ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

Structures of Alkaloid Biosynthetic Glucosidases Decode Substrate Specificity

Liqun Xia,[†,#] Martin Ruppert,[‡,#] Meitian Wang,[*,§] Santosh Panjikar,[*,∥] Haili Lin,[†] Chitra Rajendran,[§] Leif Barleben,[⊥] and Joachim Stöckigt*[,†]

[†]Institute of Materia Medica, College of Pharmaceutical Sciences, Zhejiang University, 866 Yu Hang Tang Road, Hangzhou 310058, P.R. China

[‡]Lehrstuhl für Pharmazeutische Biologie, Institut für Pharmazie und Biochemie, Johannes Gutenberg-Universität, Staudinger Weg 5, D-55099 Mainz, Germany

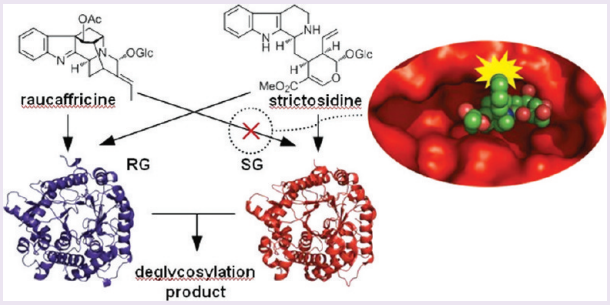
[§]Swiss Light Source, Paul Scherrer Institute, CH-5232 Villigen, Switzerland

[∥]Australian Synchrotron, 800 Blackburn Road, Clayton VIC, Australia 3168

[⊥]Department of Medical Biochemistry and Biophysics, Karolinska Institute, SE-17177 Stockholm, Sweden

Supporting Information

ABSTRACT: Two similar enzymes with different biosynthetic function in one species have evolved to catalyze two distinct reactions. X-ray structures of both enzymes help reveal their most important differences. The *Rauvolfia* alkaloid biosynthetic network harbors two *O*-glucosidases: raucaffricine glucosidase (RG), which hydrolyses raucaffricine to an intermediate downstream in the ajmaline pathway, and strictosidine glucosidase (SG), which operates upstream. RG converts strictosidine, the substrate of SG, but SG does not accept raucaffricine. Now elucidation of crystal structures of RG, inactive RG-E186Q mutant, and its complexes with ligands dihydro-raucaffricine and secologanin reveals that it is the "wider gate" of RG that allows strictosidine to enter the catalytic site, whereas the "slot-like" entrance of SG prohibits access by raucaffricine. Trp392 in RG and Trp388 in SG control the gate shape and acceptance of substrates. Ser390 directs the conformation of Trp392. 3D structures, supported by site-directed mutations and kinetic data of RG and SG, provide a structural and catalytic explanation of substrate specificity and deeper insights into *O*-glucosidase chemistry.



Alkaloids represent one of the most interesting groups of plant-derived natural products due, in part, to the enormous diversity they display in carbon frameworks. They also possess therapeutic value for the treatment of human diseases.[1] Many efforts have been made over the past decades to unravel the complex biosynthetic formation of single alkaloids at the enzyme level in order to fully comprehend their multistep pathways and their biosynthetic networks. To date, this objective has been fully satisfied for only a very few alkaloids: the isoquinoline morphine;[2] the diterpene alkaloid taxol;[3,4] and the monoterpenoid indole alkaloids, ajmalicine and ajmaline.[5] To date, characterization of the largest alkaloidal networks based on isolated enzymes has been achieved in *Papaver*, *Taxus*, and *Rauvolfia*.

The biosynthetic enzymes of the alkaloid ajmaline belong to a variety of enzyme families including synthases, esterases, glucosidases, reductases, oxidases, and transferases. It is only since 2004, through structural biology techniques leading to the successful crystallization and 3D X-ray analysis, that a much more thorough understanding of the major enzymes participating in the *Rauvolfia* ajmaline biosynthetic network has been gained (Figure 1). The enzyme strictosidine synthase (STR1, EC 4.3.3.2) initiates ajmaline biosynthesis by catalyzing the Pictet–Spengler reaction between the biosynthetic precursors, tryptamine and the monoterpene secologanin. It is this step that is central to the biosynthesis of about 2000 plant-derived monoterpenoid indole alkaloids.[6] Because of its fundamental importance, STR1 has become the most extensively characterized enzyme of the network, with much being known about its 3-dimensional structure and mechanism.[7−9]

A second enzyme, polyneuridine aldehyde esterase (PNAE, EC 3.1.1.78), shows exceptionally high substrate specificity. The structural and mechanistic characteristics of PNAE have also been recently determined.[10] While this enzyme belongs to the well-known α/β hydrolase superfamily, it also functions as the "enzymatic gate" to the biosynthesis of C9 monoterpenoid indole alkaloids from their C10 progenitor, polyneuridine aldehyde (PNA).

The PNAE product epi-vellosimine serves as the substrate for the second synthase of the metabolic network, vinorine

Received: July 28, 2011
Accepted: October 17, 2011
Published: October 17, 2011

 © 2011 American Chemical Society

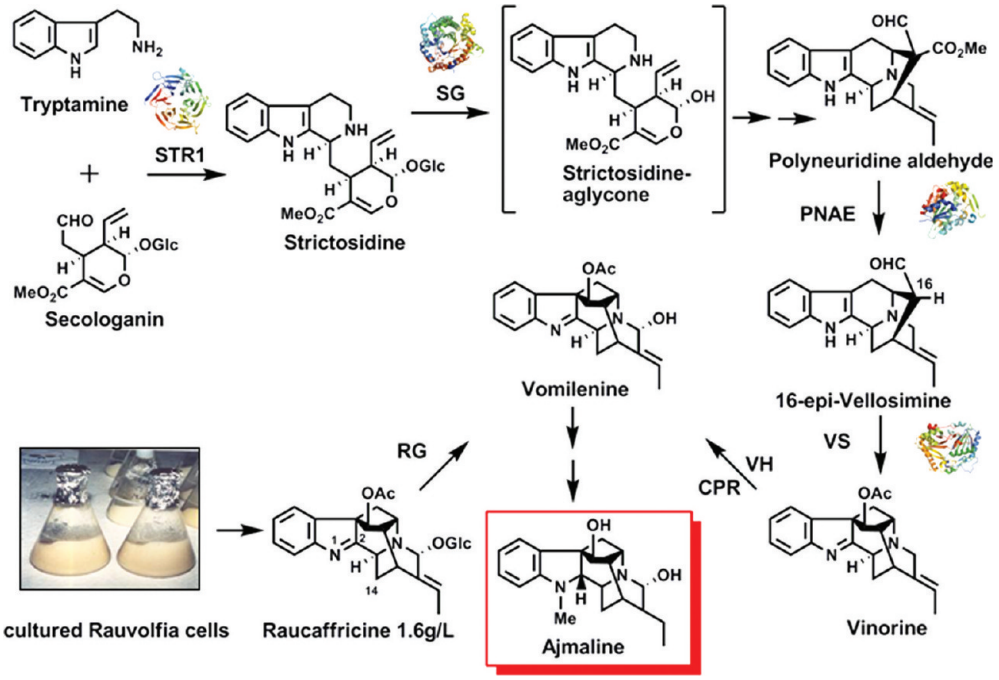


Figure 1. Enzymatic biosynthesis of monoterpenoid indole alkaloids in cell suspension cultures of the Indian medicinal plant *Rauvolfia*. Single steps were elucidated by isolation, characterization and partial sequencing of individual enzymes, cloning and expression of corresponding cDNAs by "reverse genetics" followed by crystallization and 3D X-ray analysis of the major enzymes. (STR1, strictosidine synthase; SG, strictosidine glucosidase; PNAE, polyneuridine aldehyde esterase; VS, vinorine synthase; VH, vinorine hydroxylase; CPR, cytochrome P 450 reductase; RG, raucaffricine glucosidase.).

synthase (VS, EC 2.3.1.160), whose reaction product, vinorine, already exhibits the complete carbon skeleton of the target alkaloid, ajmaline. Elucidation of the crystal structure of VS has not only provided significant insight into its binding pocket and catalytic center but also represented the first 3D example of the small BAHD enzyme family.[11,12] This superfamily is particularly important for the biosynthesis of several naturally derived therapeutics, such as pain killers (morphine), anticancer drugs (taxol and vinblastine/vincristine), and antiarrhythmia drugs (ajmaline), as well as a number of nonalkaloidal plant products.

Two additional *O*-β-D-glucosidases are present in the *Rauvolfia* alkaloid proteome, but they exhibit quite different functions and substrate specificities. The first, strictosidine *O*-β-D-glucosidase (SG, EC 3.2.1.105),[13] acts at the beginning of the ajmaline route by deglucosylation of the glucoalkaloid strictosidine. The second glucosidase, raucaffricine *O*-β-D-glucosidase (RG, EC 3.2.1.125),[14−16] hydrolyzes the glucoalkaloid raucaffricine forming the aglycone vomilenine, an intermediate that appears in the middle of the ajmaline pathway. RG can also hydrolyze strictosidine, the substrate of SG, but in sharp contrast, SG does not accept raucaffricine. The structural data presented here provide an explanation of the different functional behavior of each of these enzymes.

The X-ray structure of SG and its complex with the substrate strictosidine has recently been elucidated.[17] The completion of the 3D structure of wild-type RG and its ligand complexes together with site-directed mutagenesis studies allow a much clearer understanding of the catalyzed reaction. These structures enable comparisons of the structural and functional relationships of both glucosidases. The extensive, high quality 3D X-ray data reported here provide rare structural examples of two different but closely related enzymes of plant secondary metabolism that in principle catalyze the same reaction type in the same biosynthetic network and originate in the same plant but show remarkable differences in their substrate recognition and function (Figure 1).

RESULTS AND DISCUSSION

Metabolic Function and Significance of RG. The β-glucosidase RG likely performs a special metabolic function within the *Rauvolfia* alkaloid metabolic network. The hydrolysis of its substrate, raucaffricine, is only one reaction step away from the major biosynthetic route to ajmaline (Figure 1). Soluble RG directly delivers the key ajmalicine pathway intermediate, vomilenine. This aglycone is also the substrate for the raucaffricine-synthesizing vomilenine glucosyltransferase, which is responsible for the accumulation of raucaffricine. Concentrations of this glucoalkaloid can reach 1.6 g L^{-1} of cultured *Rauvolfia* cells, which is the highest amount of a monoterpenoid indole alkaloid ever observed in plant cell suspension cultures. This accumulation exceeds the highest concentration found in plants by >60-fold.[15]

RG-catalyzed reutilization of raucaffricine could be a crucial metabolic step in ajmaline biosynthesis. Raucaffricine accumulation could be important for future production in cell culture systems by blocking the raucaffricine-forming enzyme or by improving the RG-catalyzed reaction *in vivo*. Furthermore, RG is a promising candidate for the generation of libraries of novel structurally related alkaloids through chemo-enzymatic approaches starting with raucaffricine (Supplementary Figure S1). Both objectives require a better insight into the structure of the enzyme and the mechanism of the reaction that it catalyzes. Hence comparison of RG and SG at the structural level presents an interesting challenge from a mechanistic point of view and may offer significantly more detailed insight into higher plant β-glucosidases.

dx.doi.org/10.1021/cb200267w | *ACS Chem. Biol.* 2012, 7, 226−234

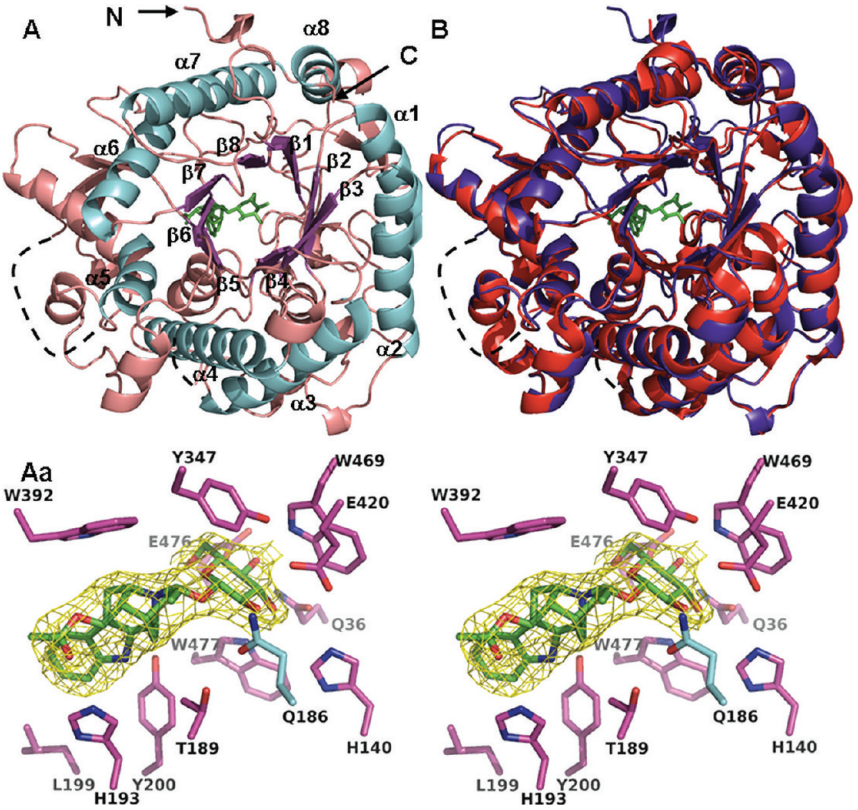


Figure 2. Structural comparison of the overall architecture of RG substrate complex and SG structure. (A) Ligand complex of RG-E186Q mutant with dihydro-raucaffricine (DHR) in stick presentation is illustrated in green and the N- and C-termini are marked. The figure resembles the $(\beta/\alpha)_8$ barrel fold, the binding site, and the secondary structure elements. The α-helices, β-sheets of the barrel, and loops are in cyan, magenta, and salmon, respectively. Loops with missing density are marked as "- - - - -". (Aa) Stereoview of raucaffricine binding site of RG. The $2F_o - F_c$ SIGMAA-weighted electron density of DHR contoured at 1.0 σ is shown in yellow and DHR in green. In complex with the inactive mutant (E186Q), residues within 4.0 Å distance from DHR are shown in magenta, and Gln186 is in cyan. Supplementary Figure S3 better displays the aglycone part of DHR following rotation of Aa by about 90°. (B) RG-DHR complex superimposed with the overall structure of SG (in red) indicating the high degree of structural similarity of both glucosidases. Both structures superimposed with a rms deviation of <0.52 Å.

Table 1. Structural Alignment of Plant β-Glucosidases of the GH-1 Family[a]

protein name	EC no.	plant	PDB[chain D]	rmsd	Cα
raucaffricine β-glucosidase	3.2.1.125	*Rauvolfia serpentina*	4A3Y[A]	0	472
strictosidine β-glucosidase	3.2.1.105	*Rauvolfia serpentina*	2JF7[A,B]	0.513	423
β-glucosidase 2 (cyanogenic)	3.2.1.21	*Trifolium repens*	1CBG[A]	0.711	426
β-glucosidase (Os04g0474800;Os4bglu12)	3.2.1.21	*Oryza sativa Japonica Group*	3PTK[A,B]	0.600	425
β-glucosidase (Bglu1;Os03g0703000)	3.2.1.21	*Oryza sativa Japonica Group*	3F5L[A,B]	0.701	417
β-glucosidase (Os03g0212800;Os3bglu6)	3.2.1.21	*Oryza sativa Japonica Group*	3GNR[A]	0.673	408
zeatin β-glucosidase (p60.1;Zm-p60.1)	3.2.1.21	*Zea mays*	1HXJ[A,B]	0.850	409
β-glucosidase 1 (Glu1)	3.2.1.21	*Zea mays*	1V08[A,B]	0.737	406
cyanogenic β-glucosidase (dhurrinase 1)	3.2.1.21	*Sorghum bicolor P721N*	1V02[A,B,C,D,F]	0.713	408
β-glucosidase (TaGlu1b)	3.2.1.21	*Triticum aestivum*	2DGA[A]	0.764	408
β-glucosidase (ScGlu)	3.2.1.21	*Secale cereale*	3AIU[A]	0.760	396

[a]3D structures of all of these *O*-β-glucosidases were elucidated. They are from plant sources and belong to the GH-1 family; myrosinases (*S*-β-glucosidase) are excluded. Total compared backbone atoms are 472. The rmsd values are obtained by PYMOL program (July 6, 2011 in CAZy).

Overall Structure of RG. Extensive investigations into plant-derived glucosidases of the GH-1 family, to which RG and SG belong, have been published over the past decade. The occurrence of both of these enzymes in plant cells of *Rauvolfia* and their closely related properties make their detailed comparison interesting from both functional and structural perspectives. Their shared physical, biochemical, and molecular properties (Supplementary Table S1) are proof of a tight link between SG and RG. This similarity is now also supported by the newly obtained 3-dimensional structure of RG. The refined crystal structure of RG consists of 13 α-helices and 13 β-strands. The overall structure of the wild-type RG enzyme possesses the expected TIM $(\beta/\alpha)_8$ barrel fold belonging to GH1 (http://www.cazy.org[18,19]), as seen earlier for SG.[17] RG now represents the 11th (excluding myrosinase) structural example from a total of ~380 identified GH1 members of plant origin and the 32nd example from all structurally elucidated GH1 proteins to date. The $(\beta/\alpha)_8$ barrel of RG hosts binding sites for the natural substrate, raucaffricine, and its derivative, dihydro-raucaffricine (DHR). The groove leading to the

dx.doi.org/10.1021/cb200267w|*ACS Chem. Biol.* 2012, 7, 226–234

catalytic center is formed by irregular loops between the secondary structures located on the surface of the enzyme. Similarly to SG, RG contains the catalytic residues Glu186 and Glu420. The most striking difference is at the catalytic center of RG and SG. The identical positions of RG-Trp392 and SG-Trp388 but different conformations of the tryptophan side chains in the two glucosidases most likely help each protein to recognize its natural substrate (see next subtitle, paragraph 4). Figure 2A and B illustrate the overall fold of the inactive mutant RG-E186Q in complex with its substrate dihydro-raucaffricine (see Figure 2Aa and Supplementary Figure S3 for the electron density) and its superimposition with the fold of the SG-E207Q mutant. These figures illustrate the high degree of structural similarity between the enzymes.

Sequence-based alignment and a phylogenetic tree of the 11 structurally characterized GH1 enzymes (Supplementary Figure S2A and B) displays their phylogenetic distribution as well as the conserved residues during the evolution process. Further comparison of the structure of RG together with SG and other β-glucosidases by structural alignment demonstrates, on the one hand, the excellent correspondence of RG with SG by indicating smaller rmsd values (0.52 Å) for the 423 superimposed backbone carbons and, on the other hand, nearly identical but higher rmsd values (from 0.6−0.9 Å) for the backbone carbons ranging from 396 to 426 for other plant glucosidases (Table 1).

Substrate Recognition and Active Site Relationship of RG and SG. Although RG and SG are structurally closely related, surprisingly their substrate specificity differs considerably. Whereas in addition to its natural substrate raucaffricine RG also hydrolyses the glucoalkaloid strictosidine (rel activity ∼1.2%), SG does not exhibit any measurable conversion of raucaffricine (Table 2). The reason for this high specificity has not been known but can now be explained by rigorous structural comparison of the overall binding pockets, their 3D ligand complexes, and especially the shape of the entrance to the active site of both enzymes.

As is the case for SG and for other β-glucosidases,[17,20] the gate to the active site should host the binding region of the aglycone part of the substrate. This part is always solvent-exposed, whereas the glucose moiety extends into and is mainly localized deep within the active center. As illustrated in Figure 3A this assumption is clearly supported by the 3D structure of the complex of RG with its ligand dihydro-raucaffricine (DHR). The aglycone binding site is believed to be responsible for the substrate specificity of β-glucosidases (and also for catalysis) as shown by structural and computational analysis.[20−23] The shape and space of the entrance of RG and SG differ remarkably, providing insight into their substrate recognition. Whereas RG leaves significant space for the ligand DHR (Figure 3A), there is much less space in the slot-like entrance of SG (Figure 3B) in which the ligand strictosidine shows a "tighter" fit. Moreover, the aglycones of both ligands occupy different positions, particularly when compared to the appropriate tryptophans 392 in RG and 388 in SG. The parallel arrangement of Trp388 and the aglycone of strictosidine in SG will fix the substrate strictosidine more tightly by stronger hydrophobic, sandwich-like interactions compared to RG, its Trp392, and the substrate raucaffricine, which are present in more of a perpendicular orientation. The above findings help to explain the variability in substrate affinity of RG and SG, whereby wild-type RG has significantly weaker binding of its natural substrate raucaffricine, with a K_m of ∼1.3 mM, compared to the approximately 10-fold higher affinity of SG for its substrate strictosidine (K_m = 0.12 mM) and ∼15-fold lower for RG toward strictosidine (K_m = 1.8 mM) compared to SG (K_m = 0.12 mM). A similar result (∼7-fold lower RG affinity) was obtained with the His_6-tagged RG and SG (Supplementary Table S2).

Table 2. Relative Enzyme Activity of RG Wild-Type and RG Mutants[a]

substrates	raucaffricine (%) (dihydro-raucaffricine %)	strictosidine (%)	secologanin (%)	arbutin (%)
RG-WT[b]	100 (98[c])	1.18	0.24	nd
RG-W392A	1.21	2.92	0.04	
RG-H193A	43.66	1.72		
RG-Y200A	4.60	0.02		
RG-T189A	63.94	0.77		
RG-F485Y	24.24	0.08		
RG-S390G	30.91	5.37		
RG-E476L	0.62	nd		
RG-E476A	0.57	nd		
RG-E186D	0.1			
RG-E186Q	0.1			
RG-E420Q	0.5			
RG-E186Q/E420Q	nd			
SG-WT	nd	53.01		

[a]Comparison of the relative enzyme activity of RG-WT (0.27 mM min^{-1} μg^{-1}, decrease measured of substrate concentrations under the conditions described in Supplemental Experimental Procedures, activity set at 100%), and mutants for putative substrates raucaffricine, 1,2-dihydro-raucaffricine (DHR), strictosidine, secologanin, and arbutin; detection limit 0.02%; nd = not detectable; blank = not determined. [b]WT, wild-type enzymes with N-terminal His_6-tag. [c]Measured with RG isolated from *R. serpentina* cell suspension culture.

If Trp392 is important in allowing RG to distinguish between the substrates raucaffricine and strictosidine, mutation of this residue to a smaller amino acid should result in a mutant with increased SG activity. Indeed, the RG-W392A mutant exhibited a large decrease in relative activity against the substrate raucaffricine from 100% for the wild-type to about 1.2% for the mutant, probably due to loss of hydrophobic interaction by Trp392. However, when activity for the substrate strictosidine was measured, a 2.5-fold increase from ∼1.2% to ∼2.9% was observed for the mutant, indicating that creation of additional space for strictosidine in RG by the substitution described above changes substrate acceptance in favor of the natural substrate of SG (Table 2). This finding is also supported by the catalytic activity k_{cat} and k_{cat}/K_m of the RG-W392A mutant, which was about 3-times higher for strictosidine compared to raucaffricine (Supplementary Table S2), a result emphasizing even more so the prominent structural but also functional role of this particular Trp392 residue.

Except in SG, the conformation of Trp392, which we called plant glucosidase-typical conformation,[17] is the same in all

dx.doi.org/10.1021/cb200267w|ACS Chem. Biol. 2012, 7, 226−234

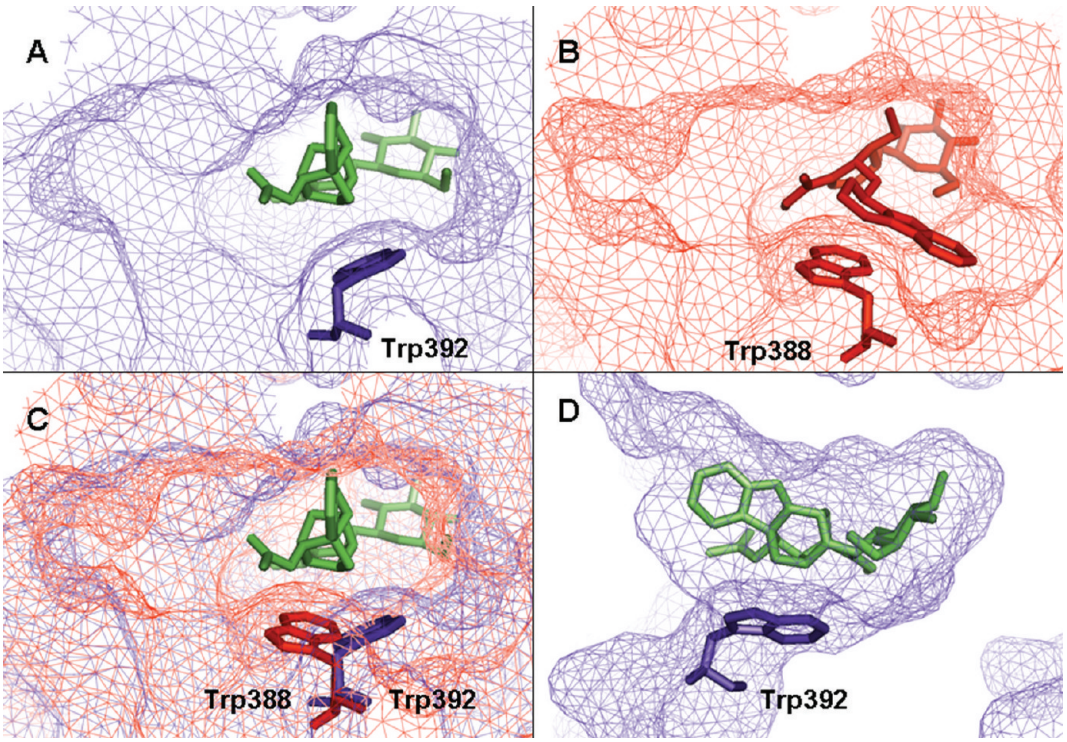


Figure 3. Surface of the cavity representing the binding pockets of RG-E186Q (blue) and SG-E207Q (red). Ligand complexes with their substrates DHR and strictosidine (STR), respectively, in stick presentation. (A) The whole binding pocket of RG-E186Q with particular emphasis on the shape of the entrance to the binding site and the aglycone part of DHR (green), which is somewhat sandwiched with Trp392. (B) The different shape of the slot-like entrance of SG-E207Q mutant compared to panel A and the ligand STR in sandwich-like $\pi-\pi$ interaction with Trp388. (C) Superimposed meshes of entrances of the RG-DHR ligand complex (blue) with that of SG (substrate STR not shown), which illustrates that the RG substrate can move easily through the RG entrance but is hindered with regards to the SG entrance (red). In addition, different conformations of both Trp392 of RG and Trp388 of SG, respectively, are shown. (D) The whole binding pocket of the RG-DHR complex after rotation of A around the X axis by 90°. Trp392 of RG is also illustrated (in blue, stick representation) at the bottom of the figure.

structurally characterized plant-derived β-glucosidases. This conformation is obviously due to interactions with second sphere residues Ser390 in RG and Asn or Thr in the other enzymes (Supplementary Figure S2). An early publication put forward the idea that the conformation of the conserved Trp may vary depending on the nature of neighboring amino acid, *e.g.*, of Tyr473 in the enzyme Glu1 (Figure 4G and Table 1) because it will form a hydrogen bond with the nitrogen of that Trp.[20] Tyr is replaced by Phe in RG (Phe485) and 7 other enzymes, while SG with its variant Trp conformation still features Tyr (Tyr481). This observation therefore indicates that at least for 8 of the 11 GH-1 enzymes Tyr is not essential for plant glucosidase-typical conformation of Trp. In contrast, in SG the Ser residue is replaced by Gly386, which is unable to form a hydrogen bond with Trp388, resulting in the SG-typical conformation of Trp388.[17] We thus suggest that the crucial conformation of this particular Trp residue is influenced more by the neighboring Ser (Asn or Thr) than by Tyr. The corresponding mutations in RG support this suggestion, for which the relative activity of RG-S390G (SG-Gly386) for strictosidine increased ~4.5 times compared to that of RG-WT, while the relative activity of RG-F485Y (SG-Tyr481) for strictosidine decreased. Replacement of Ser390 by Gly leads to a more flexible conformation of Trp392, whereas Phe485 mutated to Tyr results in a more fixed Trp392 conformation, in turn resulting in a less productive accommodation of strictosidine.

When the mesh of the RG complex with DHR was superimposed with the mesh for SG (Figure 3C), it became obvious that for steric reasons alone DHR can relatively easily pass through the entrance into the RG binding pocket but cannot enter the SG pocket. This structural result is in agreement with both observations that (i) raucaffricine is not converted by SG and (ii) SG activity is not inhibited by raucaffricine (data not shown), since raucaffricine cannot access the active site of SG. These results provide strong evidence that the wider entrance of the binding pocket of RG contributes to the lower substrate specificity compared to that of SG. This structural arrangement is important in the modulation of glucosidase specificity.[24] This result also points to the significant structural role played by residues Trp392 and Trp388. The shape of the entrance to the catalytic pocket determines which glycoalkaloid will enter and be converted.

Aglycone Binding Site. The RG aglycone binding site is basically very similar to that of other β-glucosidases, since interactions are nearly all of a hydrophobic nature. Surprisingly, the amino acids surrounding the aglycone are not identical to those in SG, although both enzymes are specific for the conversion of indole alkaloid glucosides. The conformation of Trp392 in relation to Trp388 has already been discussed above. When complexes of RG and its substrate and SG and its substrate, respectively, are superimposed, only a few residues are found to be identical. Tyr347 of RG occupies the same position as Tyr345 in SG. The same situation is observed for Thr189 and Thr210, but their substrate binding is different. Whereas Thr189 interacts with C-14 of DHR, the second Thr210 of SG interacts with the exocyclic double bond of the secologanin part of strictosidine. It is evident that both

dx.doi.org/10.1021/cb200267w | *ACS Chem. Biol.* 2012, 7, 226–234

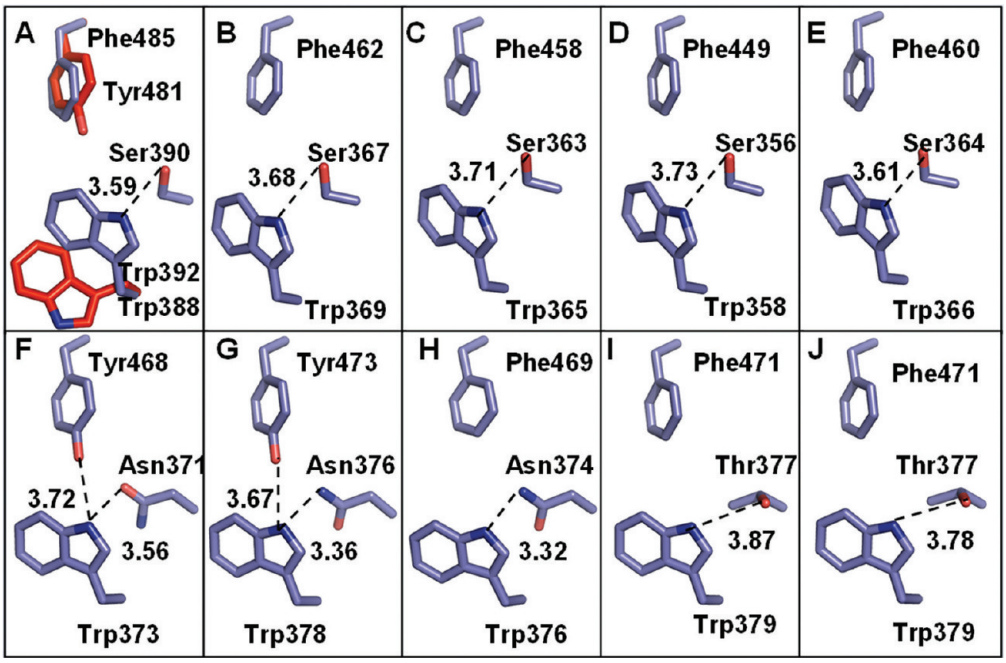


Figure 4. Conformation of conserved Trp and its neighboring residues in GH1 enzymes of higher plants. (A) The superimposed residues of RG (purple, PDB 4A3Y) and SG (red, PDB 2JF7). The side chains of the two Trp point to the opposite direction. Panels B–J represent other plant GH1 enzymes with plant glucosidase-typical Trp conformation (see Table 1 and Supplementary Figure S2A); their PDB codes are 1CBG, 3PTK, 3F5L, 3GNR, 1HXJ, 1V08, 1V02, 2DGA, and 3AIU, respectively. The residues were marked according to their position in each enzyme. In panel A–E and H–J, Phe is >4.1 Å from Trp to form strong interaction. It is the hydrogen binding of Trp to the neighboring residues (≤3.9 Å distance) that illustrates the essential role of Ser/Asn/Thr for directing the conformation of Trp.

threonines help to fix the aglycone and indirectly keep the glycone moiety in a favorable position for hydrolysis. In agreement with this observation, following site-directed mutation of Thr189 into Ala, the mutant shows only 64% activity toward raucaffricine compared to that of RG-WT (Table 2). Leu199 in RG and Gly386 of SG are completely different residues; however, it appears that they have the same function of shielding the 10 and 11 positions of the indole moiety of the corresponding substrate by the Leu199 side chain and by the loop region in which Gly386 is located, respectively.

A similar case of two different residues is observed for Tyr200 in RG and its counterpart Phe220 in SG. While they overlap perfectly in both complexes, binding of the former is likely to be more efficient due to the presence of two interactions: one sandwich-like hydrophobic interaction with the indole part, and the second being the H-bonding to the indole nitrogen (Figure 5A). Further assays showed that the hydrolytic activity of the RG-Y200A mutant decreases dramatically for both substrates, proving the significant role of the two interactions of Tyr200.

The biggest difference between both ligand complexes is represented by His193 in RG, a residue that, in fact, is replaced by Asn in SG and exceeds 4 Å distance to the substrate strictosidine. It has three binding opportunities represented by fixing the acetyl group of DHR and keeping the indole part fixed by steric, sandwich-like interaction and a hydrogen bridge with the indole nitrogen as illustrated in Figure 5A (see also Figure 2Aa and Supplementary Figure S5). Consequently, the RG-H193A mutant shows only low transformation of raucaffricine (44%). The relative activity toward strictosidine increases slightly, probably because of the absence of the interactions. It is clear that the indole nitrogen of the two substrates have different orientations when overlapping with the raucaffricine complex of RG and the strictosidine complex of SG (Figure 5Bb). As a result, strictosidine can interact more freely with this RG mutant.

In conclusion, although both glucosidases are strongly related, their aglycone binding sites are very different, particularly when compared with published structures of other β-glucosidases, indicating that the aglycone must have a significant influence on the enzyme activity.

Insight into the Catalytic Site of RG. Mechanistic insights into glycosidase chemistry have been reviewed several times in the past.[25,26] In order to illustrate the hydrolytic mechanism of RG, the structure in Figure 3A was rotated, allowing much improved visualization of the pocket for the glucosidic part (Figure 3D). Interactions of amino acid residues <4.0 Å in distance are two-dimensionally illustrated in Figure 5A. Two glutamic acids constitute the catalytic residues as in all structures of β-glucosidases of the GH-1 family. In RG, these residues are represented by Glu186 and Glu420, as verified by the corresponding mutants E186D, E186Q, E420Q, and the double mutant E186Q/E420Q, in which hydrolase activity is ≤0.5% of RG-WT (Table 2). Glu186 and Glu420 are separated by a distance of ∼5.1 Å between their carboxylate carbons, which provides the space for reacting with the glucosidic bond. This situation is typical for enzymes of the GH1 retaining the β-configuration of the anomeric carbon of glucose during catalysis. Whereas in SG a total of 11 hydrophilic residues interact with the glucose part of its substrate strictosidine,[17] in RG there appear to be only seven residues responsible for keeping the glucose in the catalytically desired position, all being identical with those in the SG complex. Glu476 is one of the seven residues and participates in an important hydrophilic interaction with both O6 and O4 of the glycone part of the substrate (Figure 5). The activity of mutants RG-E476L and

dx.doi.org/10.1021/cb200267w|ACS Chem. Biol. 2012, 7, 226–234

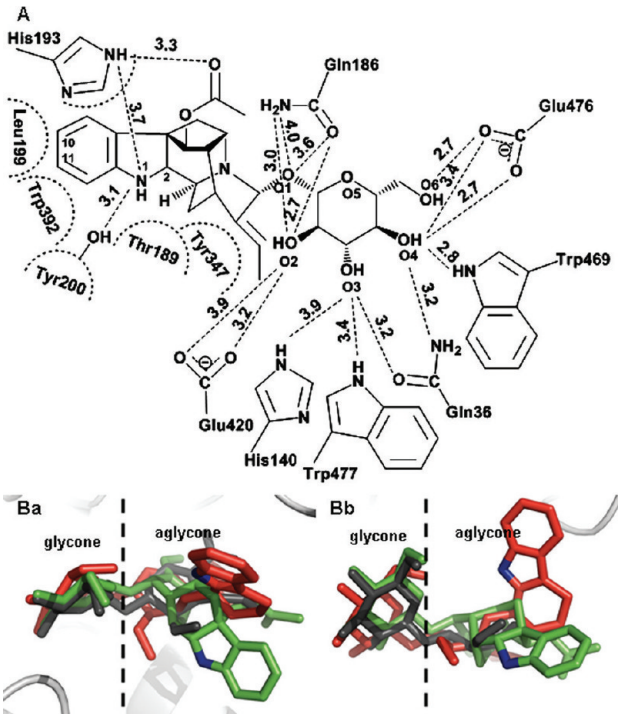


Figure 5. Detail of the binding mode of different substrates for RG and SG. (A) Hydrogen binding network between 1,2-(*S*)-dihydro-raucaffricine and residues in a distance of ≤4.0 Å in the ligand structure of His_6-RG inactive mutant E186Q, illustrating the interactions of amino acids with both the aglycone and the glucosidic part of the substrate. (Ba) Superimposed glucose part of RG complexes with DHR (green), SG with strictosidine (red), and RG with secologanin (black), displaying the identical position of the glucose moiety in the RG complexes and slightly different binding in the SG complex. (Bb) Panel Ba is rotated in order to better display the different direction of indole nitrogen (blue) and the slight shift in the glucose part.

RG-E476A dropped to less than 1% compared to the wild-type, clearly demonstrating the importance of Glu476's hydrophilic interaction with the substrate. The remaining four residues, though conserved in SG (except Tyr481, which is replaced in RG by Phe485, which cannot form a hydrogen bridge) and in other *β*-glucosidases (Supplementary Figure S2A), seem not to be essential for binding (>4.0 Å) the glucose unit of the substrate of RG. This might be due to the fact that the glucose unit has slightly shifted (∼1.5 Å) in the RG complex (compared to the SG complex), indicating a higher structural flexibility within the active site. This shift has been observed several times in different crystals and data sets of RG complexes, which might be based on the bigger pocket size and/or the different aglycone structures of RG and SG substrates. As previously discussed, the aglycone may indeed have an influence on the binding properties of the glucose in the catalytic center.[27] The superimposed ligand complexes of RG and SG (Figure 5Ba and Bb) disclose the structural differences in the catalytic pocket of both enzymes.

A more detailed inspection and comparison of the binding sites of RG and SG wild-type and their substrate complexes has also revealed several additional structural differences. In the RG and SG structures, the side chains of Glu476 and Glu472, respectively, can occupy different conformations (Supplementary Figure S4). The distance between the carboxy carbons in the two positions is about 2.4 Å. In RG wild-type and in RG/DHR and secologanin ligand complexes each Glu476 side chain points away from the oxygen of its backbone peptide bond (C-terminal). The same observation was made for the SG wild-type structure. Different conformations of Glu476 seem to be rare in other *β*-glucosidase structures, with the only exception being the zeatin *β*-glucosidase from *Zea mays* (PDB 1HXJ) and a rice glucosidase (PDB 3GNO[28]). In the SG complex with strictosidine the Glu472 points to the oxygen of that peptide bond, which is the N-terminal direction (Supplementary Figure S4). Independently from these observations, the binding distance of Glu476 and Glu472 residues to oxygens O4 and O6 of the glucose moiety is in all cases very similar. The results, therefore, suggest that the positions of the appropriate Glu side chain in *β*-glucosidases differ; however, this difference appears not to influence the interactions with the glucose part very much, since distances to O4 and O6 are nearly the same.

When all three ligand complexes, RG with DHR and secologanin and SG with strictosidine, were compared by superimposition (Figure 5B), it was found that the binding site for the glucose unit is identical in the first two cases. This observation is in contrast to the rate of catalytic conversion, which is the highest for the substrate raucaffricine and its derivative DHR (100% and 98% rel activity, respectively) but much lower for the glucoside secologanin, showing only 0.24% relative activity (Table 2). The finding again supports the influence of the structure of the aglycone on enzyme conversion rates.

Our results indicate the importance of the nature of the aglycone for glycosidase action. The catalytic site of each enzyme has to adapt not only to the structure of the glycosyl residue but also to that of the aglycone, which can occur in highly diverse and unexpected ways. Using structural differences in order to understand differences in substrate acceptance by closely related members of the GH1 is exceptionally complexed, as has been recently discussed.[28]

In sharp contrast, comparison of RG and SG substrate complexes has now provided a clear structural explanation of their distinct substrate specificities and the basis for molecular recognition of their intrinsic ligands.

METHODS

RG Crystallization and Preparation of Its Complexes with Its Substrates. The crystallization of wild-type RG was described by Ruppert, et al.[16] Crystallization of the inactive mutant RG-E186Q and preparation of its complex with the substrate 1,2-(*S*)-dihydro-raucaffricine and secologanin were carried out with the optimal crystallization conditions found to be in the range of 0.01−0.3 M ammonium sulfate, 0.1 M sodium acetate, pH 4.5−5.0, 9−12% (w/v) PEG 4000 as precipitant buffer, and 20 °C for incubation temperature. RG complexes with raucaffricine showed always insufficient density of the ligand in contrast to 1,2-dihydro-raucaffricine, which exhibits nearly identical conformation. After stable crystals were formed (≥10 d), freeze-dried ligands 1,2-(*S*)-dihydro-raucaffricine and secologanin were added directly to the drops that contained appropriate crystals. These soaking experiments were carried out for 10 min to 3 h. Thereafter crystals were measured at RT.

Structure Determination of Wild-Type RG, Model Building and Refinement. Prior to X-ray data collection, the crystals were treated with cryoprotectant [10 mM calcium acetate, 10 mM Tris pH 7.5, 7% DMSO, 20% (v/v) glycerol] and flash-cooled to100 K. Diffraction data were then collected on the X11 beamline of EMBL-Hamburg, Germany; 720 images were measured with 0.5° rotation per frame using aMAR CCD detector at wavelength 0.8051 Å. The data set statistics are summarized in Supplementary Table S3.

The 3D-structure of wild-type RG was determined by molecular replacement. The applied software pipeline was AUTO-RICKSHAW.[29] Packing considerations[30] suggested that the content of one asymmetric unit of RG crystals was two monomers. The initial models were solved using the program MOLREP[31] with the structure of homologous SG (PDB 2JF7) as a target model. The output model was subjected to iterative refinement using CNS[32] and REFMAC5.[33] A random set containing 2.4% of the total data was excluded from the refinement, and the agreement between calculated and observed structure factors corresponding to these reflections (R_{free}) was taken to follow the progress of the refinement.[34] The model phase was subjected to bias reduction and density modification using the program PIRATE and continued with automated model building program ARP/wARP.[35]

At this stage, ARP/wARP generated 2mfo-Dfc maps[36] were used for careful examination of the maps and allowed corrections and incorporation of missing residues into the model. The graphics program COOT[37] was used for rebuilding of the model. Manual building alternated with additional REFMAC cycles, which included bulk solvent correction, anisotropic scaling, and with each molecule defined as a TLS group in the modeling of anisotropy in the program REFMAC5.[38] Overall geometric quality of the model was assessed using PROCHECK.[39]

Structure Determination of RG-E186Q Complexes, Model Building and Refinement. All RT crystal mounting and data collection were carried out at Swiss Light Source (SLS) macromolecular crystallography beamline X06SA.

The MiTegen RT system[40] was used for the crystal mounting. Crystals were measured in continuous data acquisition mode with a noise-free and high dynamic range pixel-array detector - PILATUS 6 M.[4142] All room-temperature data were collected in continuous mode with 5 Hz frame rate. After mounting, crystals were quickly transferred for diffraction data collection in order to reduce possible dehydration. X-ray beam was defocused to about 100 μm × 100 μm to match the size of RG crystals, which are about 100 μm × 200 μm in general. Collection protocol for RG crystal is 0.5° oscillation per 0.2 s, 1 Å radiation, and attenuated flux of about 60 × 10^9 photons/s. All diffraction data were processed with automated data processing script GO.COM, which combines XDS robust indexing, integration, scaling, and merging processes[43,44] with high-performance computing facility at X06SA beamline. Typical data processing time for a data set of 270 images is about 1−2 min, which is proven to be very valuable for real-time data quality assessment. Structures were determined with molecular replacement package MOLREP using refined RG model from cryo-data as search model. Atomic coordinates and atomic displacement parameters were refined with REFMAC package. Manual and semiautomated model building of protein, solvent, and substrate were carried out with COOT. PROCHECK and MolProbity were used for structure validation. All figures were made with PyMOL.[45]

■ ASSOCIATED CONTENT

Supporting Information

This material is available free of charge *via* the Internet at http://pubs.acs.org.

Accession Codes

Coordinates and structure factors have been deposited in Protein Data Bank with the codes 4A3Y for RG-WT-glycerol complex, 3U5U for RG-E186Q, 3U57 for RG-E186Q-dihydroraucaffricine complex, and 3U5Y for RG-E186Q-secologanin complex.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: meitian.wang@psi.ch; panjikar.santosh@embl-hamburg.com; joesto2000@yahoo.com.

Author Contributions

#These authors contributed equally to this work.

■ ACKNOWLEDGMENTS

Staff members at the EMBL beamline at the DORIS storage ring (DESY, Hamburg, Germany), beamline X06SA at Paul Scherrer Institute (Villigen, Switzerland) and beamline BL17U1 at Shanghai Synchrotron Radiation Facility (China) are kindly appreciated for their help. We thank Deutsche Forschungsgemeinschaft (Bonn-Bad Godesberg, Germany), Fonds der Chemischen Industrie (Frankfurt/Main, Germany) and K.P.Chao Foundation (Zhejiang University, Hangzhou, China) for support. We also acknowledge H. Zou (Zhejiang University, Hangzhou, China) for samples of secologanin and D. E. Cane (Brown University, USA) for critical advices.

■ REFERENCES

(1) Nicolaou, K. C. Montagnon, T. (2008) *Molecules That Changed the World*, Chapter 10: Morphine, pp 67−78, Wiley-VCH, Weinheim.

(2) Zenk, M. H., and Juenger, M. (2007) Evolution and current status of the phytochemistry of nitrogenous compounds. *Phytochemistry 68*, 2757−2772.

(3) Chau, M., Jennewein, S., Walker, K., and Croteau, R. (2004) Taxol biosynthesis. *Chem. Biol. 11*, 663−672.

(4) Kaspera, R., and Croteau, R. (2006) Cytochrome P450 oxygenases of Taxol biosynthesis. *Phytochem. Rev. 5*, 433−444.

(5) Ruppert, M., Ma, X, and Stöckigt, J. (2005) Alkaloid biosynthesis in *Rauvolfia*−cDNA cloning of major enzymes of the ajmaline pathway. *Curr. Org. Chem. 9*, 1431−1444.

(6) Stöckigt, J., and Zenk, M. H. (1977) Strictosidine (isovincoside): Key intermediate in biosynthesis of monoterpenoid indole alkaloids. *J. Chem. Soc., Chem. Commun.*, 646−648.

(7) Ma, X., Panjikar, S., Koepke, J., Loris, E., and Stöckigt, J. (2006) The structure of *Rauvolfia serpentina* strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins. *Plant Cell 18*, 907−920.

(8) Maresh, J., Giddings, L., Friedrich, A., Loris, E., Panjikar, S., Trout, B., Stöckigt, J., Peters, B., and ÒConnor, S. (2008) Strictosidine synthase: Mechanism of a Pictet-Spengler catalyzing enzyme. *J. Am. Chem. Soc. 130*, 710−723.

(9) Stöckigt, J., Barleben, L., Panjikar, S., and Loris, E. (2008) 3D-Structure and function of strictosidine synthase−the key enzyme of monoterpenoid indole alkaloid biosynthesis. *Plant Physiol. Biochem. 46*, 340−355.

(10) Yang, L., Hill, M., Wang, M., Panjikar, S., and Stöckigt, J. (2009) Structural basis and enzymatic mechanism of the biosynthesis of C(9)-from C(10)-monoterpenoid indole alkaloids. *Angew. Chem., Int. Ed. 48*, 5211−5213.

(11) Ma, X., Koepke, J., Panjikar, S., Fritzsch, G., and Stöckigt, J. (2005) Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. *J. Biol. Chem. 280*, 13576−13583.

(12) ÒAuria, J. C. (2006) Acetyltransferases in plants: a good time to be BAHD. *Curr. Opin. Plant Biol. 9*, 331−340.

(13) Gerasimenko, I., Sheludko, Y., Ma, X., and Stöckigt, J. (2002) Heterologous expression of a *Rauvolfia* cDNA encoding strictosidine glucosidase, a biosynthetic key to over 2000 monoterpenoid indole alkaloids. *Eur. J. Biochem. 269*, 2204−2213.

(14) Warzecha, H., Obitz, P., and Stöckigt, J. (1999) Purification, partial amino acid sequence and structure of the product of raucaffricine-*O*-beta-D-glucosidase from plant cell cultures of *Rauvolfia serpentina*. *Phytochemistry 50*, 1099−1109.

(15) Warzecha, H., Gerasimenko, I., Kutchan, T. M., and Stöckigt, J. (2000) Molecular cloning and functional bacterial expression of a plant glucosidase specifically involved in alkaloid biosynthesis. *Phytochemistry 54*, 657−666.

(16) Ruppert, M., Panjikar, S., Barleben, L., and Stöckigt, J. (2006) Heterologous expression, purification, crystallization and preliminary X-ray analysis of raucaffricine glucosidase, a plant enzyme specifically involved in *Rauvolfia* alkaloid biosynthesis. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun. 62*, 257−260.

(17) Barleben, L., Panjikar, S., Ruppert, M., Koepke, J., and Stoeckigt, J. (2007) Molecular architecture of strictosidine glucosidase: the gateway to the biosynthesis of the monoterpenoid indole alkaloid family. *Plant Cell 19*, 2886−2897.
(18) Henrissat, B., and Davies, G. (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol. 7*, 637−644.
(19) Ketudat Cairns, J. R., and Esen, A. (2010) β-Glucosidases. *Cell. Mol. Life Sci. 67*, 3389−3405.
(20) Czjzek, M., Cicek, M., Zamboni, V., Bevan, D. R., Henrissat, B., and Esen, A. (2000) The mechanism of substrate (aglycone) specificity in β-glucosidases is revealed by crystal structures of mutant maize beta-glucosidases-DIMBOA, -DIMBOAGlc, and -dhurrin complexes. *Proc. Natl. Acad. Sci. U.S.A. 97*, 13555−13560.
(21) Dopitova, R., Mazura, P., Janda, L., Chaloupkova, R., Jerabek, P., Damborsky, J., Filipi, T., Kiran, N. S., and Brzobohaty, B. (2008) Functional analysis of the aglycone -binding site of the maize β-glucosidase Zm-p60.1. *FEBS J. 275*, 6123−6135.
(22) Hill, A. D., and Reilly, P. J. (2008) Computational analysis glycoside hydrolase family 1 specificities. *Biopolymers 89*, 1021−1031.
(23) Mendonca, L. M., and Marana, S. R. (2008) The role in the substrate specificity and catalysis of residues forming the substrate aglycone-binding site of a β-glucosidase. *FEBS J. 275*, 2536−2547.
(24) Sue, M., Nakamura, C., Miyamoto, T., and Yajima, S. (2011) Active-site architecture of benzoxazinone-glucoside β-D-glucosidases in Triticeae. *Plant Sci. 180*, 268−275.
(25) Vocadlo, D. J., and Davies, G. J. (2008) Mechanistic insights into glycosidase chemistry. *Curr. Opin. Chem. Biol. 12*, 539−555.
(26) Vasella, A., Davies, G. J., and Bohm, M. (2002) Glycosidase mechanisms. *Curr. Opin. Chem. Biol. 6*, 619−629.
(27) Verdoucq, L., Moriniere, J., Bevan, D. R., Esen, A., Vasella, A., Henrissat, B., and Czjzek, M. (2004) Structural determinants of substrate specificity in family 1 beta-glucosidases−Novel insights from the crystal structure of sorghum dhurrinase-1, a plant beta-glucosidase with strict specificity, in complex with its natural substrate. *J. Biol. Chem. 279*, 31796−31803.
(28) Seshadri, S., Akiyama, T., Opassiri, R., Kuaprasert, B., and Cairns, J. K. (2009) Structural and enzymatic characterization of Os3BGlu6, a rice beta-glucosidase hydrolyzing hydrophobic glycosides and (1→3)- and (1→2)-linked disaccharides. *Plant Physiol. 151*, 47−58.
(29) Panjikar, S., Parthasarthy, V., Lamzin, V. S., Weiss, M. S., and Tucker, P. A. (2005) Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 61*, 449−457.
(30) Matthews, B. W. (1968) Solvent content of protein crystals. *J. Mol. Biol. 33*, 491−497.
(31) Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr. 30*, 1022−1025.
(32) Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 905−921.
(33) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 53*, 240−255.
(34) Brünger, A. T. (1993) Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 49*, 24−36.
(35) Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol. 6*, 458−463.
(36) Read, R. J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr., Sect. A: Found. Crystallogr. 42*, 140−149.
(37) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2126−2132.
(38) Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 57*, 122−133.
(39) Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283−291.
(40) Kalinin, Y., Kmetko, J., Bartnik, A., Stewart, A., Gillilan, R., Lobkovsky, E., and Thorne, R. (2005) A new sample mounting technique for room-temperature macromolecular crystallography. *J. Appl. Crystallogr. 38*, 333−339.
(41) Henrich, B., Bergamaschi, A., Broennimann, C., Dinapoli, R., Eikenberry, E. F., Johnson, I., Kobas, M., Kraft, P., Mozzanica, A., and Schmitt, B. (2009) PILATUS: A single photon counting pixel detector for X-ray applications. *Nucl. Instrum. Methods Phys. Res., Sect. A 607*, 247−249.
(42) Broennimann, C., Eikenberry, E. F., Henrich, B., Horisberger, R., Huelsen, G., Pohl, E., Schmitt, B., Schulze-Briese, C., Suzuki, M., Tomizaki, T., Toyokawa, H., and Wagner, A. (2006) The PILATUS 1M detector. *J. Synchrotron Rad. 13*, 120−130.
(43) Kabsch, W. (2010) XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 125−132.
(44) Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 133−144.
(45) DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.