



## Substrate specificity in hydrolysis and transglucosylation by family 1 $\beta$ -glucosidases from cassava and Thai rosewood

Prachumporn T. Kongsaree<sup>a,b,\*</sup>, Khakhanang Ratananikom<sup>a</sup>, Khuanjarat Choengpanya<sup>b</sup>, Nusra Tongtubtim<sup>a,b</sup>, Penporn Sujiwattanarat<sup>b</sup>, Chompoonuth Porncharoenpob<sup>b</sup>, Amornrat Onpium<sup>a</sup>, Jisnusun Svasti<sup>c</sup>

<sup>a</sup> Department of Biochemistry, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

<sup>b</sup> Interdisciplinary Graduate Program in Genetic Engineering, Faculty of Graduate School, Kasetsart University, Bangkok 10900, Thailand

<sup>c</sup> Department of Biochemistry and Center for Excellence in Protein Structure and Function, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

### ARTICLE INFO

#### Article history:

Received 13 May 2010

Received in revised form 29 July 2010

Accepted 9 September 2010

Available online 17 September 2010

#### Keywords:

Substrate specificity

$\beta$ -Glucosidase

Dalcochinase

Linamarase

Transglucosylation

### ABSTRACT

Thai rosewood (*Dalbergia cochinchinensis* Pierre) dalcochinase and cassava (*Manihot esculenta* Crantz) linamarase are glycoside hydrolase family 1  $\beta$ -glucosidases with 47% amino acid sequence identity. Each enzyme can hydrolyze its natural substrate, dalcochinin-8'-O- $\beta$ -D-glucoside and linamarin, respectively, but not the natural substrate of the other enzyme. Linamarase can transfer glucose to primary, secondary and tertiary alcohols with high efficiency, while dalcochinase can transglucosylate primary and secondary alcohols at moderate levels. In this study, eight amino acid residues in the aglycone binding pocket of dalcochinase were individually replaced with the corresponding residues of linamarase, in order to identify residues that may account for their catalytic differences. The residues I185 and V255 of dalcochinase appeared important for its substrate specificity, with their respective mutations resulting in 24- and 12-fold reductions in  $k_{cat}/K_m$  for the hydrolysis of dalcochinin-8'-O- $\beta$ -D-glucoside. Transglucosylation activity was improved when I185, N189 and V255 of dalcochinase were replaced with A201, F205 and F271 of linamarase, respectively, suggesting these residues support transglucosylation in linamarase. Among these three mutants, only the N189F mutant showed significant increases in the rate constants for the reactivation of trapped glucosyl-enzyme intermediates by all alcohols. Together, our results suggest that both hydrophobicity and geometry are important determinants for substrate specificity in hydrolysis and transglucosylation by these family 1  $\beta$ -glucosidases.

© 2010 Elsevier B.V. All rights reserved.

### 1. Introduction

$\beta$ -Glucosidases (EC 3.2.1.21) catalyze the hydrolysis of the  $\beta$ -O-glycosidic linkage between a glucose moiety and an aglycone (aryl or alkyl) or another sugar [1]. Most enzymes belong to glycoside hydrolase families 1 and 3 (GH1 and GH3), and hydrolyze their substrates via a double-displacement mechanism [2]. Isozymes of  $\beta$ -glucosidases differ considerably in their specificities for the aglycone moiety linked to the glucosyl group, and show different capabilities in catalyzing reverse hydrolysis and glucosyl transfer reactions. These differences presumably result from specific interactions between unique amino acid residues in their active site pockets and their specific substrates.

Dalcochinase is an isoflavonoid  $\beta$ -glucosidase, purified from the seeds of Thai rosewood (*Dalbergia cochinchinensis* Pierre) [3]. Linamarase is a cyanogenic  $\beta$ -glucosidase present in cassava (*Manihot esculenta* Crantz) [4]. Both enzymes belong to GH1 family, and contain the highly conserved TL/FNEP and I/VTENG motifs that are present in all GH1  $\beta$ -glucosidases [5,6]. The two glutamate residues, E182 and E396 in dalcochinase and E198 and E413 in linamarase, present in the conserved motifs act as the catalytic acid/base and nucleophile residues, respectively. Both enzymes are predicted to have tertiary structures consisting of an  $(\alpha/\beta)_8$  barrel similar to other GH1 enzymes. While the two enzymes share 47% sequence identity, their catalytic capabilities are distinct. In hydrolysis, dalcochinase and linamarase can each efficiently hydrolyze its own natural substrate, dalcochinin-8'-O- $\beta$ -D-glucoside (Dal-Glc) and linamarin, respectively (Fig. 1), but cannot hydrolyze the natural substrate of the other enzyme [7]. Furthermore, these two enzymes showed distinct capabilities in transglucosylation and reverse hydrolysis reactions. Dalcochinase can transglucosylate primary alcohols well, and secondary alcohols moderately, but not tertiary alcohols [8]. However, linamarase

Abbreviations: Dal-Glc, dalcochinin-8'-O- $\beta$ -D-glucoside; GH, glycoside hydrolase family; pNP-Glc, p-nitrophenyl- $\beta$ -D-glucoside.

\* Corresponding author. Tel.: +66 2 562 5555x2026; fax: +66 2 561 4627.

E-mail address: [fscipt@ku.ac.th](mailto:fscipt@ku.ac.th) (P.T. Kongsaree).

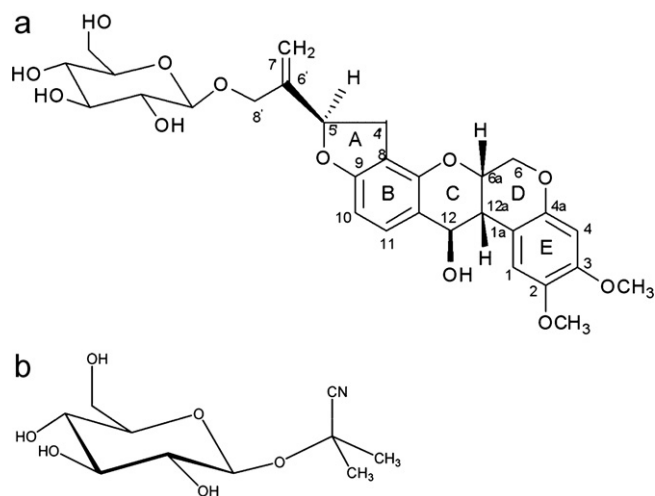


Fig. 1. Structures of dalcochinin-8'-O- $\beta$ -D-glucoside (a) and linamarin (b).

is the only enzyme studied thus far that can efficiently catalyze transglucosylation reactions using primary, secondary and tertiary alcohols as acceptors [9]. On the other hand, dalcochinase can also synthesize oligosaccharides and glycosides by reverse hydrolysis [10], while linamarase cannot. The distinct preferences shown by Thai rosewood dalcochinase and cassava linamarase for the leaving aglycone moieties in hydrolysis reactions and for the incoming alcohol groups in transglucosylation reactions presumably result from differences in the amino acids found in the aglycone binding sites.

We have previously reported the cloning and expression of dalcochinase in the yeast *Pichia pastoris* [11]. Purified recombinant dalcochinase exhibits enzymatic properties that are similar to those of natural dalcochinase. This allows us to examine the identity of amino acids conferring substrate specificity in hydrolysis and transglucosylation in dalcochinase and linamarase via site-directed mutagenesis. In this study, eight residues in the aglycone binding pocket of dalcochinase were individually replaced with the corresponding residues of linamarase. While detailed crystal structures of both enzymes are still lacking, this report provides information on the structure-function relationships of dalcochinase, and has implications for other GH1  $\beta$ -glucosidases.

## 2. Experimental

### 2.1. Materials

*Escherichia coli* strain DH5 $\alpha$  was used as a cloning host. *P. pastoris* strain GS115 [*his4*] (Invitrogen, Carlsbad, CA, USA) was used as an expression host. The recombinant plasmid pPICZ-His8-trncTRBG, containing the coding sequence of dalcochinase with an N-terminal truncation following the  $\alpha$  mating factor propeptide and 8 histidine residues [11], was used for the expression of recombinant wild-type dalcochinase. Natural dalcochinase, natural linamarase, Dal-Glc and linamarin were purified from their natural sources as described previously [3,7,12]. *p*-Nitrophenyl- $\beta$ -D-glucoside (pNP-Glc), 2',4'-dinitrophenyl 2-deoxy-2-fluoro- $\beta$ -D-glucoside, 4-methylumbelliferyl- $\beta$ -D-glucoside, and glucose-oxidase/peroxidase mixture were purchased from Sigma Chemical (St. Louis, MO, USA). 2,2'-Azonobis-3-ethylbenz-thiazolinesulfonic acid was purchased from Roche (Mannheim, Germany). Silica gel 60 F254 aluminum sheets for TLC were purchased from Merck (Darmstadt, Germany).

### 2.2. Sequence alignment, homology modeling and molecular docking

Since we have not been able to crystallize natural or recombinant dalcochinase for structure determination, despite efforts over several years, residues likely to be involved in substrate binding were identified by molecular modeling. Alignment of amino acid sequences of *D. cochinchinensis* (Thai rosewood) dalcochinase (gi:6118076), *M. esculenta* (cassava) linamarase (gi:249261), *Zea mays* (maize)  $\beta$ -glucosidase ZmGlu1 (gi:13399866) and *Sorghum bicolor* (sorghum)  $\beta$ -glucosidase Dhr1 (gi:49259431) was performed with ClustalW 2.0.10 [13]. The three-dimensional model of wild-type dalcochinase was created with Geno3D [14], using the structure of maize  $\beta$ -glucosidase 1 complexed with a natural substrate DIMBOA- $\beta$ -D-glucoside (PDB code 1E56A [15]) as a template (45% identity to dalcochinase). The overall structure of the model was checked by the PROCHECK, ProSA, Verify-3D and WHATIF programs [16–19]. The structure of Dal-Glc was built using the Sybyl 7.2 molecular modeling package, and optimized with the Tripos force field with 100 iterations of Simplex followed by Powell minimization algorithm with a 0.05 kcal/mol energy gradient convergence criterion; partial atomic charges were attributed with the Gasteiger–Huckel method (Tripos Inc., St. Louis, MO, USA). The active site pocket of modeled enzymes was determined from the coordinates of DIMBOA- $\beta$ -D-glucoside [15]. Dal-Glc was docked into the active site pocket of modeled enzymes with GOLD 3.1 [20].

### 2.3. Construction of dalcochinase mutants

Eight single mutations were made via site-directed mutagenesis to replace amino acid residues in the aglycone binding pocket of dalcochinase to the corresponding residues of linamarase. These mutations are I185A, N189F, M195V, H253F, V255F, N323Q, A454N and E455I (the numbers indicate their positions in the sequence of dalcochinase as reported in Ref. [5]). The recombinant plasmid pPICZ-His8-trncTRBG [11] was used as a template. Site-directed mutagenesis was performed according to the method published previously [21]. Each 50  $\mu$ L mutagenesis reaction used a sense/antisense mutagenic primer pair (only the sequences of the sense mutagenic primers are shown in Table 1) and three units *Pfu* DNA polymerase (Promega, Madison, WI, USA). The reactions were then incubated with 10 units *DpnI* (Promega) at 37 °C overnight to remove the parental double-stranded DNA. The *DpnI*-treated DNA was transformed into competent *E. coli* by electroporation, and selected on LB-agar plates containing 25  $\mu$ g/mL zeocin at 37 °C. The plasmids containing correctly mutated dalcochinase sequences were identified by DNA sequencing. Subsequently, the mutant plasmids were linearized with *SacI*, transformed into *P. pastoris* by electroporation, and selected on YPDS plates with 100  $\mu$ g/mL zeocin, following the protocols from Invitrogen.

### 2.4. Expression and purification of recombinant dalcochinase in *P. pastoris*

Colonies of *P. pastoris* harboring wild-type and mutant constructs of dalcochinase were grown in BMGH with  $4 \times 10^{-3}$ % (w/v) histidine and  $4 \times 10^{-5}$ % (w/v) biotin overnight at 30 °C, 200 rpm. The culture was changed into BMMH with  $4 \times 10^{-3}$ % (w/v) histidine,  $4 \times 10^{-5}$ % (w/v) biotin and 0.5% (w/v) casamino acid, and induced with 0.5% (v/v) methanol every day. Recombinant dalcochinase (wild-type and mutant) was secreted into the culture media, and its  $\beta$ -glucosidase activity was assayed every 2 days until stable. Then, recombinant dalcochinase was purified from the culture media by hydrophobic interaction chromatography followed by immobilized metal-ion affinity chromatography as described previously [11].

**Table 1**

Sequences of the sense mutagenic primers (5' → 3') used in the generation of eight single mutants. The underlined bases indicated the mutation site. Sequences of the antisense mutagenic primers are the reverse complements of the sequences shown.

Dalcochinase mutant	Primer sequences
I185A	GGATTACACTAAATGAGCCATCAGCTTTACCCGGAATGGGTATGC
N189F	GAGCCATCAATCTTACCCGCGTTTGGGTATGCATACGG
M195V	GGGTATGCATACGGTGTTTTGCACCAAGGTCGATGTTCTCC
H253F	CAGAAAGGTACAATAGGCATTCTTGTGTGTAGTTGGGTATACCGC
V255F	GGCATTCCTTGACAGTATTTGGGTATACCGC
N323Q	GGGTTCATTGATTTTATGGACTACAATATTACACCACTAACTATGCTACC
A454N	GGTCATTGTTGGACAACCTTGAATGGAATGAGGGTTATACATCACG
E455I	CATTGTTGGACAACCTTGAATGGGCTATTGGTTATACATCACGATTGG

Protein concentration was determined by the Bio-Rad Protein Assay (Bio-Rad, Hercules, CA, USA). SDS-PAGE (7.5% resolving gel) was performed to check the purity of protein [22]. Western blot analysis and chemiluminescent detection was done with mouse monoclonal antibody against natural dalcocinase (a gift from Dr. Watchara Kasinrer, Chiangmai University, Thailand), horseradish peroxidase-conjugated rabbit polyclonal antibody against mouse immunoglobulins (Dako, Glostrup, Denmark) and ECL Plus Western Blotting Detection reagents (GE Healthcare, Buckinghamshire, UK). Non-denaturing PAGE and activity staining with 1 mM 4-methylumbelliferyl- $\beta$ -D-glucoside was performed as reported previously [3].

### 2.5. Hydrolytic activities and kinetic measurements

$\beta$ -Glucosidase activity of recombinant dalcocinase (wild-type and mutant) in the culture media and during purification was assayed with 1 mM pNP-Glc in 0.1 M sodium acetate, pH 5.0, in a 0.5 mL reaction at 30 °C for 30 min. The reaction was stopped by adding 1 mL 2 M sodium carbonate, pH 10, and the *p*-nitrophenol released was measured by its absorbance at 400 nm.

To determine the level of enzyme activity of purified enzymes for kinetic measurement and transglucosylation reactions,  $\beta$ -glucosidase activity was assayed with 15 mM pNP-Glc in 0.1 M sodium acetate, pH 5.0, in a 0.1 mL reaction at 30 °C for 5 min. The reaction was stopped by adding 1 mL 2 M sodium carbonate, pH 10, and *p*-nitrophenol released was measured at 400 nm. One unit of activity is defined as the amount of enzyme used to release 1  $\mu$ mol of *p*-nitrophenol in 1 min.

Kinetic parameters for the hydrolysis of pNP-Glc, Dal-Glc and linamarin were determined by incubating 0.1 unit of purified enzymes with substrates at various concentrations (0–45 mM pNP-Glc and linamarin, or 0–25 mM Dal-Glc) in 0.1 M sodium acetate, pH 5.0, in a 50  $\mu$ L reaction at 30 °C for 5 min. The reaction was stopped by boiling for 5 min and immediately cooled to room temperature. Subsequently, the reaction was incubated with 50  $\mu$ L 2 mg/ml 2,2'-Azonobis-3-ethylbenz-thiazolinesulfonic acid and 100  $\mu$ L 5 unit/ml glucose oxidase reagent for 15 min at 37 °C, and glucose released was calculated by comparison of the absorbance at 410 nm to that from a glucose standard curve. Kinetic parameters ( $K_m$ ,  $k_{cat}$  and  $k_{cat}/K_m$ ) were calculated by nonlinear regression of the Michaelis–Menten equation with KaleidaGraph (Synergy Software, Reading, PA, USA).

### 2.6. Transglucosylation activities and rate constants of reactivation by alcohols

Transglucosylation activities of various enzymes were assessed by quantification of alkyl glucoside formation on TLC as described previously [8]. Purified enzymes (0.1 unit) were incubated with 10 mM pNP-Glc (acting as a glucosyl donor) and 0.9 M alkyl alcohols of chain length C<sub>1</sub>–C<sub>4</sub> (acting as a glucosyl acceptor) in 0.1 M sodium acetate, pH 5.0, in a 100  $\mu$ L reaction for 20 h at 30 °C

with shaking. The alcohols tested were methanol, ethanol, 1-propanol, 2-propanol, 2-methyl-1-propanol, 2-methyl-2-propanol, 1-butanol and 2-butanol. The reaction was stopped by boiling for 2 min. Aliquots of the reactions (8  $\mu$ L each) with the same alcohol acceptors were spotted on the same TLC sheet. The TLC was developed twice in ethyl acetate/methanol/water (16:6:1 by vol) for 4.25 cm and twice in 2-propanol/ethanol/water (5:1:2 by vol) for 1.75 cm, and visualized by dipping into 20% (v/v) sulfuric acid in ethanol and incubating at 125 °C for 10 min. All TLC plates contained standard markers, consisting of 20, 40, 60 and 80 nmol of glucose, methyl glucoside and pNP-Glc. The chromatograms were scanned with a GS-8000 imaging densitometer (Bio-Rad) and the amounts of alkyl glucoside products were quantified with Quality One version 4.2.1 software. Standard curves between density and amounts of standard markers with correlation coefficient ( $r^2$ ) of at least 0.95 were used. Results were expressed in percent of moles of alkyl glucoside out of the total moles of free glucose, pNP-Glc and alkyl glucoside present in each reaction.

The rate constants of reactivation of the trapped glucosyl-enzyme intermediates by various alcohols were determined following the methods described previously [23]. Purified enzymes were inactivated with 50  $\mu$ M 2',4'-dinitrophenyl 2-deoxy-2-fluoro- $\beta$ -D-glucoside at 30 °C for 30 min in 0.1 M sodium acetate, pH 5.0, giving a trapped 2-deoxy-2-fluoro-glucosyl-enzyme intermediate. The excess inhibitor was removed by centrifugal ultrafiltration, and fresh buffer was added to make up the same protein concentration. The trapped inactive enzyme intermediate was then reactivated by incubating in 0.1 M sodium acetate, pH 5.0, alone (hydrolysis), or in the presence of 0.5 M alcohols (transglucosylation) at 30 °C. The regained activity was measured by the hydrolysis of 200  $\mu$ M 4-methylumbelliferyl- $\beta$ -D-glucoside, and the release of 4-methylumbelliferone was monitored by a continuous fluorometric spectrophotometer (excitation wavelength = 350 nm and emission wavelength = 450 nm). The observed rate constants of reactivation ( $k_{re,obs}$ ) of each enzyme by a particular alcohol was determined from the slope of the plot of  $\ln[(V_0 - V)/V_0]$  versus time of incubation with alcohol, when  $V$  and  $V_0$  represent the initial velocity of the reaction in the presence and absence of inhibitor, respectively.

## 3. Results and discussion

### 3.1. Sequence alignment, homology modeling and molecular docking

Sequence alignment and structural modeling were performed in order to identify amino acid residues in the aglycone binding pockets that differ between dalcocinase and linamarase. The amino acid sequences of dalcocinase and linamarase were aligned with those of maize  $\beta$ -glucosidase ZmGlu1 and sorghum  $\beta$ -glucosidase Dhr1 (Fig. 2). A three-dimensional homology model of dalcocinase was generated with the crystal structure of maize  $\beta$ -glucosidase ZmGlu1 in complex with its natural substrate DIMBOA- $\beta$ -D-

TRDC	-----IDFAKEVRETI-----TEVPP-FNRSFCPSDFIFGTASSSYQYEGE---GRVP	44
CVLM	MLVLFISLLALTRPAMGTDDDDNIPDDFSRKYFPDDFIFGTATSAYQIEGEATAKGRAP	60
ZmGlu1	--SARVGSQ-NGVQMLSP----SEIPQ---RDWFPSDFTFGAATSAYQIEGAWNEDGKGP	50
SbDhr1	--AQTTSSSAGIHRLLSP----WEIPR---RDWFPPSFLFGAATSAYQIEGAWNEDGKGP	51
TRDC	SIWDFNTHQYPEKIADRSNGDVAVDQFHRYKKDIAIMKDMNLDAYRMSISWPRILPTGRV	104
CVLM	SVWDIFSKEPTDRI LDGSDGDAVDVFNRYIQDKNVKMGFNAFRMSISWSRVIPSGRR	120
ZmGlu1	SNWDHFCNHNPERILDGSDSDIGANSYHMYKTDRVLLKEMGMDAYRFSISWPRILPKGTK	110
SbDhr1	STWDHFCNHNFEWIVDRSNGDVAADS YHMYAEDVRLLKEMGMDAYRFSISWPRILPKGTL	111
TRDC	SGGINQTVGDYNNRLINESLANGITPFVTIFHWDLPQALEDEYGGFLNHS---VVNDFQD	161
CVLM	REGVNEEGIQFYNDVINEIISNGLEPFVTIFHWDTPQALQDKYGGFLSRD---IVYDYLQ	177
ZmGlu1	EGGINPDGKYYRNLIINLLLENGIEPYVTIFHWDVPQALEEKYGGFLDKSHKSIVEDYTY	170
SbDhr1	AGGINEKGVEYYNKLIDLLLENGIEPYITIFHWDTPQALVEAYGGFLDER---I IKDYTD	168
TRDC	YADLCFQLFGDRVKHWITLNEPSIFTANGYAYGMFAPGRCSPSYNPTCTGGDAGTETYL	221
CVLM	YADLLFERFGDRVKPMMTFNEPSAYVGFADDDGVFAPGRCSWVNQRCLAGDSATEPYIV	237
ZmGlu1	FAKVCDFNFGDKVKWNLTFTNEPQTFTSFSYGTGVFAPGRCSPLDCAYPTGNSLVEPYTA	230
SbDhr1	FAKVCFEKFGKTVKNWLTFTNEPETFCVSYSYGTGVLPAPGRCSPGVSCAVPTGNSLSEPYIV	228
TRDC	AHNLLSHAATVQVYKRKYQEHQKGTIGISLHVWVVIPLSNSTDQNATQRYLDFTCGWF	281
CVLM	AHNLLSHAAAVHQYRKYQGTQKKGIGITLFTFWYEPLSDSKVDVQAAKTALDFMFLGW	297
ZmGlu1	GHNILLAAEAVDLYNKHYKRDD-TRIGLAFDVMGRVPYGTSLDKQAERSWDINLWGF	289
SbDhr1	AHNLLRAHAETVDIYNKYHKGAD-GRIGLALNVFGRVPYNTFLDQQAQERSMDKCLGW	287
TRDC	MDPLTAGRYPDSMQYLVGDRLPKFTTDQAKLVKGSFDFIGLNYTTNYATKSDASTCCPP	341
CVLM	MDPMTYGRYPRTMVDLAGDKLIGFTDEESQLLRGSYDFVGLQYTYAYAEPIPPVDPKFR	357
ZmGlu1	LEPVVRGDYFFSMRSLARERLPFFKDEQKEKLAGSYNMLGLNYTSTRFSKNIDISPNYSP	349
SbDhr1	LEPVVRGDYFFSMRVSARDRVYPFKEKEQEKLVGSYDMIGINYTTSTFSKHIDLSFNNSP	347
TRDC	SYLTDPQVTLQ--QRNGVFIGPVTSPGWMCIYPKGLRDLLEYFKEKYNPLVYITENG	399
CVLM	RYKTDGSGVNATPY-DLNGNLIGPQAYSSWFYIFPKGIRHFLNYTKDITYNDPVIYVTE	416
ZmGlu1	VLNTDDAYASQEVNPGDPKPIGPPMGNPWYIMYPEGLKDLMLMKNKYGNPPIYITENG	409
SbDhr1	VLNTDDAYASQETKGPDGNAIGPPTGNAWNIMYPKGLHDILMTMKNKYGNPPMYITENG	407
TRDC	DEKNDAS--LSLEESLIDTYRIDSYRHLFYVRYAIR-SGANVKGFFAWSLLDNFEWAEG	456
CVLM	DNYNNS--QPIEEALQDDFRISYKHKHMNALGSLKNYGKLVKGYFAWSYLDNFEWNIG	474
ZmGlu1	GDVDTKETPLPMEALNDYKRLDYIQRHIATLKESID-LGSNVQGYFAWSLLDNFEWFAG	468
SbDhr1	GDIDKGD--LPKPVALEDHTRLDYIQRHLSVLKQSID-LGADVGRGYFAWSLLDNFEWSSG	464
TRDC	YTSRFGLYFVNYTT-LNRYPKLSATWFKYFLARDQESAKLEILAPKARWSLSTMIKEET	515
CVLM	YTSRFGLYYVDYKNNLTRYPKKSAHWFTKFLNISVNANNIYELTSKDSRKVGKFFVVM---	531
ZmGlu1	FTERYGIVYVDRNNNCTRYMKESAKWLKEFN-TAKKP-SKKILTTPA-----	512
SbDhr1	YTERFGIVYVDRENGCTMRKRSARWLQEFNGAAKKVENNKILTAPAGQLN-----	514
TRDC	KPKRGIEGF	524
CVLM	-----	
ZmGlu1	-----	
SbDhr1	-----	

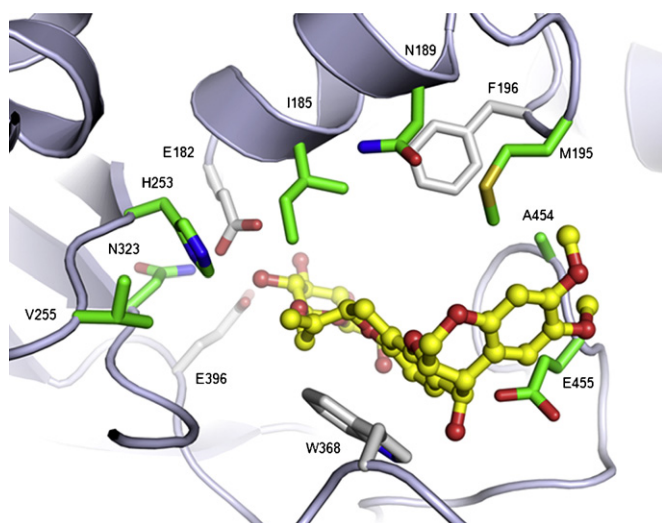
**Fig. 2.** Alignment of amino acid sequences of Thai rosewood dalcocinase, cassava linamarase and maize  $\beta$ -glucosidase. The alignment is generated with ClustalW 2.0.10 using the reported sequences of Thai rosewood dalcocinase (TRDC, gi:6118076), cassava linamarase (CVLM, gi:249261) maize  $\beta$ -glucosidase Glu1 (ZmGlu1, gi:13399866) and sorghum  $\beta$ -glucosidase Dhr1 (SbDhr1, gi:49259431). The sequence of Thai rosewood dalcocinase is shown without the first 23 N-terminal residues, MLAMTSKAIIILGLLALVSTAS, predicted to be a signal peptide [5]. The highly conserved (L/F)NEP and (I/V)TENG motifs are underlined, and the two active site glutamate residues are shown in BOLD. The residues that were targeted for site-directed mutagenesis are marked with ▼ above the sequences. The residues shown to be located in the active site or involved in aglycone binding in dalcocinase, cassava linamarase, maize  $\beta$ -glucosidase Glu1 and sorghum  $\beta$ -glucosidase Dhr1 are marked with ■, □, ● and ○ below the sequences, respectively [15,24–26].

glucoside (PDB code 1E56A) as a template. This structure showed the highest identity (45%) to dalcocinase among all structures of  $\beta$ -glucosidases with complexed natural substrates, which would simplify subsequent docking studies.

The quality of the homology model was evaluated by the PROCHECK, ProSA, Verify-3D and WHATIF programs. PROCHECK showed 97% of the residues in the homology model were located in the most favorable or allowed regions of the Ramachandran plot and only 1.5% of them in the disallowed region (compared with 98.5% and 0%, respectively, for the template structure). The G-factor values in PROCHECK (in which a score above  $-0.5$  indicates a reliable model) were obtained to be 0.22 and 0.34 for the model and the template structures, respectively. The ProSA z-scores

of the selected dalcocinase model and the template structure were  $-8.15$  and  $-9.63$ , respectively, which were within the acceptable range. Assessment of the compatibility of the residues with their environment by Verify-3D (in which a score over 0.2 were considered reliable) showed 91.2% of residues in the model had a score over 0.2, compared with 93.9% for the template. The packing quality of each residue was also evaluated by the WHATIF program (in which a score above  $-5.0$  is deemed satisfactory). In this case, the selected dalcocinase model showed an average score of  $-1.357$  (95.7% of residues with scores above  $-5$ ), while the template structure obtained an average score of  $-0.774$  (98.0% of residues with scores above  $-5$ ). The root mean square deviation of the C $\alpha$  atoms was calculated to be  $1.68 \text{ \AA}$  with respect to the





**Fig. 3.** Three-dimensional model of the active site pocket of dalcochinase with a docked molecule of Dal-Glc. The homology model of wild-type dalcochinase was created with Geno3D [14], using the structure of maize  $\beta$ -glucosidase 1 complexed with a natural substrate DIMBOA- $\beta$ -D-glucoside (PDB code 1E56A [15]) as a template. The model is shown as viewed from the exterior of the enzyme, looking into the active site pocket, with the glucose moiety buried behind the structure of the aglycone moiety. The main chain conformation is shown in gray ribbon. The side chains that are targeted for site-directed mutagenesis are shown as green stick models. Dal-Glc is shown as yellow ball-and-stick model. The catalytic residues (E182 and E396) and the conserved residues (F196 and W368) are shown in gray. The picture was generated with PyMOL version 0.99 (DeLano Scientific, Palo Alto, CA, USA).

template. All evaluations suggest that a reliable homology model for Thai rosewood dalcochinase has been obtained for subsequent docking study. The model showed an  $(\alpha/\beta)_8$  barrel structure as expected for GH1 enzymes (not shown). The two catalytically active site residues (E182 and E396 in dalcochinase) were separated by about 5 Å (as measured between the O $\epsilon$ 1 atoms of both residues) and were located near the bottom of the binding pocket, which appeared as a 25 Å long and 7 Å wide slot-like pocket.

The structure of Dal-Glc, the natural substrate of dalcochinase, was then docked into the active site pocket of the homology model of dalcochinase (Fig. 3). A fitness score of docked Dal-Glc into the dalcochinase model was obtained to be 61.41. Re-docking DIMBOA- $\beta$ -D-glucoside into the active site pocket of maize  $\beta$ -glucosidase ZmGlu1 was performed as a control docking, and gave a fitness score of 69.15. In the aglycone binding pocket of the homology model, the position of the Dal-Glc was directed mainly through stacking  $\pi$ - $\pi$  interactions between ring B (phenyl ring) of Dal-Glc and the indole ring of the conserved W368. There appeared to be a hydrophobic interaction of Dal-Glc with I185. Also, the methylene group at C-7' of Dal-Glc appeared to fit between H253 and W368 of dalcochinase. Further inside the active site pocket, N323 interacted with the carboxyl groups of both acid/base catalyst (E182) and nucleophile (E396). Q339, the corresponding position in linamarase, has been shown via site-directed mutagenesis to coordinate the catalytic diad [24]. At the mouth of the pocket, Rings D and E of Dal-Glc appeared almost perpendicular to rings A, B and C, and were exposed to solvent, with no apparent specific interaction with any residues of dalcochinase. The distances between Dal-Glc and amino acid residues in binding pocket of dalcochinase model as predicted by molecular docking were summarized in Table 2.

From the sequence alignment and structural modeling, eight amino acid residues in the aglycone binding pocket of dalcochinase, namely I185, N189, M195, H253, V255, N323, A454 and E455, were chosen as targets for site-directed mutagenesis to investigate their roles in directing substrate specificity. Each residue in dalcochinase would be replaced with the corresponding residue in

linamarase, namely A201, F205, V211, F269, F271, Q339, N472 and I473, respectively. Amino acid residues corresponding to all of these positions in other GH1 enzymes, except M195 in dalcochinase, have been previously shown to be involved in enzyme–substrate interactions. Crystal structures of maize  $\beta$ -glucosidase ZmGlu1 [15] and sorghum  $\beta$ -glucosidase Dhr1 [25] in complex with their natural substrates, DIMBOA- $\beta$ -D-glucoside and dhurrin, respectively, have shown the amino acid residues corresponding to N189, H253, V255, A454 and E455 of dalcochinase included in this study, to be in direct contact with the aglycone groups of the substrates. Previous mutational analyses in linamarase [24] and dalcochinase [26] have identified the amino acid residues corresponding to I185, H253, N323, A454 and E455 of dalcochinase included in this study, as essential active site residues for catalysis and substrate specificity. The residue corresponding to M195 of dalcochinase, which falls within the motif GM/VFAPGRCS between residues 194 and 202 of dalcochinase, is replaced with V in linamarase, maize  $\beta$ -glucosidase ZmGlu1 and sorghum  $\beta$ -glucosidase Dhr1. As the side chain of M195 pointed toward the aglycone group of Dal-Glc in our model, its mutation to V was included in this study.

### 3.2. Construction and production of dalcochinase mutants

Eight single mutants of dalcochinase, which are I185A, N189F, M195V, H253F, V255F, N323Q, A454N and E455I, were constructed via site-directed mutagenesis. All wild-type and mutant forms of dalcochinase were expressed in *P. pastoris*, with levels of total  $\beta$ -glucosidase activity in culture media at the time of harvest varying between 9.2 units (I185A) and 102 units (M195V) per 1-L culture. Purification yields were between 15 and 40%, and specific activities ranged from 0.15 unit/mg (I185A) to 1.30 unit/mg (H253F).

All wild-type and mutant forms of dalcochinase showed a broad band with an apparent molecular mass of about 66 kDa on SDS-PAGE (Supplementary data 1, left panels), which migrated more slowly than the apparent 63 kDa band of natural dalcochinase. The broadening effect and the greater apparent molecular weight of recombinant dalcochinase are most likely due to greater glycosylation in yeast than in plants [27]. Indeed, we have previously shown that the natural and recombinant wild-type dalcochinase contained approximately 13.6 and 16.1% sugar (by weight), respectively [11]. All wild-type and mutant forms of dalcochinase could be detected equally well by a mouse monoclonal antibody against natural dalcochinase (Supplementary data 1, right panels), and migrated to the same distance on non-denaturing gels (Supplementary data 2, left panels), suggesting that the overall conformation of the wild-type enzyme was retained in the mutants. However, they exhibited variable activities toward 1 mM 4-methylumbelliferyl- $\beta$ -D-glucoside

**Table 2**

Distances between Dal-Glc and amino acid residues in binding pocket of dalcochinase model as predicted by molecular docking.

Dal-Glc position	Amino acid position	Distance (Å)
Sugar ring		
O2	Y325-HH	2.49
O3	N181-HD22	2.89
O4	W453-He1	1.48
O6	W445-He1	2.16
H at O2	E396-O $\epsilon$ 1	1.72
C'		
Methylene	H235-He1	2.34
Methylene	W368-Hz3	2.51
Ring-A		
H	I185-HD11	2.41
Ring-B		
Ring-B	W368	$\pi$ - $\pi$ (3.52–4.09)
Ring-C		
OH	W368-He1	2.35

**Table 3**  
Kinetic parameters for the hydrolysis of pNP-Glc and Dal-Glc of wild-type and mutant forms of dalcocinase and linamarase.  $k_{\text{cat}}$  values were estimated assuming a subunit molecular weight of 66,000.

Enzymes	pNP-Glc			Dal-Glc		
	$K_m$ (mM)	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$k_{\text{cat}}/K_m$ ( $\text{s}^{-1}\text{mM}^{-1}$ )	$K_m$ (mM)	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$k_{\text{cat}}/K_m$ ( $\text{s}^{-1}\text{mM}^{-1}$ )
Wild-type dalcocinase	$3.91 \pm 0.31$	$237.0 \pm 5.2$	$60.6 \pm 5.0$	$5.67 \pm 0.33$	$62.3 \pm 1.7$	$11.0 \pm 0.7$
Natural linamarase	$0.72 \pm 0.04$	$131.9 \pm 1.3$	$183.2 \pm 10.3$	N.D. <sup>a</sup>	N.D. <sup>a</sup>	N.D. <sup>a</sup>
I185A	$14.33 \pm 1.07$	$12.3 \pm 0.4$	$0.86 \pm 0.07$	$1.62 \pm 0.21$	$0.74 \pm 0.04$	$0.46 \pm 0.06$
N189F	$0.59 \pm 0.07$	$172.5 \pm 2.9$	$292.4 \pm 35.0$	$0.63 \pm 0.05$	$36.0 \pm 0.9$	$57.1 \pm 4.7$
M195V	$1.77 \pm 0.16$	$36.4 \pm 0.6$	$20.6 \pm 1.9$	$6.25 \pm 0.76$	$31.4 \pm 2.2$	$5.0 \pm 0.7$
H253F	$2.69 \pm 0.30$	$58.0 \pm 1.7$	$21.6 \pm 2.5$	$4.35 \pm 0.42$	$21.9 \pm 0.9$	$5.0 \pm 0.5$
V255F	$2.67 \pm 0.17$	$10.6 \pm 0.1$	$4.0 \pm 0.3$	$1.74 \pm 0.16$	$1.5 \pm 0.1$	$0.89 \pm 0.09$
N323Q	$1.84 \pm 0.18$	$36.2 \pm 0.6$	$19.7 \pm 2.0$	$4.94 \pm 0.49$	$33.5 \pm 1.5$	$6.8 \pm 0.7$
A454N	$1.34 \pm 0.13$	$157.9 \pm 3.1$	$117.8 \pm 11.7$	$3.63 \pm 0.60$	$14.7 \pm 1.4$	$4.0 \pm 0.8$
E455I	$11.83 \pm 0.88$	$49.8 \pm 1.3$	$4.2 \pm 0.3$	$4.69 \pm 0.45$	$120.7 \pm 5.2$	$25.7 \pm 2.7$

<sup>a</sup> N.D., no activity detected.

in activity staining of non-denaturing gels (Supplementary data 2, right panels), reflecting altered activities as a result of mutations.

### 3.3. Kinetic measurement for hydrolytic activities

The kinetic properties of recombinant wild-type dalcocinase, dalcocinase mutants and natural linamarase for the hydrolysis of pNP-Glc and Dal-Glc were determined (Table 3). The recombinant wild-type dalcocinase showed kinetic properties toward pNP-Glc and Dal-Glc that are similar to those reported for natural dalcocinase [3,7]. Cassava linamarase hydrolyzed pNP-Glc with a 6-fold lower  $K_m$  and a 2-fold lower  $k_{\text{cat}}$ , resulting in a 3-fold higher efficiency compared with the recombinant wild-type dalcocinase, suggesting that residues in the substrate binding pocket of linamarase are more suited for the binding of pNP-Glc than those in dalcocinase. Only linamarase could hydrolyze linamarin with a  $K_m$  of  $0.84 \pm 0.07$  mM, a  $k_{\text{cat}}$  of  $79.8 \pm 1.6 \text{ s}^{-1}$ , and a  $k_{\text{cat}}/K_m$  of  $95 \pm 8 \text{ s}^{-1} \text{ mM}^{-1}$ . In agreement with previous studies, neither the recombinant wild-type dalcocinase could appreciably hydrolyze linamarin, nor could linamarase hydrolyze Dal-Glc [7].

In hydrolysis of pNP-Glc, the I185A and E455I mutations showed 3-fold increases in  $K_m$ , and 70- and 14-fold decreases in  $k_{\text{cat}}/K_m$ , respectively, compared with the recombinant wild-type enzyme. Presumably, I185 and E455 in dalcocinase could form hydrophobic and ionic interactions to the pNP moiety of pNP-Glc, respectively, so they were more favorable for the binding and hydrolysis of pNP-Glc than the corresponding positions in linamarase, A201 and I473. On the other hand, the N189F and A454N mutations yielded 5- and 2-fold increases in  $k_{\text{cat}}/K_m$ , respectively, resulting from 7- to 3-fold decreases in  $K_m$ , respectively, and a 1.5-fold decrease in  $k_{\text{cat}}$ , compared with the recombinant wild-type enzyme. So their corresponding positions in linamarase, F205 and N472, may contribute to better binding to pNP-Glc compared with dalcocinase by offering suitable hydrophobic and polar interactions with the substrate, respectively. Four other mutations in dalcocinase resulted in lower hydrolytic efficiencies (from 3- to 15-fold) compared with the recombinant wild-type enzyme, mostly due to 4- to 22-fold reductions in the  $k_{\text{cat}}$ . Among all mutants, V255F showed the greatest decrease in  $k_{\text{cat}}$  (22-fold) compared with the recombinant wild-type enzyme, but little change in  $K_m$ , suggesting that F271 in linamarase could disturb the hydrolysis but not binding of pNP-Glc.

In hydrolysis of Dal-Glc, the natural substrate of dalcocinase, all dalcocinase mutants exhibited less than an order of magnitude reduction in  $K_m$  compared to the recombinant wild-type enzyme, indicating small effects of mutations on the binding properties. On the other hand, two mutants, I185A and V255F, showed 84- and 40-fold decreases in  $k_{\text{cat}}$  compared with the recombinant wild-type dalcocinase, resulting in 24- and 12-fold decreases

in  $k_{\text{cat}}/K_m$ , respectively. So, the reduction in substrate specificity toward Dal-Glc by these two mutations, as judged by the values of  $k_{\text{cat}}/K_m$ , appeared to result from reduced catalytic activity rather than changes in binding affinity. These two mutations also showed large decreases in the  $k_{\text{cat}}$  for the hydrolysis of pNP-Glc. A201 has been shown to be in the active site of linamarase [24], and here we showed the role of the corresponding position in dalcocinase, I185, in directing substrate specificity. Molecular modeling of Dal-Glc in the active site pocket of dalcocinase suggested a hydrophobic contact between side chain of I185 and the aglycone moiety of Dal-Glc (Fig. 3). The I185A mutation could create a void space in the substrate binding pocket of dalcocinase leading to incorrect positioning of the aglycone moiety and placing the glucosidic bond in an unfavorable orientation for hydrolysis. Similar reasons were proposed for the reduction in  $k_{\text{cat}}$  of F193A and F193I mutations in Zm-p60.1 maize  $\beta$ -glucosidase, corresponding to position N189 of dalcocinase in this study, for the hydrolysis of pNP-Glc [28,29]. In addition, the I185A mutation resulted in a decreased hydrophobic environment close to the catalytic site that might affect the  $\text{pK}_a$ s of the catalytic glutamates. Indeed, an A201V mutation (increasing hydrophobicity) in linamarase, corresponding to position I185 of dalcocinase in this study, increased the value of  $\text{pK}_{\text{E}2}$  from 7.22 to 7.44 for the hydrolysis of pNP-Glc [24]. The possibilities of mis-positioning of the glucosidic bond for hydrolysis and changing the  $\text{pK}_a$ s of the catalytic glutamates may also account for the reduction in  $k_{\text{cat}}$  caused by the V255F mutation.

Other interesting mutants, N189F and E455I, showed unexpected 5- and 2-fold increases in substrate specificity for Dal-Glc, respectively, as judged by the values of  $k_{\text{cat}}/K_m$ , compared with the recombinant wild-type dalcocinase. The increase in  $k_{\text{cat}}/K_m$  of the N189F mutant was due to a 9-fold reduction in  $K_m$ , suggesting favorable hydrophobic interaction between phenylalanine and Dal-Glc at this position. On the other hand, the increase in  $k_{\text{cat}}/K_m$  of the E455I mutant was due to a 2-fold increase in  $k_{\text{cat}}$ . So, it appeared that increased hydrophobicity as a result of isoleucine substitution at this position may lead to better placement of the substrate for hydrolysis. Four other mutations led to 2- to 3-fold reductions in  $k_{\text{cat}}/K_m$ , as a result of 2- to 4-fold reductions in  $k_{\text{cat}}$  but small changes in  $K_m$ , compared with the recombinant wild-type enzyme, indicating that their substitutions affected catalysis rather than binding. However, it should be noted that a reduction in  $K_m$  may not always correspond to improved substrate binding, but may also result from a reduced efficiency of the deglycosylation step by water in a double-displacement mechanism.

However, none of single mutants of dalcocinase could hydrolyze linamarin. Combinations of some of these mutations in the aglycone binding pocket may offer cumulative effects that are needed to overcome steric hindrance imposed by the three substituents of the 1-cyano-1-methylethyl moiety of linamarin. Also,

**Table 4**

Mole percent of alkyl glucoside products from transglucosylation reactions of wild-type and mutant forms of dalcocinase and linamarase. Each value is an average from duplicate TLC plates. Alkyl glucosides were quantitated using methyl glucoside as standard.

Enzymes	Primary alcohols					Secondary alcohols		Tertiary alcohol
	Methanol	Ethanol	1-Propanol	1-Butanol	2-Methyl-1-propanol	2-Propanol	2-Butanol	2-Methyl-2-propanol
Wild-type dalcocinase	22	42	68	79	71	4	12	N.D. <sup>a</sup>
Natural linamarase	52	51	73	71	74	60	65	35
I185A	20	40	73	92	90	18	25	N.D. <sup>a</sup>
N189F	43	64	86	96	90	7	11	N.D. <sup>a</sup>
M195V	24	46	68	90	71	3	10	N.D. <sup>a</sup>
H253F	29	44	66	87	75	6	11	N.D. <sup>a</sup>
V255F	40	54	80	92	84	15	21	N.D. <sup>a</sup>
N323Q	29	43	69	88	74	6	12	N.D. <sup>a</sup>
A454N	29	50	68	85	75	4	15	N.D. <sup>a</sup>
E455I	26	40	67	87	80	7	10	N.D. <sup>a</sup>

<sup>a</sup> N.D., no product detected.

other amino acid residues further away from the aglycone binding pocket may contribute to favorable environment for the hydrolysis of linamarin.

#### 3.4. Transglucosylation activities and rate constants of reactivation by alcohols

In the transglucosylation reactions, the enzyme cleaves pNP-Glc to form a glucosyl-enzyme intermediate complex and transfers the glucose to an alcohol acceptor, yielding alkyl glucoside, detectable as a spot on TLC. Free glucose could result from transfer of glucose to water, or secondary hydrolysis of alkyl glucoside products. The spots of alkyl glucoside products and glucose varied in size and intensity. An example of TLC from transglucosylation reactions with methanol as acceptors is shown in [Supplementary data 3](#). It should be noted that transglucosylation is a kinetic process, and the amounts of alkyl glucoside product may vary with time. However, previous studies showed that in reactions of dalcocinase, hydrolysis of pNP-Glc took place rapidly with little hydrolysis of alkyl glucosides over 24 h [8]. So, in this study, only single-point measurements of transglucosylation yield were performed in order to provide quick assessment of the transglucosylation activities of dalcocinase mutants toward various alcohols, using an incubation time of 20 h to maximize cleavage of pNP-Glc donor and minimize secondary hydrolysis of alkyl glucoside products. The mole percents of alkyl glucoside products from all transglucosylation reactions are shown in [Table 4](#). The recombinant wild-type dalcocinase showed transglucosylation activities to primary and secondary alcohol acceptors that are comparable to those previously reported for the natural enzyme [8]. In agreement with previous studies, linamarase was the only enzyme that could transfer glucose to the tested primary, secondary as well as tertiary alcohol acceptors [9,23]. However, our product yields from linamarase were lower than those previously reported [9], probably due to lower starting concentrations of pNP-Glc and alcohols.

For transglucosylation reactions to primary alcohols, only the N189F and V255F mutants showed improved product yields compared with the recombinant wild-type dalcocinase with all primary alcohols tested, while the I185A mutant gave better yields only with longer chain primary alcohols. For transglucosylation to secondary alcohols, only the I185A and V255F mutants were better than the wild-type enzyme. However, none of our mutants could use a tertiary alcohol as a glucosyl acceptor, indicating that single mutations were insufficient to change the architecture of the aglycone binding pocket of dalcocinase to accept aglycone groups with three substituents at the carbinol carbon, such as the 1-cyano-1-methylethyl moiety of linamarin and 2-methyl-2-propanol. Nonetheless, our transglucosylation results

using primary and secondary alcohols suggested that transglucosylation may be improved by creating a more hydrophobic environment (in the case of the N189F mutation), and providing suitable geometry for enzyme–substrate interactions (in the cases of the I185A and V255F mutations). Thus, the positions A201, F205 and F271 in linamarase were favorable for transglucosylation activity. This is in agreement with their significant roles in substrate binding in cassava linamarase, maize  $\beta$ -glucosidase ZmGlu1 and sorghum  $\beta$ -glucosidase Dhr1, respectively [15,24,25].

To evaluate the transglucosylation activities more accurately, the method of Hommalai et al. [23] was applied to test the I185A, N189F and V255F mutants that showed higher transglucosylation activities than the recombinant wild-type enzyme. The trapped 2-deoxy-2-fluoro-glucosyl-enzyme intermediates were reactivated by the transfer of a 2-deoxy-2-fluoro-glucosyl moiety from the enzymes to the alcohol acceptors. In all cases, the plots of  $\ln[(V_0 - V)/V_0]$  versus time of incubation with alcohol were linear, suggesting pseudo-first-order reactions (an example is shown in [Supplementary data 4](#)). The observed rate constants of enzyme reactivation ( $k_{re,obs}$ ) in the absence and presence of alcohol represented the progress of the hydrolysis and transglucosylation reactions, respectively ([Table 5](#)).

Under the conditions used, our values of  $k_{re,obs}$  for the recombinant wild-type dalcocinase were similar to those obtained previously for the natural enzyme with all alcohols tested, except for 1-butanol and 2-methyl-1-propanol, where our values were about half of those reported [23]. Of the dalcocinase mutants, the I185A mutant gave statistically significant increases in the values of  $k_{re,obs}$  compared with the recombinant wild-type enzyme with only primary alcohol acceptors and 2-propanol, while the V255F mutant showed significant increases in the values of  $k_{re,obs}$  with only 1-butanol and 2-methyl-1-propanol.

However, the N189F mutant showed statistically significant increases in the values of  $k_{re,obs}$  with all alcohols tested compared with the recombinant wild-type enzyme. Surprisingly, this mutant also gave a 3-fold increase in the value of  $k_{re,obs}$  with 2-methyl-2-propanol, compared to its own hydrolysis reaction and to the recombinant wild-type dalcocinase, indicating improved transglucosylation activity to tertiary alcohol acceptor. This lower effectiveness of water compared to alcohol in cleaving the glucosyl-enzyme, might also be partly responsible for the lower  $K_m$  value observed in hydrolysis of pNP-Glc and Dal-Glc with the N189F mutant. However, the  $k_{re,obs}$  with 2-methyl-2-propanol was still very much lower than that reported for cassava linamarase ( $47 \pm 6 \times 10^{-4} \text{ min}^{-1}$ ) [23], which was 8-fold faster than simple hydrolysis. The transglucosylation rate constants and dissociation constants of alcohols were not determined in this study since most  $k_{re,obs}$  values were less than 10-fold different from the rates of hydrolysis.



**Table 5**  
Observed rate constants of enzyme reactivation ( $k_{\text{re,obs}} \times 10^{-4} \text{ min}^{-1}$ ) of wild-type and mutant forms of dalcocinase for a series of 0.5 M alcohols. The reported values are averages from three independent experiments with standard deviations.

Alcohols	Wild-type dalcocinase	I185A	N189F	V255F
Primary alcohols				
Methanol	2.4 ± 0.3	3.7 ± 0.2 <sup>a</sup>	4.7 ± 0.5 <sup>a</sup>	1.7 ± 0.3
Ethanol	3.1 ± 0.2	3.6 ± 0.1 <sup>a</sup>	6.9 ± 0.9 <sup>a</sup>	1.6 ± 0.5
1-Propanol	4.6 ± 0.8	7.3 ± 0.2 <sup>a</sup>	36.0 ± 7.8 <sup>a</sup>	3.9 ± 0.8
1-Butanol	6.3 ± 1.1	9.7 ± 0.8 <sup>a</sup>	98.9 ± 6.5 <sup>a</sup>	7.5 ± 0.3 <sup>a</sup>
2-Methyl-1-propanol	5.4 ± 1.0	20.4 ± 2.4 <sup>a</sup>	84.9 ± 4.7 <sup>a</sup>	7.7 ± 1.4 <sup>a</sup>
Secondary alcohols				
2-Propanol	2.5 ± 0.5	3.6 ± 0.2 <sup>a</sup>	8.1 ± 1.0 <sup>a</sup>	1.5 ± 0.4
2-Butanol	2.4 ± 0.3	2.7 ± 0.1	12.2 ± 2.1 <sup>a</sup>	2.5 ± 0.3
Tertiary alcohol				
2-Methyl-2-propanol	2.7 ± 0.3	3.1 ± 0.3	8.9 ± 1.3 <sup>a</sup>	2.3 ± 0.3
Control (no acceptor)	2.2 ± 0.1	2.6 ± 0.3	3.4 ± 0.7 <sup>a</sup>	1.3 ± 0.4

<sup>a</sup> Statistically significant increase compared to the recombinant wild-type dalcocinase ( $p$ -value  $\leq 0.05$ ).

Discrepancies between the TLC results and the values of  $k_{\text{re,obs}}$  may partly be due to secondary product hydrolysis that occurred during the 20-h incubation in the former method. In some reactions, the mutant enzymes gave lower alkyl glucoside yields, but exhibited significantly higher  $k_{\text{re,obs}}$  values compared with the recombinant wild-type dalcocinase, suggesting better alkyl glucoside cleavage by the mutant than the wild-type enzyme. On the other hand, in some reactions, the mutant enzymes gave higher alkyl glucoside yields, but exhibited similar or lower  $k_{\text{re,obs}}$  values compared with the recombinant wild-type dalcocinase, suggesting lower alkyl glucoside cleavage by the mutant than the wild-type enzyme.

Since the proton removal in the deglycosylation step appeared to be the major step governing transition state formation in the transglucosylation reaction of linamarase [23], we also used 1,1,1-trifluoromethyl-2-propanol, which has a similar structure to 2-methyl-2-propanol but a lower  $\text{pK}_a$  value (11.6 versus 18.0), to reactivate the trapped 2-deoxy-2-fluoro-glucosyl-enzyme intermediates of all enzymes. However, the resulting  $k_{\text{re,obs}}$  values of the wild-type dalcocinase and all three mutants did not differ significantly from simple hydrolysis ( $\text{pK}_a$  of water  $\sim 15.7$ ) (results not shown). So, the nucleophilic strength appeared less important than the complementarity of size, shape and the hydrophobicity of the incoming alcohol acceptor and the aglycone binding pocket of the enzyme.

#### 4. Conclusions

Taken together, our mutational analysis in the aglycone binding pocket of dalcocinase has identified the amino acid residues that may help explain the distinct aglycone specificities in hydrolysis and transglucosylation. In particular, residues I185 and V255 of dalcocinase play important roles in substrate specificity as judged by decreases in the  $k_{\text{cat}}/K_m$  ratios of their mutants for hydrolyzing Dal-Glc. On the other hand, residues N189 and E455 of dalcocinase may not be optimal for Dal-Glc hydrolysis, as their mutations led to increases in the  $k_{\text{cat}}/K_m$  ratios. Since both side chains were polar, our results suggested that these positions might prefer hydrophobic residues. Alkyl glucoside synthesis by transglucosylation was improved by the I185A, N189F and V255F mutations, suggesting the positive roles played by the corresponding residues, A201, F205 and F271 of linamarase. However, only the N189F mutant showed significant increases in the reactivation rate constants with all alcohols, compared to the wild-type enzyme, while the other two mutants showed improvement only with longer-chain primary alcohols. So it appeared that the yield of alkyl glucoside products depended on the efficiencies of both glucose transfer and secondary product cleavage. Our results suggest that substrate specificity in hydrolysis and transglucosylation is accomplished as a result of

complementarity in both geometry and hydrophobicity of the substrate and the binding pocket of enzyme.

#### Acknowledgements

The authors especially thank Palangpon Kongsaree for help in the preparations of Dal-Glc and linamarin, Patchreenart Saparpakorn for help with modeling programs, and Watchara Kasinrerk for providing the mouse monoclonal antibody against natural dalcocinase. The comments and suggestions from James Ketudat-Cairns during the preparation of manuscript are much appreciated. The project was financially supported by the Thailand Research Fund and the Commission on Higher Education (Grant numbers MRG4980131 to P.T.K. and MRG485S015 to P.S.), and by the Faculty of Science, Kasetsart University. K.R. and K.C. are recipients of the Ph.D. scholarships from the Commission on Higher Education, Thailand.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.molcatb.2010.09.003.

#### References

- [1] A. Esen, in: A. Esen (Ed.),  $\beta$ -Glucosidases Biochemistry and Molecular Biology, American Chemical Society, Washington, DC, 1993, pp. 1–14.
- [2] B. Henrissat, G.D. Davies, *Plant Physiol.* 124 (2000) 1515–1519.
- [3] C. Srisomsap, J. Svasti, R. Surarit, V. Champattanachai, K. Boonpuan, P. Sawan-gareetrakul, P. Subhasitanont, D. Chokchaichamnankit, *J. Biochem.* 119 (1996) 585–590.
- [4] Y. Eksittikul, M. Chulavatanatol, *Arch. Biochem. Biophys.* 266 (1988) 263–269.
- [5] J.R. Ketudat Cairns, V. Champattanachai, C. Srisomsap, B. Wittman-Liebold, B. Thiede, J. Svasti, *J. Biochem.* 128 (2000) 999–1008.
- [6] M.A. Hughes, K. Brown, A. Pancoro, B.S. Murray, E. Oxtoby, J. Hughes, *Arch. Biochem. Biophys.* 295 (1992) 273–279.
- [7] J. Svasti, C. Srisomsap, S. Techasakul, R. Surarit, *Phytochem* 50 (1999) 739–743.
- [8] K. Lirdprapamongkol, J. Svasti, *Biotechnol. Lett.* 22 (2000) 1889–1894.
- [9] J. Svasti, T. Phongsak, R. Sarntima, *Biochem. Biophys. Res. Commun.* 305 (2003) 470–475.
- [10] C. Srisomsap, P. Subhasitanont, S. Techasakul, R. Surarit, J. Svasti, *Biotechnol. Lett.* 21 (1999) 947–951.
- [11] P. Toonkool, P. Metheenukul, S. Sujiwattanarat, P. Paiboon, N. Tongtubtim, M. Ketudat-Cairns, J. Ketudat-Cairns, J. Svasti, *Protein Expr. Purif.* 48 (2006) 195–204.
- [12] P.M. Dewick, in: J.B. Harborne (Ed.), *The Flavonoids: Advances in Research Since 1986*, Chapman and Hall, London, 1993, pp. 117–238.
- [13] D.G. Higgins, J.D. Thompson, T.J. Gibson, *Nucleic Acids Res.* 22 (1994) 4673–4680.
- [14] C. Combet, M. Jambon, G. Deléage, C. Geourjon, *Bioinformatics* 18 (2002) 213–214.
- [15] M. Czjzek, M. Cicek, V. Zamboni, D.R. Bevan, B. Henrissat, A. Esen, *Proc. Natl. Acad. Sci. U.S.A.* 97 (2000) 13555–13560.
- [16] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, *J. Appl. Crystallogr.* 26 (1993) 283–291.
- [17] M. Wiederstein, M.J. Sipp, *Nucleic Acids Res.* 35 (2007) 407–410.



- [18] R. Luthy, J.U. Bowie, D. Eisenberg, *Nature* 356 (1992) 83–85.
- [19] G. Vriend, *J. Mol. Graph.* 8 (1990) 52–56.
- [20] M.L. Verdonk, J.C. Cole, M.J. Hartshorn, C.W. Murray, R.D. Taylor, *Proteins* 52 (2003) 609–623.
- [21] J. Braman, C. Papworth, A. Greener, *Methods Mol. Biol.* 57 (1996) 31–44.
- [22] U.K. Laemmli, *Nature* 227 (1970) 680–685.
- [23] G. Hommalai, P. Chaiyen, J. Svasti, *Arch. Biochem. Biophys.* 442 (2005) 11–20.
- [24] Z. Keresztessy, K. Brown, M.A. Dunn, M.A. Hughes, *Biochem. J.* 353 (2001) 199–205.
- [25] L. Verdoucq, J. Morinière, D.R. Bevan, A. Esen, A. Vasella, B. Henrissat, M. Czjzek, *J. Biol. Chem.* 279 (2004) 31796–31803.
- [26] P. Chuankhayan, T. Rimlumduan, W. Tantanuch, N. Mothong, P.T. Kongsaree, P. Metheenukul, J. Svasti, O.N. Jensen, J.R. Ketudat Cairns, *Arch. Biochem. Biophys.* 468 (2007) 205–216.
- [27] J.L. Cereghino, J.M. Cregg, *FEMS Microbiol. Rev.* 24 (2000) 45–66.
- [28] R. Dopitová, P. Mazura, L. Janda, R. Chaloupková, P. Jeřábek, J. Damborský, T. Filipi, N.S. Kiran, B. Brzobohatý, *FEBS J.* 275 (2008) 6123–6135.
- [29] J. Zouhar, J. Vévodová, J. Marek, J. Damborský, X.D. Su, B. Brzobohatý, *Plant Physiol.* 127 (2001) 973–985.