Roles of the N-terminal domain and remote substrate binding subsites in activity of the debranching barley limit dextrinase

Andersen, Susan; Svensson, Birte; Møller, Marie Sofie

Published in:
B B A - Proteins and Proteomics

Link to article, DOI:
10.1016/j.bbapap.2019.140294

Publication date:
2020

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
Roles of the N-terminal domain and remote substrate binding subsites in activity of the debranching barley limit dextrinase

Susan Andersen, Birte Svensson, Marie Sofie Møller

PII: S1570-9639(19)30180-3
DOI: https://doi.org/10.1016/j.bbapap.2019.140294
Reference: BBAPAP 140294

To appear in: BBA - Proteins and Proteomics

Received date: 16 July 2019
Revised date: 15 September 2019
Accepted date: 10 October 2019

Please cite this article as: S. Andersen, B. Svensson and M.S. Møller, Roles of the N-terminal domain and remote substrate binding subsites in activity of the debranching barley limit dextrinase, BBA - Proteins and Proteomics (2019), https://doi.org/10.1016/j.bbapap.2019.140294

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier.
Title page

Roles of the N-terminal domain and remote substrate binding subsites in activity of the debranching barley limit dextrinase

Susan Andersen, Birte Svensson, Marie Sofie Møller*

Enzyme and Protein Chemistry, Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads, building 224, DK 2800 Kgs. Lyngby, Denmark
*Corresponding author: Marie Sofie Møller. E-mail: msmo@dtu.dk; Tel. +45 45 25 27 41
Abstract
Barley limit dextrinase (HvLD) of glycoside hydrolase family 13 is the sole enzyme hydrolysing α-1,6-glucosidic linkages from starch in the germinating seed. Surprisingly, HvLD shows 150- and 7-fold higher activity towards pullulan and β-limit dextrin, respectively, than amylpectin. This is investigated by mutational analysis of residues in the N-terminal CBM-21-like domain (Ser14Arg, His108Arg, Ser14Arg/His108Arg) and at the outer subsites +2 (Phe620Ala, Asp621Ala, Phe620Ala/Asp621Ala) of the active site. The Ser14 and His108 mutants mimic natural LD variants from sorghum and rice with elevated enzymatic activity. Although situated about 40 Å from the active site, the single mutants had 15–40% catalytic efficiency compared to wild type for the three polysaccharides and the double mutant retained 27% activity for β-limit dextrin and 64% for pullulan and amylpectin. These three mutants hydrolysed 4,6-O-benzylidene-4-nitrophenyl-63-α-D-maltotriosyl-maltotriose (BPNPG3G3) with 51–109% of wild-type activity. The results highlight that the N-terminal CBM21-like domain plays a role in activity. Phe553 and the highly conserved Trp512 sandwich a substrate main chain glucosyl residue at subsite +2 of the active site, while substrate contacts of Phe620 and Asp621 at subsite +3 are less prominent. Phe553Gly showed 47% and 25% activity on pullulan and BPNPG3G3, respectively having a main role at subsite +2. By contrast at subsite +3, Asp621Ala increased activity on pullulan by 2.4-fold, while Phe620Ala/Asp621Ala retained only 7% activity on pullulan albeit showed 25% activity towards BPNPG3G3. This outcome supports that the outer substrate binding area harbour preference determinants for the branched substrates amylpectin and β-limit dextrin.

Key words
Glycoside hydrolase family 13 subfamily 13; carbohydrate binding module 21-like; pullulanase; amylpectin; β-limit dextrin; α-limit dextrin

Abbreviations
BPNPG3G3, 4,6-O-benzylidene-4-nitrophenyl-63-α-D-maltotriosyl-maltotriose; CBM, carbohydrate binding module; CD, cyclodextrin; GG23G32, 63-α-D-glucosyl-maltotriosyl-maltotriose; G2SG4, 63-α-D-maltosyl-6-thiomaltotetraose; GH13, glycoside hydrolase family 13; GH13_13, GH13 subfamily 13; HvLD, Hordeum vulgare limit dextrinase; LD, limit dextrinase; OsLD, Oryza sativa subsp. indica 93-11 (rice) limit dextrinase; PDB, Protein Data Bank; PUL, pullulanase; PULI, type I pullulanase; PULII, type II pullulanase; SbLD, Sorghum bicolor (L.) Moench limit dextrinase; SbLD-GD, Sorghum bicolor frequent allele limit dextrinase; SbLD-RA, Sorghum bicolor rare allele limit dextrinase; SBS, surface binding site; SPR, surface plasmon resonance.
1. Introduction

Starch is the major seed storage polysaccharide in cereal crops (rice, maize, wheat, barley, and sorghum) and the most abundant α-glucan in human diet [1]. Notably, starch holds great potential in carbohydrate-based foods with beneficial qualities in human nutrition as resistant starch or prebiotics [2] as well as in feedstocks for bioethanol production [3]. Starch is deposited as semi-crystalline supramolecular structures referred to as granules composed of two polysaccharides, amylose and amylopectin, of a typical weight ratio of 1:3. While amylose is an almost linear α-1,4-glucan, amylopectin carries α-1,6-branch chains at about 5% of the glucosyl residues. The degree of polymerisation, branch chain-length distribution, and branching pattern in amylopectin are plant-specific [4].

Despite the simplicity of the chemical structure of the α-glucan substrates of glycoside hydrolase family 13 (GH13) more than 20 enzyme specificities catalyse their conversions as reflected also in the division of GH13 into 42 subfamilies [5–7] (Carbohydrate-Active enZymes database, CAZy; http://www.cazy.org/; [5]). Among GH13 members, α-glucan debranching enzymes cover a surprising variety of substrate specificities characterised by specific protein structural elements and sequence motifs [8]. Limit dextrinase from barley (Hordeum vulgare) (HvLD) belongs to the type I pullulanases (PUL) (EC 3.2.1.41), which exclusively act on α-1,6-glucosidic linkages, in contrast to the type II pullulanases (PULII) hydrolysing both α-1,4- and α-1,6-glucosidic linkages [8]. HvLD is categorized into GH13 subfamily 13 (GH13_13) [5,8] containing HvLD-like debranching enzymes and closely related bacterial pullulanases (PULs). HvLD-like enzymes, are commonly referred to as plant PULs due to their high activity towards pullulan [9,10], a water-soluble linear polysaccharide composed of α-1,6-connected maltotriose repeats from the yeast-like Aureobasidium pullulans [11]. Thus pullulan is not a natural substrate of plant PULs. These enzymes are therefore named limit dextrinases (LDs) after their natural substrates limit dextrins formed in the starch catabolism during germination.

HvLD is the sole debranching enzyme in germinating barley seeds and acts on α-1,6-linkages in α-glucans, primarily α- and β-limit dextrins released from storage starch by concerted action of the amylolytic enzymes; α-amylase, β-amylase, and α-glucosidase [12–16]. HvLD has high activity on pullulan, as opposed to the other debranching enzyme in plants isoamylase of GH13_11, that is confined to diurnal starch metabolism [17].

Surprisingly, HvLD has about two orders of magnitude lower activity for amylopectin than pullulan [18,19]. HvLD possesses the typical core multi-modular architecture of GH13 of a catalytic (β/α)8-barrel fold with an intimately associated C-terminal antiparallel β-sandwich domain [21,24,25]. In addition it contains an N-terminal family 21-like carbohydrate binding module (CBM21-like) found in PULs [20,21] followed by a CBM48 (Fig. 1). No carbohydrate binding function has yet been demonstrated for CBM48 from PULs or the CBM21-like N-terminal domain although these both are starch binding domain CBM families [22,23].

The HvLD CBM21-like domain structurally resembles the CBM21 of Rhizopus oryzae glucoamylase [26], but residues from the two binding sites in this glucoamylase starch binding domain are not conserved in HvLD [21]. Indeed for other debranching enzymes, which have an N-terminal CBM41, CBM48, and/or CBM68 starch binding domains, confirmed ligand binding is rare, albeit seen to CBM41 of Streptococcus pneumoniae PULI from GH13_12 [27] and Klebsiella pneumoniae PULI [28], and CBM68 in Anoxybacillus sp. LM 18-11 PUL of GH13_14 [29]. The consequence of these remote binding sites for enzyme function is not understood. It is tempting to categorise these sites as surface binding sites (SBS), which are situated outside of the active site region and commonly contribute to function of GH13 enzymes [30–32]. In LDs a possible involvement of the CBM21-like
domain in activity was highlighted as evidenced by sorghum (*Sorghum bicolor* (L.) Moench) and rice (*Oryza sativa* subsp. *indica* 93-11) GH13_13 LDs. Firstly, a low-frequency allele of the drought-tolerant cereal sorghum was associated with the starch metabolic gene encoding LD (*SbLD*) and a homozygote of this sorghum line compared to the line carrying the frequent allele conferred 67% higher activity towards pullulan and 41% higher *in vitro* digestibility of starch [33]. These effects were suggested to stem from glycine and aspartic acid in

**Fig. 1.** Overall structure of *HvLD* (PDB ID: 4A1O) with 6'-α-D-glucosyl-maltotriosyl-maltotriose (GG$_2$G$_3$; purple sticks; form PDB ID: 4J3X) and a branched oligosaccharide (G3G$_3$; green sticks; from PDB ID: 4J3W) in the active site. The CBM21-like N-terminal domain is in red; CBM48 in teal; the catalytic domain in grey; and the C-terminal domain in orange. Ser14, Ala90 and His108 in the N-terminal domain are shown as green spheres; catalytic residues Asp473, Gln510, Asp642 as yellow spheres; the Asp600–Leu622 loop in pink; Phe620 and Asp621 as violet-purple spheres; Trp512 and Phe553 as deep blue spheres. The N-terminal domain of the frequent allele LD, referred to as *SbLD*-GD, being mutated to arginine and alanine in the rare allele, *SbLD*-RA. These positions correspond to Ser14 and Ala90 in the *HvLD* CBM21-like domain [33]. Secondly, examination of enzyme activities in developing rice seeds using zymology revealed two polymorphic LD (*OsLD*) variants which differ at nine amino acid positions, of which one aligns with His108 in the *HvLD* CBM21-like domain [34].

Specificity of GH13 enzymes beyond direct protein residue–substrate contacts is defined by the topology of the active site, which in debranching enzymes has two characteristic parallel crevices [35]. Also short canonical regions extending from β-strands of the catalytic (β/α)$_8$-barrel display conserved sequence motifs correlated with different substrate specificities [6]. *HvLD* Trp512 belongs to conserved region III associated with the fifth β-strand of GH13 [6] and forms an aromatic sandwich with Phe553, situated in a short loop between the sixth β-strand and sixth α-helix of the catalytic (β/α)$_8$-barrel. This structural element flanks the substrate main chain at subsite +2 [19] (Fig. 2A and 2B) and is generally critical for substrate binding [36]. In different *HvLD* complexes Trp512 stacks onto specific glucosyl residues of cyclodextrins (CDs) and α-limit dextrins, i.e. maltooligosaccharides containing one or more α-1,6-linkages [19,24]. Phe553 in the three-dimensional space superimposes with tyrosine or phenylalanine in GH13_12–14 members but adopts different conformations to optimize substrate interactions [19]. Its tight substrate sandwich with Trp512 is understood to control substrate positioning to assure productive accommodation of only α-1,6- and not α-1,4-linkages for hydrolysis between subsites +1 and −1 (Fig. 2A) [19]. *HvLD* has high affinity for CD inhibitors (*K$_d$* for β-CD is 0.7 μM) [24]. In crystal structures (PDB ID: 2Y4S, 2Y5E, and 4J3U) Phe553 intrudes into the central cavity of the CD ring interacting with subsites 0', +1, +2 and +3 [19,24] (Fig. 2C). Recently, an important role of the corresponding Phe746 of *K. pneumoniae* PULI has been shown in inhibition by β-CD, as a 1,700-fold increase of

**Fig. 2.** Nomenclature and overview of subsites in the *HvLD* active site. (A) *HvLD* (PDB ID: 4A1O) superimposed with 6'-α-D-glucosyl-maltotriosyl-maltotriose (GG$_2$G$_3$; purple sticks; from PDB ID: 4J3X) and G3G$_3$ (green sticks; from PDB ID: 4J3W) in the active site. Catalytic site, yellow; Asp600–Leu622 loop, pink; Phe620 and Asp621, violet-purple; Trp512 and Phe553, deep blue; CBM21-like N-terminal domain, red; CBM48, teal; C-terminal domain, orange. The scissile α-1,6 bond of the substrate is placed between subsites −1 and +1 and marked with an arrow. (B) 90° right turned view of (A). (C) the same as (A), but with a β-CD inhibitor (cyan sticks; from PDB ID: 2Y4S) included. (D) schematic representation of the subsites accommodating G3G$_3$ and GG$_2$G$_3$. The grey hexagon indicates the reducing end of the oligosaccharide. *K*$_i$ for the Phe746Ala mutant and notably a reduction of both *K*$_{in}$ and *k*$_{cat}$ were seen [28].

GH13_13 enzymes possess a characteristic long loop (Asp600–Leu622 in *HvLD*) as part of the connection between the sixth α-helix and the seventh β-strand of the catalytic
(β/α)8-barrel. This loop is lacking in the two other PULI subfamilies GH13_12 and GH13_14 [19] both harbouring PULs with higher activity on amylpectin than members of GH13_13 [8]. Structural evaluation of Phe620 and Asp621 on this long loop suggests they reach into the substrate binding area and narrow the main chain binding crevice beyond subsite +3, thus restricting substrate main chain binding towards the reducing end to three glucosyl residues from the branch point [19] (Fig. 2). This long loop also moved towards the ring of maltosyl-S-β-cyclodextrin (G2Sβ-CD) and Phe620 stacks onto one of its glucose residues at subsite +3 [19]. However, in earlier solved α- and β-CD complex structures this loop appeared dynamic based on absence of electron density for the side chain of Phe620 and of hydrogen bonds between Asp621 and the CDs [24]. The present work investigates the outer substrate aglycon binding region. Previously, an HvLD mutant Met440Gly addressed substrate branch chain binding beyond subsites −3/−4, but was without effect on pullulan and resulted in 2.6-fold decrease in activity for amylpectin [19].

Here two aspects of polysaccharide binding and specificity of HvLD are explored. First, effects are analysed of Ser14Arg, His108Arg, Ser14Arg/His108Arg mutations in the N-terminal CBM21-like domain as guided by variants of sorghum and rice LDs. Secondly, based on comparison of crystal structures of oligosaccharide complexes of different debranching enzymes, active site mutants Phe553Gly, Phe620Ala, Asp621Ala, and Phe620Ala/Asp621Ala at the outer reducing-end main chain binding area probe specificity determinants in PULs.

2. Materials and methods

2.1. Bioinformatic analysis

Sequence comparison and structural modelling tools were applied to investigate the function of the N-terminal CBM21-like domain (residues 2-124) of HvLD. Protein sequence searching by BLASTP 2.7.1+ [37], after removal of fragment and redundant sequences (>99% identity), resulted in 42 sequences (Table S1) subjected to multiple alignment performed in MAFFT [38] and visualised by ESPript using default settings (http://espript.ibcp.fr; [39]). The 42 sequences were moreover investigated through the Conserved Domains Database [40], the CAZY database (CAZy; http://www.cazy.org/; [5]), and dbCAN [41]. Structural positions of selected residues in N-terminal domain of HvLD were compared with two homology models of SbLD-RA (ABK63626.1; sequence only covering the N-terminal domain) and OsLD (EEC76742.1). For both enzymes one model was made using HHpred [42] in combination with Modeller [43] with HvLD in complex with the branched thio-linked hexasaccharide 63-α-D-maltosyl-6-thiomaltotetraose (G2SG24) (PDB ID: 4J3V) as template and the second model by using SWISS-MODEL [44] and as template HvLD in complex with an amylpectin fragment G3G13 (PDB ID: 4J3W). The quality of the obtained models was evaluated by ProQ [45] and PyMol 2.0 (Schrödinger, LLC, New York, NY, USA) (Table S2). Two structure-based alignment tools FATCAT [46] and DALI against all PDB entries [47] were applied to investigate if the best match to the N-terminal domain is still CBM21. Additionally, a structure-based sequence alignment of HvLD and structure-determined GH13_12–14 together with GH13_11 members (isomylases and glycogen debranching enzymes) and sequences of characterised members of GH13_39 (PULs with dual α-1,4 and α-1,6-hydrolytic activity) was conducted by PROMALS3D [48] (Table S3).

2.2. Mutagenesis and Pichia pastoris transformation

Established methods for expression, mutagenesis, screening and purification of HvLD [18] were optimized to increase production efficiency and introduce His-Tag purification. Briefly, the HvLD gene was codon optimised for expression in P. pastoris (Life Technologies, Carlsbad, USA) and cloned into the pPICZαA vector (Invitrogen) for easy selection on zeocine and encoding an C-terminal His-tag. This was particularly important as
HvLD mutated at subsites +2 and +3 was expected not to bind to β-CD-Sepharose used before for HvLD affinity chromatographic purification [18]. The Q5® Site-Directed Mutagenesis Kit (New England Biolabs Ltd., Ipswich, USA) was used for higher yield and fidelity of variant plasmids. Production and transformation of competent \textit{P. pastoris} strain X-33 cells was achieved using the Pichia EasyComp™ Kit (Invitrogen, Carlsbad, USA).

Overall screening and production time of recombinant HvLD variants was reduced by 50%. Yields were typically 0.5 mg L$^{-1}$ culture or less for mutants.

\textit{HvLD} mutants Ser14Arg, His108Arg, Ser14Arg/His108Arg, Phe553Gly, Phe620Ala and Phe620Ala/Asp621Ala were obtained (Q5® Site-Directed Mutagenesis Kit; New England Biolabs Ltd.) using 10 ng pPICzaA-HvLD double stranded (ds) DNA template. For Phe620Ala/Asp621Ala and Ser14Arg/His108Arg dsDNA templates were pPICzaA-HvLD-Phe620Ala and pPICzaA-HvLD-Ser14Arg, respectively. Mutagenic primers are given in Table S4. The site directed mutagenesis PCR protocol included an initial denaturation (30 s, 98°C), followed by 25 cycles (each 10 s at 98°C, 30 s at a primer specific annealing temperature (Table S4), and 95 s at 72°C) and a final extension (2 min, 72°C). pPICzaA-HvLD wild type and variants were transformed into \textit{E. coli} DH5α [49] using heat shock [50] and selected with 25 µg mL$^{-1}$ zeocin (Invitrogen). Mutant plasmids were verified by sequencing (GATC Biotech, Konstanz, GER), linearized by PremI (20 units; New England Biolabs), purified (GeneJet PCR Purification Kit; Fermentas, Waltham, USA), and precipitated (Pellet Paint® Co-Precipitant; Merck Millipore, Billerica, USA) yielding about 3 µg linearized plasmid. Electrocompetent \textit{P. pastoris} X-33 cells were transformed using Pichia EasyComp™ Kit (Invitrogen) according to the manufacturer’s recommendations, plated on yeast extract peptone dextrose with sorbitol (YPDS) supplemented with 100 µg mL$^{-1}$ zeocin (Invitrogen) and incubated at 30°C until colonies were visible. HvLD production by \textit{P. pastoris} transformants was assessed for small scale cultures (10 mL buffered glycerol-complex medium, BMGY; 30°C, 24 h); cells were harvested (3,000 g, 10 min, 22°C) and resuspended to a final OD$_{600}$ = 1 in BMGY buffered methanol-complex medium (BMMY) (22°C, methanol supplemented to 0.5% (v/v) every 24 h) with induction maintained for 72 h. Supernatants were 10-fold concentrated (3000 g, Centricron, 30 kDa cut-off; Millipore) prior to assaying activity (Limit-Dextrizyme tablets; Megazyme, Irishtown, IRL) essentially as previously described, but using 50 mM instead of 100 mM sodium acetate pH 5.5 [18,51]. In addition, immunoblotting was done to confirm the presence of HvLD as described [18,52]. Based on the activity assay and immunoblot results well-expressing transformants were selected for protein production.

2.3. Production and purification of recombinant HvLD

HvLD wild type and variants were produced by high cell-density fermentation according to the Pichia Fermentation Process Guidelines in 2 L Fermentation Basal Salts Medium (Invitrogen) with a glycerol batch phase, a glycerol fed-batch phase, and a methanol fed-batch maintained for approximately 90 h [18] using a 5 L bioreactor (Biostat B; B. Braun Biotech International, Melsungen, Germany). Cultures were centrifuged (8,000 g, 30 min, 4°C). The supernatants were filtered (0.45 µm Durapore membrane filters; Millipore), concentrated 10-fold (10-kDa cut-off filter, Pellicon ultra-filtration unit; Millipore), adjusted to pH 7.4 with Na$_2$HPO$_4$ (solid), filtered, loaded (3.5 mL min$^{-1}$) onto a 5 mL HisTrap Excel column (GE Healthcare, Little Chalfont, UK), washed with 25 column volumes (CV) 20 mM sodium phosphate, 500 mM NaCl, pH 7.4, and eluted by a linear imidazole gradient (0–320 mM; 30 CV; 1 mL min$^{-1}$). Fractions (1.92 mL) were collected directly into 80 µL 1 M MES/NaOH, 125 mM CaCl$_2$, 0.125% Triton X-100, pH 6.0. The resulting pH 6.5 ensured HvLD stability. HvLD-containing fractions were pooled, concentrated (4000 g, 30 kDa cut-off, Amicon Ultra-15 centrifugal filter units; Millipore) and gel filtrated (Hiload 26/60
Superdex 200 column; GE Healthcare) in 50 mM MES/NaOH, 250 mM NaCl, 0.5 mM CaCl₂, pH 6.6 (0.5 mM min⁻¹). Fractions containing pure HvLD (monitored by SDS-PAGE) were pooled and concentrated as above. Protein concentration was determined spectrophotometrically using a theoretical molar extinction coefficient, ε₂₈₀ = 130180 M⁻¹ cm⁻¹ (ProtParam tool, ExPASy server; [53]). HvLD wild type and variants stocks were added NaN₃ to 0.02% (w/v) and stored at 4°C.

2.4. Enzyme activity

Initial rates of hydrolysis by 2.1–81 nM HvLD wild type and variants (110 µL) of 0.02–0.6 mg·mL⁻¹ pullulan (Megazyme), 0.23–10 mg·mL⁻¹ potato amylopectin (Sigma-Aldrich) and 0.2–5 mg·mL⁻¹ β-limit dextrin (Megazyme) (990 µL) in 20 mM sodium acetate pH 5.5, 5 mM CaCl₂, 0.005% Triton X-100 at 37°C were determined using a reducing sugar assay [18,19]. Protein concentration was determined by fitting the Michaelis–Menten equation (Eq. (1)); amylopectin and β-limit dextrin, respectively) were removed after 3, 6, 9, 12, and 15 min, mixed with 0.4 M sodium carbonate pH 10.7, 2.5 mM CuSO₄, 2.5 mM 4,4-dicarboxy-1,2-biquinoline, 6 mM L-serine (500 µL) to stop the reaction [54], followed by addition of Milli-Q water to a final volume of 1 mL, incubation (80°C, 30 min) and absorbance measurement at 540 nm (Epoch Microplate Spectrophotometer, BioTek). Released reducing sugar was quantified using maltoose (0–55.5 µM) as standard. Kinetics were analysed in duplicates for wild type and variants except for the Ser14Arg, His109Arg and Phe620Ala/Asp621Ala variants obtained in very low yields and measured in single experiments. Kᵢ,ₘ and kₗ were determined by fitting the Michaelis–Menten equation (Eq. (1); amylopectin and β-limit dextrin) or the equation for uncompetitive substrate inhibition (Eq. (2); pullulan) to initial velocities [18,19] using the Enzyme Kinetics Module of the program Sigmaplot 14.0 (Systat Software, Chicago, USA). Kᵢ,ₘ is the dissociation constant for the inhibitory substrate–enzyme–substrate ternary complex.

\[
V = \frac{v_{\text{max}}}{1 + \frac{k_m}{[S]}} \quad (1)
\]

\[
V_{i,\text{sub}} = \frac{v_{\text{max}}}{1 + \frac{k_m}{[S]} + \frac{[I]}{K_{i,s}}} \quad (2)
\]

Specific activity of HvLD wild type and variants (3.6–27 nM) was determined on 0.4 mg·mL⁻¹ pullulan essentially as described above. One Unit (U₁) is defined as the amount of enzyme releasing one micromole of reducing sugar per min under the assay conditions. Specific activity of HvLD wild type and variants (50–107 nM) was also determined for 4,6-O-benzylidene-4-nitrophenyl-6'α-D-maltotriosyl-maltotriose (BNPG3G3; Megazyme). Enzyme (30 µL) in 100 mM sodium acetate, pH 5.5, 0.05% BSA, 0.02% NaN₃ was pre-incubated (40°C, 5 min). The reaction was initiated by addition of 30 µL PULL6G reagent containing substrate and stopped after 10 min by 900 µL 2% Tris Base (w/v), pH 9.0. Substrate consumption was quantified spectrophotometrically at 400 nm using 4-nitrophenol (0.1–1 mM) as standard. One Unit (U₂) is defined as the amount of enzyme hydrolysing one micromole of substrate under the assay conditions.

2.5. Surface plasmon resonance binding analysis of HvLD-Phe553Gly

The affinity of HvLD-Phe553Gly for β-CD was determined by surface plasmon resonance (SPR) (BIACore®T100; GE Healthcare, Sweden) as previously described [18]. HvLD-Phe553Gly was immobilized in the presence of 1 mM β-CD on a CM5 sensor chip (GE Healthcare, Sweden) by random amine coupling (1800 response units, RU). Binding to β-CD was tested by passing 1 mM β-CD in 20 mM sodium acetate, pH 5.5, 100 mM NaCl, 0.005% surfactant P-20 (running buffer) over the surface at 25°C using 3 min association and 2 min dissociation.

3. Results and discussion
3.1. Bioinformatics and homology modelling of the N-terminal domain

Sequence- and structure-based alignments and modelling were applied to evaluate which substitutions to do in the HvLD N-terminal CBM21-like domain (residues 2–124) to assess the function of this domain for activity. One hundred homologs were identified using BLASTP having E-values between 2·10^{-77} and 5·10^{-31}, 95–100% query coverage, and 51–96% sequence similarity. Forty-two non-redundant sequences from seeds, fruits and flowers of plants and trees (Fig. 3) were annotated as PULs, LDs or hypothetical proteins, the majority being marked as predicted (Table S1). According to the CAZy database [5] only four proteins – sorghum, rice, wheat (Triticum aestivum L.) and HvLD – have been characterised. Comparison of the HvLD, SbLD and OsLD sequences revealed that the rare allele variant of sorghum, SbLD-RA, found in grains containing highest amount of digestible starch [33] has arginines at positions corresponding to HvLD Ser14 and His108 (6_ABK63626, Fig. 3). Notably, the most frequent sorghum allele, SbLD-GD, has glycine and aspartic acid (7_ABK63617, Fig. 3), which are not conserved in CBM21-like domains, corresponding to serine and alanine at HvLD positions 14 and 90. In fact lysine and arginine commonly occur at the position corresponding to Ser14 in HvLD (1_AAF98802, Fig. 3). SbLD-RA and HvLD both have alanine at position 90 (HvLD numbering). It is therefore hypothesized that arginine in SbLD-RA, matching Ser14 HvLD, is responsible for higher PUL activity and more digestible starch in SbLD-RA compared to SbLD-GD, as measured using porcine enzymes as an in vitro proxy for monogastric digestion of starch from the rare allele sorghum homozygote [33]. Also a variant of the rice OsLD line (15ACY56102, Fig. 3), which was associated with increased activity on branched substrates in developing seeds [34], has arginine aligned to HvLD His108 similarly to SbLD-RA (Fig. 3).

The structural positioning of the residues in SbLD-RA and OsLD equal to Ser14, Ala90 and His108 of HvLD was evaluated using the homology model, which showed only minor differences of SbLD-RA and OsLD with respect to the overall structure of HvLD (PDB ID: 4AIO) and were evaluated as good to very good (Table S2). Alignment of all models with HvLD (PDB ID: 4AIO) using PyMol 2.0 thus showed overall rmsd for Cα between 0.251 and 0.423 Å (Table S2) and overlaying the SbLD-RA and OsLD models with the HvLD structure confirmed that counterparts of HvLD Ser14, Ala90 and His108 are exposed on the surface of the N-terminal CBM21-like domain (Fig. 4) approving our choice of the Ser14 and His108 for mutational analysis. The structural similarity of the N-terminal domain of HvLD and CBM21 in R. oryzae glucoamylase was confirmed by the FATCAT database search (Table S5) and a DALI search (Table S6) although key residues of its two ligand binding sites were not conserved in HvLD [21].

3.2. Impact of remote residues on the catalytic activity

3.2.1. N-terminal CBM21-like domain mutants

Kinetic parameters towards pullulan, amylopectin and β-limit dextrin confirmed that HvLD has much higher (150-fold) catalytic efficiency towards pullulan than amylopectin (Table 1). Moreover, for β-limit dextrin, which is an important natural substrate for cereal LDs/PULs, lower K_M and higher k_cat yield 6.5-fold higher catalytic efficiency than for amylopectin. This is in excellent agreement with LDs having preference for branched maltooligosaccharides, also referred to as α-limit dextrans. The function and possible involvement of the N-terminal domain in the activity towards amylopectin was investigated using three different HvLD mutants. Firstly, the previously

**Fig. 3.** Multiple sequence alignment of the N-terminal domain (residues 2–124) of HvLD and homologue sequences obtained from a BLAST search using the HvLD N-terminal CBM21-like domain as query (top line). Strictly conserved, white letters with red background; >70% conservation, red and framed blue. Position
numbers correspond to the HvLD sequence. Note the numbering is shifted by one in the HvLD structures referred in the text due to one extra residue starting at the N-terminal sequence originating from the used vector for recombinant protein expression [18]. Hence residue 13 in the sequence is identical to residue 14 in the structures and so forth. Sequences from sorghum and rice are highlighted in pink and green, respectively. The positions of the HvLD residues Ser14, Ala90, and His108 mutated in rice and sorghum LD variants are marked by asterisks and coloured according to their origin. Table S1 shows GenBank accessions and organisms. The alignment was visualised using ESPript [39].

Fig. 4. Comparison of homology models of SbLD-Ra and OsLD 93-11 with the HvLD structure. (A) Superimposition of N-terminal domains of HvLD (PDB ID: 4AIO, grey), and homology models of SbLD-Ra (NCBI: ABK63626.1; violet) and OsLD 93-11 (NCBI: EEC76742.1; split pea). Residues corresponding to HvLD Ser14, Ala90 and His108 (observe that the side chain of this residue is not completely solved in the crystal structure due to lack of electron density [21]) in the sorghum and rice models are shown as sticks and labelled according to origin following the same colour code. (B) Surface representation of HvLD with the three N-terminal domain residues highlighted in orange and active site residues in blue.

<table>
<thead>
<tr>
<th>HvLD</th>
<th>Pullulan</th>
<th>Amylopectin</th>
<th>β-limit dextrin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K_m</td>
<td>k_cat</td>
<td>k_cat/K_m</td>
</tr>
<tr>
<td>wild type</td>
<td>0.08±0.008</td>
<td>24±1.3</td>
<td>300±34</td>
</tr>
<tr>
<td>Ser14Arg*</td>
<td>0.1±0.02</td>
<td>6.3±0.6</td>
<td>63±14</td>
</tr>
<tr>
<td>His108Arg*</td>
<td>0.2±0.03</td>
<td>9.0±1.0</td>
<td>45±8</td>
</tr>
<tr>
<td>Ser14Arg/His108Arg*</td>
<td>0.1±0.01</td>
<td>19±1.5</td>
<td>190±24</td>
</tr>
<tr>
<td>Phe620Ala</td>
<td>0.09±0.02</td>
<td>18±1.7</td>
<td>200±48</td>
</tr>
<tr>
<td>Asp621Ala</td>
<td>0.12±0.01</td>
<td>56±2.7</td>
<td>467±45</td>
</tr>
<tr>
<td>Phe620Ala/Asp621Ala*</td>
<td>0.09±0.01</td>
<td>3.8±0.3</td>
<td>42±6</td>
</tr>
</tbody>
</table>

Table 1 Kinetic parameters of HvLD wild type and mutants in the CBM21-like N-terminal domain (Ser14 and His108) and at remote main chain binding subsite +3 (Phe620 and Asp621) of the active site towards pullulan, amylopectin and β-limit dextrin. *Uncompetitive substrate inhibition, †Classical Michaelis-Menten equation. Due to very low HvLD mutant protein yields, the activity was not determined for some substrates (n.d.) or only determined from single experiments (*). Standard deviations refer to deviations from the fitted model.

reported sorghum SbLD-Ra from seeds with high debranching activity and more digestible starch [33], was mimicked by HvLD-Ser14Arg, which, compared to wild type, lost 65–80% catalytic efficiency for the three polysaccharides mostly due to reduced k_cat (Table 1). Moreover, HvLD-His108Arg lost 70–85% catalytic efficiency for the polysaccharide substrate also mostly due to reduced k_cat and emulated OsLD displaying increasedzymograph degradation of amylopectin [34]. For both mutants the decrease in catalytic efficiencies (Table 1) and specific activities (Fig. 5A) was similar for amylopectin, β-limit dextrin and pullulan, while the activity of the double mutant HvLD-Ser14Arg/His108Arg decreased by only 35–37% activity for pullulan and amylopectin, though by 73% for β-limit dextrin similarly to the single mutants (Fig. 5A). The possibility of Ser14 and His108 exerting a long-range impact on the active site in HvLD was assessed from their reduced mutant activity on the hexasaccharide BPNPG3G3 presumed to bind at subsites +1 through +3 and –1 through –3, placing the scissile α-1,6-bond between subsites +1 and –1. Relative to wild type, HvLD-Ser14Arg and HvLD-His108Arg were 49% and 32% less active on BPNPG3G3 compared to loss of 71% and 47% activity, respectively for pullulan (Fig. 5A). The double mutant HvLD-Ser14Arg/His108Arg compensated and had the same specific activity (109%) as wild type on BPNPG3G3, but lost some, albeit less activity than the single mutants on the polysaccharides (Fig.)
debranching enzymes are lacking the loop where Phe553 is situated (Fig. 6). Phe553 intrudes sequence alignments show glycine or tyrosine at this position. However, a superposition of the structures from GH13_12 and GH13_11 shows that Phe553 occupies the same three-dimensional space (Fig. 6). The alignment was visualised using ESPript [39].

3.2.2. Substrate main chain binding subsites +2 and +3 at the active site

Insights into substrate interactions at the active site of debranching enzymes have been particularly difficult to obtain and only one structure, namely that of HvLD was determined of a branched substrate complex [19]. Knowledge still lacks, however, on how glucosyl residues of the substrate main chain at some distance of the subsites +1 and −1, encompassing the site of catalysis, contribute to substrate specificity. In HvLD Trp512, Phe553, Phe620 and Asp621 appear critical at subsites +2 and +3 (Fig. 2). Thus to compare side chains accommodating substrate main chain at these subsites, a structure-based multiple sequence alignment was carried out for debranching enzymes of GH13_11, i.e. isoamylases, GH13_12–14 PULs, and GH13_39 PULII (Fig. 6; for full sequence alignment see Fig. S1 and for corresponding organisms see Table S3).

HvLD Trp512 at subsite +2 is invariant in the structural alignment of GH13_11–14 and GH13_39 (Fig. 6) in agreement with earlier findings [6]. Therefore Trp512 is not considered useful for mutational analysis of specificity determinants. Phe553, however, is not generally conserved in GH13, and previously, it was shown that a structure-based multiple sequence alignment does not bring the amino acid residues of other PULs equivalent to Phe553 in perfect alignment (Fig. 6).

Fig. 6. Excerpt of structure-based sequence alignment of isoamylases of GH13_11, PULs of GH13_12–14 and PULII of GH13_39 (See Fig. S1 for full alignment and Table S3 for organisms). Protein structures were used for the alignment except for GH13_39, where GH13_39 protein sequences from selected characterized enzymes were included as no structures were available from the subfamily. The numbering above the alignment corresponds to the HvLD structure numbering (PDB ID: 4AIO). Fully conserved residues are in white letters on red background and 70% conservation is in red letters in a blue box. Residues corresponding to HvLD Trp512, Phe553, Phe620 and Asp621 are in green boxes. The alignment was visualised using ESPript [39].

Fig. 7. Comparison of HvLD (GH13_13) and isoamylases (GH13_11). HvLD (PDB ID: 4AIO; grey) shown with 6′-α-D-glucosyl-maltotriosyl-maltotriose (GG3G3; pink sticks) from the HvLD complex structure (PDB ID: 4J3X). Asp600-Leu622 loop (pink); Phe620 and Asp621 (violet-purple sticks); Trp512 and Phe553 (deep blue sticks). Isoamylase 1 (GH13_11) from Chlamydomonas Reinhardtii (PDB ID: 4J7R; green); isoamylase/4-α-glucanotransferase (GH13_11) from Sulfolobus Solfataricus (PDB ID: 2VNC; bright orange).

However, a superposition of the structures from GH13_12–14 showed that the residues corresponding to Phe553 occupied the same three-dimensional space [19]. In a few cases sequence alignments show glycine or tyrosine at this position [19]. GH13_11 and GH13_39 debranching enzymes are lacking the loop where Phe553 is situated (Fig. 6). Phe553 intrudes
into the central cavity of CD inhibitors in HvLD complexes [19,24] and HvLD-Phe553Gly was made to interrogate the importance of this aromatic side chain. The structure-based alignment furthermore showed that the long loop from α6 to β7 in HvLD (Asp600–Leu622) of GH13_13 is absent in GH13_12 and 14 [19]. The loop, however, exists in one GH13_11 isoamylase from Chlamydomonas reinhardtii (PDB ID: 4J7R) [59] (Fig. 6), but adopts a different position in that enzyme than in HvLD (Fig. 7). Both Phe620 and Asp621 from this loop are conserved in GH13_13 PULs; Phe620 can be Phe, Tyr or His, while Asp621 is an Asn [19]. Phe620 and Asp621 are situated near subsite +3 in the complex of HvLD with G2SβCD (PDB ID: 4J3U) [20] and therefore assigned a role in substrate binding, presumably accommodating the third glucosyl residue towards the reducing end from the α-1,6 scissile bond. The electron density of Phe620 and Asp621 at subsite +3 is in some of the available HvLD structures lacking for both or just one of the side chains. Only in the structure with G2SβCD (PDB ID: 4J3U) Phe620 is seen to interact with the CD ring. Asp621 is interacting with ligand in the maltotetraose containing structures (PDB ID: 4J3S and 4J3T), while it is present in the structure with G2SβCD (PDB ID: 4J3U), but it is not interacting with the ligand. In the complex with G2SβCD, the loop backbone is twisted differently than in the rest of the available HvLD structures. Based on these different observations, the loop is thought to be flexible [19]. The loop and Phe620 are not conserved in subfamilies GH13_12 and GH13_14 and generally lack in most of the GH13 subfamilies included in the structure-based alignment in Fig. 6. Thus HvLD Phe620Ala, Asp621Ala and the double mutant Phe620Ala/Asp621Ala, were made to explore the role of these loop residues for activity on poly- and oligosaccharide substrates.

3.2.3. Subsite +2 and +3 mutants

HvLD-Phe553Gly targeting the substrate main chain binding at subsite +2 lost 53% specific activity towards pullulan and 75% towards BPNPG3G3 (Fig. 5B). The oligosaccharide substrate obviously depends more strongly on the specific binding at subsite +2. SPR analysis confirmed that Phe553Gly was unable to bind β-CD. When binding to 1 mM β-CD was tested no significant signal, hence no binding was detected, while the affinity of wild-type HvLD is excellent with $K_d = 0.7 \mu M$ [18], thus the affinity must have been lost by a factor of more than 3 orders of magnitude. Recently, an important role in activity was also reported for Phe746 in K. pneumoniae PULi [28] that is equivalent to HvLD Phe553. Thus $K_i$ for inhibition of Phe746Ala of K. pneumoniae PULi by β-CD increased 1,700-fold compared to the wild type enzyme.

The catalytic efficiency of the subsite +3 mutant HvLD-Phe620Ala was reduced by only 33, 20, and 33% on pullulan, amylpectin, and β-limit dextrin, respectively (Table 1). The $k_{cat}$ value on pullulan was reduced by 25% while the $K_m$ was unchanged, whereas on amylpectin $k_{cat}$ increased by 28% and $K_m$ increased by 58%. For HvLD-Asp621Ala, noticeably the catalytic efficiency on pullulan was increased by 58%, while it was reduced about 30% on amylpectin and β-limit dextrin. These changes in catalytic efficiency in case of pullulan are due to increased $k_{cat}$, while $K_m$ on amylpectin and β-limit dextrin increased by 133% and 40%, respectively, and $k_{cat}$ by 60% and 6%, respectively (Table 1). Hence, Asp621 is a substrate preference determinant discriminating between the natural branched substrates and the linear model substrate pullulan. The double mutant Phe620Ala/Asp621Ala, however, retained only 14% residual activity on pullulan and even less, 7 and 9%, towards amylpectin and β-limit dextrin (Table 1 and Fig. 5). The relatively higher activity on BPNPG3G3 of 57% and 25% seen for Phe620Ala and Phe620Ala/Asp621Ala, respectively (Fig. 5B), moreover, emphasises the different importance of subsites +2 and +3 for the activity on oligosaccharide substrate. The combination of Phe620 and Asp621 is particularly...
important for activity on branched polysaccharides and loss of activity was less dramatic if only Phe620 is substituted.

4. Conclusion

Using HvLD as a model and implementing mutations imitating positions tentatively associated with high activity sorghum and rice variants, it was demonstrated that the N-terminal CBM21-like domain of HvLD indeed plays a role in activity and possibly possesses a characteristic high level of dynamics similarly to structure-determined CBM21 domains. The results indicated that the targeted surface exposed residues of the N-terminal CBM21-like domain have a role for hydrolysis of polysaccharide substrates and also of a branched oligosaccharide. In the latter case one may envisage a conformational change as a result of the mutations resulting in a long distance effect on the active site or an allosteric stimulation of activity by substrate binding at a remote site on the protein surface. However, carbohydrate binding to the domain has not been demonstrated experimentally. HvLD is the only LD or PUL for which a structure of a substrate complex has been determined, and the present mutational analysis strongly supports that the Trp512-Phe553 aromatic sandwich at subsite +2 is especially critical for activity on oligosaccharides, while both Phe620 and Asp621 at subsite +3 are important in particular for polysaccharide degradation and discriminate between pullulan and the two branched polysaccharides. These latter characteristics may also have functional relevance for LDs (GH13_13) as opposed to PULs (GH13_12 and GH13_14) in agreement with the low degree of sequence conservation between these enzymes in this active site region. Finally, this first kinetic analysis on HvLD degrading β-limit dextrin, that presumably is a natural substrate for LD, showed a preference over amylopectin, although $K_m$ and $k_{cat}$ were still 25-fold higher and 9-fold lower, respectively, than for pullulan.

Acknowledgements

Karina Jansen and Julia Tanas Tanasi are thanked for expert technical assistance and Tobias Hegelund Olsen for experimental work during his bachelor thesis.

Funding

This work was supported the Independent Research Fund Denmark | Natural Sciences (FNU; 6108-00476A) and Independent Research Fund Denmark | Technology and Production Sciences (FTP; DFF-1337-00158) and by a PhD scholarship at The Technical University of Denmark (S.A.).

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Author contributions

SA: Investigation, Formal Analysis, Writing – Original draft preparation, Writing – Reviewing and Editing, Visualisation. BS: Conceptualization, Writing – Reviewing and Editing, Supervision, Funding Acquisition. MSM: Conceptualization, Investigation, Formal Analysis, Writing – Reviewing and Editing, Visualisation

Appendix A. Supplementary data

Supplementary data to this article can be found online at XXX.
References


[34] Y. Chen, J. Bao, Underlying mechanisms of zymographic diversity in starch synthase I and pullulanase in rice developing endosperm, J. Agric. Food Chem., 9 (2016) 2030-


[68] M.V. Ramesh, S.M. Podkoyev, S.E. Lowe, J.G. Zeikus, Cloning and sequencing of the Thermoanaerobacterium saccharolyticum B6A-RI apu gene and purification and


Highlights

- *HvLD* is 7–150-fold more active on pullulan than on β-limit dextrin and amylopectin
- N-terminal CBM-21-like domain mutants affect activity on poly- and oligosaccharides
- Phe620 and Asp621 at subsite +3 are critical for activity on branched substrates
- Phe553 at subsite +2 is important for enzymatic activity and inhibitor binding

Graphical abstract
Figure 2

[Diagram showing molecular structures with labeled amino acids and positions.

Legend: G3G'3, GG^2G^23.]

Main chain: +2, +1, 0', +3.
Branch chain: -1, -2, -3.
Figure 5

A

![Bar chart showing relative activity of different enzymes for different substrates.]

B

![Bar chart showing relative activity of different enzymes for different substrates.]

Figure 5