The unique branching patterns of Deinococcus glycogen branching enzymes are determined by their N-terminal domains.

Palomo, M.; Kralj, S.; van der Maarel, Marc; Dijkhuizen, Lubbert

Published in:
Applied and Environmental Microbiology

DOI:
10.1128/AEM.02141-08

IMPORTANT NOTE: You are advised to consult the publisher’s version (publisher’s PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2009

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
The Unique Branching Patterns of Deinococcus Glycogen Branching Enzymes Are Determined by Their N-Terminal Domains

M. Palomo, S. Kralj,† M. J. E. C. van der Maarel,* and L. Dijkhuizen

Microbial Physiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands

Received 16 September 2008/Accepted 27 December 2008

Glycogen branching enzymes (GBE) or 1,4-α-glucan branching enzymes (EC 2.4.1.18) introduce α-1,6 branching points in α-glucans, e.g., glycogen. To identify structural features in GBEs that determine their branching pattern specificity, the Deinococcus geothermalsis and Deinococcus radiodurans GBE (GBEDg and GBEDr, respectively) were characterized. Compared to other GBEs described to date, these Deinococcus GBEs display unique branching patterns, both transferring relatively short side chains. In spite of their high amino acid sequence similarity (88%) the D. geothermalsis enzyme had highest activity on amylose while the D. radiodurans enzyme preferred amylopectin. The side chain distributions of the products were clearly different: GBEDg transferred a larger number of smaller side chains; specifically, DP5 chains corresponded to 10% of the total amount of transferred chains, versus 6.5% for GBEDr. GH13-type GBEs are composed of a central (β/α) barrel catalytic domain and an N-terminal and a C-terminal domain. Characterization of hybrid Deinococcus GBEs revealed that the N2 modules of the N domains largely determined substrate specificity and the product branching pattern. The N2 module has recently been annotated as a carbohydrate binding module (CBM48). It appears likely that the distance between the sugar binding subsites in the active site and the CBM48 subdomain determines the average lengths of side chains transferred.

Glycogen is an energy reserve polymer of many animals and microorganisms. It is composed of a backbone of glucose residues linked by α-1,4-glycosidic bonds with α-1,6-linked side chains (7, 31). In bacteria, the linear α-1,4-glucan is synthesized from ADP-glucose by the enzyme glycogen synthase, which is thought to be involved in both initiation and elongation of the chain (40). Side chains are introduced by glycogen branching enzyme (GBE) or 1,4-α-glucan branching enzyme (EC 2.4.1.18). This enzyme catalyzes formation of α-1,6 branch points by cleaving an α-1,4-glycosidic linkage in the donor substrate and transferring the nonreducing end-terminal fragment of the chain to the C-6 hydroxyl position of an internal glucose residue that acts as the acceptor substrate (4). Depending on its source, GBEs have a preference for transferring different lengths of glucan chains (1, 23). Most GBEs are members of subfamily 8 (Eukarya) or 9 (Bacteria) of glycoside hydrolase family 13 (GH13) (34). Recently, the first GBE from family GH57 was described (28) (http://www.cazy.org).

GH13-type GBEs are composed of three major domains of secondary structure, a central (β/α) barrel catalytic domain or A domain, an N-terminal domain, and a C-terminal domain (1). Domain A is present in all members of family GH13 and consists of a highly symmetrical fold of eight parallel β-strands encircled by eight α-helices. However, some variations occur in GBEs (a missing α-helix 5 and insertion of extra α-helices) (1). Domain A contains the four conserved amino acid regions (I to IV) typical for enzymes of family GH13 (35). In most GH13 enzymes, an extra domain is present, inserted between β-strand 3 and α-helix 3 (domain B), which affects their catalysis and product specificity (16). In GBEs, the length of this loop is only 40 residues, not long enough to be considered a separate domain (1). Domain C is found in most GH13 enzymes and is believed to shield the hydrophobic residues of the catalytic domain from contacts with the solvent. Domain C has also been suggested to be involved in substrate binding (25).

Domain N is typical for GH13 enzymes cleaving or forming endo-α-1,6 linkages (17), namely, isoamylase (EC 3.2.1.68; subfamily 11) (18), pullulanase (EC 3.2.1.41; subfamilies 12 to 14) (27), and both starch (subfamily 8) and glycogen branching enzymes (subfamilies 8 and 9). An exception is the 4-α-D-((1→4)-α-D-glucan)trehalose trehalohydrolase (EC 3.2.1.141; subfamily 10), which hydrolyzes linear maltodigosaccharide-like substrates (6) (39). The crystal structures of most of these enzymes with their N domains (all or part) have been published previously (1, 6, 18, 27, 39). The exact function of this N domain has remained unclear, and the similarity between the N domains in these different enzymes is low. They vary in length, and some of them consist of two or three modules. However, they all possess one common module that was recently classified as a family 48 carbohydrate-binding module (CBM48) (19) (http://www.cazy.org/).

In GBEs domain N comprises a module of 150 amino acids, termed the N2 module, that contains the putative CBM48. In some branching enzymes, it is preceded by a module of 100 to 150 amino acids, termed the N1 module. It has been proposed that the N1 module has originated from a DNA duplication of...

* Corresponding author. Mailing address: Microbial Physiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands. Phone: 31 50 3632150. Fax: 31 50 3632154. E-mail: m.j.e.c.van.der.maarel@rug.nl.

† Present address: Genencor International B.V., Archimedesweg 30, 2333 CN Leiden, The Netherlands.

Published ahead of print on 9 January 2009.
the N2 module (24). Based on the architecture and length of the N domain, GBEs can be divided into group 1, containing both the N1 and the N2 modules, and group 2, containing only the N2 module (12). A 112-amino-acid truncation of the N1 module transferred longer glucan chains (5). No studies have been reported thus far investigating the role of the N2 module (containing the putative CBM48 domain) in GBEs as well as in other GH13 members.

Here, we report a detailed biochemical characterization of two GH13 GBEs from the extremophile bacteria *Deinococcus geothermalis* and *Deinococcus radiodurans*. These two GBEs (GBE<sub>Dg</sub> and GBE<sub>Dr</sub>, respectively) generate unique branching patterns by transferring glucosidic chains that are shorter than those of other GBEs reported to date (9, 36, 38, 41). To investigate the role of the different domains in these enzymes, chimeras of GBE<sub>Dg</sub> and GBE<sub>Dr</sub> were constructed. Their characterization revealed that substrate and chain length specificity of these *Deinococcus* GH13 GBEs are largely determined by the putative CBM48 part of the N domain.

### MATERIALS AND METHODS

**Phylogenetic analysis.** Amino acid sequences of all GH13 GBEs reported in the Carbohydrate Active Enzymes database (http://www.cazy.org/) were used to construct a phylogenetic tree; these were the following: *Agrobacterium tumefaciens* (AF033856), *Aerococcus ruber* (AJ414166), and *Chlamydomonas reinhardtii* (AB060080). The complete *glgB* gene and the N-terminal, catalytic A domain, and C-terminal domains of each GBE were amplified by PCR using as primers *Dg*<sub>NT</sub>-FP/Dg-RP for the NT <sup>1</sup>Dg chimeric gene, *Ctdg*-FP/CTD<sub>Dg</sub>-RP for the CTD<sub>Dg</sub> chimeric gene, and *Dr*<sub>NT</sub>-FP/Dr-RP for the NT Dr chimeric gene. They were all cloned into the pET15b vector (Novagen) using NdeI and BglII restriction sites, resulting in plasmids pDg and pDr with N-terminal His tags.

**Construction of truncated and chimeric genes.** The 3’ 207-nucleotide extension in domain C of *glgB<sub>Dg</sub>* was truncated (resulting in CTD<sub>Dg</sub>) by PCR using plasmid pDr as a template and the primers Dr-FP and CTT-RP (Table 1). Similarly, the 5’ and 3’ truncated forms of *glgB<sub>Dr</sub>* (477 nucleotides from 5’ for the N terminus [NT<sub>Dr</sub>] and 387 nucleotides from 3’ for the C terminus [CTD<sub>Dr</sub>]) were amplified by PCR using as primers DgNT-FP and Dr-RP for the NT<sub>Dr</sub> chimeric gene and Dg-RP and DrCTT-RP for the CTD<sub>Dr</sub> chimeric gene. They were all cloned into the pET15b vector yielding pD<sub>Dr</sub>C<sub>Dg</sub>TT, pD<sub>Dg</sub>N<sub>Dr</sub>T, and pD<sub>Dg</sub>C<sub>Dr</sub>T plasmids.

**Construction of the GGR, RRR, RGG, and GRR chimeric genes was achieved by overlap extension PCR (43). They were named according to the *glgB* source of their N, A, and C domains (e.g., the GGR gene consists of the GGR<sub>Dg</sub> N domain, the GGB<sub>Dr</sub> A domain, and the GRR<sub>Dr</sub> C domain in respective order).** Briefly, separate PCR products were performed to yield fragments encompassing the N domain, catalytic A domain, and/or the C-terminal domain for each *glgB* gene. Appropriate templates and oligonucleotides were combined in subsequent PCRs (Table 2) to yield the respective hybrids. Templates, primers, and amplified fragments are listed in Table 2. The GGR, RRR, RGG, and GRR chimeric genes were cloned into pNel and pBglII sites of the pET15b vector, yielding plasmids pG<sub>GGR</sub>, pRR<sub>RR</sub>, pRG<sub>G</sub>, and pG<sub>R</sub>&<sub>G</sub>RR.

**Expression and purification.** Overexpression of all proteins, except for NT<sub>Dg</sub>-CTD<sub>Dg</sub> and GRR, was achieved by overnight growth of E. coli BL21(DE3) Star cells containing the corresponding plasmid at 37°C and 210 rpm in Luria-Bertani medium supplemented with 50 µg/ml ampicillin. NT<sub>Dg</sub>-CT<sub>Dg</sub> and GRR were expressed successfully by induction with 1 mM isopropyl-p-D-thiogalactopyranoside at an optical density of 600 nm of ~0.6, followed by 5 h of additional growth. For all proteins, the cells were harvested by centrifugation (10 min at 10,000 g), washed with 50 mM sodium phosphate buffer, pH 8.0, and centrifuged again (10 min at 10,000 g). The pellets were resuspended in 50 mM sodium phosphate buffer, pH 8.0, and the cells were disrupted by sonication (seven times for 15 s at 7 µm with 30-s intervals) and centrifuged (30 min at 15,000 x g). All proteins were found as soluble proteins in the supernatants except for NT<sub>Dg</sub>-CTD<sub>Dg</sub> and GRR, which were found in the pellets as inclusion bodies. Only GRR was successfully isolated from the inclusion bodies by resuspension of the pellet in 8 M urea, followed by centrifugation (30 min at 15,000 x g). The supernatant containing unfolded GRR was subjected to dialysis against sodium phosphate

### Table 1. Oligonucleotides used to amplify *glgB<sub>Dg</sub>* and *glgB<sub>Dr</sub>* genes and to construct truncated and chimeric mutants by PCR

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’→3’)</th>
<th>Position (nt)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dg-FP</td>
<td>GGAATTCCCATATGGAATCATGCTTGACGGCCGGG</td>
<td>1–20</td>
</tr>
<tr>
<td>Dg-RP</td>
<td>GCTAGTTCTAGAAAAGCTTACCGCGGCCGGAC</td>
<td>1959–1938</td>
</tr>
<tr>
<td>Dr-FP</td>
<td>GATGACTCATATGACGTTTCTCTGGCCACCGAC</td>
<td>1–34</td>
</tr>
<tr>
<td>Dr-RP</td>
<td>GCTAACGATGCCTTATGACGGCTTACCGACGGG</td>
<td>2124–2099</td>
</tr>
<tr>
<td>CTT-RP</td>
<td>GTAAGGATCTTACTACGGCCGTAAGACGGCCGAC</td>
<td>1917–1986</td>
</tr>
<tr>
<td>H1</td>
<td>GTTCAAGATCTGTGCGGCGCTGGAA</td>
<td>1531–1556</td>
</tr>
<tr>
<td>H2</td>
<td>TCTAGGGCCGCCAACGATGGTACGGCGAC</td>
<td>1556–1531</td>
</tr>
<tr>
<td>H3</td>
<td>GTGACGCGGCGCGACCAGATGGTACG</td>
<td>521–546</td>
</tr>
<tr>
<td>H4</td>
<td>GTAATGCAATGAAACGGCGCCTGCTAGC</td>
<td>546–521</td>
</tr>
<tr>
<td>DgNT-FP</td>
<td>CGATGCGATATTGGGCACCGCGCCTTACGATC</td>
<td>478–505</td>
</tr>
<tr>
<td>DgCT-RP</td>
<td>GCTACTAGATGCTCTAGCCTGGAAGGGACCTGG</td>
<td>1571–1540</td>
</tr>
</tbody>
</table>

* nt, nucleotide.
buffer (25 mM; pH 8), leading to the refolding of GRR, as confirmed by circular dichroism.

All proteins were purified by His tag affinity chromatography using a HiTrap chelating column (Amersham Pharmacia, Uppsala, Sweden) charged with nickel sulfate. The proteins were eluted with a linear gradient of 0 to 500 mM imidazol in sodium phosphate buffer at pH 8.0 and 20°C using a concentration of protein (Bio-Rad, Germany).

Circular dichroism. Multiple far-UV spectra (195 to 260 nm) were recorded in sodium phosphate buffer at pH 8.0 and 20°C using a concentration of protein of 0.1 mg/ml on a Jasco J-715 spectropolarimeter (Jasco, Tokyo, Japan) in quartz cells; the following parameters were used: optical path length, 1 mm; data interval, 0.2 mm; bandwidth, 1.0 nm; sensitivity, 20 millidegrees; response time, 0.125 s. All recorded spectra were corrected by subtraction of the blank spectrum.

Enzyme activity assays. (i) Iodide assay. The total activity of the enzyme was measured using the iodine assay (10), which is based on formation of a blue complex between iodine-iodide and a linear -1,4-glucan of a certain length. The sum of the transglycosylation and the hydrolytic activity of the enzyme can be measured by monitoring the decrease in absorbance. The reaction mixtures contained 150 μl of appropriately diluted GBE (1, 2, or 4 μg) in 1,400 μl of 0.125% amyllose V (Avebe, Foxhol, The Netherlands) at pH 7.0. Samples (200 μl) were taken at different time points (0 to 20 min), and the reaction was stopped by boiling for 3 min. The numbers of reducing ends before and after debranching were measured at each time point by a 2,2′-bicinchoninate assay (42) using a molar extinction calibration curve. Debranching was performed as follows: 50 μl of reaction mixture was debranched for 20 h at 40°C in 129 μl of 25 mM sodium phosphate buffer (pH 8.0), 20 μl of 1 M acetate buffer (pH 3.8), and 1 μl (0.175 U) of Pseudomonas sp. isomylase (Megazyme, Wicklow, Ireland). One unit of GBE branching activity is defined as 1 μmol of α-1,6 linkages synthesized per min.

Side chain distribution. To determine the side chain distribution of the final products, 10 ml of 0.125% (wt/vol) amyllose V (Avebe) solution was incubated with 400 units (determined by iodine assay) of GBEDr or GBEDg for 72 h. The reaction was terminated by boiling the mixture for 10 min. The reaction mixture was divided in two equal parts; one part was treated with 10 units of isomylase from Pseudomonas sp. (Megazyme) after the addition of 750 μl of 1 M citrate buffer, pH 4.0, for 20 h at 40°C to hydrolyze the α-1,6 glucosidic linkages. Both parts of the reaction mixture, before and after debranching, were freeze-dried, dissolved in 90% dimethyl sulfoxide (0.25 mg/ml) and analyzed by high-performance anion-exchange chromatography (HPAEC) analysis. Mixtures of oligosaccharides with degrees of polymerization ranging from 1 to 7 (DPI to DP7) and a debranched waxy maize starch solution containing a broad mix of oligosaccharides of known composition were used as standards. Separation of oligosaccharides was achieved with a CarboPac PA1 anion exchange column (250 mm by 4 mm; Dionex) coupled to a CarboPac PA4 guard column (Dionex, Sunnyvale, CA). The following gradient of eluent A (1 ml/min) was used: 95% (10 min), 65% (10 min), 55% (30 min), 35% (4 min), 0% (7 min), and 95% (14 min). Eluent A was sodium hydroxide (0.1 M), and eluent B was sodium acetate (0.6 M) in sodium hydroxide (0.1 M). Detection was performed with an ED40 electrochemical detector ( Dionex) equipped with an Au working electrode and an Ag/AgCl reference electrode with a sensitivity of 300 nC. The pulse program used was the following: +0.1 V (0 to 0.41 s), −2.0 V (0.41 to 0.43 s), +0.6 V (0.43 to 1.44 s), and −0.10 V (0.44 to 0.50 s). The integration time was 0.20 to 0.40 s. Data were integrated using a TotalChrom (PerkinElmer, Waltham, MA) data integration system.

RESULTS AND DISCUSSION

Amino acid sequence analysis of the glycogen branching enzymes of Deinococcus. At present there is no proof for actual synthesis of glycogen in the extremophilic bacteria D. geothermalis and D. radiodurans. However, the key enzymes involved in glycogen metabolism, such as glycogen branching enzyme (including a demonstration of its activity [this paper]), ADP-glucose pyrophosphorylase, and glycogen synthase, have been annotated in both species (28).

D. geothermalis and D. radiodurans both contain a single gene in their genomes coding for a putative glycogen branching enzyme. The D. geothermalis gene, Dgeo_0981, consists of 1,959 nucleotides and encodes an enzyme of 652 amino acids (GBE<sub>Dr</sub>) that has a predicted molecular mass of 74.4 kDa. The DR1848 gene from D. radiodurans consists of 2,124 nucleotides and encodes an enzyme of 705 amino acids (GBE<sub>Dg</sub>) that has a theoretical molecular mass of 80 kDa. Both enzymes belong to group 2 of GH13 GBEs. Furthermore, they contain the four conserved amino acid regions (I to IV) typical for enzymes of family GH13 (25) plus the four conserved amino acid residues specific for bacterial GBEs, namely W<sup>298</sup>, Q<sup>301</sup>, S<sup>466</sup>, and H<sup>530</sup> (E. coli GBE numbering) (Fig. 1) (1). Their N2 modules, containing the CBM48, possess the two consensual aromatic residues crucial for starch or glycogen binding (26) (W<sup>77</sup> and Y<sup>123</sup> for GBE<sub>Dr</sub> and W<sup>101</sup> and Y<sup>107</sup> for GBE<sub>Dg</sub>). GBE<sub>Dr</sub> and GBE<sub>Dg</sub> show an overall identity in 622 amino acids of 81% (88% similarity). The similarity between both N
domains is in the same order (145 to 158 amino acids long, 81% identity, and 88% similarity in 136 amino acids). The identity in their catalytic A domains is relatively high (364 amino acids; 87% identity and 93% similarity). The two enzymes differ most in their C domains (130 amino acids in GBE(Dg) and 195 in GBE(Dr) with only 70% identity and 77% similarity in 113 amino acids). Unique for the GBE(Dg) is a 66-amino-acid extension in the C domain, which is not present in any other GBE known to date. BLAST searches with the 66-residue-long extension did not give a significant result (data not shown).

Characterization of two Deinococcus branching enzymes. Both Deinococcus GBE genes were cloned, and the enzymes were expressed in E. coli and purified to homogeneity by His tag purification (data not shown), yielding up to 30 mg of purified GBE per liter of E. coli culture. The molecular mass of the proteins as determined by SDS-PAGE was 75 kDa for GBE(Dg) and 80 kDa for GBE(Dr), in accordance with the calculated values. Both enzymes were active in a relatively broad range of pH (7 to 9) and temperature (30 to 40°C) values and showed maximum activity at pH 8.0 and 34°C, as measured by the iodine assay. Their specific activities on amylose V were 621.3 U/mg for GBE(Dg) and 404.2 U/mg for GBE(Dr) (Table 3). These are relatively high values compared to specific activities previously reported for GH13 GBEs from G. stearothermophilus (33). Specific activities measured by the branching assay were 7.2 U/mg for GBE(Dg) and 4.7 U/mg for GBE(Dr) (values are not directly comparable to iodine assay-specific activities because of a different definition of the unit) (see Materials and Methods). This assay allowed us to discriminate between branching activity and hydrolytic activity; the latter represented less than 1% of total activity (data not shown). Therefore, the iodine assay was used for subsequent characterizations.

The two enzymes showed clear differences in their thermostability, as expected from their optimal growth temperatures (25 to 30°C for D. radiodurans and 50 to 55°C for D. geothermalis). GBE(Dg) remained fully active after a 1-h incubation at 60°C and lost less than 20% of activity after a 1-h incubation at 65°C. GBE(Dr) exhibited a 10°C lower thermostability than GBE(Dg) and remained fully active at 50°C (1 h incubation).

The two Deinococcus GBEs clearly differ in substrate specificity. The specific activities of both GBEs were determined on amylpectin using the iodine assay (Table 3) and were compared to their specific activities on amylose as previously measured (see above). In contrast to the results with amylose, the specific activity on amylpectin was higher for GBE(Dg) (538.0 U/mg protein) than for GBE(Dr) (474.3 U/mg protein). The ratio of the specific activities of amylase to amylpectin was 1.31 for GBE(Dg) and 0.75 for GBE(Dr). E. coli GBE had an amylase/amylpectin ratio of 1.78, which is higher than observed for the two Deinococcus GBEs (2). Thus, in spite of their high overall sequence identity and similarity, GBE(Dg) prefers amylase over amylpectin whereas the opposite is true for GBE(Dr).

The two Deinococcus GBEs have a preference for transferring short side chains. The products obtained after prolonged incubations of both enzymes with amylose showed a similar side chain distribution. Most side chains were between 4 and 17 glucose residues long, with side chains of 6 and 7 glucose residues (35% of total) predominantly (Fig. 2). The branching pattern is significantly different from that of the other GH13 GBEs for which a quantitative analysis has been described to date, namely, E. coli GBE (group 1 GBE) wild type.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Specific activity (U/mg) on:</th>
<th>Amylase/amylpectin activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Amylose</td>
<td>Amylpectin</td>
</tr>
<tr>
<td>GBE(Dg)</td>
<td>621.3</td>
<td>474.3</td>
</tr>
<tr>
<td>GBE(Dr)</td>
<td>404.2</td>
<td>538.0</td>
</tr>
<tr>
<td>CTT(Dr)</td>
<td>442.3</td>
<td>556.5</td>
</tr>
<tr>
<td>GGR</td>
<td>468.8</td>
<td>359.2</td>
</tr>
<tr>
<td>RRG</td>
<td>411.7</td>
<td>505.6</td>
</tr>
<tr>
<td>RGG</td>
<td>624.1</td>
<td>342.3</td>
</tr>
<tr>
<td>GRR</td>
<td>&lt;0.5</td>
<td>&lt;0.5</td>
</tr>
</tbody>
</table>

* ND, not determined.
and three N1 module-truncated variants (5) and *A. aeolicus* GBE (group 2 GBE) (41). In these GBEs, the preferred DP of chains transferred (~40 to 50%) lies between DP10 and DP14, with only 20 to 30% of chains shorter than 10 glucose residues, a value increasing to 70% for *Deinococcus* GBEs (Fig. 2). Side chain distributions of other GH13 GBEs have been reported, e.g., for *G. stearothermophilus* GBE (36) and *A. gottschalkii* GBE (38), both group 2 GBEs. Even though no quantitative data are available for these enzymes, their branching patterns are clearly different from those of the *Deinococcus* GBEs, with a preference for the transfer of longer chains, as in the case of *E. coli* and *A. aeolicus* GBEs. This unique side chain distribution of the *Deinococcus* GBEs may be a consequence of the fact that they are distantly related to other known GH13 GBEs. The two *Deinococcus* enzymes form a separate cluster, and although they both belong to group 2 of GH13-type GBEs, they appear equally related to group 1 and group 2 at the amino acid level (Fig. 3).

**Construction of chimeric *Deinococcus* enzymes showed the importance of the N domain in substrate specificity and branching pattern.** Comparison of the branching patterns of the two *Deinococcus* GBEs revealed that the branched product of GBE$_{Dg}$ contained more short side chains overall, especially of DP5, which corresponded to about 10% of the total (only

![FIG. 2. HPAEC-pulsed amperometric detection analysis of the end products obtained by incubation of GBE$_{Dg}$ or GBE$_{Dr}$ with amylose, followed by debranching with isoamylase. For comparison, the distribution of DPs of the debranched end products of the *E. coli* GBE (5) and *A. aeolicus* GBE (41) enzymes are also shown.](image)

![FIG. 3. Neighbor-joining tree of known GH13 GBEs (http://www.cazy.org) constructed using MEGA, version 3.1 (20). Bootstrap values are indicated on the nodes of each branch. The bar indicates 10% amino acid sequence difference. See Materials and Methods for accession numbers. *, GBE from group 1.](image)
6.5% in the case of GBE_{Dr} (Fig. 2). Thus, strikingly, even though both enzymes are highly similar (88% similarity and 81% identity), they exhibit distinct substrate specificities and clear differences in their branching patterns.

The relatively high overall amino acid sequence identity between GBE_{Dg} and GBE_{Dr} made them ideal candidates for investigating a possible role of the different domains in causing their differences in substrate preferences and branching patterns. In order to gain more insight into which domain (N, catalytic, or C domain) in the *Deinococcus* GBEs influenced activity and specificity, a number of chimeric enzymes were constructed and biochemically characterized (Fig. 4).

The characterization of chimeric enzymes has also been used to study the role of the N and C domains in starch branching enzymes, with different and sometimes contradictory results. Domain C was proposed to be involved in substrate specificity (14, 15, 21), in maximal activity (14, 21), and in side chain distribution (15) while domain N was suggested to affect the catalytic efficiency (11, 13, 21) and the branching pattern (13, 21). However, these results cannot be compared directly to GBEs since they show less than 30% identity (less than 50% similarity) to starch branching enzymes, and the differences are even larger in the noncatalytic domains.

Here, we exchanged domain N (N2 module in this case) and domain C between the two *Deinococcus* GBEs, yielding GGR, RRG, RGG, and GRR chimeric mutants (Fig. 4). Furthermore, the extra 66-amino-acid extension in domain C of GBE_{Dg} was deleted (resulting in CTT_{Dg}) to analyze its possible effects on enzyme properties. In addition, a truncation of the entire C domain (CT_{Dg}) and a truncation of the entire N domain (NT_{Dg}) were made in GBE_{Dg} (Fig. 4).

The hybrid and truncated enzymes were expressed in *E. coli* BL21(DE3) Star and purified to homogeneity by His tag affinity chromatography. The GBE_{Dg} C- and N-domain-truncated enzymes, as well as the GRR chimeric enzyme, ended up in inclusion bodies. Only the GRR enzyme was successfully purified from the aggregates, and proper folding of this chimera was confirmed by circular dichroism (data not shown). All chimeras (as well as CTT_{Dg}) showed enzyme activity although GRR activity was very low, less than 1% of the wild type (Table 3). An optimum pH of 8.0 and an optimum temperature of 34°C were observed for all chimeric enzymes (data not shown). GRR was not further characterized due to its low activity.

**Domain C does not influence substrate specificity or branching pattern.** The specific activity of CTT_{Dg} was 442.3 U/mg on amylose and 468.8 U/mg on amylpectin. Both specific activities as well as their ratio of 0.79 were comparable to those of the wild-type GBE_{Dn} (Table 3). Furthermore, the side chain distribution produced by CTT_{Dg} was very similar to that made by GBE_{Dn} (data not shown). The 66-amino-acid extension at the C terminus of GBE_{Dr} thus does not affect substrate specificity or branching pattern.

**Domain N is involved in determining substrate specificity and branching pattern.** Of the N-domain exchange mutants, only RGG (but not GRR) had enough activity to enable further tests. Therefore, substrate specificity and side chain distribution were measured only for RGG. It exhibited a specific activity on amylose of 624.1 U/mg and on amylpectin of 342.3 U/mg. Its amylose/amylpectin activity ratio was 1.9. Interestingly, this is much higher than the value measured for either parent (Table 3). This result clearly shows that domain N (N2 module) influences the GBE substrate specificity.

The product synthesized by RGG had a branching pattern that was clearly different from the product of GBE_{Dn} (Fig. 5C). Interestingly, the side chain distribution of RGG closely resembled that of GBE_{Dn}, which was the source of only the N domain of RGG. From this result, it is concluded that N domain plays a critical role in determining the length of the chains transferred by these GBEs.

It has previously been shown that the N1 module of domain N of GBEs of group 1 influences their branching patterns (2, 3, 5). However, no such information is available for GBEs belonging to group 2, which lack the N1 module. The N2 module has recently been annotated as a CBM48 and is most likely involved in binding of the substrate. The CBM48 found in AMP-activated protein kinase has been shown to bind glycogen as well as β-cyclodextrin and oligosaccharides such as maltoheptaose (19, 30). There are clear examples of the importance of CBMs for enzyme activity, e.g., in the case of CGTase, which has a carbohydrate binding site in domain C anchoring the enzyme to the starch granule (22), and the raw starch binding domain E (designated CBM20), which has two carbohydrate binding sites binding raw starch and guiding the linear starch chains toward the active site (29). In the crystal
structures of isoamylase and 4-\(\alpha\)-\(\alpha\)-t-{(1\(\rightarrow\)4)-\(\alpha\)-\(\alpha\)-glucano}trehalose trehalohydrolase, the N domain (both with a putative CBM48) is in close proximity to the B domain [loop between the third \(\beta\)-strand and the third \(\alpha\)-helix of the \(\beta\)/\(\alpha\) barrel] (25), which is known to be involved in catalysis and product specificity of GH13 enzymes (16). Moreover, it has been suggested that the N domain of \(D.\) radiodurans 4-\(\alpha\)-\(\alpha\)-t-{(1\(\rightarrow\)4)-\(\alpha\)-\(\alpha\)-glucano}trehalose trehalohydrolase is involved in the initial steps of sugar binding by contributing to hydrophobic stacking interactions at the entrance of the channel leading to the active site (39). This channel in \(D.\) radiodurans 4-\(\alpha\)-\(\alpha\)-t-{(1\(\rightarrow\)4)-\(\alpha\)-\(\alpha\)-glucano}trehalose trehalohydrolase can accommodate seven sugar units in subsites \(-1\) to \(-7\). Since GBE uses substrates longer than DP7, it is indeed possible that the N2 module accommodates the rest of the minus subsites (from subsite \(-8\) on), thus determining the number of glucose units binding in the donor subsites (corresponding to the length of the chains transferred). It is also possible that by interacting with the \(-7\) subsite, N2 influences substrate binding at the \(-6\) subsite. A weaker binding of the donor substrate at the \(-6\) and \(-7\) subsites would result in an increase of transferred chains of DP5.

In the case of group 1 GBEs, the branching patterns are most likely determined by the combined action of both modules N1 and N2 of domain N. A dynamic mechanism in which a flexible N1 would move and restrict exposure of longer chains to the active site has been suggested (5). We propose that the N2 module helps in controlling the length of the chains transferred in these enzymes and plays a main role in determining them in group 2 GBEs. At present, we are studying this region of the protein in more detail to identify amino acids that are crucial for the fine-tuning of the activity of these enzymes.

Conclusions. Here, we report the first detailed characterization of the highly similar GH13 GBEs from \(D.\) geothermalis and \(D.\) radiodurans. The side chain distributions of their final products are unique and differ significantly from that of other GBEs described thus far in possessing a relatively large percentage of short chains. This exceptional branching pattern makes their products structurally unique. The physicochemical properties of such new, modified products remain to be determined, and they may have new food, cosmetics, and pharmaceutical applications.

Although both enzymes are highly similar (88% similarity and 81% identity), their specificities for amylose and amylpectin differ remarkably. Moreover, they display clear differences in their branching patterns. The variations observed apparently were not caused directly by differences in their catalytic domains (93% similar and 87% identical). Therefore, the influence of their noncatalytic domains was investigated. By constructing chimeric enzyme variants of the two Deinococcus GBEs, we demonstrated that the N domain (N2 module, recently annotated as a CBM48) is involved in substrate specificity and in determining the length of chains transferred by GBEs. Further research will focus on the influence of this putative carbohydrate binding domain on GBE enzyme activity.

ACKNOWLEDGMENTS

We thank Peter Sanders for HPAEC analysis and Hans Leemhuis for critically reading the manuscript.
REFERENCES


