Substrate (aglycone) specificity of human cytosolic β-glucosidase

Jean-Guy BERRIN*†, Mirjam CZJZEK‡, Paul A. KROON*, W. Russell MCLAUCHLAN*, Antoine PUIGSERVER†, Gary WILLIAMSON*[1] and Nathalie JUGE*†[2]

*Institute of Food Research, Colney Lane, Norwich NR4 7UA, U.K., †Institut Méditerranéen de Recherche en Nutrition, UMR INRA 1111, Faculté des Sciences et Techniques de Saint-Jérôme, 13397 Marseille Cedex 20, France, and ‡AFMB, CNRS-UMR 6098, 31 Ch. J. Aiguier, 13402 Marseille Cedex 20, France

Human cytosolic β-glucosidase (hCBG) is a xenobiotic-metabolizing enzyme that hydrolyses certain flavonoid glucosides, with specificity depending on the aglycone moiety, the type of sugar and the linkage between them. Based upon the X-ray structure of *Zea mays* β-glucosidase, we generated a three-dimensional model of hCBG by homology modelling. The enzyme exhibited the $(\beta/\alpha)_8$-barrel fold characteristic of family 1 β-glucosidases, with structural differences being confined mainly to loop regions. Based on the substrate specificity of the human enzymes, sequence alignment of family 1 enzymes and analysis of the hCBG structural model, we selected and mutated putative substrate (aglycone) binding site residues. Four single mutants (Val^{168}→Tyr, Phe^{225}→Ser, Tyr^{308}→Ala and Tyr^{308}→Phe) were expressed in *Pichia pastoris*, purified and characterized. All mutant proteins showed a decrease in activity towards a broad range of substrates. The Val^{168}→Tyr mutation did not affect K_m on *p*-nitrophenyl (*p*NP)-glycosides, but increased K_m 5-fold on flavonoid glucosides, providing the first biochemical evidence supporting a role for this residue in aglycone-binding of the substrate, a finding consistent with our three-dimensional model. The Phe^{225}→Ser and Tyr^{308}→Ala mutations, and, to a lesser degree, the Tyr^{308}→Phe mutation, resulted in a drastic decrease in specific activities towards all substrates tested, indicating an important role of those residues in catalysis. Taken together with the three-dimensional model, these mutation studies identified the amino-acid residues in the aglycone-binding subsite of hCBG that are essential for flavonoid glucoside binding and catalysis.

Key words: binding subsite, flavonoid glycosides, glycosyl hydrolase family 1, site-directed mutagenesis, three-dimensional model.

INTRODUCTION

β-Glucosidases (β-D-glucoside glucohydrolase; EC 3.2.1.21) are a widespread group of enzymes that hydrolyse a broad variety of glycosides including aryl- and alkyl-β-D-glycosides. The physiological function of β-glucosidases varies greatly depending upon their origin (plants, fungi, animals or bacteria) and substrate specificity.

A distinguishing feature of the human cytosolic β-glucosidase (hCBG) is its ability to hydrolyse many common dietary xenobiotics, including glycosides of phytoestrogens, flavonoids, simple phenolics and cyanogens [1,2]. The hCBG shows high specificity for 4′- and 7-glucosides of isoflavones, flavonols, flavones and flavanones, but does not hydrolyse 3-linked flavonoid glucosides [2]. It is a 53-kDa monomeric protein with a pI of 4.7–4.8 [1,2], present in the liver, kidney, intestine and spleen of humans [3]. This enzyme has been suggested to be involved in the metabolism of xenobiotics but its function *in vivo* has yet to be confirmed. The hCBG belongs to family 1 of glycosyl hydrolases [4,5] (http://afmb.cnrs-mrs.fr/CAZY/index.html) and catalyses the hydrolysis of O-linked β-glycosidic bonds at the non-reducing end of carbohydrates with retention of anomeric configuration. This classification is based on similarities in the amino-acid sequence and reflects more the structural features of enzymes than their substrate specificities. The three-dimensional structure of nine family 1 members has been solved [6–14], but no crystal structures of animal β-glucosidases have been reported. All nine enzymes are oligomers and have essentially the same $(\beta/\alpha)_8$ barrel fold structure, although they only share between 17 and 44 % sequence identity. They contain the highly conserved peptide motifs Thr-Phe-Asn-Glu-Pro (TFNEP) and Ile-Thr-Glu-Asn-Gly (ITENG) including the two catalytic glutamic acids located at the C-terminal end of β-strands 4 and 7 [15,16], which make up part of a crater-shaped active site [14]. The role of these residues in catalysis has been well established by mutagenesis and inhibitor studies with family 1 β-glucosidases [17–19]. However, there is little information on the interaction of β-glycosidases with their substrates, specifically the aglycone moiety, which is the basis of the diversity in natural substrates. Lactase-phlorizin hydrolase (LPH) is also a human β-glucosidase with broad substrate specificity. The enzyme is anchored in the mucosal membrane in the brush-border of the small intestine and contains two separate active sites located on domains III and IV [20]. The lactase site is selective towards glycosides with hydrophilic moieties such as lactose, and the phlorizin site is more selective for hydrophobic substrates such as phlorizin and glycosylceramides [21]. LPH can also hydrolyse (iso)flavonoid glycosides [22,23] and pyridoxine-5′-β-D-glucopyranoside, a common dietary form of vitamin B6 (K. Nemeth and P. A. Kroon, unpublished work). Defining the factors that govern the aglycone specificity for these different β-glucosidases will help unravel the details of the catalytic mechanism and, in the long term, provide

us with the biotechnological tools for the hydrolysis and synthesis of specific glucosides [24].

We recently cloned and heterologously expressed hCBG in *Pichia pastoris*, allowing a detailed enzymic and biochemical characterization of the recombinant enzyme [2]. The enzyme shows a preference for substrates with rigid, planar and hydrophobic aglycone moieties [2]. In the present study, site-directed mutagenesis was used together with molecular modelling to identify the key residues involved in aglycone recognition and substrate specificity of hCBG.

MATERIALS AND METHODS

Materials and strains

The pHIL-S1/*cbg*-1 expression vector expressing wild-type CBG was from our in-house collection [2]. Restriction endonucleases and DNA-modifying enzymes were purchased from Promega (Madison, WI, U.S.A.) and were used according to the manufacturer's recommendations. *Escherichia coli* DH5α cells (*supE*44, hsdR17, *recA*1, *endA*1, *gyrA*96, *thi*-1, *relA*1) and *E. coli* XL10-Gold® cells (*endA*1, *supE*44, *thi*-1, *recA*1, *gyrA*96, *relA*1, *lac*Hte; Stratagene Europe, Amsterdam Zuidoost, The Netherlands) were used for DNA manipulation. Oligonucleotides were synthesized by Sigma-Genosys (Cambridge, U.K.). Flavonoids and their conjugates were purchased in the purest form available from Extrasynthèse (Genay, France) or Apin Chemicals (Abingdon, Oxon., U.K.) except for quercetin 7-glucoside which was a gift from Paul Needs (Institute of Food Research, Norwich, U.K.). *p*-Nitrophenyl (*p*NP)-glycosyl derivatives were obtained from Sigma–Aldrich (Poole, Dorset, U.K.).

Cloning and site-directed mutagenesis

Mutations were introduced into the pHIL-S1/*cbg*-1 using the QuikChange™ XL site-directed mutagenesis kit (Stratagene Europe), according to the manufacturer's instructions. Briefly, two overlapping complementary oligonucleotides (SDS/PAGE-purified) for each mutation were designed to contain the corresponding nucleotide changes (see Table 1). Each primer (125 ng) was annealed to pHIL-S1/*cbg*-1 (10 ng) and both strands of the plasmid were amplified using 1 unit of *Pfu* polymerase in a linear extension reaction carried out by PCR under the following conditions: 1 min denaturation at 95 °C, and 18 cycles consisting of 1 min denaturation at 95 °C, 1 min annealing at 60 °C and 25 min extension at 68 °C. The reactions were cooled to room temperature (20 °C) and were treated with *Dpn*I restriction enzyme for 1 h at 37 °C to hydrolyse the methylated parental plasmid strands. *Dpn*I-treated DNA (1 μl) was used to transform *E. coli* DH5α and XL10-Gold® ultracompetent cells (50 μl) and the transformed bacteria were grown in 0.5 ml of NZY

Plus broth, pH 7.5 [1 % (w/v) NZ amine (casein hydrolysate), 0.5 % (w/v) yeast extract, 0.5 % (w/v) NaCl, 12.5 mM MgCl$_2$, 12.5 mM MgSO$_4$, 0.4 % (w/v) glucose] for 1 h at 37 °C. Transformants were selected on Luria–Bertani ('LB') medium plates containing 50 μg/ml ampicillin. Recombinant plasmids were isolated using Qiagen columns (Mini-Prep kit), and checked for integrity by restriction mapping. All transformants were subjected to DNA sequencing using the ABI Prism Big Dye™ Terminator Cycle Sequencing kit to confirm the presence of the single mutation and ensure that no errors were generated during the PCR.

Transformation of the *P. pastoris* strain (*his4*)/GS115 [25] and screening were carried out as described in [2]. Briefly, pHIL-S1/*cbg*-1 mutants (approx. 1 μg) and pHIL-S1 vector (negative control) were digested with *Bgl*II before transformation by the spheroplast method [26]. After screening for methanol-sensitive colonies, Mut^s colonies were used to inoculate 10 ml of buffered minimal glycerol-complex medium (BMGY) pH 6. After 2 days at 250 rev./min and 30 °C, the cells were pelleted and resuspended in 2 ml buffered minimal methanol-complex medium (BMMY). Following another 4 days at 30 °C, the culture was centrifuged at 4000 *g* for 5 min at room temperature and the amount of CBG mutant proteins in the supernatant was estimated by activity measurement assays using *p*NP-β-D-glucopyranoside (*p*NPGlc) and quercetin-4′-glucoside as substrates.

Protein production and purification

Cells for large-scale expression studies were grown in 1.25 litres of BMGY to a density of 20–25 at D_{680} in 5-litre baffled flasks at 30 °C. Cells were harvested, resuspended in 250 ml of BMMY and incubated with shaking (180 rev./min) for 4 days in 1-litre baffled flasks at 30 °C. Mutant proteins were purified in a single step using hydrophobic-interaction chromatography. Culture supernatant (250 ml) was loaded on to a column (1.5 cm × 5 cm) of Octyl Sepharose previously equilibrated with 20 mM sodium phosphate buffer (pH 6.5) containing 1 mM EDTA, the column was washed with 20 % (v/v) ethylene glycol in sodium phosphate buffer and unbound material was discarded. Bound material was eluted with 50 % (v/v) ethylene glycol at a flow rate of 0.8 ml · min^{-1} over 1 h. β-Glucosidase-containing fractions were pooled, concentrated using a 50 ml stirred ultrafiltration cell (Amicon, Stonehouse, Gloucestershire, U.K.) and a YM3 ultrafiltration membrane (Millipore, Watford, Hertfordshire, U.K.) and were then checked for purity by SDS/PAGE.

Enzyme assays

Fractions generated after purification of CBG mutants from culture supernatant were assayed for CBG activity as described previously for the wild-type enzyme [2]. Briefly, the release

Table 1   Mutagenic primer sequences

Bold residues indicate the sites of mutation.

Mutation	Direction	Primer sequence
Val^{168}→Tyr	Forward	5′-GG ATC ACC ATA AAT GAA GCT AAT **TA**T CTT TCT GTG ATG TCA TAT GAC-3′
Val^{168}→Tyr	Reverse	5′-GTA ATA TGA CAT CAC AGA AAG A**TA** ATT AGC TTC ATT TAT GGT GAT CC-3′
Phe^{225}→Ser	Forward	5′-G TAT GGT GTC TCT CTA TCA CTT T**C**T GCG GTC TGG TTG GAA C-3′
Phe^{225}→Ser	Reverse	5′-G TTC CAA CCA GAC CGC A**G**A AAG TGA TAG AGA CAC ATA C-3′
Tyr^{308}→Ala	Forward	5′-GCT GAT TTT TTT GCT GTG CAA T**T**T TAT ACA ACT CGC TTA ATC AAG-3′
Tyr^{308}→Ala	Reverse	5′-CTT GAT TAA GCG AGT TGT ATA A**A**A TTG CAC AGC AAA AAA ATC AGC-3′
Tyr^{308}→Phe	Forward	5′-GCT GAT TTT TTT GCT GTG CAA **GC**T TAT ACA ACT CGC TTA ATC AAG-3′
Tyr^{308}→Phe	Reverse	5′-CTT GAT TAA GCG AGT TGT ATA A**GC** TTG CAC AGC AAA AAA ATC AGC-3′

of 4-nitrophenol from *p*NPGlc was monitored at 400 nm using the molar absorption coefficient for 4-nitrophenol of 18 300 $M^{-1} \cdot cm^{-1}$. The activity of purified CBG mutants towards other *p*NP-glycosides [*p*NP-β-D-galactopyranoside (*p*NPGal), *p*NP-α-L-arabinopyranoside (*p*NPAra) and *p*NP-β-D-fucopyranoside (*p*NPFuc)] was determined using the same method. Activities towards flavonoid glycosides were determined using HPLC with on-line diode array detection by measuring the amount of aglycone released from the substrate (10–500 μM in 50 mM sodium phosphate buffer), as described previously in [2], with particular care taken to ensure complete solubility of substrates [27].

Protein assays and protein sequencing

For purified wild-type CBG, total protein was calculated using an absorption coefficient at 280 nm (122 120 $M^{-1} \cdot cm^{-1}$) derived from the CBG amino-acid composition and also confirmed by CBG specific activity towards *p*NPGlc (10 units $\cdot$ mg^{-1}). The concentration of purified CBG mutant proteins was estimated using the Pierce Protein Assay Reagent (Rockford, IL, U.S.A.) with wild-type CBG as standard. Protein sequencing was performed at the Protein Sequencing and Peptide Synthesis Facility (John Innes Centre, Norwich, U.K.) using an ABI 491 Procise sequencer.

Production of anti-hCBG polyclonal antibodies

Anti-hCBG polyclonal antibodies were produced by injecting 4-month-old male New Zealand White rabbits with 100 μg of pure recombinant CBG (reCBG) [2] mixed 1:1 with Freund's complete adjuvant in a final volume of 100 μl. Similar injections were given at 4, 8 and 12 weeks, again using 100 μg of pure reCBG per injection. Rabbits were bled 10 days after the final injection. Blood was collected in heparinized tubes and was immediately centrifuged (5000 ***g***, 10 min) to recover plasma, which was stored in 100 μl aliquots at -20 °C.

Gel electrophoresis and immunoblotting

SDS/PAGE was routinely performed using 12 % homogeneous Tris/glycine gels (Novex, Frankfurt, Germany) according to the manufacturer's instructions before staining with Coomassie Blue. For immunodetection of CBG mutants, proteins were transferred on to nitrocellulose membranes by semi-dry blotting and blots were probed with a 1:2000 dilution of rabbit anti-hCBG polyclonal antiserum followed by a 1:2000 dilution of horseradish-peroxidase-conjugated goat anti-rabbit secondary antibody. Immunoblots were developed using enhanced chemiluminescent detection reagents (ECL® Plus Detection Kit, Amersham Biosciences, Little Chalfont, Bucks., U.K.) and the immunoreactive proteins visualized on X-ray film (Hyperfilm™ ECL; Amersham Biosciences.).

Molecular modelling

A molecular model of the hCBG was constructed using atomic co-ordinates from the known crystal structure of the β-glucosidase from *Zea mays* (ZMGlu) [14]. The sequence alignment of CBG and ZMGlu was structure-based, initially using ClustalW [28] and refined by hand when necessary: insertions or deletions in the aligned sequences were displaced into loop regions, when possible, rather than in the β-sheet or α-helical regions. The graphic program Turbo-Frodo [29] was used to carry out the modelling

of the structure. Insertions or deletions that occurred in the sequence were modelled using a fragment database [30] to search for short sequences that overlapped with the structure either side of the unknown region and which contained the same number of residues. The best-fit fragment was then used in the model structure. The non-identical amino acids were modelled by their most likely rotameric conformers, thus avoiding major side-chain clashes. Another homology model was generated for the enzyme using the MODELLER program [31] and the overall structure of the enzyme found to be very similar [the overall root mean square distance between the two models for all atoms is 4.39 Å (1 Å = 0.1 nm); the root mean square distance of 405 Cα atoms is 0.44 Å]. The model for the *p*NPGlc located in the active site of hCBG was generated using the equivalent position of the *p*NP-thio-glucoside model of ZMGlu [32].

RESULTS

Model of hCBG and design of mutants

No X-ray crystallographic structures of hCBG are available. In order to design catalytic mutants for hCBG, a molecular model of the human enzyme was generated using ZMGlu as a template (Figure 1). The modelled hCBG is very similar to other members of family 1 β-glucosidases, forming an eight-stranded β/α-barrel with the four loops (A–D) being the sites of highest variability. The loops B, C and D of the hCBG are shorter than ZMGlu whereas loop A is similar in size (Figure 1). The connections between the $\beta\alpha$ repeats at the bottom of the barrel are very short, with the exception of the region between α-helix 5 and β-strand 6, which is ten amino acids longer than ZMGlu (Figure 1). A similar extended connection is found in the three-dimensional structure of *Lactococcus lactis* 6-phospho-β-galactosidase [7] and is expected to be present in LPH (Figure 1). As expected, the active-site architectures of the hCBG model and the three-dimensional structure of ZMGlu are similar (Figure 2). The active-site aglycone-binding pocket of the hCBG model indicates the presence of a cluster of hydrophobic residues (Phe[225], Tyr[308], Tyr[309], Trp[345] and Phe[433]) representing one wall of the binding pocket and with Val[168], Met[172] and Phe[179] forming the opposite wall (Figure 2). The side chain of Tyr[308] is turned away from the pocket and probably has an influence on the overall structural arrangement in this region, including the precise form of the active-site aglycone pocket. All of these hydrophobic residues are located above (towards the molecular surface) the two catalytic residues and four more hydrophobic amino acids (His[120], Phe[121], Trp[417] and Trp[425]), highly conserved in family 1 enzymes, and involved in the interaction with the glycoside (Figure 2). The acid/base catalyst glutamate (Glu[165]) and the nucleophilic glutamate (Glu[373]) are positioned within the active site, as expected. The side chain of Gln[307], a variant residue within family 1 enzymes, is protruding between the two catalytic residues and appears to interact with the OH in position 2 of the glucoside moiety (Figure 2).

Interestingly, the model predicts a different form of the substrate-binding cleft compared with ZMGlu, less deep and more open. Phe[225] and Val[168] form a supplementary step or plateau within the pocket, forcing the aglycone outwards. The aglycone-binding pocket in hCBG is lined by Trp[345], Met[172], Phe[179] and Phe[433], which form a hydrophobic environment, but the pocket is less constrained than the 'pinch' formed by the residues Trp[378], Phe[198], Phe[205] and Phe[466] in ZMGlu [14].

In order to investigate the aglycone specificity of hCBG, non-conserved residues of the mammalian β-glucosidases were selected from the hydrophobic cluster in the vicinity of the catalytic

```
                                                    β1                      ←—
hCBG     : --------------------MAFPAGFGWAAATAAYQVEGGWDADGKGPCVW          32
LPH dIII : --------------ERDLFYHGTFRDDFLWGVSSSAYQIEGAWDADGKGPSIW          39
LPH dIV  : ---------------DEFLYGRFPEGFIWSAASAAYQIEGAWRADGKGLSIW           37
ZMBGlu   : SARVGSQNGVQMLSPSEIPQRDWFPSDFTEGAATSAYQIEGAWNEDGKGESNW          53
                                   *    *    *** ** *  ****   *

– LOOP A ———————→              α1               β2                  α2
DTFTHQGGERVFKNQTGDVACGSYTLWEEDLKCTKQLGLTHYRFSLSWSRLLPDGT-TGFINQKGIDYYNKIIDDLLKN   110
DNFTHTPGSNVKDNATGDIACDSYHQLDADLNMLRALRVKAYRFSISWSRILPDGT-TRYINEAGLNYYVRLIDTLLAA   117
DTFSHTP-LRVENDAIGDVACDSYHKIAEDLVTLQNLCVSHYRFSISWPRILPDGT-TRYINEAGLNTYVRLIDTLLAA   114
DHFCHNHPERILDGSNSDIGANSYHHYKTDVRLLKEMGMDAYRFSISWPRILPKGTKEGGINPDGIKYYRNLINLLLEN   132
 * *  *          *     **   **     *    ****  **  *   *     *  *  **    *  *

    β3               FDLPQTLEDQ-GGW---LSE                  β4  ▼        ←——— LOOP B
GVTPIVTLYHFDLPQTLEDQ-GGW---LSEAIIESFDKYAQFCFSTFGDRVKQWITINEANVLSVMSYDLGMFPPG----  182
NIFPMVTLFHWDLPQALQDI-GGW---ENPALIDLFDSYADFCFQTFGDRVKFWMTFNEPMYLAWLGYGSGEFPPG----  189
SIQPQVTIYHWDLPQTLQDV-GGW---ENETIVQRFKEYADVLFQRLGDKVKFWITLNEPFVIAYQGYGYGTAAPG----  186
GIEIYVTIFHWDVPQALEEKYGGFLDKSHKSIVEDYTYFAKVCFDNFGDKVKNWLTFNEPQTFTSFSYGTGVFAPGRCSP  212
   * **  * *  ** *      **          *  ** ** * *  ●     *    *   **

———————→              α4                        β5                    α5
-IPH------FGTGGYQAAHNLIKAHARSWHSYDSLFRKKQKGMVSLSLSFAVWLEPADPNSVSDQFAAKRAITFHLDLFA   255
-VKD------PGWAPYRIAHTVIKAHARVYHTYDEKYRQEQKGVISLSLSTHFAEPPSPGVPRDVFAADRMLQFSLGLFA   262
-VSN-----RPGTAPYIVGHNLIKAHAEAVHLYNDVYRASQGGVISITISSDWAEPFDPSNQEDVEAAERYVQFHGGLFA   260
GLDCAYPTGNSLVEPYTAGHNILLAHAEAVDLYNKHYK-RDDTRIGLAFDVBGRVPYG-TSFLDKQAEFRSYDINLGLFL   290
   *    *    *** *      * *               ●   *         *        *         *

                                              β6              ←——————
KPIFIDGDYPEVVKSQIASMSQKQGYPSSRLPEFTEEEKKMIKGTADFFAVQYYTTRLIKYQENKKGELGILQ--DAEI   332
HPIFRNGDYPDTMKWKVGNRSELQHLATSRLPSFTEEEKRFIRATADVFCLNTYYSRIVQHKTPRLNPPSYED--DQEM   339
HPIFKNGDYNEVMKTRIRDRSLAAGLNKSRLPEFTESEKRRINGTYDFFGFNIYTTVLAYNLNYATAISSFDA--DRGV   337
IPV-VRGDYPFSMRS---------LARERLPFFKDEQKEKLAGSYNMLGLNYYTSRFSKNIDISPNYSPVLNTDDAYA   358
  *     ***        *** *      *           ●*                         *

– LOOP C ———————→         α6                β7 ▼  ←—— LOOP D ———→        α7
-EFFPDPSWKNVD------WIYVVPWGVCKLLKYIKDTYNNPVIYITENGFPQSDPA--------PLDDTQRWEYFRQTF   397
-AEEEDPSWPSTA------MNRAAPEGTRRLLNWIKEFYGDIPIYITENGVGLTNP--------NTEDTDRIFYHKTYI   403
-ASIADRSWPDSG----SFWLKMTPFGFRRILNWLKEFYNDPPIYVTENGVSQREET-------DLNDTARIYYLRTYI   404
SQEVNGPDGKPIGPPMGNPWIYMYPEGLKDLLMIMKNKYGNPPIYITENGIGDVDTKETPLPMEAALNDYKRLDYIQRHI  438
                         *           *   ** ****    *             *     *

    β8                                             α8
QELFKAIQLDKVNLQVYCAWSLLDNFEWNQGYSSRFGLFHVDFEDPARPRVPYTSAKEYAKIIRNNGLE------   466
NEALKAYRLDGIDLRGYVAQSLMDNFEWLNGYTVKFGLYHVDFNNTNRPRTARASARYYTEVITNNGMP------   462
NEALKAVQ-DKVDLRGYTVESAMDNFEWATGFSERFGLHFVNYSDPSLPRIPKASAKFYASVVKCNGFP------   472
ATLKESIDL-GSNVQGYFAWSLLDNFEWFAGFTERYGIVYVDRNN-NCTRYMKESAKWLKEFNTAKKPS------  505
   *  **  ***** *       *    *       *  *           *  **
```
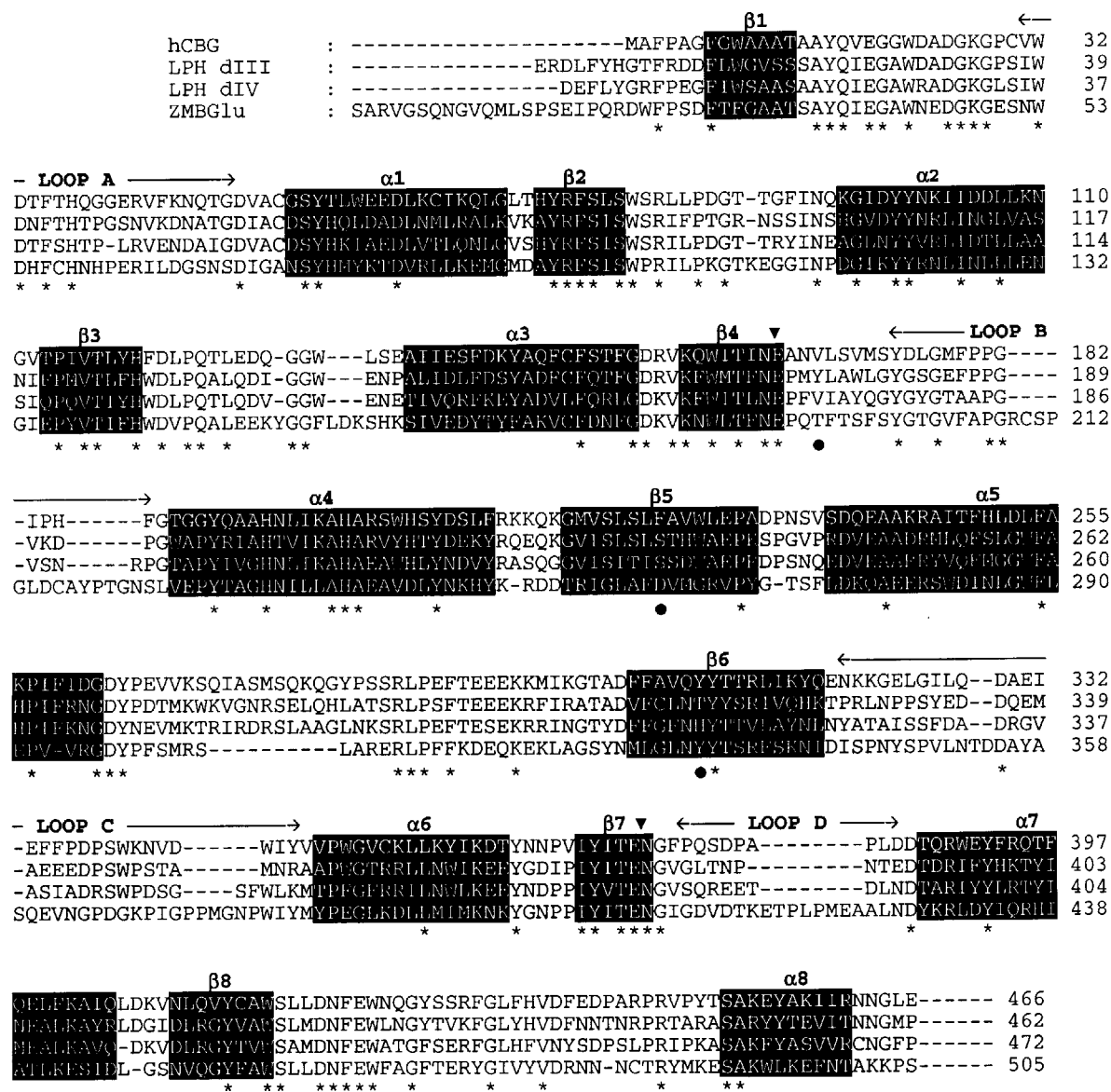
Figure 1   Sequence alignment of human family 1 β-glucosidases with ZMGlu

LPH dIII, domain III of the human LPH; LPH dIV, domain IV of the human LPH. The secondary-structure elements of the $(\beta/\alpha)_8$-barrel structure are highlighted in black. Residues that are identical in all enzymes are marked with an asterisk. The mutated residues Val[168], Phe[225] and Tyr[308] are marked with closed circles (●). Arrowheads (▼) indicate the two catalytic glutamates. The four loops A–D are delimited by arrows above the aligned sequences.

residues. In the model, Val[168] is located close to the catalytic nucleophile between β-strand 4 and loop B; Phe[225] and Tyr[308] are in the middle of β-strands 5 and 6 respectively (Figure 1). Val[168] was replaced by tyrosine, the corresponding amino acid in LPH domain III, Phe[225] by serine, the corresponding amino acid in other mammalian family 1 β-glucosidases, and Tyr[308] by both phenylalanine and alanine to remove either the hydroxyl group or the hydrophobic side chain respectively.

Production and characterization of the hCBG variants

The wild-type and hCBG mutants were produced in *P. pastoris* and secreted into the culture medium. Mutant Val[168]→Tyr was secreted with a yield similar to that of the wild-type (up to $10 \text{ mg} \cdot l^{-1}$), whereas Phe[225]→Ser, Tyr[308]→Phe and Tyr[308]→Ala were produced with a secretion yield of approx. $0.3 \text{ mg} \cdot l^{-1}$.

All proteins were purified to apparent homogeneity from the culture supernatant in one chromatographic step. The purified hCBG variants migrated on SDS/PAGE as a single band with a molecular mass of 53 kDa, identical with that of the recombinant wild-type enzyme, and cross-reacted with polyclonal antibodies raised against hCBG (results not shown). Isoelectric focusing revealed that hCBG mutants consisted of two molecular isoforms with pIs of approx. 3.7 and 3.8, similar to the wild-type enzyme (results not shown). The N-terminus, Arg-Gly-Ala-Phe-Pro, of each hCBG mutant was identical with that of the recombinant wild-type hCBG, indicating correct processing of the hCBG signal sequence.

In order to evaluate the effect the mutations had on enzyme activity, pH and temperature optimum were determined using aryl-β-D-glucoside as a substrate. The specific activity of all mutants decreased on this substrate as compared with the
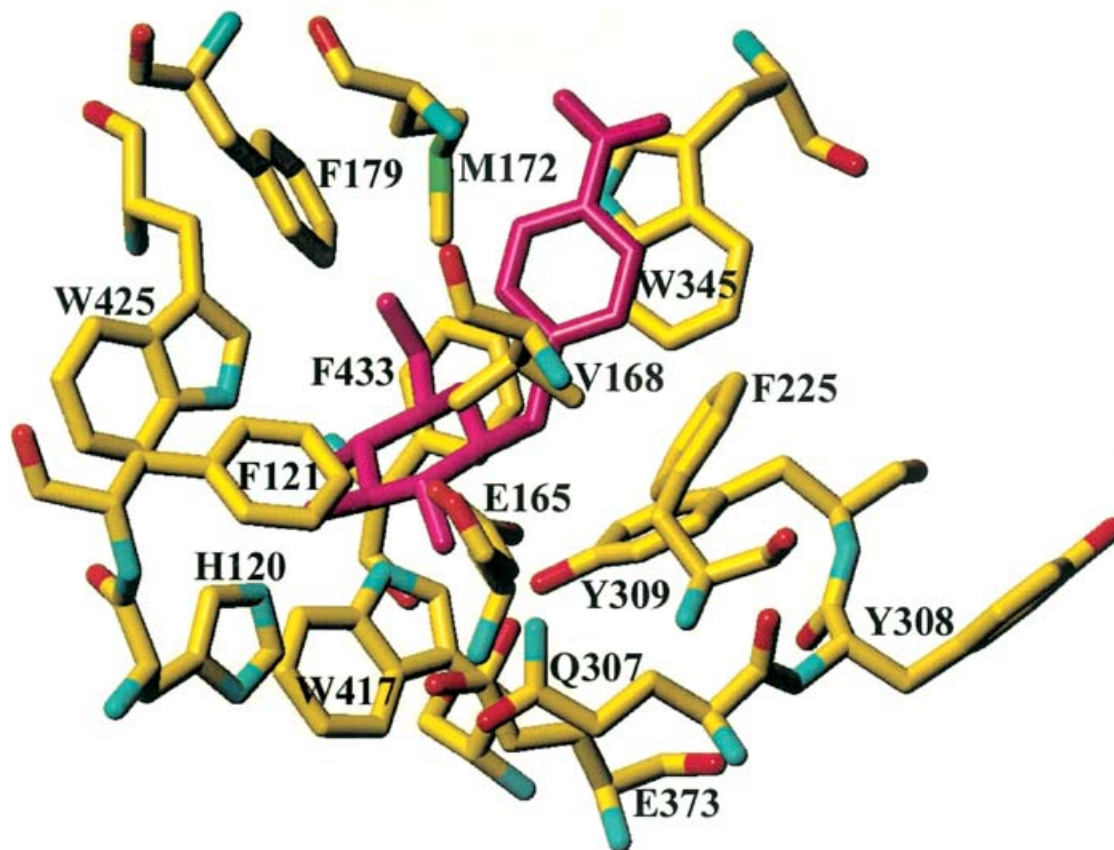
Figure 2   Close-up view of the active site of hCBG model structure, with the modelled *p*NP-glucoside (chair conformation) in purple

Aromatic and hydrophobic residues surrounding the glucose (His[120], Phe[121], Trp[417] and Trp[425]) and the aglycone moieties (Val[168], Met[172], Phe[179], Phe[225], Tyr[308], Tyr[309], Trp[345] and Phe[433]), as well as the two catalytic glutamates, Glu[165] and Glu[373], and Gln[307], are labelled. All residues are coloured in standard atom-type colours. The *p*NP-glucoside molecule is coloured in purple and is here modelled in the energetically more favourable chair conformation. The figure was drawn with Turbo-Frodo [28].

Table 2   Specific activities of wild-type and mutant hCBGs on *p*NP-glycosides and flavonoid glucosides

*Specific activities are mean data ($n \geqslant 2$) with typical standard deviation below 10 % when applicable. Specific activities were determined with a substrate concentration of 10 mM for *p*NP-glycosides and 500 μM for flavonoid glucosides. ND, not determined.

Substrate	Specific activity (units · mg^{-1})*				
	Wild-type	Val[168]$\rightarrow$Tyr	Phe[225]$\rightarrow$Ser	Tyr[308]$\rightarrow$Phe	Tyr[308]$\rightarrow$Ala
*p*NPFuc	9	1.6	0.033	1.1	0.011
*p*NPGlc	10	1.2	0.065	0.6	0.031
*p*NPGal	15	1.7	0.033	0.7	0.008
*p*NPAra	6	0.5	ND	0.6	ND
Daidzein-7-Glc	2.75	0.76	0.172	1.31	0.078
Quercetin-4'-Glc	1.19	0.24	0.014	0.39	0.020
Quercetin-7-Glc	0.77	0.15	0.045	0.31	0.030
Luteolin-4'-Glc	1.30	0.68	0.017	0.70	0.017
Luteolin-7-Glc	2.85	0.37	0.180	0.82	0.080
Naringenin-7-Glc	0.93	0.31	0.060	0.59	0.047
Eriodictyol-7-Glc	0.90	0.19	0.041	0.27	0.019
Apigenin-7-Glc	1.30	ND	0.065	ND	0.065

wild-type hCBG (Table 2). However, the pH activity curves of Val[168]$\rightarrow$Tyr and Tyr[308]$\rightarrow$Phe were identical with that of the wild-type hCBG [2], with the optimum pH at 6.5. The temperature curves of Val[168]$\rightarrow$Tyr and Tyr[308]$\rightarrow$Phe mutants were also identical with 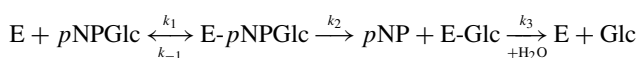that of the wild-type hCBG [2], with the optimum activity at 50 °C, and > 80 % of the activity remaining after incubation for 24 h at 37 °C (results not shown). The low specific activities for Phe[225]$\rightarrow$Ser and Tyr[308]$\rightarrow$Ala (Table 2) precluded determination of pH and temperature optima for these mutants.

Table 3   Kinetic parameters of wild-type and mutant hCBGs on *p*NP-glycosides and flavonoid glucosides

Results are means $\pm$ S.E.M. ($n \geqslant 20$). ND, not determined.

Substrate	Wild-type			Val$^{168}\to$ Tyr			Tyr$^{308}\to$ Phe		
	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM$^{-1}\cdot$ s^{-1})	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM$^{-1}\cdot$ s^{-1})	K_m (μM)	k_{cat} (sec^{-1})	k_{cat}/K_m (mM$^{-1}\cdot$ s^{-1})
*p*NPFuc	0.37 $\pm$ 0.01	10.7 $\pm$ 0.0	28.9	0.32 $\pm$ 0.02	1.71 $\pm$ 0.02	5.34	0.18 $\pm$ 0.02	1.18 $\pm$ 0.01	6.55
*p*NPAra	0.57 $\pm$ 0.05	5.97 $\pm$ 0.45	10.4	0.18 $\pm$ 0.03	0.52 $\pm$ 0.01	2.96	ND	ND	ND
*p*NPGlc	1.76 $\pm$ 0.05	12.1 $\pm$ 0.03	6.9	0.80 $\pm$ 0.05	1.27 $\pm$ 0.02	1.59	0.37 $\pm$ 0.05	0.69 $\pm$ 0.02	1.86
*p*NPGal	3.14 $\pm$ 0.15	17.6 $\pm$ 0.3	5.6	2.28 $\pm$ 0.10	1.85 $\pm$ 0.02	0.81	0.77 $\pm$ 0.09	0.76 $\pm$ 0.02	0.99
Daidzein-7-Glc	118 $\pm$ 11	3.55 $\pm$ 0.16	30	470 $\pm$ 12	0.69 $\pm$ 0.01	1.46	177 $\pm$ 8.5	1.42 $\pm$ 0.04	8
Quercetin-4′-Glc	31.8 $\pm$ 2.9	1.08 $\pm$ 0.02	34	118 $\pm$ 13	0.19 $\pm$ 0.01	1.6	39.3 $\pm$ 4.7	0.42 $\pm$ 0.01	10
Quercetin-7-Glc	42.2 $\pm$ 3.2	0.69 $\pm$ 0.02	16	158 $\pm$ 9	0.12 $\pm$ 0.00	0.75	73.6 $\pm$ 4.8	0.33 $\pm$ 0.01	4.5
Luteolin-4′-Glc	10.0 $\pm$ 0.1	1.17 $\pm$ 0.01	117	68 $\pm$ 5	0.61 $\pm$ 0.01	8.9	29.6 $\pm$ 4.4	0.72 $\pm$ 0.04	24
Luteolin-7-Glc	50 $\pm$ 3.2	3.05 $\pm$ 0.07	61	108 $\pm$ 8	0.31 $\pm$ 0.01	2.9	63.1 $\pm$ 4.0	0.84 $\pm$ 0.02	13
Naringenin-7-Glc	432 $\pm$ 33	2.60 $\pm$ 0.01	6.0	1240 $\pm$ 150	0.33 $\pm$ 0.03	0.26	453 $\pm$ 58	0.55 $\pm$ 0.02	1.2
Eriodictyol-7-Glc	253 $\pm$ 13	1.26 $\pm$ 0.03	5.0	1142 $\pm$ 130	0.18 $\pm$ 0.02	0.16	282 $\pm$ 18	0.29 $\pm$ 0.02	1.0

Kinetic analyses and substrate specificity of hCBG variants

The hydrolysis of *p*NPGlc by (guinea pig) CBG [33] follows the scheme:

$$\text{E} + p\text{NPGlc} \underset{k_{-1}}{\overset{k_1}{\longleftrightarrow}} \text{E-}p\text{NPGlc} \overset{k_2}{\longrightarrow} p\text{NP} + \text{E-Glc} \overset{k_3}{\underset{+\text{H}_2\text{O}}{\longrightarrow}} \text{E} + \text{Glc}$$

where E is enzyme. For all enzymic reactions, $K_m \approx K_s$ (i.e. the affinity constant) only when $k_2 \ll k_3$. The rate-determining step for CBG was k_2 only under some conditions. Therefore our measured changes in K_m(app) may be directly related to changes in affinity [$K_S = (k_1/k_{-1})$], but may also have some component of k_2 (formation of glucosyl–enzyme intermediate) and/or k_3 (deglucosylation). Therefore the measurements of K_m reported below are apparent affinities, but the detailed mechanism is not sufficiently elucidated to predict exactly how changes in K_m relate to binding energy.

The ability of hCBG variants to hydrolyse a range of *p*NP-glycosides and flavonoid glucosides was assessed to investigate the effect of the mutations on the enzyme specificity (Tables 2 and 3). All variants showed a significant decrease in activity towards all substrates tested. Mutant proteins Val$^{168}\to$Tyr and Tyr$^{308}\to$Phe retained enough activity for accurate evaluation of their kinetic parameters (Table 3), whereas only specific activities could be measured for Phe$^{225}\to$Ser and Tyr$^{308}\to$Ala (Table 2). Val$^{168}\to$Tyr and Tyr$^{308}\to$Phe exhibited normal Michaelis–Menten kinetics on all substrates tested.

The specific activity of wild-type hCBG on *p*NPGal, *p*NPGlc and *p*NPFuc was 15, 10 and 9 units $\cdot$ mg^{-1} (Table 2) respectively, but the substrate with the highest catalytic efficiency was *p*NPFuc (Table 3). A similar trend was observed for Val$^{168}\to$Tyr and Tyr$^{308}\to$Phe, although the catalytic efficiency (k_{cat}/K_m) was reduced by 5-fold compared with the wild-type hCBG. The apparent affinity (K_m) of Tyr$^{308}\to$Phe and Val$^{168}\to$Tyr for *p*NP-glycosides was similar or slightly better than that of wild-type hCBG with a 2–5-fold decrease in K_m (Table 3). The overall decrease in catalytic efficiency was mostly accounted for by the 10-fold decrease in k_{cat} (Table 3). Phe$^{225}\to$Ser and Tyr$^{308}\to$Ala showed a drastic decrease in specific activity (99.3–99.8 % and 99.7–99.9 % respectively) with no significant variations between the different types of sugar tested (Table 2).

Wild-type hCBG efficiently hydrolysed many flavonoid β-D-glucosides, 4′- and 7-conjugated (Figure 3), the best substrate being luteolin-4′-Glc with a k_{cat}/K_m of 117 mM$^{-1}\cdot$ sec^{-1} (Table 3). Phe$^{225}\to$Ser and Tyr$^{308}\to$Ala showed a drastic reduction

A
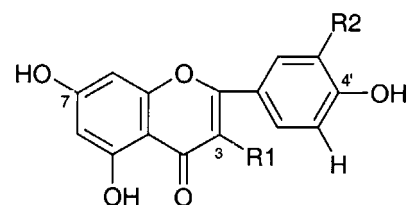
B
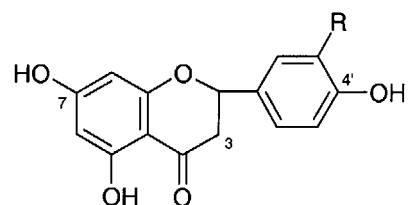
C
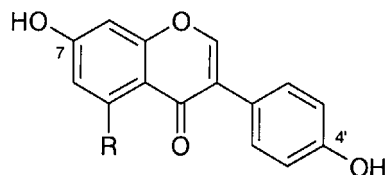
Figure 3   Structures of flavonoid aglycones

(**A**) Quercetin (R1, OH; R2, OH), apigenin (R1, H, R2, H), luteolin (R1, H, R2, OH). (**B**) Naringenin (R, H), eriodictyol (R, OH). (**C**) Daidzein (R, H), genistein (R, OH).

(93–99 %) in specific activities towards flavonoid glucosides. For example, specific activities obtained for both mutants using apigenin-7-Glc and luteolin-4′-Glc were 20- and 76-fold lower than those obtained for the wild-type hCBG (Table 2). The large decrease in the catalytic efficiency of Val$^{168}\to$Tyr towards flavonoid glucosides was due to a combination of increased K_m and decreased k_{cat} (Table 3). In contrast, the apparent affinity of the Tyr$^{308}\to$Phe mutant towards flavonoid glucosides was unchanged

and the decrease in catalytic efficiency was entirely accounted for by a decrease in k_{cat} (Table 3).

The activity of the mutants towards flavonoid glucosides was not affected by the position of conjugation of the glucose residue. Flavonoid-3-glucosides, phlorizin and lactose were not hydrolysed by the mutant enzymes, as previously found for the wild-type hCBG [2]. Similarly, pNP-β-D-mannopyranoside was found to be an effective inhibitor of Val[168]$\rightarrow$Tyr and Tyr[308]$\rightarrow$Phe [2]; the hydrolysis of 10 mM pNPGlc was drastically reduced ($> 98\%$) in the presence of 10 mM pNP-β-D-mannopyranoside (results not shown).

DISCUSSION

Three-dimensional model of hCBG

The active site of the hCBG model forms an oval shaped pocket with a cluster of hydrophobic amino-acid side chains lining the walls of the active site. In mammalian CBGs, several observations support the idea of the presence of a hydrophobic cluster of amino acids located in the vicinity of the active site: (i) the high affinity of the enzyme for hydrophobic matrices [1,27], (ii) the inhibition of β-glucosidase activity by amphipathic compounds such as glucosylsphingosine [34], and (iii) the strong interaction of linear alkyl glycosides with the putative hydrophobic site of hCBG [35]. In family 1 enzymes, the cleft-like gate to the active site is formed by four extended solvent-exposed loops (loops A–D). In the hCBG model, loops B–D are shorter, resulting in a small entrance to the pocket, which could explain the inability of hCBG to hydrolyse large substrates such as glycosphingolipid and mucopolysaccharides [3]. The structure of the aglycone-binding subsite of the hCBG model is in agreement with the substrate specificity of hCBG, which exhibits a preference for rigid, planar, hydrophobic moieties compared with long, flexible alkyl chains [2,36]. As expected, the nature of the amino acids forming the glucose binding site of hCBG is well conserved among family 1 β-glucosidases [37]. One difference is the presence of Gln[307] protruding between the two catalytic residues. In all known structures of family 1 β-glucosidases, this amino acid is aspartic acid, and the shorter side chain prevents it from being involved in a hydrogen bond with the glucose OH in position 2 (OH2). In hCBG, the elongated side chain of Gln[307] allows hydrogen bonding with OH2 and this enzyme is therefore more sensitive to the stereochemistry of the OH2 group, in contrast with the other family 1 β-glucosidases. This might explain why mannose, which has an axial OH2 group, is such a good inhibitor of hCBG.

Other differences in the overall structure of hCBG are found in the regions involved in the dimerization of ZMGlu [32] with, notably, a proline residue in the region between $\beta 5$ and $\alpha 5$, a longer $\alpha 5–\alpha 6$ connection and a shorter loop C. It should be noted that eight family 1 β-glucosidases with known three-dimensional structures occur as oligomers. However, the regions involved in the dimerization vary from one enzyme to the other [32]. Dimerization and other levels of quaternary associations of β-glucosidases have been shown to be important for stability and/or activity of these enzymes [38]. The hCBG is active as a monomer [1], suggesting that oligomerization is not a strict requirement for stability/activity of all family 1 β-glucosidases.

The hCBG substrate (aglycone) specificity

Based on the hCBG model, we predicted that mutations of Val[168], Phe[225] and Tyr[308] would significantly affect the substrate (aglycone)-binding capacity of the enzyme. The catalytic activity of the mutants was decreased on both flavonoid glucosides and pNP-glycosides. Since the mutations were targeted towards amino acids in the aglycone subsite of CBG, the binding of both the pNP moiety of aryl glycosides and the aglycone moiety of flavonoid glucosides was affected. This suggests that the major binding determinant for substrate hydrolysis by hCBG is the aglycone moiety, in accordance with inhibition studies reported on pig kidney CBG [39]. The importance of the aglycone in substrate binding is also reflected in the enzyme structure as the major differences in the active site of family 1 β-glucosidases occur at the aglycone-binding site [37]. The Val[168]$\rightarrow$Tyr mutation decreased the catalytic efficiency more effectively on flavonoid glucosides than on pNP-glycosides. The mutation affected the K_m and the k_{cat} of the enzyme towards flavonoids, whereas only the k_{cat} was affected towards pNP-glycosides, suggesting that Val[168] is a key residue to determine aglycone specificity. The Phe[225]$\rightarrow$Ser mutation had a dramatic effect upon activity, irrespective of the nature of the substrate. This was surprising since serine is present on both domain III and IV of LPH, which catalyses the hydrolysis of both flavonoid glucosides and pNP-glycosides. In addition, the Phe[225] residue is not found in other mammalian β-glucosidases (Figure 1). Taken together, these findings suggest that this residue, uniquely for family 1 β-glucosidases, is essential for catalysis in hCBG. The catalytic efficiency of Tyr[308]$\rightarrow$Phe towards both flavonoid glucosides and pNP-glycosides was reduced compared with wild-type hCBG. The removal of the hydrophobic chain in Tyr[308]$\rightarrow$Ala had an even more dramatic effect upon activity, indicating that Tyr[308] is a key residue for structural integrity of the binding pocket. This hydrophobic residue might be important for the correct packing of the β-sheets and loops that form the base of the pocket. In the model of hCBG, Tyr[308] is found in close proximity to Tyr[309] with the aromatic side chain facing the active site. However this residue is conserved in all other active β-glucosidases, and is thus less likely to be involved in hCBG substrate specificity.

Although the three amino-acid residues Val[168], Phe[225] and Tyr[308] are part of the cluster of hydrophobic amino acids in the aglycone-binding subsite close to the glycosidic bond, Val[168] was clearly essential for binding of the aglycone moiety whereas Phe[225] and Tyr[308] seemed to be more involved in global catalysis of artificial aryl-glycosides and flavonoid glucosides. Tailoring the aglycone (substrate) specificity of this type of enzyme will require mutations of more than a single residue in the aglycone subsite. A close inspection of the hCBG model shows that a second selectivity further away from the glycosidic linkage might be present, involving Met[172], Phe[179], Trp[345] and Phe[433]. Met[172] is unique to hCBG, while phenylalanine residues are also present in the binding pocket of other β-glucosidases in the same position as Phe[179] and Phe[433] in hCBG. Domain IV (lactase site) and domain III (phlorizin site) of LPH contribute 75% and 25% respectively to the total activity of this β-glucosidase on flavonoid glucosides [23]. Trp[345] is conserved in domain IV, but not in domain III, suggesting that Trp[345] could also be a candidate to engineer the aglycone specificity of human β-glucosidases.

In conclusion, a molecular basis has been provided for the aglycone specificity of hCBG. Moreover, the hCBG structural model provides a valuable tool for our further understanding of the fundamental difference in substrate specificity between members of family 1 glycosyl hydrolases and the design of engineered enzymes with new specificities.

REFERENCES

1   Daniels, L. B., Coyle, P. J., Chiao, Y. B., Glew, R. H. and Labow, R. S. (1981) Purification and characterization of a cytosolic broad specificity β-glucosidase from human liver. J. Biol. Chem. **256**, 13004–13013

2   Berrin, J. G., McLauchlan, W. R., Needs, P., Williamson, G., Puigserver, A., Kroon, P. A. and Juge, N. (2002) Functional expression of human liver cytosolic β-glucosidase in *Pichia pastoris*: insights into its role in the metabolism of dietary glucosides. Eur. J. Biochem. **269**, 249–258

3   Glew, R. H., Gopalan, V., Forsyth, G. W. and VanderJagt, D. J. (1993) The mammalian cytosolic broad-specificity β-glucosidase. In β-Glucosidases: Biochemistry and Molecular Biology (Esen, A., ed.), pp. 83–112, American Chemical Society (ACS), Washington, DC

4   Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. **280**, 309–316

5   Henrissat, B. and Davies, G. (1997) Structural and sequence-based classification of glycoside hydrolases. Curr. Opin. Struct. Biol. **7**, 637–644

6   Hakulinen, N., Paavilainen, S., Korpela, T. and Rouvinen, J. (2000) The crystal structure of β-glucosidase from *Bacillus circulans* sp. *alkalophilus*: ability to form long polymeric assemblies. J. Struct. Biol. **129**, 69–79

7   Wiesmann, C., Beste, G., Hengstenberg, W. and Schulz, G. E. (1995) The three-dimensional structure of 6-phospho-β-galactosidase from *Lactococcus lactis*. Structure **3**, 961–968

8   Barrett, T., Suresh, C. G., Tolley, S. P., Dodson, E. J. and Hughes, M. A. (1995) The crystal structure of a cyanogenic β-glucosidase from white clover, a family 1 glycosyl hydrolase. Structure **3**, 951–960

9   Kaper, T., Lebbink, J. H., Pouwels, J., Kopp, J., Schulz, G. E., van der Oost, J. and de Vos, W. M. (2000) Comparative structural analysis and substrate specificity engineering of the hyperthermostable β-glucosidase CelB from *Pyrococcus furiosus*. Biochemistry **39**, 4963–4970

10 Burmeister, W. P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S. and Henrissat, B. (1997) The crystal structures of *Sinapis alba* myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase. Structure **5**, 663–675

11 Sanz-Aparicio, J., Hermoso, J. A., Martinez-Ripoll, M., Lequerica, J. L. and Polaina, J. (1998) Crystal structure of β-glucosidase A from *Bacillus polymyxa*: insights into the catalytic activity in family 1 glycosyl hydrolase.J. Mol. Biol. **275**, 491–502

12 Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M. and Pearl, L. H. (1997) Crystal structure of the β-glycosidase from the hyperthermophilic archeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability. J. Mol. Biol. **271**, 789–802

13 Chi, Y. I., Martinez-Cruz, L. A., Jancarik, J., Swanson, R. V., Robertson, D. E. and Kim, S. H. (1999) Crystal structure of the β-glycosidase from the hyperthermophile *Thermosphaera aggregans*: insights into its activity and thermostability. FEBS Lett. **445**, 375–383

14 Czjzek, M., Cicek, M., Zamboni, V., Bevan, D. R., Henrissat, B. and Esen, A. (2000) The mechanism of substrate (aglycone) specificity in β-glucosidases is revealed by crystal structures of mutant maize β-glucosidase-DIMBOA, -DIMBOAGlc, and -dhurrin complexes. Proc. Natl. Acad. Sci. U.S.A. **97**, 13555–13560

15 Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J. P. and Davies, G. (1995) Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. Proc. Natl. Acad. Sci. U.S.A. **92**, 7090–7094

16 Davies, G. and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. Structure **3**, 853–859

17 Henrissat, B. and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. Biochem. J. **316**, 695–696

18 Keresztessy, Z., Brown, K., Dunn, M. A. and Hughes, M. A. (2001) Identification of essential active-site residues in the cyanogenic β-glucosidase (linamarase) from cassava (*Manihot esculenta* Crantz) by site-directed mutagenesis. Biochem. J. **353**, 199–205

19 Hays, W. S., VanderJagt, D. J., Bose, B., Serianni, A. S. and Glew, R. H. (1998) Catalytic mechanism and specificity for hydrolysis and transglycosylation reactions of cytosolic β-glucosidase from guinea pig liver. J. Biol. Chem. **273**, 34941–34948

20 Zecca, L., Mesonero, J. E., Stutz, A., Poiree, J. C., Giudicelli, J., Cursio, R., Gloor, S. M. and Semenza, G. (1998) Intestinal lactase-phlorizin hydrolase (LPH): the two catalytic sites; the role of the pancreas in pro-LPH maturation. FEBS Lett. **435**, 225–228

21 Arribas, J. C., Herrero, A. G., Martin-Lomas, M., Cañada, F. J., He, S. and Withers, S. G. (2000) Differential mechanism-based labeling and unequivocal activity assignment of the two active sites of intestinal lactase/phlorizin hydrolase. Eur. J. Biochem. **267**, 6996–7005

22 Nemeth, K., Plumb, G., Berrin, J. G., Juge, N., Jacob, R., Naim, H. Y., Williamson, G., Swallow, D. M. and Kroon, P. A. (2003) Deglycosylation by small intestinal epithelial cell β-glucosidases is a critical step in the absorption and metabolism of dietary flavonoid glycosides in humans. Eur. J. Nutr. **42**, 29–42

23 Day, A. J., Cañada, F. J., Diaz, J. C., Kroon, P. A., McLauchlan, R., Faulds, C. B., Plumb, G. W., Morgan, M. R. and Williamson, G. (2000) Dietary flavonoid and isoflavone glycosides are hydrolysed by the lactase site of lactase phlorizin hydrolase. FEBS Lett. **468**, 166–170

24 Mayer, C., Zechel, D. L., Reid, S. P., Warren, R. A. and Withers, S. G. (2000) The E358S mutant of *Agrobacterium* sp. β-glucosidase is a greatly improved glycosynthase. FEBS Lett. **466**, 40–44

25 Cregg, J. M., Barringer, K. J., Hessler, A. Y. and Madden, K. R. (1985) *Pichia pastoris* as a host system for transformations. Mol. Cell. Biol. **5**, 3376–3385

26 Hinnen, A., Hicks, J. B. and Fink, G. R. (1978) Transformation of yeast. Proc. Natl. Acad. Sci. U.S.A. **75**, 1929–1933

27 Lambert, N., Kroon, P. A., Faulds, C. B., Plumb, G. W., McLauchlan, W. R., Day, A. J. and Williamson, G. (1999) Purification of cytosolic β-glucosidase from pig liver and its reactivity towards flavonoid glycosides. Biochim. Biophys. Acta **1435**, 110–116

28 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**, 4673–4680

29 Roussel, A. and Cambillau, C. (1991) TURBO-FRODO. In Silicon Graphics Geometry Partners Directory, pp. 86, Silicon Graphics, Mountain View, CA

30 Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, Jr, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) The Protein Data Bank: a computer-based archival file for macromolecular structure. J. Mol. Biol. **112**, 535–542

31 Sali, A. and Blundell, T. J. (1993) Comparative protein modelling by satisfaction of spatial restraints. J. Biol. Chem. **234**, 779–815

32 Czjzek, M., Cicek, M., Zamboni, V., Burmeister, W. P., Bevan, D. R., Henrissat, B. and Esen, A. (2001) Crystal structure of a monocotyledon (maize ZMGlu1) β-glucosidase and a model of its complex with *p*-nitrophenyl β-D-thioglucoside. Biochem. J. **354**, 37–46

33 Gopalan, V., VanderJagt, D. J., Libell, D. P. and Glew, R. H. (1992) Transglucosylation as a probe of the mechanism of action of mammalian cytosolic β-glucosidase. J. Biol. Chem. **267**, 9629–9638

34 LaMarco, K. L. and Glew, R. H. (1985) Galactosylsphingosine inhibition of the broad-specificity cytosolic β-glucosidase of human liver. Arch. Biochem. Biophys. **236**, 669–676

35 Gopalan, V., Daniels, L. B., Glew, R. H. and Claeyssens, M. (1989) Kinetic analysis of the interaction of alkyl glycosides with two human β-glucosidases. Biochem. J. **262**, 541–548

36 Legler, G. and Bieberich, E. (1988) Active site directed inhibition of a cytosolic β-glucosidase from calf liver by bromoconduritol B epoxide and bromoconduritol F. Arch. Biochem. Biophys. **260**, 437–442

37 Marques, A. R., Coutinho, P. M., Videira, P., Fialho, A. M. and Sa-Correia, I. (2003) *Sphingomonas paucimobilis* β-glucosidase Bgl1: a member of a new bacterial subfamily in family 1 of glycoside hydrolases. Biochem. J. **370**, 793–804

38 Esen, A. and Gungor, G. (1993) Stability and activity of plant and fungal β-glucosidases under denaturing conditions. In β-Glucosidases: Biochemistry and Molecular Biology (Esen, A., ed.), pp. 214–239, American Chemical Society (ACS), Washington, DC

39 Pocsi, I. and Kiss, L. (1988) Kinetic studies on the broad-specificity β-D-glucosidase from pig kidney. Biochem. J. **256**, 139–146