 14850 Ithaca, NY, USA*

[c]*AnalytiCon Discovery, Hermannswerder Haus 17, D-14473 Potsdam, Germany*

Received 19 November 2001; accepted 14 January 2002

Dedicated to Professor Leopold Horner on his 90th birthday.

Abstract—Plant glucosyltransferases (GTs) play a crucial role in natural product biosynthesis and metabolization of xenobiotics. We expressed the arbutin synthase (AS) cDNA from *Rauvolfia serpentina* cell suspension cultures in *Escherichia coli* with a 6xHis tag and purified the active enzyme to homogeneity. The recombinant enzyme had a temperature optimum of 50 °C and showed two different pH optima (4.5 and 6.8 or 7.5, depending on the buffer). Out of 74 natural and synthetic phenols and two cinnamyl alcohols tested as substrates for the AS, 45 were accepted, covering a broad range of structural features. Converting rates comparable to hydroquinone were not achieved. In contrast to this broad acceptor substrate specificity, only pyrimidine nucleotide activated glucose was tolerated as a donor substrate. Nucleotide and amino acid sequence analysis revealed AS to be a new member of the NRD1β family of glycosyl transferases and placed the enzyme into the group of plant secondary product GTs. Arbutin synthase is therefore the first example of a broad spectrum multifunctional glucosyltransferase.

Introduction

For natural products, the conjugation of one or more sugar residues — glycosylation — constitutes an essential component of functional diversity. Glucosylation, the transfer of activated glucose to an aglycone substrate, is the predominant modification in plants, compared to microorganisms or animals. Since this enzymatic conjugation is in most cases the last step in the biosynthesis of plant natural products it plays an important role in stabilizing the product or in altering its physiological activity. Plant systems typically retain secondary metabolites intracellularly, therefore their accumulation might interfere with basic physiological functions, especially if the compounds are toxic or precipitate due to low solubility. Glucosylation helps circumvent these detrimental side effects by converting barely soluble aglycones into water soluble glucoconjugates and by enabling storage of the formed compounds within the vacuole. Moreover, once de-toxified and compartmentalized these products can subsequently be hydrolyzed by specific enzymes. Upon cell destruction, glucosidases break down these stable conjugates yielding the toxic aglycones which might act as deterrents against herbivores. Prominent examples of such a plant defense mechanism are glucosinolates and cyanogenic glucosides which are converted into highly reactive isothiocyanates and the toxic hydrogen cyanide.[1,2] An even larger group of structurally diverse secondary metabolites consists of phenolic glucosides whose aglycones may serve as efficient enzyme inhibitors after oxidation.[3]

The biosynthetic formation of all these glucosides is catalyzed by a large enzyme family, named glycosyltransferases (GTs). Although a variety of GTs of plant origin have been identified of which several have been expressed heterologously, their characterization is still very preliminary, especially regarding their substrate specificities and active domains.[4,5]

*Corresponding author. Tel.: +49-6131-392-5751; fax: +49-6131-392-3752; e-mail: stoeckig@mail.uni-mainz.de

†The cDNA sequence reported in this paper is deposited with Genbank under accession number AJ310148

0968-0896/02/$ - see front matter
PII: S0968-0896(02)00029-9

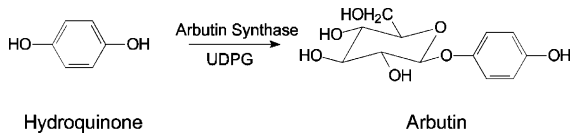


Figure 1. Enzymatic formation of arbutin catalyzed by the novel enzyme arbutin synthase.

In the present paper, we report on the properties of a recently detected glucosyltransferase from plant cell suspension cultures of the Indian medicinal plant *Rauvolfia serpentina*, catalyzing the biosynthesis of arbutin (Fig. 1).[6] The enzyme is named arbutin synthase (AS) (EC 2.4.1...) and has been actively expressed in *Escherichia coli*.[7] It is a novel member of the Class IV GTs belonging to the Nucleotide Recognition Domain type 1β (NRD1β) family of glycosyltransferases.[8] The transferase exhibits the broadest acceptor substrate specificity compared to any of the glucosyl- or glycosyltransferring enzymes previously described. The multifunctional nature of AS sets this enzyme apart from other enzymes of plant secondary metabolism. In addition based on sequence alignment studies AS occupies a special position within the GT enzyme super family.

Results and Discussion

Arbutin, the *O*-β-D-glucoside of hydroquinone, is widely distributed among plants especially of the families Ericaceae, Rosaceae and Saxifragaceae. The glucoside has slight antibiotic activity based on the release of the aglycone hydroquinone during its intestinal metabolism. Arbutin also inhibits the melanin biosynthetic pathway and is therefore applied as a skin lightener in specific cosmetics.[9] Moreover, the compound is one of the most structurally simple plant-derived glucoconjugates. This glucoside is biosynthesised by an UDP-glucose dependent hydroquinone:*O*-glucosyltransferase (named arbutin synthase, AS), recently isolated from cell suspension cultures of the plant *R. serpentina* (Fig. 1).[6] This enzyme belongs to the large enzyme family of glycosyltransferases (GTs), presently consisting of more than 200 members.

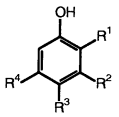
As previously described, we subcloned the AS cDNA from *Rauvolfia* into the pQE60 expression vector and transferred it into M15 *E. coli* cells.[7] These cells were used to produce sufficient enzyme for the extensive characterization presented here, starting with rigorous identification of the enzyme product by enzymatic and spectroscopic analyses. The enzyme product was subjected to enzymatic hydrolysis by α- or β-glucosidase; the latter split the product, arbutin, quantitatively into hydroquinone and glucose. There was, however, no hydrolysis observed in presence of the α-glucosidase, which clearly demonstrates the formed arbutin was a β-glucoside. The ^{1}H and ^{13}C NMR spectra of the enzyme-product coincided in every respect with those of commercially available arbutin. The large coupling constant (7.6 Hz in the ^{1}H NMR spectrum) of the anomeric H-1′ at 4.63 ppm is characteristic for the β-glucosidic linkages which also identifies the formed arbutin as an *O*-β-glucoside.

The simple purification of a crude protein extract based on one-step chromatography on Ni-nitrilotriacetic acid (Ni-NTA) delivers AS in very pure form (~98%), estimated by Coomassie-blue stained SDS-polyacrylamide gel electrophoresis (SDS–PAGE), which was directly used for determination of the enzyme properties. The thermal stability of AS was surprisingly high, as demonstrated by a temperature optimum of 50 °C. It is possible to perform assays for 1 h at this temperature without significant loss of enzyme activity. At 60 °C, the enzyme shows still 20–50% of its activity depending on the length of pre-incubation.

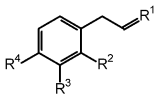
When the pH-dependence of AS was investigated two optima were measured with nearly the same enzyme activity, one optimum at pH 4.5, the other at 6.8 in phosphate buffer (KPi) or 7.5 in citric acid/phosphate buffer, respectively. To exclude temperature effects we verified the result by repeating the experiment at 30 °C (data not shown), which yielded the same results. Although this feature has been observed for other enzymes like an endopeptidase from *Bacillus intermedius*[10] or phytase from *Aspergillus niger*[11] there is no rational explanation for this rather unusual phenomenon. One possibility is that the two pH optima would enable the enzyme to function in intracellular compartments where pH can vary, such as the vacuole.[12] Beside the very broad substrate specificity (see below) this assumption would corroborate the putative physiological role of the enzyme — the detoxification of harmful phenolics. Since most of the conjugated phenolics are stored in vacuoles, a co-localization of the GT would assure that the enzyme encounters its substrates and does not interfere with other primary or secondary pathways, as hypothesized recently.[5] However, with the data obtained so far the subcellular localization of the enzyme remains unknown and needs further investigation. It might be noteworthy that other known UDPG dependent glucosyltransferases from plants exhibit an optimum between pH 7.5 and 8.5[13–15] or around pH 9.0.[16] Even the luteolin glucuronosyl transferase from rye, which was detected in the vacuolar fraction[17] exhibited an pH-optimum of 7.0.[18]

To reveal the function of the AS within *Rauvolfia* cells a careful and extensive analysis of the endogenous substrate was performed. The occurrence of endogenous arbutin, together with the low K_m (<1 μM) provides evidence that hydroquinone (also identified by MS analysis) indeed is the natural substrate for the AS. In contrast, many GTs from various plants like *Gardenia jasminoides* and *Alfalfa* exhibit much higher K_m values for their natural substrate,[16,19] but have not been reported to catalyze the reaction of AS. In the 1960s, it was shown that *Vicia faba* seedlings and wheat germ extracts were capable of synthesizing arbutin when treated with hydroquinone,[20,21] but so far no data have been published regarding the responsible enzyme or the presence of endogenous arbutin in those plants.

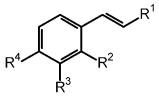
Table 1. Chemical structures of compounds accepted as a substrate by heterologously expressed AS and comparison of relative enzyme activities



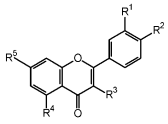
	R^1	R^2	R^3	R^4	%
1. Hydroquinone	–H	–H	–OH	–H	100
2. 4-Chloro-2- methylphenol	$–CH_3$	–H	–Cl	–H	19.0
3. 2-Methoxy- hydroquinone	$–OCH_3$	–H	–OH	–H	16.7
4. 3-Methoxyphenol	–H	$–OCH_3$	–H	–H	10.9
5. Vanillin[a]	$–OCH_3$	–H	–CHO	–H	9.4
6. Saligenin[a]	$–CH_2OH$	–H	–H	–H	9.4
7. Resorcinol[a]	–H	–OH	–H	–H	8.7
8. 4-Hydroxy- benzaldehyde	–H	–H	–CHO	–H	8.6
9. Thymol[a]	$–CH(CH_3)_2$	–H	–H	$–CH_3$	7.6
10. 4-Chlorophenol	–H	–H	–Cl	–H	7.5
11. 4-Nitrophenol	–H	–H	$–NO_2$	–H	6.2
12. Phenol[a]	–H	–H	–H	–H	5.1
13. 4-Methylphenol	–H	–H	$–CH_3$	–H	2.8
14. Methylvanillate[a]	$–OCH_3$	–H	$–COOCH_3$	–H	2.2
15. 4-Hydroxyaceto-phenone[a]	–H	–H	$–COCH_3$	–H	1.8
16. 4-Benzyloxy-phenol	–H	–H	$–O–C_6H_5$	–H	1.4
17. 4-Hydroxy- thiophenol	–H	–H	–SH	–H	1.2
18. Biphenyl-2-ol	$–C_6H_5$	–H	–H	–H	1.1
19. Vanillic acid[a]	$–OCH_3$	–H	–COOH	–H	1.0
20. Carvacrol	$–CH_3$	–H	–H	$– CH(CH_3)_2$	1.0
21. 4-Methoxy- phenol[a]	–H	–H	$–OCH_3$	–H	0.8
22. 3.4-Dimethoxy- phenol	–H	$–OCH_3$	$–OCH_3$	–H	0.8
23. Isovanillic acid	$–OCH_3$	–H	–H	–COOH	0.2
24. 4-Hydroxy- benzoic acid	–H	–H	–COOH	–H	0.1



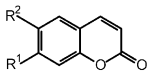
	R^1	R^2	R^3	R^4	%
1. Eugenol	$=CH_2$	–H	$–OCH_3$	–OH	6.2



	R^1	R^2	R^3	R^4	%
1. Coniferyl alcohol[a]	$–CH_2OH$	–H	$–OCH_3$	–OH	1.2
2. *trans*-Cinnamyl alcohol	$–CH_2OH$	–H	–H	–H	0.6
3. *o*-Coumaric acid*	–COOH	–OH	–H	–H	0.4
4. Ferulic acid[a]	–COOH	–H	$–OCH_3$	–OH	0.1

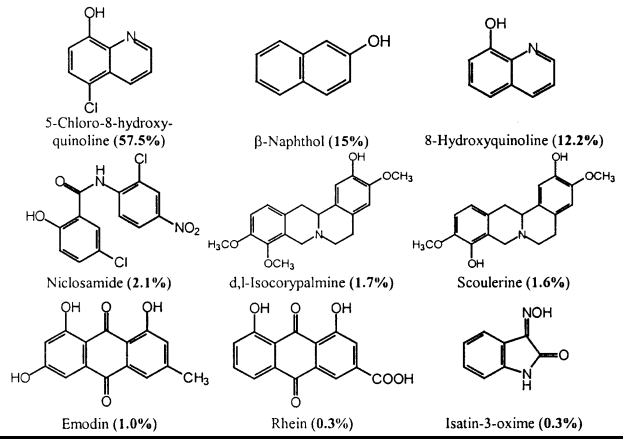


	R^1	R^2	R^3	R^4	R^5	%
1. Apigenin	–H	–OH	–H	–OH	–OH	5.8
2. 3-Hydroxyflavone	–H	–H	–OH	–H	–H	5.4
3. 7-Hydroxyflavone	–H	–H	–H	–H	–OH	3.0
4. Quercetin	–OH	–OH	–OH	–OH	–OH	2.6
5. Kaempferol	–H	–OH	–OH	–OH	–OH	1.0



	R^1	R^2	%
1. Umbelliferone	–OH	–H	3.6
2. Scopoletin	–OH	$–OCH_3$	3.0

Table 1 (*continued*)



[a]Substrates, recently described.[7]

However, the most important property of this novel enzyme is its exceptionally low substrate specificity which could be characterized in a simple way by HPLC analyses. The applied HPLC method is suitable to separate all the formed glucosides from the corresponding aglycones. The difference in retention times between all substrates and products was between 1.5 and 2.5 min depending on the structure of the tested compounds, with the glucosides being the first to elute, providing a simple test system for glucosylation.

From 74 tested natural and synthetic phenolics and two cinnamyl alcohols, 45 were glucosylated in a range from 0.1 to 100% relative activity compared to hydroquinone. This broad range of acceptor substrates consist of 24 simple phenolic structures, four phenylpropanoid derivatives, two coumarins, two anthraquinones, a naphthol, two hydroxyquinolines, five flavonoids including two flavones and three flavonols, and two isoquinoline alkaloids of the protoberberine type. In addition to the more simple phenolics, other substrates with substantial structural differences to the natural substrate hydroquinone are still glucosylated in varying extents (Table 1). For example the synthetic compounds 5-chloro-8-hydroxyquinoline and 4-chloro-2-methylphenol are well accepted with 57 and 19% relative activity followed by the natural occurring phenolics vanillin ($K_m = 440$ μM, Fig. 2), saligenin, thymol (~8%) or eugenol with 6% rel. act. The flavonoids such as apigenin, kaempferol, quercetin, 3-hydroxy- and 7-hydroxyflavone can also serve as an acceptor substrate of the AS in a relative range of ~1–6%. Moreover, coumarins are glucosylated at about 3% rel. act. (K_m of umbelliferone = 550 μM, Fig. 3). Even the two isoquinoline alkaloids scoulerine and D,L-isocorypalmine are glucosylated, which to our knowledge is the first example of glucosylation of this class of alkaloids. Surprisingly cinnamyl alcohol is also converted into its glucoside, because of the allylic OH-group localized in the side chain. This compound and isatin-3-oxime are the only exceptions from our finding that exclusively phenolic structures are accepted by AS. Based on the latter result, it cannot be excluded at this stage that the AS, besides the known flavanon-*O*-glucosyltransferases, could also be responsible in vivo for the biosynthesis of appropriate glucosides of quite different groups of natural products, as indicated by the determined K_m values; especially because the highly substituted apigenin, kaempferol and quercetin are converted into several glucosides. As depicted in Figure 4, up to three different glucosides are formed for instance from quercetin, indicating that the enzyme obviously exhibits not only a broad substrate acceptance but also a low regioselectivity. By LC–MS measurements it could be proved that exclusively monoglucosides are formed, for example, scoulerine is glucosylated at both phenolic OH-residues to the corresponding mono-glucosides. Due to this feature, the AS differs strongly from the betanidin glucosyltransferase previously described,[22] being much more promiscuous with respect to the accepted hydroxyl groups of the substrate.

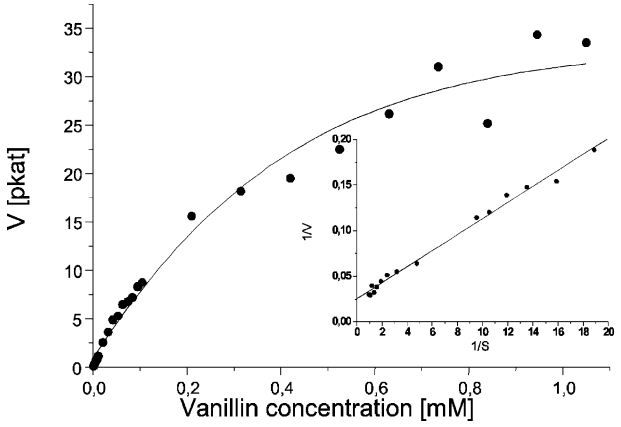


Figure 2. Dependence of arbutin synthase activity on concentration of the acceptor substrate vanillin (linear fit correlation coefficient R = 0.99).

Enzymes of secondary metabolism such as the catalysts involved in alkaloid biosynthesis usually exhibit high specificity for their substrates. In some cases, only a single substrate is known to be accepted by the enzyme,

as has been observed for (*S*)-*N*-methylcoclaurine-3′-hydroxylase[23] in isoquinoline alkaloid biosynthesis or for polyneuridine aldehyde esterase[24] which is a central enzyme in the formation of certain monoterpenoid indole alkaloids. Previous work on methyltransferases, however, points out that this class of enzymes might have varying methylating capabilities accepting for instance hydroxylated cinnamic acids and flavonoid aglycones. Surprisingly, in a recent investigation methyltransferases have been shown to act on more different natural products like simple catechols, phenylpropanoids and structurally complex isoquinoline alkaloids.[23,25] This indicates their multifunctional properties, which is also true for arbutin synthase. Our current knowledge on the substrate specificity of arbutin synthase is summarized in Table 1. Although none of these substrates reached conversion rates comparable to hydroquinone, the result indicates that the enzyme in principal tolerates substrates with very diverse chemical structures, belonging to many classes of plant natural products. Beside *O*-methyltransferases, in fact, arbutin synthase shows the greatest diversity of functions in plant secondary metabolism for any enzyme currently characterized.

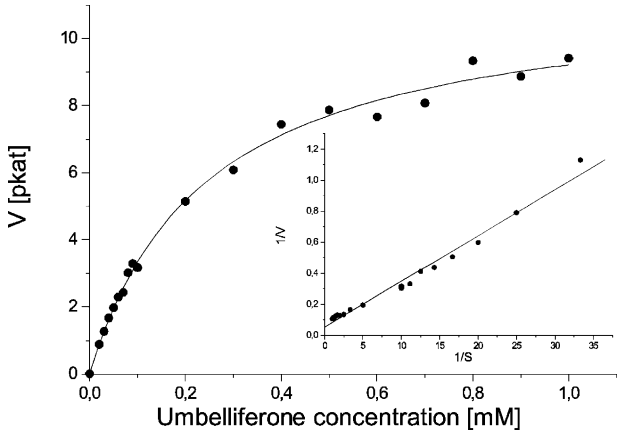


Figure 3. Dependence of arbutin synthase activity on concentration of the substrate umbelliferone (linear fit correlation coefficient R = 0.99).

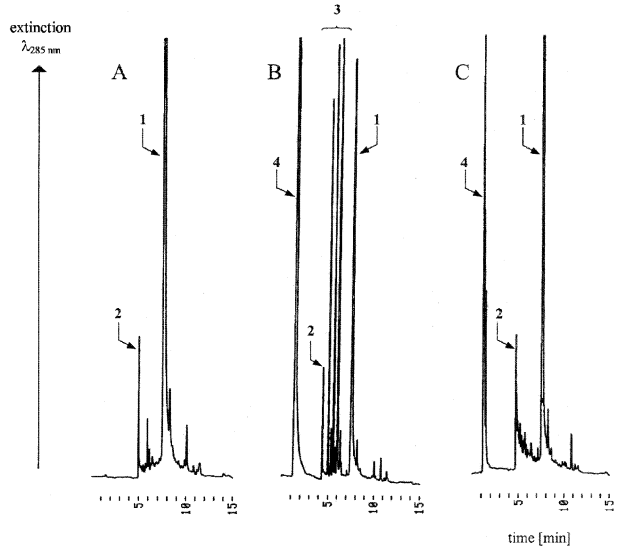


Figure 4. Formation of several glucosides from the flavonoid quercetin catalyzed by arbutin synthase in presence of UDPG. (A) Substrate (Quercetin) incubated with arbutin synthase; (B) incubated in presence of enzyme and UDPG; (C) incubation mixture of (B) re-incubated with 50 nkat β-glucosidase: (1) Quercetin; (2) Imidazole from enzyme solution; (3) mixture of enzymatically formed quercetin glucosides; (4) UDP-glucose.

Although we have, with these experiments, clearly demonstrated the broad acceptance of a variety of compounds, a number of specific structural types could not be converted (Table 2). For instance, carboxy groups were not transformed into glucose esters, a reaction which is catalyzed by the recently described hydroxycinnamate glucosyltransferases.[26] Methoxylated cinnamic and a range of benzoic acids such as trimethoxy cinnamic acid, benzoic acid, acetylsalicylic acid, gallic, and salicylic acid were not glucosylated. Non-acceptance of nicotinic acid, indole-acetic acid and tryptophan supports the above observation. Additional experiments also showed that compounds such as aniline, 4-aminophenol or tryptamine will not act as acceptor substrates for AS, indicating that glucosylation of amino groups will not occur and that a range of lignans and anthocyanines is not glucosylated. Most likely, arbutin synthase must be considered as an enzyme which still recognises specific structural elements which, however, cannot yet be defined in detail at this stage of investigation.

By testing the glucosylation activity of AS in presence of substances not recognized as a substrate, only dichlorophen was found to significantly inhibit the enzymatic formation of arbutin by the synthase. Further analysis of this phenomenon revealed dichlorophen to be a reversible, non-competitive inhibitor (Fig. 5) exhibiting a K_i of 90 μM, but several substances with related structural characteristics did not show inhibition of

Table 2. Compounds, not accepted by heterologously expressed AS as a substrate

Thiophenol	Aniline	4-Aminophenol
Catechol	Gallic acid	Benzyl alcohol
Benzoic acid	Salicylic acid	Acetylsalicylic acid
Menthol	Nicotinic acid	Dichlorophen
Tryptophan	Tryptamine	5-Hydroxytryptamine
Indole-3-acetic acid	Morphine	Sarpagine
Rhapontin, Rhaponticin	Resorufin	Podophyllotoxin
2-Demethoxy-thiocolchicin	*p*-Coumaric acid	2.4.5-Trimethoxy-cinnamic acid
5-Hydroxy-flavone	Cyanidin	Malvidin
Pelargonidin	Arctigenin	Pinoresinol

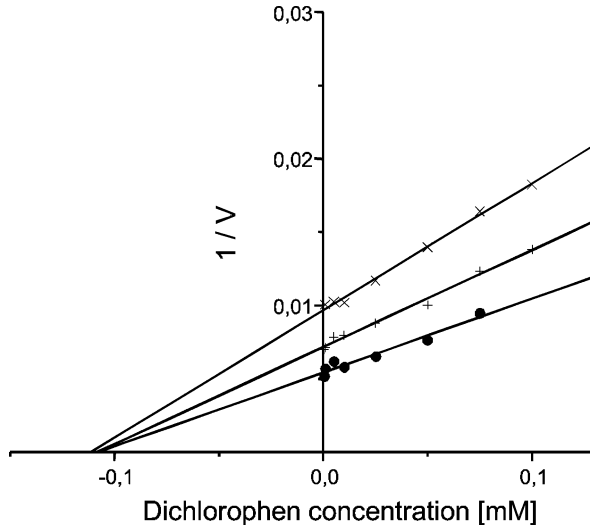


Figure 5. Determination of the inhibition type of dichlorophen and K_i by using a Dixon plot.

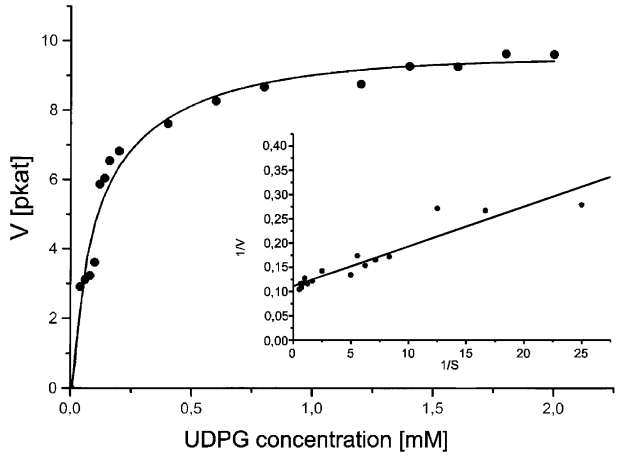


Figure 6. Dependence of arbutin synthase on the concentration of the donor-substrate UDPG (linear fit correlation coefficient R = 0.94).

arbutin formation. We also tested the influence of metal ions on the reaction catalyzed by AS. In presence of 10 mM of the appropriate salt, Mn^{2+}, Ca^{2+}, Co^{2+} and Zn^{2+} inhibited the arbutin formation, showing 75.3, 75.6, 45.8 and 12.6% of the original activity. Addition of 10 mM Mg^{2+} increased the conversion rate to 168.3%. This result is comparable to other known glucosyltransferases.[14,16,27] When checking the effect of Hg^{2+} we discovered that the enzymatic reaction was completely inhibited. This inhibition was reversible by the addition of 20 mM β-mercaptoethanol (65% activitiy recovered). These results suggest that AS has a SH group near or at the catalytic center, as recently assumed for a coumarin glucosylating enzyme.[27] By examining the donor substrate specificity it was found that only pyrimidine nucleotide activated glucose (TDPG at 68.8% and CDPG at 2.2% compared to UDPG) was accepted. The purity of CDPG was veryfied to exclude impurities, such as traces of UDPG. The K_m value of UDPG is 77 μM (K_m TDPG = 1.1 mM) demonstrating that UDPG occupies the position of the most preferred donor substrate (Fig. 6). In contrast, other activated sugars such as UDP-mannose and UDP-galactose, or UDP-glucuronic acid did not function as co-factors. Together with the data for other plant glucosyltransferases, which have been shown to exclusively accept UDP-glucose as a co-factor,[13,14,27–29] this finding is in sharp contrast to the assumed properties of bacterial glycosyltransferases involved in secondary metabolism.[30]

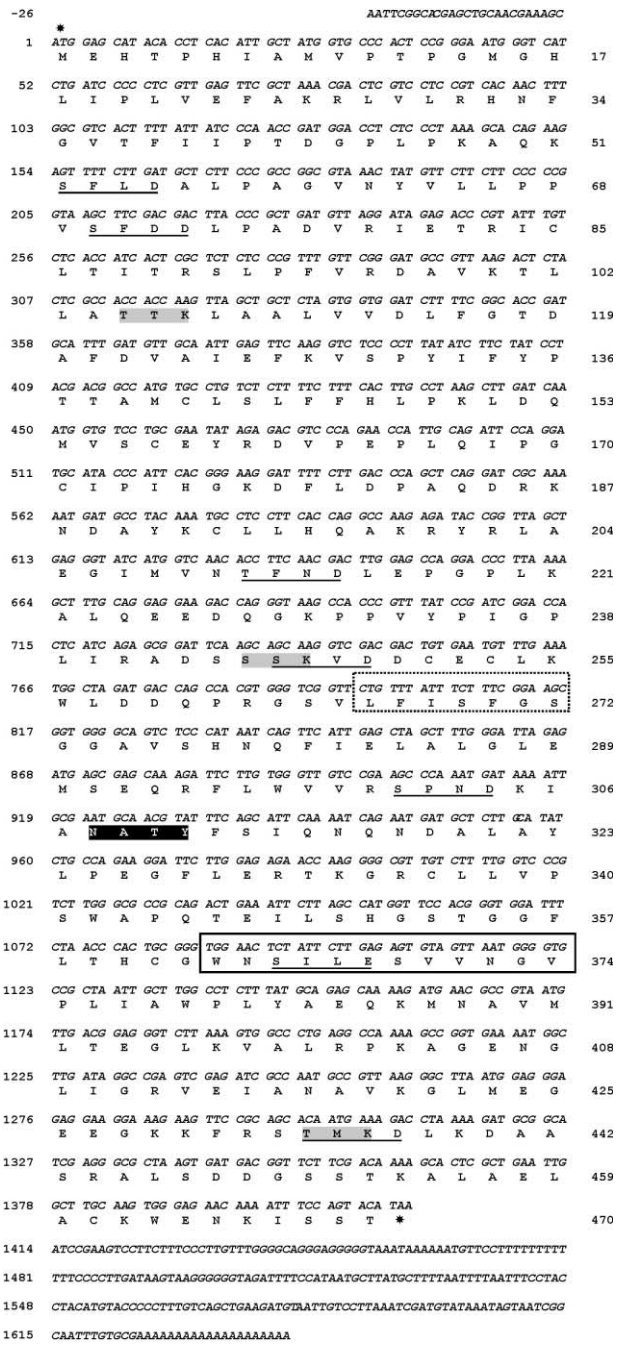


Figure 7. Nucleotide sequence and deduced amino acid sequence of AS from *R. serpentina*. Start and stop codons are marked by ⋆. Casein kinase II phosphorylation sites are underlined, protein kinase C phosphorylation sites are shaded, ASN-glycosylation site is written white on black. The small NRD1β-domain is dot-framed, the large NRD1β-domain is bold-framed.

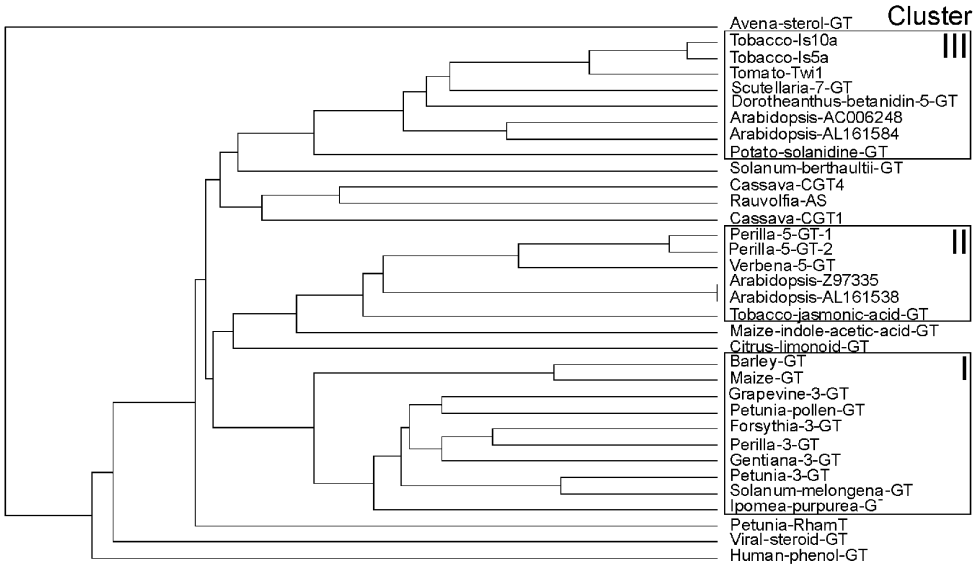


Figure 8. Dendrogram, constructed by sequence-alignment with ClustalW at the European Bioinformatics Institute server illustrates the phylogenetic position of arbutin synthase to different clusters of glucosyltransferases designed recently.[4]

From the limited amount of data available on GTs in general, it is, however, extremely difficult to draw final conclusions about their substrate specificity. In all available reports on GTs, only a small number of substrates was tested, usually having very similar structures and belonging to only one class of natural products. Moreover, in many cases only enriched enzymes were assayed without a guarantee that no other GTs were co-purified. Even for heterologously expressed transferases, no effort was made to obtain a pure enzyme preparation, a point recently criticized in the literature.[5,14] Also, for the majority of the described plant-derived transferases identity was infered only in terms of the enzymes association with the biosynthetic formation of flavonoid and anthocyanin glucosides.[15,16,31–33]

In marked contrast, the arbutin synthase preparations employed here were pure and the enzyme was rigorously characterized in terms of substrate acceptance. Jones and Vogt suggested a classification of GTs, grouping the six so far best characterized heterologously expressed plant GTs into classes with broad, intermediate, and narrow substrate specificity.[5] Adopting this ranking arbutin synthase clearly qualifies for the top rank within the broad substrate specificity group. All results now obtained seem to place this enzyme in a rather outstanding position of GTs, which to date have been assumed to be highly specific, only accepting structural very related structures.[14,31,33,34]

In order to gain more detailed insight into the relationship of AS to other GTs on a molecular basis, we analyzed the nucleotide and amino acid sequence of the AS. The full-length AS-cDNA clone expressed in *E. coli* exhibits an open reading frame of 1413 base pairs (Fig. 7). The peptide sequence deduced from this frame consists of 470 amino acids and corresponds to a calculated molecular mass of 51.8 kDa which matches well with the value of 52.5 kDa determined after purification of AS from *Rauvolfia* cell suspension cultures.[6] The catalyzed reaction of AS proceeds with inversion of the configuration at the anomeric center of the glucosyl donor compared to the formed glucoconjugate. Therefore AS should belong to the NRD1β enzyme family, representing the inverting GTs.[8] This enzyme family is termed after the 'Nucleotide Recognition Domain 1β' which is suspected to be responsible for the recognition of the UDP or TDP portion of the sugar donor. Not only the exclusive acceptance of pyrimidine nucleotide derived donor substrates (UDPG, TDPG, CDPG) but also a sequence analysis and alignment of AS supports clearly this classification. Based on an earlier classification of GTs[8] sequence analysis of AS showed two family specific sequences, a small NRD1βS domain beginning in position 266 and the large NRD1βL domain starting at position 363, both framed in Fig. 7. The sequence also denotes the presence of Ser at position 365 followed by a strictly conserved Glu three amino acids downstream, both amino acids characteristically located in the middle of the NRD1βL site. The Ser is suggested to be a potential phosphorylation site,[35] which might have a regulatory function if it would belong to the catalytic center of the enzyme. However, the catalytic site of all these sugar-transferring enzymes is not known. There are several putative phosphorylation sites, as well as a putative glycosylation site in the whole sequence of AS but their functional role is so far unknown. Eight amino acids upstream of Glu 368 a His residue is found, which again is characteristic for the NRD1β family, and is eventually involved in the catalyzed reaction as a proton donator.[8] Accepting the published classification of GTs,[8] AS would be a novel member of Class IV of the NRD1β enzyme super family.

A sequence comparison using FASTA3 at the European Bioinformatics Institute indicated highest identity [Smith-Waterman score: 62.9% identity (63.5% ungapped)] with a putative UDPG-dependent transferase

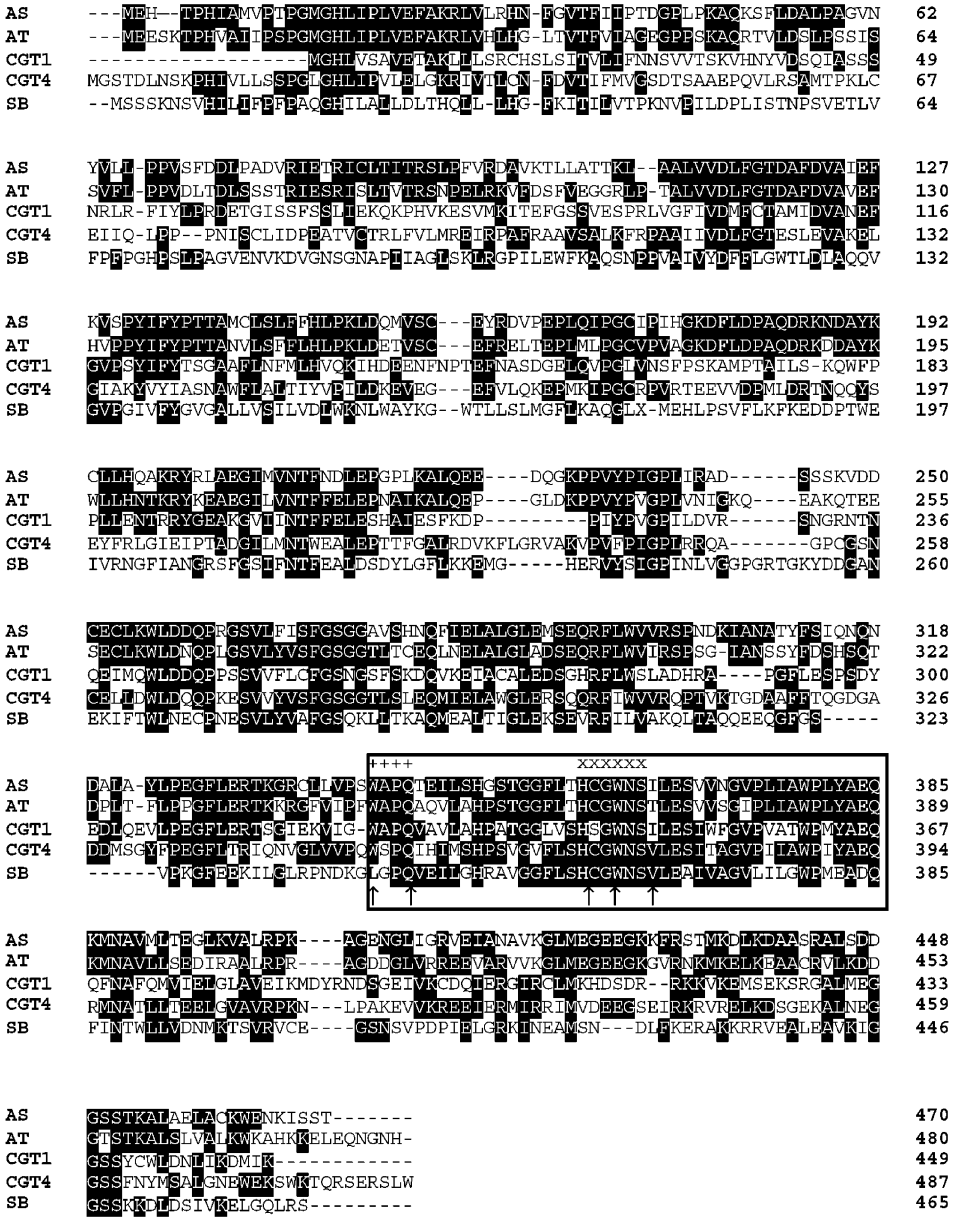


Figure 9. Alignment of deduced amino acid sequences (single-letter code) of AS with most similar glucosyltransferases from plants. The alignment was created using the ClustalW program at the server of the European Bioinformatics Institute (EBI). The 'PSPG box', representing the proposed binding site of UDP-glucose is framed. Motif A + + + +, motif B xxxx, conserved in 95% of all glucosyltransferases ↑. AS: Arbutin synthase from *R. serpentina*, AT: putative Flavonol-glucosyltransferase from *Arabidopsis thaliana* (CAB80916), CGT1: UDP-glucosyltransferase from *Manihot esculenta* (X77459), CGT4: UTP-glucosyltransferase from *Manihot esculenta* (X77462), SB: UDP-glucosyltransferase from *Solanum berthaultii* (AF006081).

(Q9M156) identified in the *Arabidopsis* genome project. The function of the latter enzyme needs, however, to be determined.

A multiple alignment using sequences of thirty three NRD1β GTs places AS between enzyme clusters II and III of a phylogenetic dendrogram (Fig. 8) created very recently.[4] Based on the presently available data, this result points to an outstanding position of the AS. Four sequences of GTs most similar to AS and placed as well between cluster II and III, were aligned with that of AS. As illustrated in Figure 9 the framed consensus sequence named 'plant secondary product GT' (PSPG-box)[36] representing the assumed binding site of UDPG can be realized. This sequence always near the C-terminus, is highly characteristic for those glycosyltransferases responsible for the formation of natural product glycosides. The two highly conserved motives A and B found in this box (WAPQV and HCGWNS, respectively) for which WQ and HGS can be identified in ~95% of *all* β-GTs[4] are again evidence for the correct classification made here for arbutin synthase.

Conclusion and Prospect

The heterologous expression, characterization, and molecular analysis of the novel arbutin synthase, which is not only a new member of Class IV of NRD1β glycosyltransferases, but obviously also the most multi-

functional enzyme of plant secondary metabolism known up to date, provides the major prerequisite to understand in greater detail the still unknown catalytic mechanism of this enzyme class. Site-directed mutagenesis will now permit identification of the proposed binding sites, the catalytic domains and the single amino acid residues which have been suggested as candidates of the catalytic center of GTs. Most probably mutant enzymes might be generated with significantly extended substrate acceptance, which could be used for efficient — also technical — production of structurally diverse and novel glucosides. Because AS is also easily available in a higher mg-scale X-ray structure analysis and/or high field NMR structure determination might become possible in the near future and would be the first example to clarify the three-dimensional structure of an eucaryotic enzyme of the NRD family. Appropriate experiments are now in progress.

Experimental

Purification of His-tagged AS

Recombinant *E. coli* were grown for 24 h at 25 °C and 100 rpm in 24 1-L Erlenmeyer flasks containing 400 mL Luria-Bertani (LB) nutrition medium each [100 mg/L ampicillin, 25 mg/L kanamycin, isopropyl-β-D-thiogalactopyranoside (IPTG) 0.3 mM]. At an OD_{600} of 0.9, the bacteria were harvested by centrifugation at 11,000*g*. A pellet of 15 g bacteria was resuspended in 250 mL buffer [KPi 50 mM, pH 8.0, 300 mM NaCl, 10 mM imidazole and 20 mM β-mercaptoethanol (BME)] and 1 mg/mL lysozyme was added. After 30 min incubation on ice, the cells were lyzed by sonication (75 W, 6 × 10 s). The lysate was centrifuged at 10,000*g* for 30 min and the enzyme-containing supernatant was pumped through a Ni-nitrilotriacetic acid (Ni-NTA) column (volume 2 mL) at a flow rate of 0.5 mL/min. The column was washed with 20 column volumes of buffer with 20 mM imidazole. Elution was achieved with an imidazole gradient (20–250 mM). The purity of AS was demonstrated by Coomassie-blue stained SDS-PAGE.

Protein determination

Enzyme concentrations were measured as described by Bradford.[37] The procedure was the same as recently published.[6]

Activity assay for AS

For monitoring the transferase activity variable enzyme amounts were added to an aqueous solution containing 1 mM hydroquinone, 2 mM UDPG, 100 mM Tris–HCl, pH 7.5 in a total assay volume of 63.8 μL (standard assay). After incubating at 50 °C, for different times, 300 μL MeOH was added and the mixture centrifuged at 18,000*g* for 5 min. The supernatant was subjected to HPLC analysis, using a 250 × 4 mm LiChrospher 60 RP-select B column (5 μm) (Merck, Darmstadt, Germany) and a solvent system consisting of 2% acetonitrile and 98% water, pH 2.3 (H_3PO_4); UDPG 1.2 min, arbutin 3.2 min, and hydroquinone 4.2 min.

Properties of the arbutin synthase

To determine the temperature optimum of AS, the standard activity assay was used (20 min incubation time) but incubation temperature varied from 5 to 90 °C. Before starting the reaction, the enzyme solution and the other components were pre-incubated at the desired temperature for 5, 10 and 15 min. The reaction was started by mixing all components. The pH-dependence of the transferase was measured at 30 and 50 °C using buffer in a range from pH 2–9: citrate/NaOH buffer pH 2–6, phosphate buffer pH 5–9. In addition, citric acid/phosphate buffer was tested in a pH range from 2.2 to 8. Incubations were performed for 20 min. For the measurements of the acceptor substrate specificity the standard assay was employed replacing hydroquinone with the putative substrates. After addition of 25 pkat AS the mixture was incubated at 50 °C for 40 min. The enzyme was precipitated with 300 μL MeOH, centrifuged at 18,000*g* for 5 min, and the supernatant analyzed by HPLC to assess conversion rates. A linear gradient program starting with 2% acetonitrile and 98% water, pH 2.3 (H_3PO_4), and ending after 8 min with 80% acetonitrile and 20% water was employed. Determination of enzyme specificities regarding the donor substrate was performed by incubating the substances (UDPG, TDPG, CDPG, GDPG and UDP-mannose, UDP-galactose, UDP-glucuronic acid) in the standard assay described above. After addition of 300 μL MeOH the supernatant was analyzed by HPLC. Purity of CDPG was checked by TLC using an EtOH/H_2O ($85/15 + 0.1 H_3PO_4$) solvent system, and thymol reagent (0.5 g thymol, 5 mL 96% H_2SO_4, 95 mL 96% EtOH); R_f of CDPG = 0.44, UDPG = 0.5) with UV (254 nm) detection. For checking the influence of metal ions on the activity of AS, standard assays were used with addition of 10 mM (final concentration) of the desired ions. Salts used were $MnSO_4$, $CaCl_2$, $CoCl_2$, $MgCl_2$, $ZnSO_4$ and $HgCl_2$. Examination of the SH dependence of arbutin synthase was performed by adding 10 mM $HgCl_2$ to a solution containing 65 pkat enzyme. After testing the activity (as described above) BME to a final concentration of 20 mM was added and arbutin formation was checked by HPLC. For determination of the inhibition type of dichlorophen and of the inhibition constant, variable concentrations of dichlorophen (0.001–0.1 mM) were incubated with 180 pkat AS in the presence of 1 mM HQ, 2 mM UDPG and 100 mM Tris–HCl, pH 7.5 in a total volume of 127.6 μL for 10 min at 50 °C. After precipitating the enzyme with 300 μL MeOH, the supernatant was analyzed as described above.

Identification of products from the AS-reaction

To verify the formation of β-glucosides, the MeOH added to the incubation mixture was removed by evaporation and 100 μL of a solution, containing 50 nkat almond-derived β-glucosidase (Sigma, Deisenhofen, Germany) in citrate buffer (100 mM, pH 5.0) was added. This mixture was incubated at 37 °C for 3 h. The

reaction was terminated with 300 μL MeOH followed by centrifugation (5 min, 18,000*g*). The supernatant was analyzed by the above described HPLC method. Enzymatically formed arbutin (0.125 μmol) was incubated for 5 h at 37 °C in presence of β-glucosidase (20 nkat) and α-glucosidase (20 nkat) from brewers yeast (Sigma, Deisenhofen, Germany) in 100 mM citrate buffer (pH 5.0) and 100 mM KPi (pH 6.0), respectively. The reaction was stopped by addition of 300 μL MeOH and centrifuged at 18,000*g* (5 min). The supernatant was analyzed by HPLC. Multiple β-glucoside formation: In a total volume of 128 μL Tris–HCl (100 mM, pH 7.5) quercetin and scoulerine, each (0.125 μmol), UDPG (0.25 μmol) and 200 pkat arbutin synthase reacted for 5 h at 37 °C. After addition of MeOH (300 μL) and centrifugation (18,000*g*), 50 μL of the supernatant was subjected to HPLC analysis. MeOH was evaporated and to the remaining incubation mixture β-glucosidase (50 nkat in 150 μL citrate buffer 100 mM, pH 5.0) was added and incubated at 37 °C for 5 h. After termination of the enzymatic reaction with 300 μL MeOH and centrifugation, HPLC analysis was performed. For determination whether different mono-glucosides or multiple glucosylated products are formed by AS the same assay was used. After terminating the reaction as described above, the complete mixture was freeze-dried. The residue was dissolved in 500 μL MeOH. This solution was subjected to LC–MS measurements on a PE SCIEX API 165 instrument (Perkin–Elmer, Langen, Germany). Column used was a 250 × 4 Select B (Merck, Darmstadt, Germany) and a linear gradient program, starting with 15% of MeOH/acetonitrile with 5 mM ammonium formate (1:1) (system A) and 85% 5 mM ammonium formate buffer, pH 4.0 (system B) and ending after 30 min with 100% of system A at a flow of 0.9 mL/min. MS conditions: electrospray ionization (ESI) positive and negative. Quercetin mono-glucoside (1) 11.7 min [*m*/*z* (−)-ESI 463; quercetin minus glucose (*m*/*z* (−)-ESI 301)], quercetin mono-glucoside (2) 12.6 min [*m*/*z* (−)-ESI 463; quercetin minus glucose (*m*/*z* (−)-ESI 301)]; scoulerine mono-glucoside (1) 6.9 min [*m*/*z* (+)-ESI 490, (−)-ESI 488; scoulerine minus glucose (*m*/*z* (+)-ESI 328, (−)-ESI 326)], scoulerine mono-glucoside (2) 8.4 min [*m*/*z* (+)-ESI 490, (−)-ESI 488; scoulerine minus glucose (*m*/*z* (+)-ESI 328, (−)-ESI 326)]. Further identifications were done by NMR or EI-mass spectrometry. ^{1}H and ^{13}C NMR spectra were obtained using either a Bruker (Karlsruhe, Germany) ARX 400 or AC 300 instrument. Solvents used were D_2O or DMSO-d_6. MS spectra were measured on a Finnigan MAT (Bremen, Germany) type MAT 44 S instrument at 70 eV.

Arbutin. ^{1}H NMR (DMSO-d_6/D_2O, 400 MHz): δ (ppm) 6.86 (2H, d, J = 9.2 Hz, H-2, H-6), 6.65 (2H, d, J = 9.2 Hz, H-3, H-5), 4.63 (1H, d, J = 7.6 Hz, H-1′), 3,67 (1H, dd, J = 1.8, 12.0 Hz, H-6A), 3.44 (1H, dd, J = 5.9, 12.0, H-6B), 3.23–3.08 (4H, m, H-2′-H-5′). ^{13}C NMR (100 MHz, DMSO-d_6): δ (ppm) 152.2 (C-1), 150.3 (C-4), 117.7 (C-2, C-6), 115.6 (C-3, C-5), 101.8 (C-1′). 76.9 (C-3′), 76.6 (C-5′), 73.3 (C-2′), 69.8 (C-4′), 60.8 (C-6′). Pentaacetyl-arbutin; EI–MS, *m*/*z* (rel. int.%): 331 (6), 271 (3), 229 (5), 169 (82), 139 (10), 127 (32), 109 (100), 97 (19), 81 (19).

Vanillinglucoside. ^{1}H NMR (D_2O, 300 MHz): δ (ppm) 9.86 (1H, s, H-7); 7.69 (1H, dd, J = 1.8, 8.3 Hz, H-6), 7.6 (1H, s, H-2), 7.40 (1H, d, J = 8.3 Hz, H-5), 5.36 (1H, d, J = 7.3 Hz, H-1), 4.01 (3H, s, H-8), 4.00 (1H, dd, partly overlapped by H-8, J = 1.8, 12.2 Hz, H-6A′), 3.86 (1H, dd, J = 5.5, 12.2 Hz, H-6B′), 3,77–3.56 (4H, m, H-2′-H-5′). ^{13}C NMR (75 MHz, D_2O): δ [ppm] 195.7 (C-7), 152.1 (C-4), 149.9 (C-3), 132.0 (C-1), 127.6 (C-6), 115.8 (C-5), 112.4 (C-2), 100.5 (C-1′), 77.2 (C-3′), 76.3 (C-5′), 73.5 (C-2′), 70.1 (C-4′), 61.3 (C-6′), 56.8 (C-8). Tetraacetyl-vanillinglucoside; EI–MS, *m*/*z* (rel.int.%): 331 (7), 271 (4), 229 (2), 211 (3), 169 (93), 152 (15), 137 (2), 127 (30), 109 (100), 97 (14), 81 (12).

Endogenous substrate of AS

To trace indirectly the endogenous substrate of arbutin synthase, hydroquinone, cell cultures from *R. serpentina* were analyzed for the occurrence of arbutin. Fresh cells of 230 g wet weight were lyophylized resulting in 9 g dry material. The obtained powder was extracted over night at 100 rpm with different solvents [petrolether, 400 mL; ether/MeOH (1:1), 400 mL; MeOH, 400 mL]. The petrolether extract was rejected. The ether/MeOH and MeOH extracts were combined and freeze-dried after solvent evaporation. The residue (1 g) was dissolved in 100 mL water and purified over a XAD-column (80 mL volume). After elution with MeOH (300 mL) and concentration under vacuum, the solution was fractionated (12 mL per fraction) using a 250 × 10 mm LiChrospher 100 RP-18 column (Merck, Darmstadt, Germany) [linear gradient program starting with 2% MeOH/H_2O (8:2) and 98% H_2O (0.05 v/v trifluoroacetic acid) and ending after 30 min with MeOH/H_2O (8:2)]. To 250 μL of selected fractions 250 nkat almond-derived β-glucosidase (Sigma, Deisenhofen, Germany) in 250 μL citrate buffer (100 mM, pH 5.0) was added. After 4 h at 37 °C the reaction was stopped with 500 μL MeOH and centrifuged at 18,000*g* for 5 min. HPLC analysis of this solution did not show the presence of arbutin. The supernatant was freeze-dried and after dissolving of the residue in 200 μL water 200 pkat of arbutin synthase and 10 μL 25 mM UDPG were added and incubated for 4 h at 50 °C. After terminating the reaction with 500 μL MeOH and subsequent centrifugation the supernatant was subjected to HPLC analysis. A peak at 3.1 min indicated the presence of arbutin. For further identification chromatography on two TLC plates (Silica gel 60 F_{254}, Merck, Darmstadt, Germany) was performed; solvent system: EtOAc/$CHCl_3$/HCOOH/H_2O — 60:19:12:9. The resulting spot (R_f = 0.44) on the plate localized by UV (254 nm) had the same R_f value as arbutin. The spot was eluted with 20 μL CH_2Cl_2/MeOH (1:1) and the solvent evaporated. The residue was dissolved in 50 μL H_2O, one fraction was incubated with 83 pkat β-glucosidase in 150 μL of the above citrate buffer for 1 h and analyzed by the above described TLC system, R_f = 0.88 (identical to hydroquinone reference). Another part was acetylated with pyridine/acetic anhydride (1:1) after treatment with β-glucosidase, dried and subjected to MS-analysis: Diacetyl-hydroquinone; EI–MS, *m*/*z* (rel.int.%): 194 (4), 152 (19), 110 (100), 81 (10). A sample of authentic arbutin and the eluate of the spot

(R_f = 0.44, described above) were also acetylated and measured by MS. The obtained spectra were identical; EI–MS, *m*/*z* (rel. int.%): 331 (6), 271 (3), 229 (5), 169 (82), 139 (10), 127 (32), 109 (100), 97 (19), 81 (19).

Acknowledgements

Financial support provided by the Deutsche Forschungsgemeinschaft (Bonn-Bad Godesberg) and by the Fonds der Chemischen Industrie (Frankfurt/Main, Germany) is very much appreciated. We also thank Mark Smith (Dept. of Chemical Engineering, Cornell University) for verifying the English version of the manuscript and Christiane Hinse, Irina Gerasimenko (Mainz, Germany) for spectroscopic analyses.

References and Notes

1. Brabban, A. D.; Edwards, C. *J. Appl. Bacteriol.* **1995**, *79*, 171.
2. Vetter, J. *Toxicon* **2000**, *38*, 11.
3. Loomis, W. D.; Battaile, J. *Phytochemistry* **1966**, *5*, 423.
4. Vogt, T.; Jones, P. *Trends Plant Sci.* **2000**, *5*, 380.
5. Jones, P.; Vogt, T. *Planta* **2001**, *213*, 164.
6. Arend, J.; Warzecha, H.; Stöckigt, J. *Phytochemistry* **2000**, *53*, 187.
7. Arend, J.; Warzecha, H.; Hefner, T.; Stöckigt, J. *Biotechnol. Bioeng.* **2001**, *76*, 126.
8. Kapitonov, D.; Yu, R. K. *Glycobiology* **1999**, *9*, 961.
9. Akiu, S.; Suzuki, Y.; Asahara, T.; Fujinuma, Y.; Fukuda, M. *Nippon Hifuka Gakkai Zasshi* **1991**, *101*, 609; *Chem. Abstr.* **1991**, *115* 126962b.
10. Leshchinskaya, I. B.; Shakirov, E. V.; Itskovitch, E. L.; Balaban, N. P.; Mardanova, A. M.; Sharipova, M. R.; Viryasov, M. B.; Rudenskaya, G. N.; Stepanov, V. M. *FEBS Lett.* **1997**, *404*, 241.
11. Han, Y.; Wilson, D. B.; Lei, X. G. *Appl. Environ. Microbiol.* **1999**, *65*, 1915.
12. Marty, F. *Plant Cell* **1999**, *11*, 587.
13. Paczkowski, C.; Kalinowska, M.; Wojciechowski, Z. A. *Acta Biochim. Pol.* **1997**, *44*, 43.
14. Ford, C. M.; Boss, P. K.; Hoj, P. B. *J. Biol. Chem.* **1998**, *273*, 9224.
15. Durren, R. L.; McIntosh, C. A. *Phytochemistry* **1999**, *52*, 793.
16. Parry, A. D.; Edwards, R. *Phytochemistry* **1994**, *37*, 655.
17. Anhalt, S.; Weissenböck, G *Planta* **1992**, *187*, 83.
18. Schulz, M.; Weissenböck, G. *Phytochemistry* **1988**, *27*, 1261.
19. Mizukami, H.; Terao, T.; Ohashi, H. *Planta Med.* **1985**, *51*, 104.
20. Pridham, J. B.; Saltmarsh, M. J. *Biochem. J.* **1963**, *87*, 218.
21. Pridham, J. B. *Annu. Rev. Plant. Physiol.* **1965**, *16*, 13.
22. Vogt, T.; Zimmermann, E.; Grimm, R.; Meyer, M.; Strack, D. *Planta* **1997**, *203*, 349.
23. Frick, S.; Kutchan, T. M. *Plant. J.* **1999**, *17*, 329.
24. Dogru, E.; Warzecha, H.; Seibel, F.; Haebel, S.; Lottspeich, F.; Stöckigt, J. *Eur. J. Biochem.* **2000**, *267*, 1397.
25. Frick, S.; Ounaroon, A.; Kutchan, T. M. *Phytochemistry* **2001**, *56*, 1.
26. Milkowski, C.; Baumert, A.; Strack, D. *FEBS Lett.* **2000**, *486*, 183.
27. Taguchi, G.; Imura, H.; Maeda, Y.; Kodaira, R.; Hayashida, N.; Shimosaka, M.; Okazaki, M. *Plant Sci.* **2000**, *157*, 105.
28. Gross, G. G. *Phytochemistry* **1983**, *22*, 2179.
29. Jackson, R. G.; Lim, E. K.; Li, Y.; Kowalczyk, M.; Sandberg, G.; Hoggett, J.; Ashford, D. A.; Bowles, D. J. *J. Biol. Chem.* **2001**, *276*, 4350.
30. Barton, W. A.; Lesniak, J.; Biggins, J. B.; Jeffrey, P. D.; Jiang, J.; Rajashankar, K. R.; Thorson, J. S.; Nikolov, D. B. *Nat. Struct. Biol.* **2001**, *8*, 545.
31. McIntosh, C. A.; Latchinian, L.; Mansell, R. L. *Arch. Biochem. Biophys.* **1990**, *282*, 50.
32. Vogt, T.; Grimm, R.; Strack, D. *Plant J.* **1999**, *19*, 509.
33. Hirotani, M.; Kuroda, R.; Suzuki, H.; Yoshikawa, T. *Planta* **2000**, *210*, 1006.
34. Marcinek, H.; Weyler, W.; Deus-Neumann, B.; Zenk, M. H. *Phytochemistry* **2000**, *53*, 201.
35. Pinna, L. A. *Biochim. Biophys. Acta* **1990**, *1054*, 267.
36. Hughes, J.; Hughes, M. A. *DNA Seq.* **1994**, *5*, 41.
37. Bradford, M. M. *Anal. Biochem.* **1976**, *72*, 248.