Functional Characterization of Genes Coding for Novel $\beta$-D-Glucosidases Involved in the Initial Step of Secoiridoid Glucosides Catabolism in Centaurium erythraea Rafn

Jelena Božunović, Milica Milutinović, Neda Anićić, Marijana Skorić, Dragana Matekalo*, Suzana Živković, Milan Dragićević, Biljana Filipović, Tijana Banjanac, Luka Petrović and Danijela Mišić*

Department of Plant Physiology, Institute for Biological Research "Siniša Stanković"- National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia

Secoiridoid glucosides (SGs) are monoterpenoids derived from the iridoid cyclopentane-C-pyran skeleton with $\beta$-D glucose linked at C1 position. Coordinated metabolic processes, such as biosynthesis and catabolism of SGs, ensure constitutive presence of these bitter tasting compounds in plant tissues, which plays a decisive role in the defense against pathogens and herbivores. These compounds are susceptible to hydrolysis mediated by enzymes $\beta$-glucosidases, and the resulting aglycones are subsequently directed toward different metabolic pathways in plants. Function of two $\beta$-D-glucosidases (named CeBGlu1 and CeBGlu2) from centaury (Centaurium erythraea Rafn; fam. Gentianaceae), belonging to the glycoside hydrolase 1 (GH1) family, was confirmed using in vitro assays with recombinant proteins, following their heterologous expression in E. coli and His-tag affinity purification. Although they show slightly differential substrate preference, both isoforms display high specificity toward SGs and the organ-specific distribution of transcripts was positively correlated with the content of SGs in diploid and tetraploid C. erythraea plants. Transient overexpression of CeBGlu1 and CeBGlu2 in C. erythraea leaves induced changes in metabolite profiles. The effectiveness of transgene overexpression has been altered by plant ploidy. UHPLC/DAD/(±)HESI – MS$^2$ profiling of leaves of diploid and tetraploid C. erythraea plants. Transient overexpression of CeBGlu1 and CeBGlu2 in C. erythraea leaves induced changes in metabolite profiles. The effectiveness of transgene overexpression has been altered by plant ploidy. Functional Characterization of Genes Coding for Novel $\beta$-D-Glucosidases Involved in the Initial Step of Secoiridoid Glucosides Catabolism in Centaurium erythraea Rafn. Front. Plant Sci. 13:914138. doi: 10.3389/fpls.2022.914138

Keywords: beta glucosidase, secoiridoid glucosides, functional characterization, agroinfiltration, Centaurium erythraea
INTRODUCTION

Secoiridoid glucosides (SGs) are a group of plant-derived natural compounds widely present in species belonging to orders Gentianales, Dipsacales, Corales, and Lamiales (family Oleaceae) (Ghisalberti, 1998; Jensen et al., 2002). Species of the genus Centaureum Hill (fam. Gentianaceae) are a rich source of SGs, among which swertioside (SW), swertiamarin (SWM), and gentiopicrin (GP) predominate. Leaves are the major site of SGs biosynthesis and accumulation in C. erythraea (Matekalo et al., 2018). Secoiridoids are named after secologanin (SEC), derived after the cleavage of the cyclopentane ring of iridoids (between C-7 and C-8). Secologanin is considered the common precursor of all secoiridoids in plant sources (Jensen and Schripsema, 2002), but is also the building block for monoterpene indole alkaloids (MIAs). Centaurium species are presumed to be a part of the secoiridoid biosynthetic pathway up to SEC with MIAs-producing Catharanthus roseus. The biosynthetic pathway of SEC starts with the synthesis of geranyl diphosphate (GPP), which is converted to geraniol and continues via a number of intermediates such as 8-hydroxygeraniol, 8-oxogeraniol, nepetalactol, iridoidal, 7-deoxylagonanic acid, 7-deoxyloganic acid, loganic acid and loganin (Figures 1A, B). The biosynthetic pathway involves a series of reactions, catalyzed by enzymes geranyl diphosphate synthase (GPPS), geraniol-8-oxidase (GBO), 8-hydroxynytagenol oxidoreductase (8HGO), iridoid synthase (IS), iridoid oxidase (IO), 7-deoxylagonanic acid glucosyltransferase (7DLGT), 7-deoxylagonanic acid hydroxide (7DLH), loganic acid O-methyltransferase (LAMT), and secologanin synthase (SLS). In C. roseus, SEC and tryptamine are precursors of monoterpene indole alkaloids (MIAs) such as strachotisidine, while in C. erythraea SEC is metabolized into secoiridoid glucosides SW, SWM and GP, as well as their derivatives. Enzymes responsible for the conversion of SEC to SW, and further to SWM and GP are not elucidated yet (Figure 1A).

The content of SGs in C. erythraea tissues is a net result of the two dynamic metabolic processes, their biosynthesis and their catabolism (Figure 1A). Catabolism of SGs starts with their deglycosylation catalyzed by β-glucosidases. Cleavage of the glucose (Glu) residue from SW, SWM, and GP results in the formation of aglycones (Figures 1A, B), which are further metabolized in plant tissues through isomerization, reduction and oxidation reactions. Gentripyric and erythrocentaurin are the common products of SWM and GP hydrolysis (Ishiguro et al., 1983; Zeng et al., 2013; Božunović et al., 2018), while SW is metabolized into naucledal and epinaucledal (Purdy and McLean, 1977; El-Sedawy et al., 1990) (Figure 1C). In order to better understand the catabolism of SGs in C. erythraea, we focused our attention on β-glucosidases (β-D-glucoside glucohydrolases, E.C. 3.2.1.21), categorized into the glycoside hydrolase family 1 (GH1), the largest GH family in plants. GH1 β-glucosidases are enzymes that hydrolyze glycosidic bonds to release non-reducing terminal glucosyl residues from various compounds – benzoazoxinoid, cyanogenic, iridoid and phenolic glucosides – as well as glucosinolates (Morant et al., 2008a). Genes coding for β-glucosidase (β-Glu) enzymes have been previously characterized in some members of the order Gentianales, including C. roseus and Rauwolfia serpentina (Geerlings et al., 2000; Warzecha et al., 2000; Barleben et al., 2007). Isolation and heterologous expression of C. roseus β-glucosidase with high specificity toward strictosidine which directly derives from SEC, provided valuable information on the function of this β-glucosidase involved in metabolic pathway of indole alkaloids (Geerlings et al., 2000). The aim of the present study was to isolate and functionally characterize C. erythraea β-glucosidase with high substrate specificity toward SGs using in vitro and in planta assays. Our presumption was that the action of C. erythraea β-glucosidase with high specificity toward SGs is essential for the catabolism of these compounds, and is indirectly related to defense against herbivores and pathogens. Secoiridoids in the form of glucosides possess remarkable antimicrobial effects (Božunović et al., 2018), and their hydrolysis/deglycosylation mediated by β-D-glucosidases release biologically active aglycones which provide more efficient antioxidant protection (Božunović et al., 2018). Moreover, during hydrolysis of SGs, glucose is also released, which may serve as an alternative source of energy under stress conditions.

MATERIALS AND METHODS

Plant Material

Plants used in experiments in vitro were obtained as previously described by Filipović et al. (2019). Briefly, mother stock shoot cultures of C. erythraea diploids and tetraploids originating from seeds collected at the locality Tjentiste (Suitjeska National Park, Bosnia and Herzegovina), in 2007 and 2016, respectively, were maintained in vitro on half-strength MS medium (1/2 MS, Murashige and Skoog, 1962) in 370 ml glass jars. Root cultures, established from root segments of 3-month-old diploid and tetraploid plants, were grown on solid 1/2 MS medium in Petri dishes. Spontaneously regenerated shoots, formed on root explants, were further transferred on fresh 1/2 MS medium in 370 ml glass jars for rooting. After 10 weeks of culturing, shoots and roots of the obtained regenerated plants were harvested and used in experiments to determine the expression patterns of CeBGlu candidates and their activities. All in vitro cultures were maintained at a temperature of 25 ± 2°C under fluorescent light of 47 µmol s⁻¹ m⁻² and a 16 h/8 h light/dark photoperiod.

Agroinfiltration experiments were performed under greenhouse conditions. C. erythraea seedlings were established in greenhouse, in pots filled with Floradur B seed substrate for multiplication (Floragard Vertriebs-GmbH, Oldenburg, Germany). Two-month-old seedlings were individually transferred into pots with Floragard growth medium (Floragard Vertriebs-GmbH, Oldenburg, Germany), and grown under greenhouse conditions at 50–85% humidity. Five-month-old diploid and tetraploid plants, displaying rosette phenotype, were used in agroinfiltration experiments. Seeds of tetraploid C. erythraea, collected in 2006, were obtained from Ecological-Botanical Garden of the University of Bayreuth (Germany). Diploid individuals were of the same origin as those used in in vitro experiments. All the C. erythraea accessions used in the present study are deposited within the seed collection at

Frontiers in Plant Science | www.frontiersin.org 2 June 2022 | Volume 13 | Article 914138
RNA Isolation and cDNA Synthesis

Total RNA from approximately 150 mg of Centaurium erythraea shoots and roots was isolated using modified CTAB method (Gasic et al., 2004). Isolated RNA was quantified using Qubit 3.0 Fluorometer (Thermo Fisher Scientific, United States), and its integrity was confirmed using gel electrophoresis. Prior to RT-PCR, isolated RNA was treated with DNase I (Thermo Fisher Scientific, United States) to deplete contaminating genomic DNA. First strand cDNA was synthesized from 300 µg RNA using the RevertAid Premium First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, United States).

Isolation and Cloning of β-Glucosidase Candidate Genes

Centaurium erythraea leaf transcriptome databases (Malkov and Simonović, 2011; Ćuković et al., 2020) was searched for homologs of Catharanthus roseus strictosidine β-glucosidase (Geerlings et al., 2000). After selecting the most promising candidate, gene specific primers were designed using Primer3Plus software1 in order to amplify the full length of the CeBGlu coding sequence. A list of primers used for cloning and expression analysis is given in Supplementary Table 1.

The full length amplification of the candidate gene was performed using Q5 Hot Start High-Fidelity DNA Polymerase (New England Biolabs, United States) and gene specific forward and reverse primers (Supplementary Table 1) following cycling conditions: one cycle of 98°C for 3 min, 35 cycles of 98°C for 30 s, 62°C for 2 min, and 72°C for 5 min followed by a final extension of 72°C for 10 min in a thermal cycler (Eppendorf, Austria). The amplicon was separated by 1% agarose gel electrophoresis and then purified by GeneJET Gel Extraction Kit (Thermo Fisher Scientific, United States) according to manufacturer’s instructions. The purified amplicon was cloned into pTZ57R/T cloning vector using T/A PCR product cloning kit (Thermo Fisher Scientific, United States) in summary, 20 µl of purified PCR product was mixed with 3 µl of pTZ57R plasmid vector (600 ng), 6 µl of 5× ligation buffer, 1 µl of T4 DNA Ligase (5 U µl⁻¹), in 1.5 ml microtube. A total of 30 µl reaction mixture

---

1http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi
was incubated for 1 h at room temperature to let the ligation reaction take place.

For transformation, 5 µl of the ligation product was added into 100 µl of Mach1 E. coli competent cells, incubated on ice for 20 min and then heat-shocked for 45 s in water bath at 42°C. The mixtures were immediately placed on ice and subsequently cultivated with 250 µl liquid Luria-Bertani (LB) media followed by 1 h incubation at 37°C. The transformed cells were cultivated on LB agar plate containing ampicillin (100 µg/ml), which was followed by overnight incubation at 37°C. Positive transformants were verified by colony PCR with CeBGlu gene-specific primers (Supplementary Table 1). Cells harboring the recombinant plasmid were cultured in ampicillin containing liquid LB medium overnight at 37°C in a shaker incubator at 220 rpm. GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, United States) was used to purify plasmids from 2 ml of Mach1 E. coli overnight culture according to the manufacturer’s instructions. Recombinant plasmids were confirmed by restriction digestion using XhoI and KpnI restriction enzymes (Thermo Fisher Scientific, United States). Subsequent transformants have confirmed the isolation of two highly similar gene variants named CeBGlu1 and CeBGlu2.

Organ-Specific Profiling of CeBGlu Expression
For qRT-PCR, due to the very high similarity between CeBGlu1 and CeBGlu2 sequences, it was possible to design only one pair of primers common to both isoforms (Supplementary Table 1). For CeBGlu expression analysis, SYBR Green I (Maxima SYBR Green/ROX Kit, Thermo Scientific, United States) was used. Amplification was conducted in Light cycler QuantStudio 3 (Thermo Fisher Scientific, United States), according to the manufacturer’s instructions. General thermocycler conditions were 95°C for 10 min; 40 cycles of 95°C for 15 s; 60°C for 30 s; 72°C for 30 s and final extension at 72°C for 10 min. Expression levels of targeted genes were calculated according to the 2−ΔΔCt method (Livak and Schmittgen, 2001). EF1 gene expression was used as endogenous control to normalize all data (primer sequences are presented within Supplementary Table 1). Presented results are obtained from three biological replicates.

Organ-Specific Profiling of Total Hydrolytic Activity Against Secoiridoid Glycosides
Shoots and roots of diploid and tetraploid C. erythraea plants were ground to a fine powder using liquid nitrogen, and proteins were isolated in 100 mM potassium-phosphate buffer (pH 6.5) supplemented with phenylmethylsulfonyl fluoride (PMSF) and 5 mM ascorbate. Protein content was determined according to Bradford (1976) using bovine serum albumin as a standard.

Activity of protein extracts to reduce the content of secoiridoid glycosides was evaluated using standards of SW, SWM and GP as substrates. The reduction in quantity of the mentioned substrates is the net result of combined enzymatic activities present in the extract including glucosyl hydrolases which produce the respective aglycones and unknown biosynthetic enzymes involved in transformations of secoiridoids (Figure 1A).

All reference compounds were diluted in dH2O (w:v = 1:1) and kept as a stock solutions. Aliquots of 30 µl were vacuum evaporated and diluted in 300 µl 50 mM citrate phosphate buffer pH 5.5 containing 10 µg of protein extracts. Reactions were incubated for 24 h at 37°C. Subsequently, 700 µl of methanol was added, and samples were centrifuged at 10,000 g for 10 min. Supernatants were filtered through 0.2 µm cellulose filters and subsequently analyzed for the content of SW, SWM and GP.

Phylogenetic Analysis
For phylogenetic tree creation, multiple sequence alignment-ta was generated using the Muscle algorithm. A neighbor-joining tree was constructed using MEGA X, Version 10.2.6 (Kumar et al., 2018). Cluster stability was estimated with 1,000 bootstrap replicates. The evolutionary distances were computed using the Poisson correction method. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 766 positions in the final dataset. The data were converted into Newick format and transferred to Dendroscope (Huson and Scornavacca, 2012) for creating the final phylogenetic tree. Refer to Supplementary Table 2 for a list of plant protein sequences used for phylogenetic analysis and their corresponding accession numbers, including those of CeBGlu1 and CeBGlu2 functionally characterized within the present study.
an RMSD threshold of 0.5 Å so that only those conformations that are at least 0.5 Å RMSD away from all retained are kept. The geometry of the resulting conformers was optimized using Merck molecular force field (MMFF94s) as implemented in RDKit 2021.09.5 (Landrum et al., 2022). The resulting ligand conformations were docked into the predicted enzyme structure using AutoDock Vina 1.2.3 (Eberhardt et al., 2021). The docking box was defined based on the coordinates of the mero secologanin ligand in the experimental structure of the raucaffricine β-D glucosidase from Rauvolfia serpentina (PDB: 3U5Y, Xia et al., 2012) after superposition of the structure to the AlphaFold model of CeBGlu1. The docking box center was defined as \( x = -1.964, y = 0.374, z = -6.148 \), while the size of the box was 18, 18.75, and 21 Å in each direction, respectively. The docking was performed using autogrid4 (Morris et al., 2009) precalculated affinity maps, with autodock4 scoring function. Flexible docking was performed where the Glu476 side chain was allowed to change conformations. The exhaustiveness of the algorithm was set to 64. The highest estimated affinity poses, as reported by AutoDock Vina, were inspected and compared to the position of secologanin in 3U5Y, and the best pose based on these two criteria for each ligand was analyzed using UCSF ChimeraX 1.4rc (Pettersen et al., 2021) and protein–ligand interaction profiler (PLIP, Adasme et al., 2021).

**Accession Numbers**
Sequence data from this article can be found in the NCBI database under the following accession numbers: ON060690 (CeBGlu1); ON060691 (CeBGlu2).

**Heterologous Expression and Purification of 6His-Tagged Proteins**
For heterologous expression, CeBGlu1 and CeBGlu2 were cloned into bacterial expression vector pRSETA, yielding the final plasmids pRSETA:CeBGlu1 and pRSETA:CeBGlu2 (Figure 2A). Amplification of the gene sequences by PCR was performed using Phusion high-fidelity DNA polymerase (Thermo Fisher Scientific, United States) and a pair of CeBGlu1/CeBGlu2 gene specific primers with added Xhol/KpnI restriction sites (Supplementary Table 1). The constructs were heat-shock transformed, as described above, into Mach1 E. coli competent cells which were further cultured on LB medium containing ampicillin. Polymerase chain reaction using gene specific primers, Xhol/KpnI double restriction digestion and sequencing were applied to verify the constructs containing transgenes of interest.

pRSETA vector contains a N-terminal polyhistidine (6×His) tag, which can be used as an affinity ligand for protein purification. The pRSETA:CeBGlu1 and pRSETA:CeBGlu2 recombinant constructs were used for heat-shock transformation of BL21 (DE3) CodonPlus-RIL cells (Stratagene, United States). Cultures were grown overnight in 20 ml 2 × Yeast Extract Tryptone (2 × YT) media broth supplemented with ampicillin (100 µg ml\(^{-1}\)), 50 µg ml\(^{-1}\) kanamycin and 17 µg ml\(^{-1}\) chloramphenicol at 37°C with shaking at 220 rpm. On the next day 200 ml culture was initiated by inoculating above mentioned overnight cultures, with the initial OD600 set to 0.1. Cultures were incubated at 37°C with shaking at 230 rpm. After reaching an OD600 of 0.4–0.5, protein expression was induced using 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Subsequently, incubation continued for 4 h at 21°C. Following centrifugation for 10 min at 5,000 g and removal of supernatant, the cell pellets were harvested. All of the purification steps were performed at 4°C. For each purification batch, Ni-NTA agarose (Qiagen, United States) beads were equilibrated with equal volumes of lysis buffer (50 mM NaH\(_2\)PO\(_4\), 300 mM NaCl, 10 mM imidazole pH 8.0). Bacterial cells were collected by centrifugation and re-suspended in 500 µl of lysis buffer. Following lysozyme (Sigma Aldrich, Germany) addition (1 mg ml\(^{-1}\) final concentration), samples were incubated for 30 min on ice. After 10 min of centrifugation (12,000 g at 4°C) supernatants were loaded into columns containing Ni-NTA resin and incubated at 4°C using FALC F205 rotary tube mixer (Falc Instruments, Treviso, Italy). After 1 h, samples were spun at 12,000 g for 5 min at 4°C. Supernatant was discarded and resin was rinsed for 3 times using washing buffer (50 mM NaH\(_2\)PO\(_4\), 300 mM NaCl, 50 mM imidazole, pH 8.0). The recombinant protein was eluted with increased imidazole concentration in elution buffer (50 mM NaH\(_2\)PO\(_4\), 300 mM NaCl, 250 mM imidazole, pH 10.2).

Eluted fractions were analyzed using 5–10% SDS-PAGE using Mini-PROTEAN II Electrophoresis Cell (Bio-Rad, United States) followed by Coomassie blue staining and immunoblotting. Proteins were electro-transferred to PVDF membrane (Amersham Biosciences, Germany), using Mini Trans-Blot Module electric transfer system (Bio-Rad, United States). Transfer on the membrane was performed at 4°C for 90 min at a constant voltage of 60 V. Membrane was blocked with 10% (v/v) non-fat dry milk (NFDM; Nestle, United States) in phosphate-buffered saline containing 0.05% Tween-20 overnight at 4°C. The presence of 6×His labeled proteins was confirmed using His-probe antibody in 1:100 dilution (Acc. No. sc-53073, Santa Cruz Biotechnology, United States), which was followed by incubation with goat anti-mouse IgG-HPR (1:5,000, Agrisera Antibodies, Sweden). The bound antibodies were visualized by enhanced chemiluminescence (ECL). The membrane was incubated for 5 min at room temperature using ECL solution containing 100 mM Tris–HCl pH 8.5, 0.2 mM p-coumaric acid, 1.25 mM 3-aminophthalhydrazide and 1.7 µl of 30% H\(_2\)O\(_2\). Detection was performed by exposure to radiographic film (Medical X-ray Green/MXG Film, Carestream Health, United States) for 10 min.

**Confirmation of CeBGlu1 or CeBGlu2 Hydrolytic Activity Using in vitro Enzymatic Assays**
Beta-glucosidase activity of CeBGlu1 and CeBGlu2 was determined in a reaction with 20 µg of recombinant proteins and 10 mM 4-nitrophenyl β-D-glucopyranoside (pNPG) as a substrate, in 50 mM citric phosphate buffer pH 5.5 at 40°C for 48 h. The reaction was stopped by the addition of ice cold 1 M Na\(_2\)CO\(_3\) (1:1 = v:v) and colorimetric detection of p-nitrophenol,
the product of pNPG hydrolysis, was spectrophotometrically measured at 410 nm (Esen and Blanchard, 2000).

Aerial parts of *C. erythraea* (100 mg) were ground into homogeneous powder using liquid nitrogen and extracted with 1 ml of 99.8% methanol (AppliChem GmbH, Germany) by vortexing for 30 s and subsequent sonication for 10 min using an ultrasonic bath (RK100, Bandelin, Berlin, Germany). After centrifugation at 10,000 × g for 10 min supernatants were filtered using 0.2 μm syringe filters (Agilent Technologies, Santa Clara, CA, United States). *C. erythraea* methanol extract (10 μl) was evaporated in a Vacuum Rotor Evaporator (Eppendorf Concentrator 5301, Germany) at room temperature. Subsequently, dried extracts were dissolved in 300 μl 50 mM citrate phosphate buffer pH = 5.5 containing 20 μg of purified recombinant enzyme (*CeBGlu1* or *CeBGlu2*) and incubated for 48 h at 37°C. Following incubation, 700 μl of methanol was added, and samples were centrifuged at 10,000 g for 10 min. Control samples contained elution buffer instead of purified recombinant protein, and the final concentration of standards was 3 μg ml⁻¹. Supernatants were filtered through 0.2 μm cellulose filters and injected into UHPLC/DAD/([±]HESI)⁻MS² instrument.

**Construction of CeBGlu Expression Plasmids and Agrobacterium-Mediated Transformation**

*CeBGlu1* and *CeBGlu2* sequences were PCR-amplified from pTZ57R/T plasmids with the addition of *PacI*/*AvrII* restriction sites. Amplicons were ligated to transient expression vector pJL-TRBO (Lindbo, 2007), and positive colonies were identified by colony PCR using vector- and gene-specific primers. All PCR reactions were performed with Phusion high-fidelity DNA polymerase (Thermo Fisher Scientific, United States) and a pair of *CeBGlu1/CeBGlu2* gene specific primers (Supplementary Table 1). After heat-shock transformation
with pJL-TRBO:CeBGlu1 and pJL-TRBO:CeBGlu2 (Figure 2B), Mach1 E. coli competent cells were cultivated overnight in LB liquid medium containing kanamycin (50 μg ml\(^{-1}\)) at 37°C in a shaker incubator at 230 rpm (IKAS K0004C control, China). GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, United States) was used to purify plasmids from 5 ml of Mach1 E. coli cultures according to the manufacturer’s instructions. After plasmid isolation, each CeBGlu sequence was confirmed by sequencing.

The recombinant pJL-TRBO:CeBGlu1 and pJL-TRBO:CeBGlu2 plasmids extracted from Mach1 cells were transferred to A. tumefaciens GV3101 strain by electroporation using “Gene Pulser” (Bio-Rad, United States) and subsequently cultured in LB plates containing kanamycin (50 μg ml\(^{-1}\)), gentamycin (25 μg ml\(^{-1}\)) and rifampicin (10 μg ml\(^{-1}\)). A single colony of recombinant bacteria was inoculated into 5 ml liquid LB media containing antibiotics and incubated overnight at 28°C with shaking at 230 rpm. On the next day, 1 ml of bacterial suspension was sub-cultivated in 10 ml of liquid LB media containing 10 mM MES-KOH (pH 5.5) and 20 μM acetosyringone. Cultures were incubated overnight at 28°C with shaking at 230 rpm. The agrobacterial cells were harvested by centrifugation for 20 min at 3,000 g, and the pellet was resuspended in the infiltration medium (10 mM MES, 10 mM MgCl\(_2\) and 100 μM acetosyringone) to a final OD600 of 1.0. After an incubation at room temperature for 4 h, cultures were introduced into abaxial surface of leaves of five-month-old C. erythraea plantlets using a blunt tipped plastic syringe and applying gentle pressure. Additionally, to prevent the silencing of transgene expression in C. erythraea, pJL-TRBO:CeBGlu1 and pJL-TRBO:CeBGlu2 were co-infiltrated in a 1:1 ratio with pBIN vector expressing the p19 silencing-suppressor gene from TBSV (pBIN:p19). After agroinfiltration, plants continued to grow for 5 days in the greenhouse. Leaves were harvested from diploid and tetraploid C. erythraea plants and immediately frozen at liquid nitrogen. Samples were stored at −80°C until further use.

**Plant Methanol Extract Preparation**

Plant material (shoots and roots) of C. erythraea was manually ground in liquid nitrogen into fine powder and diluted in 96% methanol (w:v = 10:1). Following vortexing for 1 min, extraction was performed overnight at 4°C. The next day, extraction was continued for 20 min in an ultrasonic bath (RK100, Bandelin, Germany) maintained at room temperature. Samples were centrifuged at 8,000 g for 20 min and supernatants were filtered using 15 mm RC filters with 0.22 μm pore size (Agilent Technologies, United States). Samples were stored at 4°C until use. All extractions were performed in biological triplicates.

**UHPLC/DAD/(±)HESI-MS\(^2\) Quantification of Targeted β-D Glucosides**

Samples were analyzed using Dionex Ultimate 3000 UHPLC system (Thermo Fisher Scientific, Germany) equipped with a DAD detector and connected to a triple quadrupole mass spectrometer (TSQ Quantum Access MAX, Thermo Fisher Scientific, Switzerland). Samples were chromatographically separated on a Hypersil gold C18 column (50 × 2.1 mm) with 1.9 μm particle size (Thermo Fisher Scientific, United States) thermostated at 40°C. Mobile phase, consisting of water + 0.01% acetic acid (A) and MS grade acetonitrile (B), was eluted at flow rate of 0.4 ml min\(^{-1}\) according to Banjanac et al. (2017). Injection volume was set to 10 μl. DAD absorption was acquired at λ max = 260 and 320 nm. A triple quadrupole mass spectrometer with a heated electrospray ionization (HESI) was operated with a following parameters: vaporizer temperature 300°C, spray voltage 4,000 V, sheath gas (N\(_2\)) pressure 27 AU, ion sweep gas (N\(_2\)) pressure 1.0 AU and auxiliary gas (N\(_2\)) pressure at 10 AU, capillary temperature 275°C, skimmer offset 0 V. Argon was used as the collision gas in the collision-induced fragmentation of the selected reaction monitoring (SRM) mode of the instrument, and collision energies (ce) were set as shown in Supplementary Table 3. Calibration curves of targeted compounds showed excellent linearity with correlation coefficients r = 0.999, p < 0.001. Total concentrations of targeted secoiridoids were obtained by calculating their peak areas, and were expressed as μg per 100 mg of plant fresh weight (μg 100 mg\(^{-1}\) FW). Xcalibur software (version 2.2) was used for the instrument control, data acquisition and analysis. All analyses were performed using three biological replicates.

**Statistical Analysis**

Statistical significance was determined by using Minitab Statistical Software (Minitab, State College, PA, United States). For statistical analysis of relative gene expression and compound quantification in shoots and roots of diploid and tetraploid plants one-way ANOVA was performed followed by Fisher LSD test (p < 0.05). For comparison of hydrolytic activities, as well as for comparing SGs content in leaves after agroinfiltration with CeBGlu1 and CeBGlu2 cloned into pJL-TRBO expression vector (without or with silencing inhibitor p19) Student’s t-tests were used for data analysis (p < 0.05).

**RESULTS AND DISCUSSION**

Specific hydrolytic enzymes activate many glycosylated compounds in plants (e.g., glucosinolates, alkaloids, benzoxazinoids, cyanogenic, and (seco)iridoid glucosides), which often defines their dual-defense system against herbivores. Highly active and unstable aglycones released from iridoid and secoiridoid glycosides and monoterpenoid indole alkaloids usually display more prominent biological activities toward herbivores and pathogens, by adversely affecting their enzymatic machinery. Aglycones react with nucleophilic side chains of amino acids to form covalent protein complexes (Bartholomaeus and Ahokas, 1995; Konno et al., 1999; Kim et al., 2000; Guirimand et al., 2010), and act as unspecific enzyme inhibitors (Bartholomaeus and Ahokas, 1995; Ling et al., 2003). The typical defense compounds of Plantaginaceae are iridoid glycosides, which retard growth and/or enhance mortality of non-adapted herbivores. As a part of the dual defense system, Plantago lanceolata and P. major possess β-glucosidases that hydrolyze aucubin, one of the two major iridoid glycosides in these species.
and thereby release protein-denaturing aglycones (Pankoke et al., 2013, 2015). Oleaceae species like *Ligustrum obtusifolium* and *Olea europaea* are rich sources of SGs that after tissue disruption are metabolized by endogenous plant β-glucosidases (Konno et al., 1999; Mazzuca et al., 2006). Olearopein β-glucosidase (OelGlu) from *O. europaea* is reported to have a defensive role in young organs and meristem tissues as well as in mature tissues, where it activates oleuropein into a potent protein cross-linking agent during de-compartmentalization caused by pests (Koudounas et al., 2015). Olearopein β-glucosidase (OelGlu) is highlighted as a molecular target of high biotechnology interest to regulate qualitative and quantitative content of bioactive secoiridoids in olive oils, and thus their organoleptic properties (Koudounas et al., 2021). Likewise, in MIAs-rich *C. roseus* and *R. serpentina* (fam. Apocynaceae), strictosidine is hydrolyzed by β-glucosidases (Guirimand et al., 2010). Following the disruption of strictosidine and strictosidine β-D-glucosidase (SGD) compartmentalization in *C. roseus*, initiated by cellular breakup after tissue wounding, the highly reactive strictosidine aglycone with prominent feeding-deterrent/toxic properties is released, and further conducted toward the production of cytotoxic MIAs (Guirimand et al., 2010, 2011).

Secoiridoid glucosides of *C. erythraea* are defense compounds constitutively present in tissues that, due to their extremely bitter taste, pre-ingestively deter feeding and thus reduce consumption rates of various non-adapted herbivores. The presumption of the present study was that *C. erythraea* might also possess β-glucosidases displaying high specificity toward SGs. By activating aglycones of SEC, SW, SWM, and GP, these hydrolytic enzymes could be involved in initial steps of their catabolism. Furthermore, resulting aglycones might offer multilevel protective features to plants, and, as previously shown, shape their antioxidant properties (Božunović et al., 2018).

**Isolation of Full Length of CeBGlu, Comparison With Homologs, and Phylogenetic Analysis**

By exploring transcriptomic resources of *C. erythraea* leaves, we identified a candidate for β-D-glucosidase (CeBGlu) displaying high homology with strictosidine β-D-glucosidase previously characterized from *C. roseus* (Geerlings et al., 2000). Following PCR, cloning of amplified product into *pTZ57R/T* vector and sequencing, two highly similar CeBGlu candidates (CeBGlu1 and CeBGlu2) were revealed. Full lengths of isolated CeBGlu1 and CeBGlu2 comprised open reading frames of 1,659 bp encoding polypeptides of 552 amino acids with calculated molecular masses of 62.06 kDa. The two CeBGlu gene candidates shared a high level of similarity (99.3%) and differed at 12 nucleotide sites, i.e., 4 amino acids (Figure 3 and Supplementary Table 4). Based on primary structure both enzymes were predicted to be localized in the lysosome/vacuole.

Amino acid sequences of CeBGlu1 and CeBGlu2 were aligned with previously investigated β-D-glucosidase amino acid sequences of plant species belonging to the Gentianales order (Figure 3): strictosidine β-glucosidase from *Gelsemium sempervirens* (GsSGD) (Franke et al., 2019), raucaffrinicine-O-β-D-glucosidase (RsRG) (Warzecha et al., 2000) and strictosidine β-D-glucosidase (RsSGR1) from *Rauvolfia serpentina* (Xia et al., 2012), strictosidine β-D-glucosidase from *R. verticillata* (RvSGD) (Chen et al., 2008), strictosidine β-D-glucosidase from *C. roseus* (CrSTR) (Geerlings et al., 2000), and beta-glucosidase from *Carapaichea ipecacuanha* (Citpeglu1) (Nomura et al., 2008). The highest amino acid sequence identity for CeBGlu1 and CeBGlu2 was observed with GsSGD (69.2 and 69.8%, respectively) (Supplementary Table 4).

Phylogenetic tree was constructed incorporating β-D-glucosidase amino acid sequences of diverse plant species to narrow the prediction of the function of both CeBGlu1 and CeBGlu2 (Figure 4). The neighbor-joining tree grouped the enzymes into several clusters based on their role in plants: defense response, lignification, hormone deglycosylation, β-mannosidase or myrosinase activities. Defense-related enzymes included a separate cluster that belonged to monocots, and four clusters belonging to dicots, having different types of substrates: cyanogenic glucosides, isoflavonoid conjugates, alkaloid glucosides, and terpenoid glucosides. Interestingly, even though their substrates belong to the group of terpenoids, the two *C. erythraea* candidates (CeBGlu1 and CeBGlu2) showed the highest homology with β-D-glucosidases of the order Gentianales that prefer alkaloid glucosides as substrates (GsSGD, RsRG, RsSGR1, RvSGD, CrSTR, and Citpeglu1) (Figure 4).

**Comparative Analysis of Organ-Specific Secoiridoid Glucosides Content and CeBGlu Expression and Activity**

Plant glucosidase genes are developmentally regulated (Morant et al., 2008a), and exhibit different spatial expression patterns depending on their physiological functions. In this sense, we analyzed the organ-specific distribution of transcript levels and activities of the two *C. erythraea* β-glucosidase candidates, in parallel with the content of major SGs, in both diploid and tetraploid genotypes. The presence of CeBGlu1 and CeBGlu2 transcripts in *C. erythraea* was validated by amplifying their specific fragments using qRT-PCR. Due to the high sequence similarity between the two gene candidates, a combination of primers common to both transcripts were employed (Figure 5A). CeBGlu was amplified at low level when cDNA from roots was used as a template in qRT-PCR analysis. Expression analysis revealed that the transcription of CeBGlu is regulated in an organ-specific manner (Figure 5A). The CeBGlu expression pattern revealed significantly higher transcript levels in shoots compared to roots of *C. erythraea*, in both diploid and tetraploid genotypes. However, tetraploid plants displayed significantly higher CeBGlu transcript levels in shoots and roots compared to diploids.

In parallel, content of major secoiridoids (SW, SWM, and GP) was significantly higher in shoots than in roots of both diploid and tetraploid *C. erythraea* genotypes (Figure 5B). It is well documented that aerial parts are the major site of SGs biosynthesis and accumulation in common centaury (Şiler et al., 2012, 2014; Matekalo et al., 2018; Božunović et al., 2019; Filipović et al., 2019). Similarly, as in our previous study (Božunović et al., 2019), the dominant SG in leaves of diploid and tetraploid plants was SW, followed by GP and SWM. Diploid plants displayed slightly higher GP content in shoots than tetraploid ones, which
FIGURE 3 | Multiple sequence alignment of CeBGlu1 and CeBGlu2 with β-glucosidases from species belonging to the Gentianales order. The two Glu catalytic residues are colored blue, active site residues involved in sugar specificity are colored olive, and residues differing between CeBGlu1 and CeBGlu2 are indicated in red. The secondary structure diagram shown below the alignment is based on the CeBGlu1 AlphaFold model. The pLDDT score (predicted lDDT-Cα) shown below the alignment is a per-residue estimate of AlphaFold confidence on a scale from 0 to 100. Lower values are often associated with disordered regions. It is shown for CeBGlu1 AlphaFold model. RsRG sequence corresponds to the PDB:3U5Y glucosidase from Rauvolfia serpentina.

is in accordance with our previous study (Filipović et al., 2019), and higher content of SW in roots compared to tetraploids.

To test the total hydrolytic activity of C. erythraea organs, pure SW, SWM and GP were subjected to an in vitro enzymatic assay using crude protein extract of shoots or roots, and the decrease in SG content was analyzed using UHPLC/DAD/(±)HESI-MS² analysis. SWM and GP were efficiently hydrolyzed, while no significant decrease in SW content was recorded (Figure 5C).

Shoots of both diploid and tetraploid genotypes displayed more intensive hydrolysis of SWM than corresponding roots. Both shoots and roots were efficient in hydrolyzing GP. These results generally indicate higher SG-related β-D-glucosidase activity in shoots than in roots, which corresponds to higher expression level of the two candidates identified within the present study.

Lower content of SGs in shoots and roots of tetraploid individuals when compared to diploids might be, at least partially, ascribed to higher expression level of SGs-related β-D-glucosidases and higher activity of these enzymes in tetraploids. Tetraploids most likely display more intense SGs catabolism which is reflected through lower SGs content. As tetraploid plants used in the present study are not a direct offspring of the diploid ones, the observed divergence between them is most likely also influenced by differences in their genomes.

Cloning and Heterologous Expression of CeBGlu in E. coli

Isolated BGlu candidates were sub-cloned into the pRSETA vector; pRSETA:CeBGlu1 and pRSETA:CeBGlu2 constructs were transformed into E. coli BL21 (DE3) competent cells. To achieve simple and efficient purification of the enzyme, CeBGlu1 and CeBGlu2 were expressed in fusion with the His-tag.

Following expression of recombinant proteins in E. coli, they were isolated using His-tag affinity purification and subsequently resolved in SDS-PAGE (Figure 6Aa and
FIGURE 4 | The neighbor-joining tree generated using MEGA X (Version 10.2.6) with a bootstrap of 1,000 replicates, illustrating the phylogenetic relationship of the CeBGlu1 and CeBGlu2 candidates (red letters) relatively to homologous genes from various plant species (refer to Supplementary Table 2 for a list of plant protein sequences used for phylogenetic analysis and their corresponding accession numbers). The optimal tree is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) are shown next to the branches. Groups of enzymes with experimentally defined functions are labeled with different colors.

Supplementary Figure 1). SDS-PAGE analysis of the purified recombinant proteins detected bands with a slightly higher CeBGlu1 and CeBGlu2 molecular masses of ~70 kDa than expected (62.06 kDa) (could be, at least partially, ascribed to a His-tail of recombinant proteins). Purified preparations of recombinant CeBGlu1 and CeBGlu2 were also analyzed by western blot using the anti-6×His antibody (Figure 6Ab). In both cases, a protein band with the same molecular mass as deduced from the SDS-gel was observed.

**In vitro Functional Characterization of CeBGlu1 and CeBGlu2**

Substrates such as 4-nitrophenyl-β-D-galactopyranoside, 4-nitrophenyl-β-D-thioglucoside, and 4-nitrophenyl-β-D-glucopyranoside (pNPG), are most commonly used artificial substrates in enzymatic assays for confirmation of β-glucosidase enzymes function (Czjzek et al., 2000; Geerlings et al., 2000; Warzecha et al., 2000; Morant et al., 2008b; Shaik et al., 2013; Tiwari et al., 2016). In order to examine functionality of CeBGlu1 and CeBGlu2, their hydrolytic activity against commercial substrate 4-nitrophenyl-β-D-glucopyranoside (pNPG) was analyzed. Both enzymes efficiently hydrolyze synthetic glucoside pNPG, which confirms their glucosidase activity (Figure 6B). CeBGlu2 exhibits a higher affinity for the commercial pNPG substrate (4.029 μmol pNPG min⁻¹ mg⁻¹ protein) compared to CeBGlu1 (2.267 μmol pNPG min⁻¹ mg⁻¹ protein). Similarly, enzymes isolated from Lamium galeobdolon (LgGlu2 and LgGlu4) hydrolyze pNPG, but the efficiency of the two tested enzymes differs (Hannemann et al., 2018). Purified β-glucosidases of Lotus japonicus (LjBGD2 and LjBGD4) hydrolyze pNPG, but show higher affinity for plant-synthesized substrates, such as prunasin, lotaustralin, and rhodiocyanoside (Morant et al., 2008b). However, plant β-glucosidases are not always effective in reaction with synthetic substrates. Thus, β-glucosidase of O. europaea (OeGlu) does not show hydrolytic activity against tested pNPG substrate, and its function was assigned based on a reaction in which
Božunović et al. Functional Characterization of Centaury β-D-Glucosidase

FIGURE 5 | (A) Relative expression of CeBGlu in shoots and roots of in vitro grown C. erythraea diploid (2×) and tetraploid (4×) plants. Values are normalized to the CeBGlu expression level in roots of diploid plants. Letters above the bars denote significant differences according to Fisher LSD test at $p < 0.05$. (B) Content of major secoiridoids in shoots and roots of diploid and tetraploid C. erythraea plants. Values represent means ± SE ($n = 3$). Letters above the bars denote significant differences according to Fisher LSD post hoc test at $p < 0.05$, for each compound independently. (C) Hydrolytic activity of total proteins isolated from shoots and roots of diploid and tetraploid C. erythraea genotypes, using sweroside, swertiamarin and gentiopicrin as substrates. The bars indicate the amount of substrates (%) not enzymatically hydrolyzed in an in vitro enzymatic assay. Control group (white bars), where no protein extract was added, was set to 100%, and all the values are presented relatively, in respect to control. Data are mean ± SD ($n = 3$). Red asterisks denote significantly different adjusted $p$-values, *$p < 0.001$.

OeGlu efficiently hydrolyzes a specific substrate available in planta – secoiridoid oleuropein (Koudounas et al., 2015). Also, β-glucosidases of Arabidopsis thaliana (AtBGlu21, AtBGlu22, and AtBGlu23) hydrolyze the tropan alkaloid scopolin, but do not exhibit hydrolytic activity against a synthetic substrate (Ahn et al., 2010).

CeBGlu1 and CeBGlu2 hydrolytic activity was subsequently tested in enzymatic reactions with SG-rich methanol extract of C. erythraea as a substrate, and changes in the content of major SGs (SW, SWM, and GP) were recorded (Figure 6C). CeBGlu2 was more efficient than CeBGlu1 in hydrolyzing SWM, which was the most abundant SG in C. erythraea methanol extract used in experiments. CeBGlu1 and CeBGlu2 were equally efficient against SW and GP. As the amounts of SW (2.0 µg ml$^{-1}$ extract), SWM (3.5 µg ml$^{-1}$ extract) and GP (2.1 µg ml$^{-1}$ extract) in C. erythraea methanol extract were not equal, it was not possible to make conclusions about the preferable substrates of the two CeBGlu.

To test the substrate specificity, β-D-glucosidase activity of recombinant CeBGlu1 and CeBGlu2 enzymes was determined against several plant-derived glucosides of the β-D-type (Figure 6D). Hydrolytic activity of CeBGlu1 and CeBGlu2 was tested against iridoid glucosides (1,5,9-epideoxyloganic acid-1,5,9-eDLA and loganin- LOG), secoiridoid glucosides (SEC, SW, SWM and GP), and flavonoid glucosides (apigenin- AP, isoorientin- IQ, and vitexin- VI). Pure 1,5,9-eDLA was isolated from methanol extracts of Nepeta ranjensis as described by Aničić et al. (2021), while AP and VI, as well as standards of iridoids and secoiridoids, were commercially purchased. As standards for aglycones of secoiridoids were not available, the β-D-glucosidase activity of recombinant proteins was evaluated by UHPLC/DAD/(±)HESI−MS$^2$ quantification of targeted glucosides (Supplementary Figure 2). Their content in reaction mixtures were compared to those of the negative controls where recombinant proteins were excluded. Although two recombinant proteins share 99.3% similarity of amino acid sequences, CeBGlu1 and CeBGlu2 display slightly differential hydrolytic activity and specificity against tested substrates (Figure 6D). The highest activity of CeBGlu1 was recorded when using GP and AP as substrates, which were hydrolyzed by 95.4 and 97.1%, respectively. This enzyme was also efficient in hydrolyzing LOG (76.0%). The highest hydrolytic activity of CeBGlu2 was recorded for GP (99.7%) and SW (98.3%), followed by AP (63.5%) and SWM (52.2%). Generally, both CeBGlu1 and
CeBGlu2 display higher affinity for GP, than for SW and SWM. Although significant activity of CeBGlu1 and CeBGlu2 was recorded when using AP (apigenin 7-O-glucoside) as a substrate, one should bear in mind that, up to the best of our knowledge, this compound was not previously detected in *C. erythraea* methanol extracts (Siler and Mišić, 2016; Banjanac et al., 2017; Božunović et al., 2018). Slight hydrolytic activity of CeBGlu1 and CeBGlu2 with 1,5,9-eDLA, SEC and IQ was recorded, but it was not statistically significant (Figure 6D). The two enzymes displayed no hydrolytic activity against VI, an apigenin 8-C-glycoside, which was also not previously detected in *C. erythraea*. β-D-glucosidases display substantial substrate promiscuity (Hannemann et al., 2018) so the hydrolysis of several substrates is not uncommon. Many beta-glucosidases have transglucosidase activities in addition to their hydrolase activity (Opassiri et al., 2004). However, the number of glucos conjugates in plants is likely larger than the number of beta-glucosidases and the enzymes tend to have overlapping specificities, which complicates determination of their exact functions. These cases of multiple functions have been categorized into “multitasking,” where the enzyme carries out multiple functions at the same time, or “moonlighting,” where the enzyme has two different functions in divided situations, according to Ketudat Cairns et al. (2015).

### Transient Overexpression of CeBGlu1 and CeBGlu2 in Leaves of Diploid and Tetraploid Centaurium erythraea

Transient expression of recombinant proteins in plants using tobacco mosaic virus (TMV) based viral vectors have been well documented (Chiong et al., 2021). TMV-based overexpression pJL-TRBO vector can be used for the insertion of a target gene by agroinfiltration into plants. *Agrobacterium*-mediated delivery of genes, results in highly efficient transient expression of the foreign
protein, producing up to 100-fold more recombinant protein compared to a non-viral system (Lindbo, 2007).

To verify the functional identity of the CeBGlut gene, Agrobacterium-mediated transient expression in centaury leaves was carried out. To that end, the CeBGlut coding region was subcloned into the pJL-TRBO vector. The resultant plasmids designated pJL-TRBO:CeBGlut1 and pJL-TRBO:CeBGlut2 were used to transform A. tumefaciens GV3101 for the transient in planta expression. The two CeBGlut candidates were overexpressed in leaves of five-month-old diploid and tetraploid C. erythraea plants, alone or in combination with p19 (Figure 7). Phytochemical profiling of SGs in leaves of diploid and tetraploid C. erythraea plants, harvested 5 days after agroinfiltration, was performed. Overexpression of CeBGlut1 and CeBGlut2 induced changes in the content of secoiridoid glucosides (SGs) swertiamarin (SWM), sweroside (SW) and gentiopicrocin (GP) when co-expressed with p19. No statistically significant alterations in GP content in response to agroinfiltration with CeBGlut1 and CeBGlut2 was observed, in both diploid and tetraploid plants. Viral-encoded suppressor of gene silencing, the p19 protein of tomato bushy stunt virus (TBSV), which is believed to prevent the onset of post-transcriptional gene silencing in the infiltrated tissues (Voinnet et al., 2003; Shah et al., 2013), obviously allowed high level of CeBGlut1 and CeBGlut2 transient expression, which resulted in pronounced phytochemical changes in C. erythraea leaves. Previous studies have suggested that transgene expression or RNA silencing in plants can be affected by ploidy, and is significantly less efficient in tetraploids than in diploids (Finn et al., 2011). In other words, transgenes are more prone to transcriptional inactivation in polyploids than in diploids (Mittelsten Scheid et al., 1996, 2003), which often poses a major limitation to polyploids improvement via biotechnology (Gao et al., 2013). Thus, it is not surprising that the use of p19 more efficiently increases the transient expression of recombinant CeBGlut1 and CeBGlut2 proteins in leaves of C. erythraea diploids.
than in tetraploids, which is accompanied with the more pronounced decrease in SGs content (Figure 7). Taken together, overexpression of the two CeBGlu isogenes decreased the content of SGs in C. erythraea leaves, which additionally confirmed the β-glucosidase function of these two genes. The entire centaury genome sequence is still unknown, yet it probably contains many other β-glucosidases of diverse as well as similar functions and substrate specificities, which could also contribute to shaping the secoiridoid profiles of these remarkable plants.

Comparative 3D Modeling of CeBGlu1 and CeBGlu2
To gain insights into sequence-structure relations of the two centaury β-glucosidases their 3D structure was inferred using AlphaFold (models provided as Supplementary Material 1). The obtained models had a low percentage of Ramachandran outliers and a low clash score (Supplementary Table 5). The percentage of Ramachandran favored Cα angles was between 92 and 93% for the two models. The majority of residues not within the favored parts of the Ramachandran diagram had a low pLDDT value (a per-residue estimate of AlphaFold confidence on a scale from 0 to 100) suggesting they are in disordered regions. The two predicted models corresponding to CeBGlu1 to CeBGlu2 are highly structurally similar, especially if the likely disordered regions are not taken into account – the RMSD of the superimposed pruned 475 Cα atoms in the two structures is 0.216 Å. Both obtained models showed a high degree of structural similarity to the experimental structure of the raucaffricine β-D-glucosidase from R. serpentina (PDB: 3U5Y, Supplementary Table 5; Xia et al., 2012). The models have a (β/α) 8-barrel (Figures 8A,B) fold characteristic for the CAZy GH1, consisting of eight parallel β-strands forming a barrel-like structure. These strands are connected via intricate helical regions and loops (Figure 8). Strands β5 and β6 are longer compared to other, and protrude the barrel (Figures 8A,B). Based on AlphaFold secondary structure annotation these are split into two strand regions each. Together with an antiparallel strand βa1 (Figure 8) which is not a part of the barrel, β5 and β6 form a sheet. After the (β/α) 8-barrel structure is a sheet formed from two antiparallel strands βa2 and βa3. The regions with low pLDDT are on the termini, between β4 and β5, as well as between β6 and βa1, matching closely the regions with missing density in the experimental structure of raucaffricine β-D-glucosidase from R. serpentina which strengthens the conclusion these regions are disordered.

The enzymes from GH1 family have a retaining double displacement mechanism in which one Glu residue acts as a proton donor for the leaving group, while another Glu residue acts as a nucleophile (Ketudat Cairns and Esen, 2010).
The nucleophilic attack forms a covalent glycosyl enzyme intermediate with an inverted bond (Figure 9). This covalent intermediate is hydrolyzed by nucleophilic attack of a water molecule, activated by the proton donor residue, resulting in second bond inversion and in overall anomeric retention. The catalytic proton donor and nucleophile in the two centaury
β-glucosidases correspond to Glu187 and Glu420, respectively (Figure 9). Glu187 is located in the coil just after β4, while Glu420 is situated at the end of β7, while their Ca atoms are positioned at a distance of 10.3 Å in the models. The pKa values of respective carboxylic groups were estimated to be ~8 for Glu187 and ~5.5 for Glu420 which is in concordance to their proposed role of proton donor and nucleophile, respectively. The two centaury β-glucosidases differ in four amino acids which include, from CeBGlu1 to CeBGlu2: A211D, S278C, D307E and D332E; the first two being radical replacements, while the latter two are conservative replacements. Position 211 is located in the disordered region between β4 and β5; 278 is located in β5, proximate to, with a side chain facing away from the active site, while 307 and 332 are located in the helical regions connecting β5 to β6 (Figure 8). Based on the type, position and orientation of the differing amino acids at positions 211, 307, and 332 no impact of the mentioned amino acids replacements on the overall enzyme performance can be expected. It should be noted that these amino acids are not at the dimer interface based on comparison with 3U5Y (result not shown), so not even a long-distance interaction, due to differential multimer binding, is to be expected. However, one cannot rule out the possible indirect roles of these amino acids on the overall enzyme performance. On the other hand, the position 278 is close to the substrate binding site, and the amino acids around it – Ile277 and Gln279 are directly involved in interactions with the aglycone (Figure 10). Thus, even though the predicted structures of CeBGlu1 to CeBGlu2 are highly similar, including the active site, and the orientation of Ile277 and Gln279 side chains, S278C could have an impact on substrate recognition and overall catalytic performance. It should be noted there are no nearby Cys with which potential disulfide bonds can be formed by Cys287. Thus, the observed differences in the substrate preferences of the two CeBGlu enzymes (Figure 6), might, at least partially, be ascribed to the differences in their amino acid sequences, especially at position 287.

To explore and visualize the binding of substrates to the active site ligand docking was analyzed only for CeBGlu1, because of the very high overall similarity of the predicted models for the two β-glucosidases. To paraphrase quantitatively, the RMSD between the 236 atom pairs belonging to residues Glu37, Trp142, Asn186, Glu187, Thr190, Ile277, Gln279, Asn347, Trp392, Glu420, Trp469, Glu476, and Trp477 which form the active sites including the catalytic residues between the two predicted structures is 0.625 Å. Most open-source software for ligand docking is able to perform some sort of conformer search predicted structures is 0.625 Å. Most open-source software for ligand docking was analyzed only for CeBGlu1 and CeBGlu2 via AutoDock-Vina version implements a macrocycle conformer sampling method (Eberhardt et al., 2021), however, an analogs method appropriate for small aliphatic rings is lacking. Due to this, sets of conformers were pre generated by the ETKDG3 method with small ring torsion angle preferences (Wang et al., 2020) for GP, SW and SWM and used for docking. Prior to docking, while comparing the centaury β-glucosidase models with similar experimental structures from the PDB it was noticed CeBGlu1 and CeBGlu2 side chain clashes with the glucose in the active center of superimposed experimental structures. Therefore, a flexible docking procedure was employed where Glu476 was allowed to change conformations, while all other enzyme residues were rigid. For GP the highest scoring docked pose (Table 1 and Figure 10A) seems to be very close to what could be expected in vivo, at least based on comparison to the pose of the secologanin ligand in the experimental structure of racafricine β-D-glucosidase from R. serpentina (Figure 10D, 3U5Y, Xia et al., 2012). The specificity of the active site for β-glucosides is achieved by numerous H-bonds with the sugar OH groups: Asn186, Glu187 and Asn347 form H-bonds with O2, Gln37 and Trp477 with O3, Gln37, Trp469 and Glu476 with O4 on β-glucose (Figure 10A). The aglycone binding pocket is tight and hydrophobic, enclosed from one side by Trp392 which is parallel to the aglycone, and from the others by Trp142, Thr190 and Ile277. The only polar interaction with the ligand is via an H-bond between Gln279 and the keto oxygen. The relatively tight space in the aglycone binding site allows mostly planar aglycones, while Gln279 position favors aglycons with H-acceptors or H-donors in the appropriate position. The positions of the highest scoring poses for SW and SWM were not suitable for hydrolysis, because the aglycone occupied the position of glucose. However, the 2nd best pose for SWM and 3rd best pose for SW (Table 1 and Figures 10B,C), which scored only slightly worse, were very similar both to the previously mentioned docked GP, as well as to secologanin ligand in 3U5Y (Figure 10D). The suboptimal highest scoring poses for SWM and SW might be the result of limitations of the used docking procedure, and imperfections in the scoring function.

### CONCLUSION

The present study highlights the direct link between β-D-glucosidases characterized within the present study (CeBGlu1 and CeBGlu2) and the content of SGs in analyzed C. erythraea tissues. Thus, it provides the first evidence of SGs-related β-glucosidases in centaury plants, and advances our understanding of SGs deglycosylation, as a part of their catabolism. These enzymes could serve as a molecular target of high biotechnological interest, in order to produce centaury plants with an optimal composition of secoiridoids, which are essential defense compounds in plants with beneficial effects in human and animal health. One of the major challenges ahead is to gain better understanding of how plants regulate the function of CeBGlu1 and CeBGlu2 in distinct tissues and developmental stages, and in response to biotic and abiotic stress stimuli, which is the course of our further work. Furthermore, it is imperative to assess their biological function and confirm their localization at the cellular and organelle level and define conditions under

---

**TABLE 1 | Ligand docking summary into CeBGlu1 AlphaFold model.**

| Ligand   | Chosen pose affinity (kcal/mol) | Best pose affinity (kcal/mol) | Rank of chosen pose |
|----------|---------------------------------|------------------------------|---------------------|
| Gentiiopicrin | -9.561                         | -9.561                       | 1                   |
| Swertiarmin  | -9.105                         | -9.169                       | 2                   |
| Sweroside   | -9.359                         | -9.644                       | 3                   |
which these enzymes come into contact with their physiological substrates. Alternatively, gene silencing of the CeBGlu1 and CeBGlu2 combined with metabolic profiling of silenced plants might provide more information on the role and function of these genes, and could also result in elevated amounts of SGs in tissues. In parallel, unveiling the remaining unknown steps of SGs biosynthetic pathway in *C. erythraea* is of essential importance for establishing the biotechnology-based production of these valuable bioactive compounds. Accumulating knowledge will, in the future, enable manipulation of SGs biosynthesis and catalobolism through multi-target metabolic engineering, which will further enable large-scale production of desired secoiridoids in *C. erythraea*.

**DATA AVAILABILITY STATEMENT**

The data presented in this study are deposited in the National Center for Biotechnology Information (NCBI) repository (NCBI; https://www.ncbi.nlm.nih.gov/) under accession numbers ON060690 and ON060691.

**AUTHOR CONTRIBUTIONS**

JB, DMa, SŽ, MM, MS, and DMi conceived and designed the experiments. JB, MM, NA, MS, DMa, SŽ, BF, and TB performed the experiments. DMi performed the phytochemical characterization of samples. MD was responsible for the 3D modeling and ligand docking analysis. JB, MM, DMa, and DMi organized and wrote the manuscript with editing from all the authors. All authors contributed to the article and approved the submitted version.

**REFERENCES**

Adasme, M. F., Linnemann, K. L., Bolz, S. N., Kaiser, F., Salentin, S., Haupt, V. J., et al. (2021). PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucleic Acids Res.*, **49**, W530–W534. doi: 10.1093/nar/gkab294

Ahn, Y. O., Shimizu, B. I., Sakata, K., Guntulga, D., Zhou, Z., Bevan, D. R., et al. (2010). Scopolin-hydrolyzing β-glucosidases in roots of Arabidopsis. *Plant Cell Physiol.*, **51**, 132–143. doi: 10.1093/pcp/pcp174

Anić, N., Gašić, U., Lu, F., Ćirić, A., Ivanov, M., Jevtić, B., et al. (2021). Antimicrobial and immunomodulating activities of two endemic *Nepeta* species and their major iridoids isolated from natural sources. *Pharmaceuticals* **14**(4), 1-14. doi: 10.3390/PH14050014

Banjanac, T., Dragičević, M., Siler, B., Gašić, U., Bohanec, B., Nestorović Zivković, J., et al. (2017). Chemodiversity of two closely related tetraploid *Centaurium* species and their hexaploid hybrid: metabolomic search for high-resolution taxonomic classifiers. *Phytochemistry* **140**, 27–44. doi: 10.1016/j.phytochem.2017.04.005

Barleben, L., Panjikar, S., Ruppert, M., Koepke, J., and Stöckjagt, J. (2007). Molecular architecture of strictosidine glucosidase: the gateway to the biosynthesis of the monoterpenoid indole alkaloid family. *Plant Cell Tissue Organ Cult.* **55**, 670–675. doi: 10.1134/S1282175698040069

Bartholomaeus, A., and Ahokas, J. (1995). Inhibition of P-450 by aucubin is the biological activity of aucubin due to its glutaraldehyde-like aglycone? *Toxicol. Lett.* **80**, 75–83. doi: 10.1016/0378-4274(95)03339-M

Božunović, J., Škoršić, M., Matekaldo, D., Živković, S., Dragičević, M., Anić, N., et al. (2019). Secoiridoid metabolism response to wounding in common centaury (*Centaurium erythraea* Rafn) leaves. *Plants* **8**(589). doi: 10.3390/PLANTS8120589

Božunović, J., Živković, S., Gašić, U., Glamočlija, J., Ćirić, A., Matekaldo, D., et al. (2018). In vitro and in vivo transformations of *Centaurium erythraea* secoiridoid glucosides alternate their antioxidant and antimicrobial capacity. *Ind. Crops Prod.* **111**, 705–721. doi: 10.1016/j.indcrop.2017.11.040

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254. doi: 10.1016/0003-2697(76)90527-3

Chen, R., Liao, Z. H., Chen, M., Wang, Q., Yang, C. X., and Yang, Y. J. (2008). Molecular cloning and characterization of the Strictosidine synthase Gene from *Rauwolfia verticillata*. *Russ. J. Plant Physiol.* **35**, 670–675. doi: 10.1134/S0121437008050117

Chiong, K. T., Cody, W. B., and Scholtz, H. B. (2021). RNA silencing suppressor-influenced performance of a virus vector delivering both guide RNA and Cas9 for CRISPR gene editing. *Sci. Rep.* **11**, 6769. doi: 10.1038/s41598-021-85366-4

Ćuković, K., Dragičević, M., Bogdanović, M., Paunović, D., Giurato, G., Filipović, B., et al. (2020). Plant regeneration in leaf culture of *Centaurium erythraea* Rafn. Part 3: de novo transcriptome assembly and validation of housekeeping genes for studies of in vitro morphogenesis. *Plant Cell Tissue Organ Cult.* **141**, 417–433. doi: 10.1007/s11240-020-01801-w

Czajek, M., Cicok, M., Zamboni, V., Bevan, D. R., Henrikss, B., and Esen, A. (2000). The mechanism of substrate (aglycone) specificity in β-glucosidases is revealed by crystal structures of mutant maize β-glucosidase-DIMBOA, -DIMBOAGlc, and-dhurrin complexes. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13555–13560. doi: 10.1073/pnas.97.25.13555

Eberhardt, J., Santos-Martins, D., Tillack, A. F., and Forli, S. (2021). AutoDock Vina 1.2. 0: new docking methods, expanded force field, and python bindings. *J. Chem. Inf. Model.* **61**, 3891–3898. doi: 10.1021/acs.jcim.1c00203

El-Sedawy, A. I., Hattoni Kobashi, K., and Namba, T. (1990). Metabolism of sweroside from *Swertia japonica* by human intestinal bacteria. *Shoyakugaku Zasshi* **44**, 122–126. doi: 10.1055/s-2006-961909

Esen, A., and Blanchard, D. J. (2000). A specific β-glucosidase-aggregating factor is responsible for the β-glucosidase null phenotype in maize. *Plant Physiol.* **122**, 563–572. doi: 10.1104/pp.122.2.563

Filipović, B., Šiler, B., Nestorović Živković, J., Banjanac, T., Skorić, M., Božunović, J., et al. (2019). Diploid vs. tetraploid *Centaurium erythraea* Rafn: a comparative study of regenerative in vitro potential and biosynthetic capacity. *Lekovite Sirovine* **39**, 52–59. doi: 10.5937/leksir1939053F

**ACKNOWLEDGMENTS**

We want to thank Milica Bogdanović (Institute for Biological Research “Siniša Stanković”- National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia) for providing the pBIn:p19 plasmid.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.914138/full#supplementary-material
Finn, T. E., Wang, L., Smolillo, D., Smith, N. A., White, R., Chaudhury, A., et al. (2011). Transgene expression and transgene-induced silencing in diploid and autotetraploid Arabidopsis. Genetics 187, 409–423. doi: 10.1534/genetics.110.124370

Franke, J., Lüe, J., Hamilton, J. P., Zhao, D., Pham, G. M., Wiegert-Rininger, K., et al. (2019). Gene discovery in germline highlighted conserved gene clusters in monoterpene indole alkaloid biosynthesis. ChemBioChem 20, 83–87. doi: 10.1002/cbic.201800592

Gao, S. J., Damaj, M. B., Park, J. W., Beyene, G., Buenrostro-Nava, M. T., Molina, J., et al. (2013). Enhanced transgene expression in sugarcane by co-expression of virus-encoded RNA silencing suppressors. PLoS One 8(6):e60646. doi: 10.1371/journal.pone.0060646

Gasic, K., Hernandez, A., and Korban, S. S. (2004). RNA extraction from different apple tissues rich in polyphenols and polyaccharides for CDNA library construction. Plant Mol. Biol. Rep. 22, 437–438. doi: 10.1016/S0269-704X(01)00027-6

Geelings, A., Martinez-Lozano Ibañez, M., Memelink, J., Van Der Heijden, R., and Verpoorte, R. (2000). Molecular cloning and analysis of strictidose 1-β-glucosidase, an enzyme in terpenoid indole alkaloid biosynthesis in Catharanthus roseus. J. Biol. Chem. 275, 3051–3056. doi: 10.1074/jbc.M5.0.3051

Ghisalberti, E. L. (1998). Biological and pharmacological activity of naturally occurring iridoids and secoiridoids. Phytomedicine 5, 147–163. doi: 10.1016/S0944-7113(98)00012-3

Guirmand, G., Courdavault, V., Lanoue, A., Mahroug, S., Guihur, A., Blanc, N., et al. (2010). Strictidose activation in Apocynaceae: towards a “nuclear time bomb”? BMC Plant Biol. 10:182. doi: 10.1186/1471-2229-10-182

Guirmand, G., Guihur, A., Poutrain, P., Héricourt, F., Mahroug, S., St-Pierre, B., et al. (2011). Spatial organization of the vindoline biosynthetic pathway in Catharanthus roseus. J. Plant Physiol. 168, 549–557. doi: 10.1016/j.jplph.2010.08.018

Hannemann, L., Lucaciu, C. R., Sharma, S., Rattei, T., Mayer, K. F. X., Gierl, A., et al. (2018). A promiscuous beta-glucosidase is involved in benzoazinoid deglycosylation in Lamium galeobdolon. Phytochemistry 156, 224–233. doi: 10.1016/j.phytochem.2018.10.012

Huson, D. H., and Scornavacca, C. (2012). Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Syst. Biol. 61, 1061–1067. doi: 10.1093/sysbio/sys062

Ishiguro, K., Yamaki, M., and Takagi, S. (1983). Studies on Iridoid-related compounds III: gentiopircral, the aglucone of gentiopicroside. Plant Med. 49, 208–210. doi: 10.1055/s-2007-969852

Jensen, S. R., and Schripsema, J. (2002). Chemotaxonomy and pharmacology of secoiridoids in olives. J. Agric. Food Chem. 50, 583–589. doi: 10.1021/jf010977z

Koudounas, K., Thomopoulou, M., Rigakou, A., Angeli, E., Mellou, E., Magiatis, P., et al. (2021). Silencing of oleuropein β-glucosidase abolishes the biosynthetic capacity of secoiridoids in olives. Front. Plant Sci. 12:671487. doi: 10.3389/fpls.2021.671487

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35, 1547–1549. doi: 10.1093/molbev/msy096

Landrum, G., Tosco, P., Kelley, B., Ric, Gedeck, S., Vianello, R., et al. (2022). rdkit/rdkit 2021_09_5 (Q3 2021) Release. doi: 10.5281/zenodo.630241

Li, Z., Jaroszewski, L., Iyer, M., Sedova, M., and Godzik, A. (2020). FATCAT 2.0: towards a better understanding of the structural diversity of proteins. Nucleic Acids Res. 48, W60–W64. doi: 10.1093/nar/gkaa443

Lindbo, J. A. (2007). TRBO: a high-efficiency tobacco mosaic virus RNA-based overexpression vector. Plant Physiol. 143, 1232–1240. doi: 10.1104/pp.107.106377

Ling, S. K., Tanaka, T., and Kouno, I. (2003). Effects of iridoids on lipoxigenase and hyaluronidase activities and their activation by β-glucosidase in the presence of amino acids. Biol. Pharm. Bull. 26, 352–356. doi: 10.1248/bpb.26.352

Livila, K. J., and Schmitzgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2 −ΔΔCT method. Methods 25, 402–408. doi: 10.1016/meth.2001.12.002

Malkov, S., and Simonović, A. (2011). “Shotgun assembly of Centaurium erythraea transcriptome,” in Proceedings of the 19th Symposium of the Serbian Plant Physiology Society, Banja Luka, 13–15.

Martínez-Rosell, G., Giorgio, T., and De Fabritiis, G. (2017). PlayMolecule. Proteins: Struct. Funct. Bioinf. 85, 1334–1341. doi: 10.1002/prot.25756

Morrant, A. V., Bjarnholth, N., Kragh, M. E., Kjærgaard, C. H., Jørgensen, K., et al. (2009). AutoDock4 and autodocktools4: automated docking with selective receptor flexibility. J. Comput. Chem. 30, 2785–2791. doi: 10.1002/jcc.21256

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant. 15, 473–497.

Nomura, T., Quesada, A. L., and Kutchan, T. M. (2008). The new plant model Catharanthus roseus as an experimental platform for understanding plant disease resistance. Trends Biotechnol. 26, 389–393. doi: 10.1016/j.tibtech.2007.10.010

Oliveira, J. A., Lopes, M. R., and Atlan, F. (2018). Plant Physiol. 171, 7114–7119. doi: 10.1073/pnas.93.14.7114

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. L., et al. (2021). UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci. 30, 70–82. doi: 10.1002/pro.3943
Stärk, H., Dallago, C., Heinzinger, M., and Rost, B. (2021). Light attention predicts expression vectors for transient expression of recombinant proteins in plants. Plant Mol. Biol. Rep. 31, 1529–1538. doi: 10.1007/s11105-013-0614-z

Shaik, N. M., Misra, A., Singh, S., Fatangare, A. B., Ramakumar, S., Rawal, S. K., et al. (2013). Functional characterization, homology modeling and docking studies of \( \beta \)-glucosidase responsible for bioactivation of cyanogenic hydroxynitrile glucosides from Leucaena leucocephala (subabul). Mol. Biol. Rep. 40, 1351–1363. doi: 10.1007/s11033-012-2179-6

Šiler, B., Avramov, S., Banjanac, T., Čvetković, J., Nestorović Živković, J., Patenković, A., et al. (2012). Secoiridoid glucosides as a marker system in chemical variability estimation and chemotype assignment of Centaurium erythraea Rafn from the Balkan Peninsula. Ind. Crops Prod. 40, 336–344. doi: 10.1016/j.indcrop.2012.03.026

Šiler, B., and Mišić, D. (2016). Biologically active compounds from the genus Centaurium s.l. (Gentianaceae): current knowledge and future prospects in medicine. Stud. Nat. Prod. Chem. 49, 363–397. doi: 10.1016/B978-0-444-63601-0.00011-9

Šiler, B., Živković, S., Banjanac, T., Čvetković, J., Nestorović Živković, J., Ćirić, A., et al. (2014). Centauries as underestimated food additives: antioxidant and antimicrobial potential. Food Chem. 147, 367–376. doi: 10.1016/j.foodchem.2013.10.007

Stärk, H., Dallago, C., Heinzinger, M., and Rost, B. (2021). Light attention predicts protein location from the language of life. Protein Sci. 40, 293–315. doi: 10.1002/pro.401597

Tiwari, R., Kumar, K., Singh, S., Nain, L., and Shukla, P. (2016). Molecular detection and environment-specific diversity of glycosyl hydrolase family 1 \( \beta \)-glucosidase in different habitats. Front. Microbiol. 7:1597. doi: 10.3389/fmicb.2016.01597

Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J. 33, 949–956. doi: 10.1046/j.1365-313X.2003.01676.x

Wang, S., Witek, J., Landrum, G. A., and Riniker, S. (2020). Improving conformer generation for small rings and macrocycles based on distance geometry and experimental torsional-angle preferences. J. Chem. Inf. Model. 60, 2044–2058. doi: 10.1021/acs.jcim.0c00025

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. doi: 10.1016/S0031-9422(00)00175-8

Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, I. L., Deis, L. N., et al. (2018). MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci. 27, 293–315. doi: 10.1002/pro.3330

Wright, E. S. (2015). DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. BMC Bioinformatics 16:322. doi: 10.1186/s12859-015-0749-z

Xia, L., Ruppert, M., Wang, M., Panjikar, S., Lin, H., Rajendran, C., et al. (2012). Structures of alkaloid biosynthetic glucosidases decode substrate specificity. ACS Chem. Biol. 7, 226–234. doi: 10.1021/cb200267w

Zeng, W., Han, H., Tao, Y., Yang, L., Wang, Z., and Chen, K. (2013). Identification of bio-active metabolites of gentiopicroside by UPLC/Q-TOF MS and NMR. Biomed. Chromatogr. 27, 1129–1136. doi: 10.1002/bmc.2917

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Božunović, Mihutinović, Antić, Skorić, Matekalo, Živković, Dragićević, Filipović, Banjanac, Petrović and Mišić. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.