JBC RESEARCH ARTICLE

A structural explanation for the mechanism and specificity of plant branching enzymes I and IIb

Received for publication, June 4, 2021, and in revised form, November 3, 2021 Published, Papers in Press, November 8, 2021,
https://doi.org/10.1016/j.jbc.2021.101395

Hadi Nayebi Gavgani[1,‡], Remie Fawaz[1,‡], Nona Ehyaei[1], David Walls[1], Kathryn Pawlowski[1], Raoul Fulgos[1], Sunghoon Park[2], Zahra Assar[1], Alireza Ghanbarpour[1], and James H. Geiger[1,*]

From the [1]Department of Chemistry, Michigan State University, East Lansing, Michigan, USA; [2]Department of Foodservice Management and Nutrition, College of Natural Sciences, Sangmyung University, Seoul, Korea

Edited by Joseph Jez

Branching enzymes (BEs) are essential in the biosynthesis of starch and glycogen and play critical roles in determining the fine structure of these polymers. The substrates of these BEs are long carbohydrate chains that interact with these enzymes *via* multiple binding sites on the enzyme's surface. By controlling the branched-chain length distribution, BEs can mediate the physiological properties of starch and glycogen moieties; however, the mechanism and structural determinants of this specificity remain mysterious. In this study, we identify a large dodecaose binding surface on rice BE I (BEI) that reaches from the outside of the active site to the active site of the enzyme. Mutagenesis activity assays confirm the importance of this binding site in enzyme catalysis, from which we conclude that it is likely the acceptor chain binding site. Comparison of the structures of BE from *Cyanothece* and BE1 from rice allowed us to model the location of the donor-binding site. We also identified two loops that likely interact with the donor chain and whose sequences diverge between plant BE1, which tends to transfer longer chains, and BEIIb, which transfers exclusively much shorter chains. When the sequences of these loops were swapped with the BEIIb sequence, rice BE1 also became a short-chain transferring enzyme, demonstrating the key role these loops play in specificity. Taken together, these results provide a more complete picture of the structure, selectivity, and activity of BEs.

Starch, the primary energy storage molecule of plants, is a complex biomaterial built from large α-1,4- and α-1,6-branched glucose polymers packed into a well-ordered granule, with amorphous and semicrystalline layers throughout the granule (1, 2). Two types of polymer make up these granules: Amylose, a smaller polymer (20–30% of starch weight, with a degree of polymerization (DP) of 200–500) almost devoid of σ-1,6 branches (3) (with only 5–20 branches per molecule); and Amylopectin, a much larger polymer (65–85% starch weight, with average DP 11–15 for short chains, average DP 43–50 for long chains) with approximately 5 to 6% branching (4, 5). The starch granule's size, density, and fine structure are key to its function as the primary energy storage unit for all plants, leading to more or less bioavailability and robustness, depending on the plant's energy requirements (6). The starch granule is biosynthesized by the interplay of several enzymes, including ADP-glucose pyrophosphorylase, which makes the ADP-glucose monomer building block, starch synthases, which convert ADP-glucose to α-1,4-linked glucan polymers, branching enzymes, which cleave α-1,4-linkages and transfer the resulting cleaved glucan to α-1,6 positions to form branches, debranching enzymes, which remove inappropriately placed branches, and kinases that phosphorylate the resulting polymer (7). These enzymes all work in concert to build the starch granule. Most of these enzymes have several isoforms, each with unique roles to play in constructing the granule (8, 9). Branching enzyme (BE) has at least two and often three isoforms in most plants, BEI, BEIIa, and BEIIb, that differ in both substrate and product specificity (1, 10–12). While BEI isoforms prefer amylose as a substrate and tend to transfer longer chains of 11 or more units, BEII isoforms favor amylopectin as a substrate and tend to transfer shorter chains of 6 to 7 glucose units, with BEIIa transferring a broader range of glucans while BEIIb isoforms transfer almost exclusively chains of 6 to 7 glucose units (10–13). They also differ in localization, with BEI and BEIIb almost exclusively found in the amyloplast. BEIIa is expressed throughout the plant, especially in the leaf, where it is exclusively responsible for branching the rapidly created and degraded starch (14, 15). We and others have been working to elaborate the molecular details that give rise to the unique specificity of both glycogen and starch branching enzymes (16–23). Herein we report the structure of maltododecaose (M12)-bound rice BEI (rBEI), which, we hypothesize, identifies the acceptor chain binding site on the surface of the enzyme for the first time. Mutagenesis experiments confirm the importance of this binding site for activity while also inferring that the site is not the donor chain binding site. Comparison of the rBEI structure with the M7-bound *Cyanothece* BE structure allows us to

‡ These authors contributed equally to this work.
* For correspondence: James H. Geiger, geigerj@msu.edu.
Present address for Sunghoon Park: Department of Food and Nutrition, Gangneung-Wonju National University, 25457, Gangneung, Korea.
Present address for Zahra Assar: Cayman Chemicals, 1180 E. Ellsworth Road, Ann Arbor, MI 48108, USA.
Present address for Alireza Ghanbarpour: Yale School of Medicine 333 Cedar Street New Haven, CT 06520, USA.

identify the likely donor chain binding site (22). Mutagenesis of two loops in the vicinity of this donor chain binding site converts the activity and specificity of rBEI to be very similar to that of rBEIIb, confirming the location of the donor chain binding site. Together, these results give the most complete picture to date regarding the mechanism of BE and identify the crucial regions for distinguishing the activity of the isoforms. These insights will enable the rational and systematic redesign of BEs to tailor their activities much more precisely, potentially leading to more precisely modified and optimized versions of this most critical biomaterial.

Results

Structure of M12-bound rBEI

The structures of both apo and maltopentaose-bound rBEI were previously determined (23). Three maltopentaose binding sites were identified, all distal from the active site. In the present work, a new rBEI crystal form was identified and the M12-bound rBEI structure obtained by soaking these crystals. This new crystal form was obtained using a tetra-mutant rBE1 that was truncated by 60 residues on its N-terminus (see Experimental procedures for details). Both activity and chain specificity assays showed this construct to behave identically to that of wild-type rBE1. Of the two M12 molecules bound to rBE1 (Fig. 1, *A* and *B*), one occupies the sites identified in the previously described M5-bound rBE1 structure (labeled "site 1") (23). Four glucose units are ordered in site 1 and the interactions between glucan and protein are similar to that seen in the previous structure (Fig. 2*A*). The conformation and curvature of the oligosaccharide are similar to that seen in an amylose single helix (Fig. 2*B*). The second molecule occupies a binding site heretofore not identified in BE ("site 4") and has all 12 glucose units ordered (Fig. 3*A*). It begins at a region on the catalytic domain quite far from the active site and advances toward the active site traversing the width of the catalytic domain. The glucan adopts a helical conformation (Fig. 3*B*), with six glucose units per turn, similar to one chain of a glycogen double helix (24), but it deviates from this conformation as it approaches the active site. The surface of rBE1 is predisposed toward binding the helical conformation, with aromatic stacking interactions (Tyr487 and Trp535) that serve to project the glucan away from the surface and hydrogen-bonding interactions with the glucose units that directly contact the surface (Fig. 3*A*). The residues that directly interact with M12 in this binding site are highly conserved in plant and animal BEs, with 11 of the 19 residues that directly interact with M12 identical in virtually all plant and animal BEs (Fig. 3*C*). However, little conservation is seen in bacterial BEs for this surface, indicating this binding site to be common to the eukaryotic BEs and distinct from the bacterial enzymes, including the starch-making cyanobacterial enzymes with bacteria-like BEs such as *Cyanothece* (GH13_9) (25).

When the structures of M12-bound BE1 are compared with BE1 structures with and without M5, little structural change is observed, apart from the flexible loop between residues 468 and 474 (This loop is either disordered or found in two distinct conformations in the other BEI structures). M12 binding causes this loop to adopt a conformation not seen previously. Numerous residues in the loop make interactions with M12, necessitating the loop to adopt the orientation seen in the M12-bound structure (Fig. 4*A*). This loop interacts with the

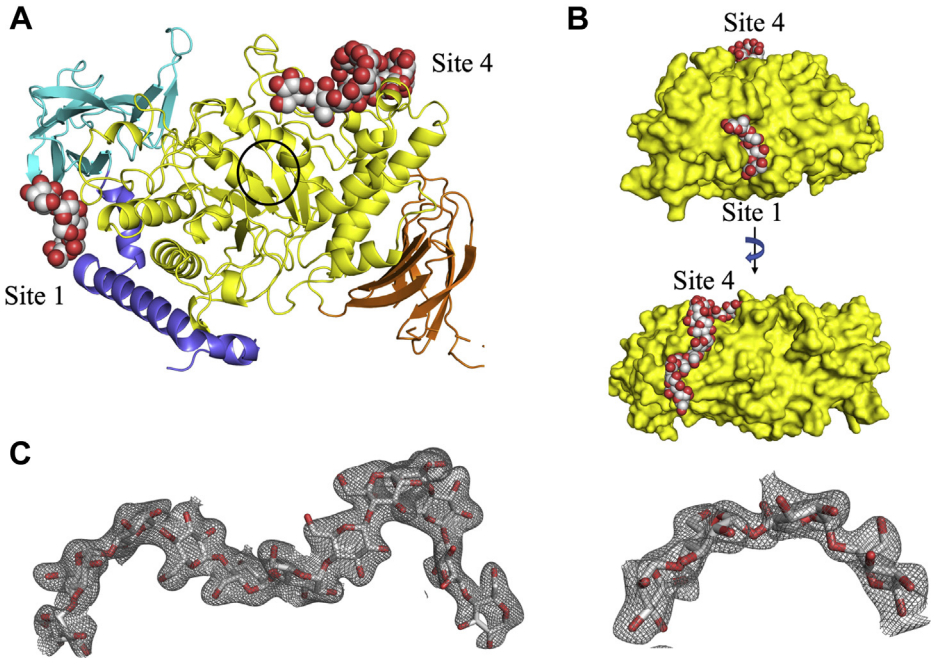


Figure 1. Overall structure of M12-bound rBEI. *A*, cartoon representation of rBEI and bound glucans. N-terminal domain (*blue*), CBM48 domain (*cyan*), catalytic core domain (*yellow*), active site (*black circle*), C-terminal domain (*orange*) and glucans (Carbon atoms, *Gray*, oxygen atoms, *Red* throughout) depicted as space filling models. *B*, surface representation of rBEI (*yellow*) with bound glucans depicted as space filling models. *C*, electron density for oligosaccharides. Electron density ($2F_{obs}$-F_{calc}) is shown in *gray* mesh contoured at 1.0 σ.

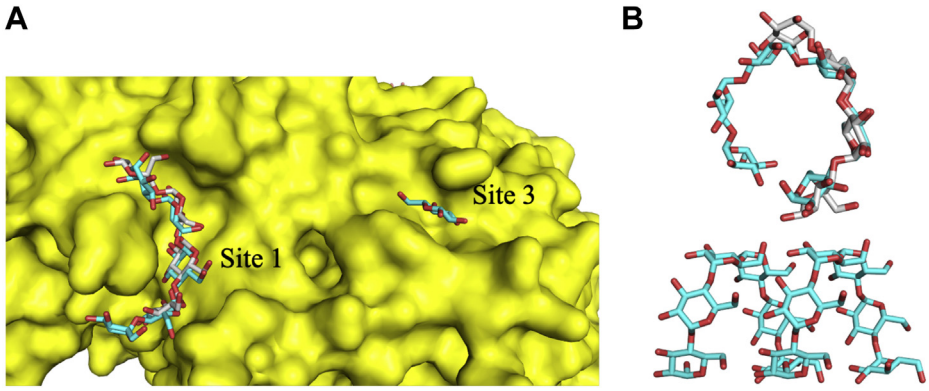


Figure 2. A comparison of binding sites 1 and 3 (*A*) between M5-bound (*cyan* carbons, PDBID 3VU2) and M12-bound (*gray* carbons and *yellow* surface) rBEI structures. The two protein structures were overlaid using PYMOL (Schrodinger Inc). The surface and glucans of the M12-bound structure and the glucans of the M5-bound structure are shown. *B*, top, M12 from site 1 (*gray*) overlaid on a model of an amylose single helix (*cyan*) (obtained from http://polysac3db.cermav.cnrs.fr). *Bottom*, 11 residues of the amylose single helix model (Based on the structure of cycloamylose 26 PDBID 1C58) (34).

glucose moiety found closest to the active site and may act as a "door" into the active site. No other large conformational changes are seen in the M12 binding site when all three rBEI structures are compared.

Mutagenesis

The strong conservation of the interacting residues and the relative proximity to the active site suggest that the M12 binding site may play a direct role in the reaction of the enzyme. This possibility was evaluated by site-directed mutagenesis, activity, and transfer chain specificity assays. Unlike the activity assays, transfer chain specificity for this enzyme is rather a complicated experiment, and it needs to be carefully designed. The duration of the transfer chain specificity assay is a critical parameter since the observed specificity is affected by the time that BE is branching a substrate. As shown in Figure 5*A*, longer reaction times invariably produce more short-branched chains and fewer longer chains. The reason for this observation is that the longer chains produced by rBEI are also substrates for secondary transfers. Secondary transfers will be limited to the branches that are still long enough to be substrates of the enzyme. Mutated enzymes with lower activity must therefore be calibrated with enzymes that have higher activity by varying the reaction time in the assay. This calibration was accomplished by terminating the transfer chain specificity assay for each mutant only when the iodine assay absorption reached 50% of the initial absorption at 660 nm.

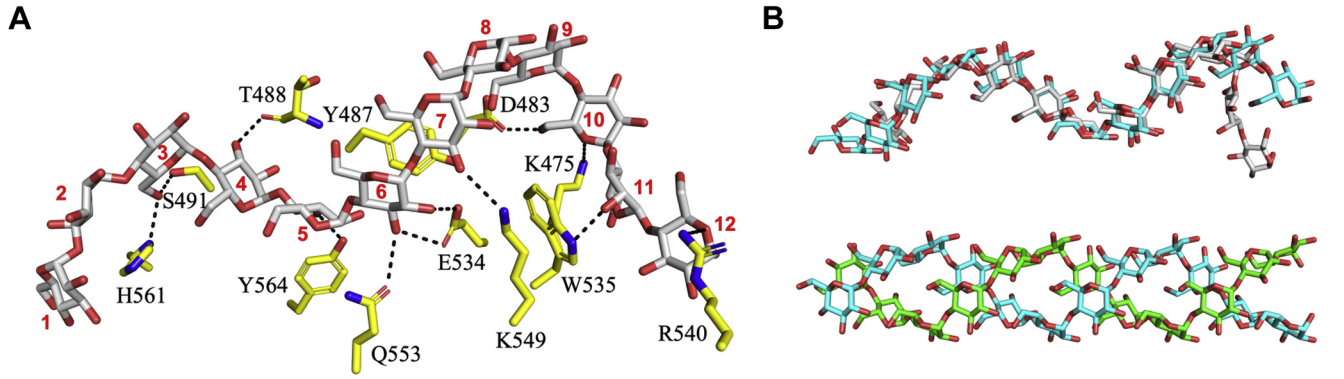
Table 1 summarizes the activities of various rBEI mutants relative to the wild type. One unit was defined as the amount of the enzyme needed to decrease the absorbance of the substrate–iodine complex by 1% per min. The wild-type rBEI specific activity on amylose substrate is 10,100 U/mg, which is comparable to 12,300 U/mg, reported previously (26). Though several point mutants in the M12 binding site (W535A, Y487A, and D483A) showed significant loss of activity, none of these mutants significantly impacted the branch chain specificity (Figs. 5*B* and S2–S9). We also identified a large insertion of 11 residues in the 525 to 553 loop (Loop 541) found in all BEII enzymes (10) (Fig. 4*B*). This loop is proximal to the M12 binding site, and five conserved residues in this loop make interactions with M12, suggesting that it plays an important role in M12 binding. To study differences between BEI and BEII isoforms, an 11-residue insertion found in rBEIIb was introduced to the 525 to 553 loop in BEI (loop 541 mutant). Though a significant loss of activity was observed, no change in branch chain specificity was identified (Fig. S9); one of the most significant differences in BEI *versus* BEII activity is the preference of BEII enzymes for the transfer of shorter (6–7 units) chains relative to BEI enzymes (10, 27).

A second loop, encompassing residues 146 to 156 (Loop 143 mutant), found proximal to the 525 to 553 loop, also had significant sequence deviation between BEI and BEII enzymes (Fig. 4*B*). When both this loop and the 525 to 553 loop were replaced in BEI with the sequence found in rBEIIb (Fig. 4*C*), dramatic differences were seen in the activity. First the overall activity was significantly decreased when amylose was used as substrate (Table 1). Second, the branch chain specificity of BEI was converted from a preference for longer (11–12 glucose units) chains to an almost exclusive preference for chains of 6 to 7 glucose units (Fig. 5, *C* and *D*), similar to that seen for rBEIIb. We therefore conclude that these two loops work together to control the branch chain specificity in rice, and likely other, plant branching enzymes.

Discussion

The unique function and specificity of branching enzymes, their role in synthesizing and modifying growing polymeric substrates, their relatively imprecise, though widely divergent transfer chain specificities, depending on species or even isoform, and their ability to properly space branch chains, make them relatively unusual enzymes, given the high specificity for substrate and product of most enzymes. Though a number of branching enzyme structures are known (16, 19, 21, 23, 28), many of which are bound to malto-oligosaccharides, the structural details that give rise to the unique characteristics of BEs remained mysterious.

M12 is the largest oligosaccharide to be observed at atomic resolution bound to a BE and elaborates a binding surface



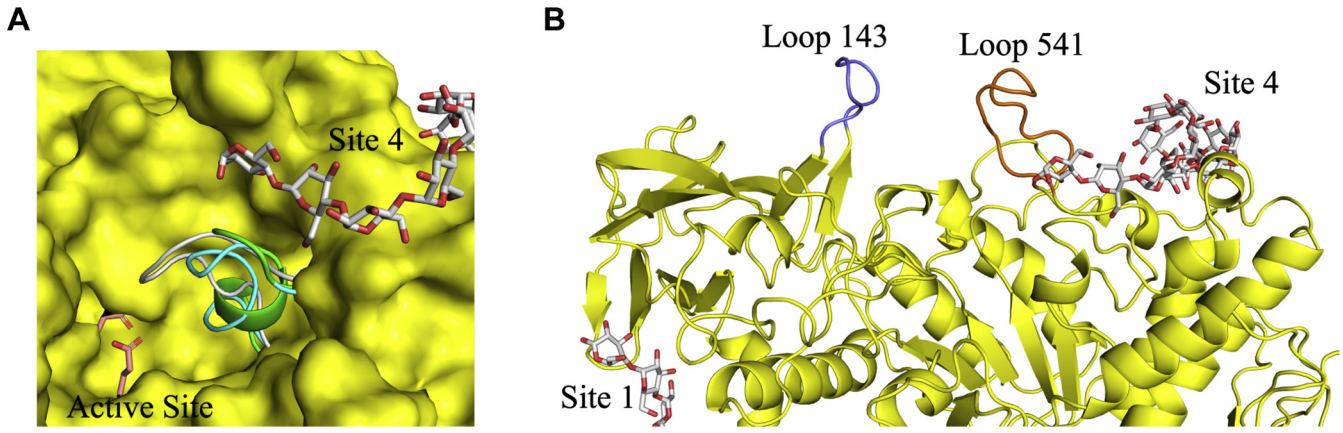
C

		Residues 469-501	Residues 528-566
Plant BEI	Oryza Sativa	QSIVGDKTIAFLLMDKEMYTGMSDLQPASPTIN	NEFGHPEWIDFPRE----------GNNWSYDKCRRQWSLVDTDHL-RYKY
	Zea mays	QSIVGDKTIAFLLMDKEMYTGMSDLQPASPTID	NEFGHPEWIDFPRE----------GNNWSYDKCRRQWSLVDTDHL-RYKY
	Solanum tuberosum	QSIVGDKTIAFLLMDKEMYSGMSCLTDASPVVD	NEFGHPEWIDFPRE----------GNNWSYDKCRRQWNLADSEHL-RYKF
	Aegilops tauschii	QSIVGDKTMAFLLMDKEMYTGMSDLQPASPTID	NEFGHPEWIDFPRE----------GNNWSYDKCRRQWSLADIDHL-RYKY
Plant BEII	Oryza sativa	QALVGDKTIAFWLMDKDMYDFMALDRPATPSID	NEFGHPEWIDFPRAPQVLPNGKFIPGNNNSYDKCRRRFDLGDADYL-RYKY
	Zea mays	QALVGDKTIAFWLMDKDMYDFMALDRPSTPTID	NEFGHPEWIDFPRGPQRLPSGKFIPGNNNSYDKCRRRFDLGDADYL-RYKY
	Solanum tuberosum	QALVGDKTIAFWLMDKDMYDFMALDRPSTSLID	NEFGHPEWIDFPRAEQHLSDGSVIPRNQFSYDKCRRRFDLGDAEYL-RYKY
	Aegilops tauschii	QALVGDKTIAFWLMDKDMYDFMALDRPSTPRID	NEFGHPEWIDFPRGPQTLPTGKVLPGNNNSYDKCRRRFDLGDADFL-RYKY
	Arabidopsis thaliana	QALVGDKTIAFWLMDKDMYDFMAVDRQATPRVD	NEFGHPEWIDFPRTDQHLPDGRVIAGNNGSYDKCRRRFDLGDAEYL-RYKY
Animal and Yeast	Human	QALVGDKSLAFWLMDAEMYTNMSVLTPFTPVID	NEFGHPEWIDFPRK----------GNNESYHYARRQFHLTDDDLL-RYKF
	Drosophila	QALVGDKTLAFWLMDKEMYTHMSTLSDSSVIID	NEFGHPEWIDFPRV----------GNNDSYHYARRQWNLVDDDLL-RYKY
	Yeast	QALVGDKTLAFWLMDAAMYTDMTVLKEPSIVID	NEFGHPEWIDFPNV----------NNGDSYKYARRQFNLADDPLL-RYQN
Prokaryotes	Cyanothece BEI	EIVHGKSNMLGKM----------PG-DEW	MEFGQWSEW----------NVW---GDLEWHLLQ---YEPHQQ
	Cyanothece BEII	EIVHLKGHLVNKM----------PG-DEW	MEFGQTSEW----------QEF---LDIDWYLLQ---YEPHKG
	Escherichia coli	EIVHGKKSILDRM----------PG-DAW	MEFAQGREW----------NHD---ASLDWHLLE-GGDNWHHG

Figure 3. Glucan binding site 4 (*A*) detailed interactions between rBEI and M12 occupying site 4. M12 and the residues from rBE1 are colored by atom (C, *yellow* for M12, C, *green* for rBE1, O, *red*, N, *blue* throughout). *B*, top, M12 from site 4 (*gray*) overlaid on one strand of a glycogen or amylopectin-like double helix (*cyan*) (obtained from http://polysac3db.cermav.cnrs.fr). *Bottom*, the original double helix (*green* and *cyan*). *C*, sequence alignment for representative plant BE1s and BEIIs, human, *drosophila*, yeast, cyanobacteria and *Escherichia coli* BE. The region of sequence that contacts M12 in site 4 is shown. Residues that directly contact or are proximal to M12 are *red*.

starting at a point distal from the active site, and stretching nearby, but not entering, the active site. Numerous mutations confirm the importance of the binding surface for the catalytic activity of the enzyme, and the residues interacting with M12 are highly conserved in eukaryotic (GH13 sub-family 8) BEs. Together this leads to the hypothesis that the M12 binding surface defines a part of either the donor or acceptor chain binding site. Notably, the Y487A mutation, though quite far from the active site, displays very low activity (0.61% ± 0.05 of wild-type rBEI). If we suppose that M12 represents the donor chain binding site, this tyrosine can be involved in interactions between the donor chain and the enzyme only if rBEI is transferring chains longer than ten glucose units. The fact that the fraction of transferred chains smaller than 11 glucose units still accounts for almost 15% of all transferred chains is inconsistent with only 0.61% ± 0.05 activity. In addition, the fact that none of the mutations to this binding surface resulted in any noticeable change in transfer chain length also argues against its involvement in donor chain binding (17, 21, 22).

Further, a recent crystal structure of Maltoheptaose (M7)-bound *Cyanothece* BE (sp. ATCC 51142) (22) shows for the first time a donor chain bound in the active site of a BE. Many of the residues that define the donor chain binding in the M7-bound *Cyanothece* BE structure are conserved in bacterial and eukaryotic BEs, including rBEI, suggesting that all BEs use a similar donor chain binding surface. The M7 binding surface follows a trajectory distinct from that of the rBEI M12 surface (Fig. 6). This leads naturally to the conclusion that the M12 binding surface represents part of the acceptor chain binding site. Using the M7-bound *Cyanothece* BE structure as a guide, an M7 was modeled into the putative donor chain binding surface of M12-bound rBEI (Fig. 6). As shown, the donor chain runs between the 525 to 553 and 146 to 157 loops, in a crevice that links the donor strand binding surface and the CBM48 domain. On the other hand, there is no overlap between the putative donor and M12 acceptor chain binding sites, though they are proximal. Interestingly, the flexible 525 to 553 loop is located between the putative donor and acceptor chain binding sites. This is the same loop that is substantially extended in BEIIb isoforms. As previously discussed, replacement of the rBEI loop sequence with that of rBEIIb reduced the activity of the enzyme and resulted in a significant change in branch chain specificity toward shorter chains, more similar to that of rBEIIb (Fig. 5). Simultaneously exchanging both this loop and



C

		Loop 143	Loop 541
Plant BEI	Oryza Sativa	IPAWIRYATFDASKFGAPY	NEFGHPEWIDFPRE----------GNNWSYDKCRR
	Zea mays	IPALIRYATVDASKFGAPY	NEFGHPEWIDFPRE----------GNNWSYDKCRR
	Solanum tuberosum	IPAWIKYATADATKFAAPY	NEFGHPEWIDFPRE----------GNNWSYDKCRR
	Aegilops tauschii	VPAWIRYATFDASKFGAPY	NEFGHPEWIDFPRE----------GNNWSYDKCRR
Plant BEII	Oryza sativa	IPAWIKYSVQAAG--EIPY	NEFGHPEWIDFPRAPQVLPNGKFIPGNNNSYDKCRR
	Zea mays	IPAWIKYSVQAPG--EIPY	NEFGHPEWIDFPRGPQRLPSGKFIPGNNNSYDKCRR
	Solanum tuberosum	IPAWINYSLQLPD--EIPY	NEFGHPEWIDFPRAEQHLSDGSVIPRNQFSYDKCRR
	Aegilops tauschii	ISAWIKFSVQAPG--EIPF	NEFGHPEWIDFPRGPQTLPTGKVLPGNNNSYDKCRR
	Arabidopsis thaliana	IPAWIKYSVQPPG--EIPY	NEFGHPEWIDFPRTDQHLPDGRVIAGNNGSYDKCRR

Figure 4. Functionally important loops in rBEI. *A*, disordered loop (residues 468 and 474) adopts a new conformation upon M12 binding. Apo rBEI (3AML, *green*), M5 bound rBEI (3VU2, *slate blue*), and M12 bound rBEI (*gray*). *B*, engineered loops of rBEI (*yellow*), Loop 143 (*blue*), loop 541 (*orange*), M12 (stick model, C, *gray*, all other atoms colored as above). *C*, sequence alignment of the two loops (loop 143 and loop 541) that distinguish BEIs and BEIIbs.

the 146 to 156 loop substantially altered branch chain specificity, essentially converting rBEI into an rBEIIb in its product specificity. Thus rBEI, which is the isoform that prefers transfer of the longest chains of any of the isoforms (preferring chains of M11 to M20–M30), is converted to an rBEIIb-like enzyme, that transfers almost exclusively only the shortest, M6 and M7 chains. Together, these results pinpoint the region of plant BEs that is responsible for the specificity of the enzyme for the first time. Notably the 146 to 156 loop represents the end of the N-terminal domain. A previous publication, using maize BEI and BEIIb chimeras, implicated the N-terminal domain in branch chain transfer specificity, consistent with our observation regarding the importance of this loop in the same activity in rBEI and rBEIIb (29). This represented one of the few studies, previous to the present one, that suggested which regions of the enzyme were responsible for this specificity. This second loop lies on the opposite side of the donor chain binding site (Fig. 6) such that the two loops surround the nonreducing end of the donor chain, exactly where they would be expected to be to play a role in branch chain specificity. This serves to confirm that the donor chain trajectory is very similar in rBEs to that of *Cyanothece* BE, making it likely that all eukaryotic BEs share a common donor chain binding surface. Further, it seems that loops on both sides of the donor chain are required for controlling donor chain length. We hypothesize that the longer 525 loop found in BEII enzymes reaches over the donor chain binding surface, interacts with the 146 to 156 loop and the end of an M6 or M7 donor chain to select for shorter donor chains. It is interesting to note that a different loop in *Cyanothece* BE occupies the space of the 525 loop, interacts with the nonreducing end of the M7 glucose unit, and likely provides some of the specificity for shorter glucan chains seen in the *Cyanothece* enzyme. This loop is not conserved in other bacterial enzymes, many of which have branch chain specificities distinct from that of *Cyanothece* BE. The proximity of the 525 loop to both donor and acceptor chains and the fact that residues in this loop make direct interactions with the putative M12 acceptor chain in the M12-bound rBEI structure suggest the possibility that there is allosteric communication between donor- and acceptor-binding sites such that binding of one does not inhibit the binding of the other in the active site. Their proximity also suggests that there may be interaction between donor and acceptor chains when both are bound, as suggested for pullulanases (30). Together, these results suggest that interaction between donor and acceptor chains may be required for acceptor chain binding within the active site. Such a requirement would prevent acceptor chain binding from inhibiting the donor chain from occupying the active site. This would explain why the M12 chain, though making numerous

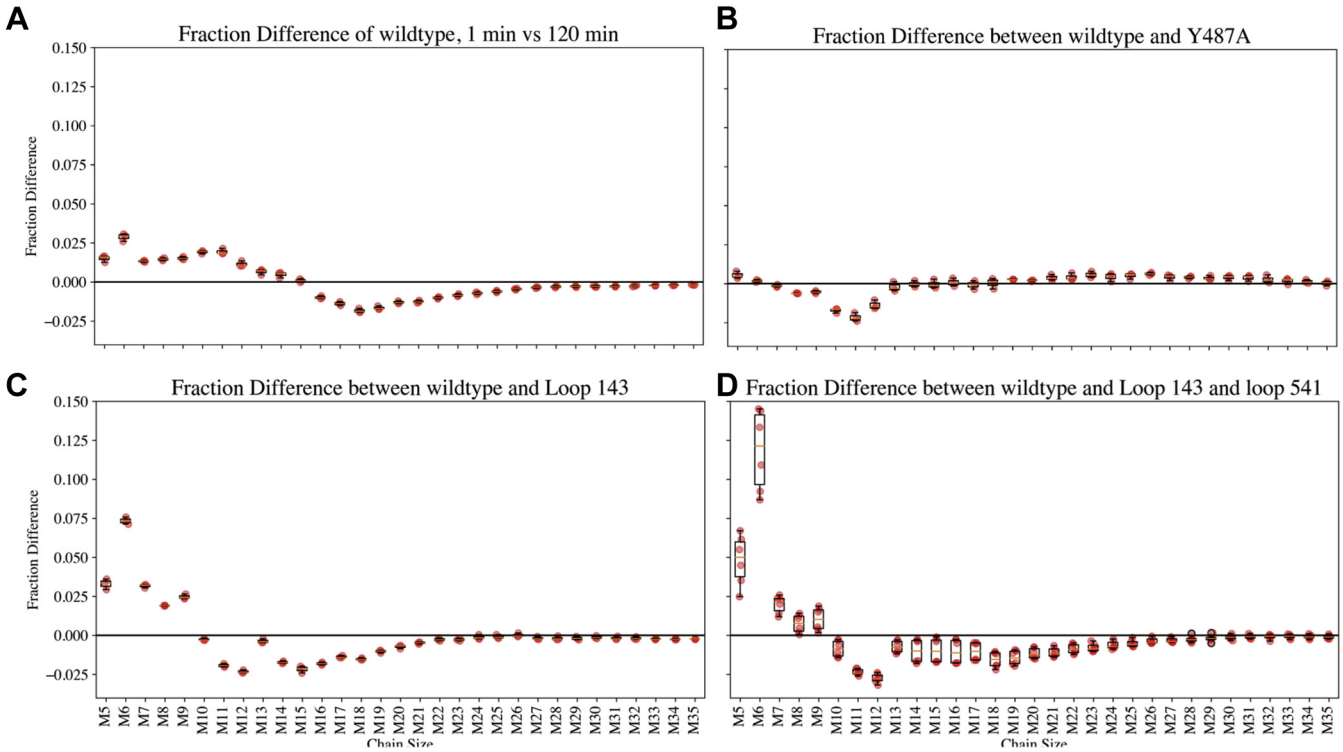


Figure 5. Transfer chain specificity assay and isoform defining loops of BEIs and BEIIs. *A*, fraction differences of transferred chains by wild-type rBEI in 2 h *versus* 1 min. This provides a baseline to account for differences in branch chain distribution caused by the length of the reaction time as opposed to the inherent selectivity of the variant. *B*, fraction differences of transferred chains by wild-type rBEI *versus* Y487A. *C*, fraction differences of transferred chains by wild-type rBEI *versus* loop 143 replacements. *D*, fraction differences of transferred chains by wild-type rBEI *versus* loop 143 and loop 541 replacement.

interactions with the surface of the enzyme, nonetheless does not reach into the active site.

Other questions arising from the structure regard the size and shape of the acceptor glucan chain–enzyme interface. Since many of the interactions occur in the groove of the helical glucan, it would appear to preclude binding of a double helical strand in this binding site. Further, while the 6-hydroxyl groups of the four glucoses closest to the active site (Fig. 7) are at least partially buried in the binding site, with many making direct interactions with rBE1, the last eight 6-hydroxyls of the malto-oligosaccharide point away from the protein surface and are not blocked by protein interactions, indicating that while the first four positions would appear to have difficulty accommodating a branch, the last eight positions would appear to easily accommodate a branched oligosaccharide. This arrangement should inhibit the binding of an oligosaccharide that contains a branch on a sugar within 4 to 6 glucan units of the acceptor site, therefore preventing rBE1 from attaching branches too close together, while still allowing reaction with an acceptor chain branched at a sugar position more than 6 units from the acceptor site, consistent with previous studies demonstrating rBE1 reactivity with branched acceptor chains (31).

In conclusion, with the insights gained from the recent donor-chain bound *Cyanothece* BE structure and the acceptor-chain bound rBEI structure described here, combined with the mutagenesis results that define the regions responsible for controlling donor chain specificity in plant isoforms, an atomic resolution picture of the critical branching process in the dynamic biosynthesis of the starch granule finally begins to emerge.

Experimental procedures

Full-length rice BEI gene was obtained from the National Institute of Agrobiological Sciences in Japan (Sciences,

Table 1
Relative activities of wild-type-rBEI and several key mutants

Mutation	Relative activity[a] (±SE)
Wild type	100% ± 2.59
Control[b] (no enzyme)	0.07% ± 0.00
D344A[b] – Active site	0.11% ± 0.01
H467A[b] – Active site	0.68% ± 0.03
E399A[b] – Active site	0.08% ± 0.04
W535A[b] – M12 Binding site	0.11% ± 0.01
Y487A[b] – M12 Binding site	0.61% ± 0.05
D483A – M12 Binding site	21.97% ± 1.47
G152W – Donor chain site	5.81% ± 1.01
Y229W – Donor chain site	16.61% ± 5.76
Y229A – Donor chain site	20.81% ± 2.28
SNN277AAA – Donor chain site	27.51% ± 1.01
A148W – Donor chain site	83.06% ± 3.22
Loop 143	31.26% ± 1.43
Loop 541	25.10% ± 1.61
Loop 143 and loop 541	2.13% ± 0.26
D156A – CBM Domain	83.06% ± 3.53
D135A – CBM Domain	81.72% ± 3.40
D147A – CBM Domain	51.80% ± 6.32
L40M-V280M-S443P-T669A	96.82% ± 5.62

(Raw data is available as supplementary data).
[a] The activities (based on iodine assay) are normalized to wild-type activity.
[b] For slower mutants, the slope of 60 min activity (decrease of absorption at 660 nm) is used to measure the activity while for faster mutants, only the first 60 s is used.

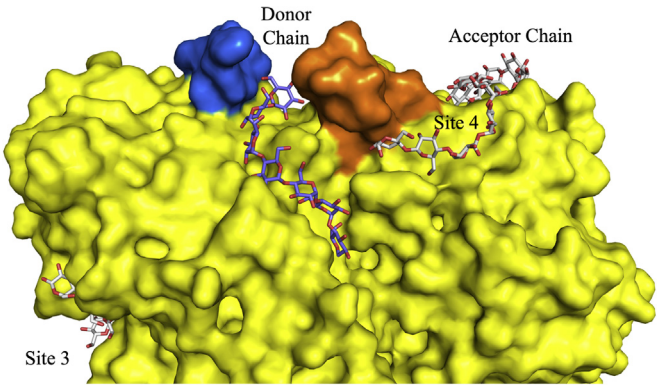


Figure 6. Model of donor strand and acceptor strand binding in rBEI. Model constructed by overlaying M7-bound *Cyanothece* BE (PDBID 5GQY) and M12-bound rBEI but displaying only the M7 from the former structure. rBEI (*yellow*), glucans bound to rBEI (C, *gray*, all other atoms as above), M7 derived from M7-bound *Cyanothece* BE structure (C, *blue*), Loop 143 (*blue*), Loop 541 (*orange*).

National Institute of Agrobiological, J. Rice Genome Resource Center, n.d.). However, the sequence of the gene delivered differed by four residues (L40M-V280M-S443P-T669A) from the complete DNA sequence found on the Rice Genome Resource Center's website http://ricexpro.dna.affrc.go.jp/GGEP/, using the keyword AK119436. The pet-28b vector and the BL21 codon plus expression cells were purchased from Novagen and Agilent Technologies, respectively. Crystallization solutions (Crystal Screen 1 & 2, PEG/Ion 1 & 2, Salt Rx 1 & 2 and Index) were commercially available from Hampton Research. Amylose substrate used in assays was purchased from Sigma-Aldrich (CAS number: 9005-82-7, Amylose from potato). Maltododecaose samples were provided by Dr Park's lab. (See supporting information for more details, section 5).

Protein preparation

The full-length rBEI* gene (with four mutations, L40M, V280M, S443P, T669A) was originally cloned into an adjusted pet-28a vector encompassing histidine and Sumo tags upstream of the gene of interest, giving rise to the following sequence: pET-28-His-Sumo-BamHI-RBEI*-XhoI. Initially, all crystallization and activity studies were carried out with rBEI*. After fixing all the mutations, activity assays (kinetics and chain length distribution assays) were repeated with the wild-type rBEI and mutants. The crystallization of wild-type rBEI did not produce diffracting crystals.

Starting from pET-28-His-Sumo-BamHI-RBEI-XhoI, looping out the N-terminal leader 65-residue sequence from the full-length cDNA generated the mature rBEI. In order to cleave the C-terminal 60 residues, a TEV-protease cut site was introduced. The TEV-protease strategy was employed because attempts at expressing the C-terminally truncated protein were unsuccessful, and as described below, crystallization of the full-length protein proved problematic. Mutations on the mature wild-type rBEI were implemented in the study. The cells were grown at 37 °C to 0.7 OD_{600}, and the expression was induced with 1 mM IPTG at 16 °C. Overnight cultures were collected at 50,000*g* at 4 °C, sonicated for 6 min (15 s on, 45 s off, 50%) at 4 °C. The supernatant was collected at 45,000*g* for 30 min at 4 °C. The protein was purified with Ni-NTA affinity resin using wash buffer [100 mM NaCl, 1 mM BME, 20 mM imidazole and 50 mM Tris (pH 8.0)] and elution buffer [100 mM NaCl, 1 mM BME, 200 mM imidazole, and 50 mM Tris (pH 8.0)]. The purified proteins were used for all activity and chain length specificity assays. Protein prepared for crystallization contained four mutations (L40M, V280M, S443P, T669A). After purification by Ni-NTA affinity chromatography, the protein to be crystallized was immediately cleaved with TEV-protease and sumo-protease (0.5 mg protease for 50 mg protein, overnight, at 4 °C) sequentially, leading to a 702-residue protein, truncated by 60 residues at its c-terminus, a final construct almost identical to that previously crystallized. This construct was found to have essentially identical activity to that of full length, wild-type rBE1. The protein was further purified by size-exclusion chromatography (Superdex 200 16/60 column from GE healthcare), in a buffer of 100 mM NaCl and 50 mM Tris-Cl (pH 8.0). If there were any impurities after this step (usually His-Sumo-protease and His-TEV-protease), they were removed by trapping them in an Ni-NTA affinity column. The PCR primers for mutagenesis are listed in Table S1.

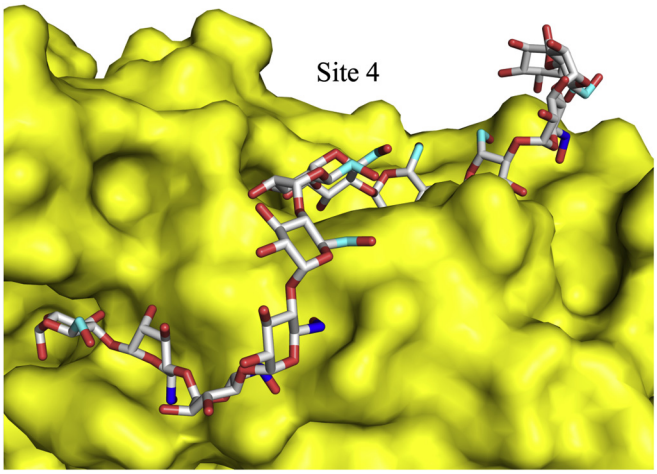


Figure 7. Orientation of C6-hydroxyl groups in the M12 bound rBEI structure. C6-carbons (*blue*) facing the enzyme surface cannot carry a branch. C6-carbons (*cyan*) facing the solvent could carry a branch.

Crystallization

The final purified proteins were concentrated to 3.0 mg/ml (Nano-drop, MW = 82,000 g/mol, ε = 133,000 M^{-1} cm^{-1}) in the size-exclusion chromatography buffer. Attempts to crystallize the full-length 820-residue protein using a multitude of conditions (from Hampton Research mentioned in the Experimental procedures section above) at 4 °C and 25 °C employing the hanging drop method failed to produce diffracting crystals. The mature BEI produced crystals that either did not survive or gave poor resolution upon soaking in dodecaose. On the other hand, soaking $BEI_{\Delta C}$ crystals in the polymer did no damage to these and the best crystals grew in 30% PEG8K, 550 mM sodium acetate, and 100 mM sodium

cacodylate (pH 6.9) in 1 week at room temperature. A maximum concentration of 45 mM dodecaose dissolved in the same buffer; and crystals were soaked in the polymer for 330 min, which was the longest soak time attempted before crystals were damaged. The crystals were then cryoprotected in conditions consisting of the growth solution in addition to 9.55% glycerol and flash frozen in liquid nitrogen.

Structure determination

The crystal structure of rBEI, soaked with M12 was determined by molecular replacement using the rice $BEI_{\Delta C}$ structure previously determined (PDB ID: 3AMK) as a search model. The data was refined to 2.35 Å, complete to 95.24%, and refined to an R_{work} of 17% and R_{free} of 23%. The X-ray diffraction data were collected from a single crystal at the Advanced Photon Source, LS-CAT 21-ID-G beamline. The diffraction images were processed using HKL2000. The structure was refined using PHENIX. The structure shows one molecule in the asymmetric unit and belongs to space group $P2_12_12_1$ with unit cell parameters a = 47.67, b = 80.11, and c = 182.72 Å. Table 2 shows detailed data statistics. X-ray structure and coordinates were deposited in the PDB as 7ML5.

Activity assay

The assay employed was the decrease in absorption of glucan–iodine complex at 660 nm (32). As the amylose substrate is branched by the enzyme, the absorption of the glucan–iodine complex decreases. One unit of activity is defined as a decrease in absorbance of 1.0 per min at 30 °C at 660 nm and is measured in U/mg. Iodine/Iodate stock solution was made by dissolving 2.6 g of KI and 0.26 g of I_2 in 10 ml water, and later it was diluted for running assays (1.95 ml Iodine/Iodate stock solution to 50 ml). Amylose stock solution was made by dissolving 50 mg of amylose in 2 ml water and 0.5 ml 10% NaOH (aq). In order to dissolve the amylose completely, the stock solution was heated by microwave for 30 s. Amylose working solution was made by taking one part freshly made amylose stock solution, one part 1 M sodium citrate buffer (pH 7.0), and eight parts waters with the final pH adjusted to 8.0 by adding HCl. After centrifugation at 11,000*g* for 5 min to separate undissolved solids, 900 µl amylose working solution was equilibrated at 30 °C. In total, 100 µl of enzymes stock solution (50 µg/ml) was added to the amylose working solution, resulting in a final enzyme concentration of 5 µg/ml, and 15 µl samples were taken into 0.985 ml Iodine/Iodate solution at 30 s intervals for 5 min and then at 6, 7, 10, 15, 20, and 60 min (and 180 min for some mutations). Activities were calculated based on the initial linear slope (1 min for relatively active mutations and the wild type, 60 min for less active mutations) and relative to the wild type.

Table 2
Crystallographic data collection and structure refinement statistics

Crystallographic data	Maltododecaose-bound rice branching enzyme I
PDB ID	7ML5
Data statistics	
Resolution range	39.68–2.35 (2.434–2.35)
Space group	P 21 21 21
Unit cell	a = 47.67, b = 80.107, c = 182.716, α = 90, β = 90, γ = 90
Molecules per asymmetric Unit	1
Total reflections	1,126,122
Unique reflections	28,616 (2155)
Completeness (%)	95.24 (73.45)
Average I/σ	14.7 (3.0)
R_{merge}(%)	9.1 (32)
Refinement statistics	
Wilson B-factor	36.88
Reflections used in refinement	28,614 (2155)
Reflections used for R-free	1457 (111)
R-work	0.1648 (0.2310)
R-free	0.2314 (0.3233)
Structure statistics	
Number of non-hydrogen atoms	5904
Macromolecules atoms	5528
Ligands atoms	178
Solvent atoms	198
Protein residues	678
RMS (bonds)	0.003
RMS (angles)	0.61
Ramachandran favored (%)	96.88
Ramachandran allowed (%)	2.82
Ramachandran outliers (%)	0.30
Rotamer outliers (%)	1.38
Clash score	5.52
Average B-factor	37.32
Macromolecules	36.76
Ligands	52.51
Solvent	39.25

Chain length specificity assay

To evaluate the branch chain specificity, amylose was initially branched by the rBEI enzymes as described in activity assay section and immediately the product of branching reaction was debranched by iso-amylase. The debranched mixture was analyzed with LC-TOF-MS to quantify the abundance of different sized oligosaccharides. For the chain length specificity assays, reaction was performed in the same way as activity assays. All reactions were stopped by adding 20 µl concentrated HCl when the absorption of glucan–iodine complex at 660 nm was 50% of its initial absorption (reaction times were estimated from the result of activity assays). After heating the reaction mixture in boiling water to denature the branching enzyme, the pH was adjusted to 3.5 to 4.0 by adding HCl. The debranching reaction was carried out overnight in 40 °C by adding 3 units of isoamylase HP (glucan 6-glucanohydolase) purchased from Megazyme Inc. The debranching reaction was stopped by adding 3× volume Methanol and centrifugation at 11,000*g* for 5 min to remove the precipitates. Solvent from the supernatant was removed using a speed-Vac. The resulting solids were dissolved in 50% methanol and desalted using HyperSep Hypercarb SPE cartridges and 25% acetonitrile as eluent. Chain length distribution was analyzed by LC-TOF-MS previously described by Vismeh *et al.* (33).

Data availability

The coordinates to the rice maltodecaose-bound BE I crystal structure have been deposited in the Protein Data Bank

(accession code 7ML5) and will be released immediately upon publication.

Supporting information—This article contains supporting information.

Acknowledgments—Generous support was provided by the NIH (R01GM101353) and DOE (DE-FG02-06ER15822). All crystallographic data were collected at the Advanced Photon Source, an Office of Science User Facility operated for the US Department of Energy (DOE) Office of Science by Argonne National Laboratory, supported by the US DOE under contract no. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation, the Michigan Technology Tri- Corridor (grant 085P1000817), and the MSU office of the Vice President for Research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions—H. N. G. and J. H. G. conceptualization; H. N. G., Remie Fawaz, and J. H. G. data curation; H. N. G., Remie Fawaz, and J. H. G. formal analysis; J. H. G. funding acquisition; H. N. G., Remie Fawaz, N. E., D. W., K. P., Raoul Fulgos, S. P., Z. A., A. G., and J. H. G. investigation; H. N. G., Remie Fawaz, N. E., S. P., and J. H. G. methodology; J. H. G. project administration; J. H. G. resources; H. N. G. and J. H. G. supervision; H. N. G., Remie Fawaz, and J. H. G. validation; H. N. G., Remie Fawaz, and J. H. G. visualization; H. N. G., Remie Fawaz, and J. H. G. writing–original draft; H. N. G. and J. H. G. writing–review and editing.

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

Abbreviations—The abbreviations used are: BE, branching enzyme; DP, degree of polymerization.

References

1. Tetlow, I. J., and Emes, M. J. (2014) A review of starch-branching enzymes and their role in amylopectin biosynthesis. *IUBMB Life* **66**, 546–558
2. Pérez, S., and Bertoft, E. (2010) The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review. *Starch - Stärke* **62**, 389–420
3. Curá, J. A., Jansson, P.-E., and Krisman, C. R. (1995) Amylose is not strictly linear. *Starch - Stärke* **47**, 207–209
4. Buleon, A., Colonna, P., Planchot, V., and Ball, S. (1998) Starch granules: Structure and biosynthesis. *Int. J. Biol. Macromol.* **23**, 85–112
5. Nakamura, Y. (2015) (Bio)Chemical and Structural Properties of Starch and Glycogen *Starch, Metabolism and Structure*. Springer, Berlin, Heidelberg
6. Blennow, A., and Engelsen, S. B. (2010) Helix-breaking news: Fighting crystalline starch energy deposits in the cell. *Trends Plant Sci.* **15**, 236–240
7. Lü, B., Guo, Z., and Liang, J. (2008) Effects of the activities of key enzymes involved in starch biosynthesis on the fine structure of amylopectin in developing rice (Oryza sativa L.) endosperms. *Sci. China C Life Sci.* **51**, 863–871
8. Rahman, S. (2001) Comparison of starch-branching enzyme genes reveals evolutionary relationships among isoforms. Characterization of a gene for starch-branching enzyme IIa from the wheat D genome donor Aegilops tauschii. *Plant Physiol.* **125**, 1314–1324
9. Denyer, K., Sidebottom, C., Hylton, C. M., and Smith, A. M. (1993) Soluble isoforms of starch synthase and starch-branching enzyme also occur within starch granules in developing pea embryos. *Plant J.* **4**, 191–198
10. Mizuno, K., Kobayashi, E., Tachibana, M., Kawasaki, T., Fujimura, T., Funane, K., Kobayashi, M., and Baba, T. (2001) Characterization of an isoform of rice starch branching enzyme, RBE4, in developing seeds. *Plant Cell Physiol.* **42**, 349–357
11. Li, C., and Gilbert, R. G. (2016) Progress in controlling starch structure by modifying starch-branching enzymes. *Planta* **243**, 13–22
12. Wu, A. C., Morell, M. K., and Gilbert, R. G. (2013) A parameterized model of amylopectin synthesis provides key insights into the synthesis of granular starch. *PLoS One* **8**, e65768
13. Nakamura, Y., Utsumi, Y., Sawada, T., Aihara, S., Utsumi, C., Yoshida, M., and Kitamura, S. (2010) Characterization of the reactions of starch branching enzymes from rice endosperm. *Plant Cell Physiol.* **51**, 776–794
14. Sun, C., Sathish, P., Ahlandsberg, S., and Jansson, C. (1998) The two genes encoding starch-branching enzymes IIa and IIb are differentially expressed in barley. *Plant Physiol.* **118**, 37–49
15. Hamada, S., Ito, H., Hiraga, S., Inagaki, K., Nozaki, K., Isono, N., Yoshimoto, Y., Takeda, Y., and Matsui, H. (2002) Differential characteristics and subcellular localization of two starch-branching enzyme isoforms encoded by a single gene in Phaseolus vulgaris L. *J. Biol. Chem.* **277**, 16538–16546
16. Abad, M. C., Binderup, K., Rios-Steiner, J., Arni, R. K., Preiss, J., and Geiger, J. H. (2002) The X-ray crystallographic structure of *Escherichia coli* branching enzyme. *J. Biol. Chem.* **277**, 42164–42170
17. Feng, L., Fawaz, R., Hovde, S., Gilbert, L., Chiou, J., and Geiger, J. H. (2015) Crystal structures of *Escherichia coli* branching enzyme in complex with linear oligosaccharides. *Biochemistry* **54**, 6207–6218
18. Feng, L., Fawaz, R., Hovde, S., Sheng, F., Nosrati, M., and Geiger, J. H. (2016) Crystal structures of *Escherichia coli* branching enzyme in complex with cyclodextrins. *Acta Crystallogr. D Struct. Biol.* **72**, 641–647
19. Sean Froese, D., Michaeli, A., McCorvie, T. J., Krojer, T., Sasi, M., Melaev, E., Goldblum, A., Zatsepin, M., Lossos, A., Álvarez, R., Escribá, P. V., Minassian, B. A., Von Delft, F., Kakhlon, O., and Yue, W. W. (2015) Structural basis of glycogen branching enzyme deficiency and pharmacologic rescue by rational peptide design. *Hum. Mol. Genet.* **24**, 5667–5676
20. Wang, Z., Xin, C., Li, C., Gu, Z., Cheng, L., Hong, Y., Ban, X., and Li, Z. (2019) Expression and characterization of an extremely thermophilic 1,4-α-glucan branching enzyme from Rhodothermus obamensis STB05. *Protein Expr. Purif.* **164**, 105478
21. Pal, K., Kumar, S., Sharma, S., Garg, S. K., Alam, M. S., Xu, H. E., Agrawal, P., and Swaminathan, K. (2010) Crystal structure of full-length mycobacterium tuberculosis H37Rv glycogen branching enzyme: Insights of N-terminal β-sandwich in substrate specificity and enzymatic activity. *J. Biol. Chem.* **285**, 20897–20903
22. Hayashi, M., Suzuki, R., Colleoni, C., Ball, S. G., Fujita, N., and Suzuki, E. (2017) Bound substrate in the structure of cyanobacterial branching enzyme supports a new mechanistic model. *J. Biol. Chem.* **292**, 5465–5475
23. Chaen, K., Noguchi, J., Omori, T., Kakuta, Y., and Kimura, M. (2012) Crystal structure of the rice branching enzyme I (BEI) in complex with maltopentaose. *Biochem. Biophys. Res. Commun.* **424**, 508–511
24. Popov, D., Buléon, A., Burghammer, M., Chanzy, H., Montesanti, N., Putaux, J.-L., Potocki-Véronèse, G., and Riekel, C. (2009) Crystal structure of A-amylose: A revisit from synchrotron microdiffraction analysis of single crystals. *Macromolecules* **42**, 1167–1174
25. Suzuki, E., and Suzuki, R. (2016) Distribution of glucan-branching enzymes among prokaryotes. *Cell. Mol. Life Sci.* **73**, 2643–2660
26. Akasaka, T., Vu, N. T., Chaen, K., Nishi, A., Satoh, H., Ida, H., Omori, T., and Kimura, M. (2009) The action of rice branching enzyme I (BEI) on starches. *Biosci. Biotechnol. Biochem.* **73**, 2516–2518
27. Nishi, A., Nakamura, Y., Tanaka, N., and Satoh, H. (2001) Biochemical and genetic analysis of the effects of amylose-extender mutation in rice endosperm. *Plant Physiol.* **127**, 459–472
28. Noguchi, J., Chaen, K., Vu, N. T., Akasaka, T., Shimada, H., Nakashima, T., Nishi, A., Satoh, H., Omori, T., Kakuta, Y., and Kimura, M. (2011) Crystal structure of the branching enzyme i (BEI) from Oryza sativa L with implications for catalysis and substrate binding. *Glycobiology* **21**, 1108–1116

29. Kuriki, T., Stewart, D. C., and Preiss, J. (1997) Construction of chimeric enzymes out of maize endosperm branching enzymes I and II: Activity and properties. *J. Biol. Chem.* **272**, 28999–29004
30. Mikami, B., Iwamoto, H., Malle, D., Yoon, H.-J., Demirkan-Sarikaya, E., Mezaki, Y., and Katsuya, Y. (2006) Crystal structure of pullulanase: Evidence for parallel binding of oligosaccharides in the active site. *J. Mol. Biol.* **359**, 690–707
31. Sawada, T., Nakamura, Y., Ohdan, T., Saitoh, A., Francisco, P. B., Suzuki, E., Fujita, N., Shimonaga, T., Fujiwara, S., Tsuzuki, M., Colleoni, C., and Ball, S. (2014) Diversity of reaction characteristics of glucan branching enzymes and the fine structure of α-glucan from various sources. *Arch. Biochem. Biophys.* **562**, 9–21
32. Boyer, C. D., and Preiss, J. (1978) Multiple forms of (1 → 4)-α-d-glucan, (1 → 4)-α-d-glucan-6- glycosyl transferase from developing Zea mays L. Kernels. *Carbohydr. Res.* **61**, 321–334
33. Vismeh, R., Humpula, J. F., Chundawat, S. P. S., Balan, V., Dale, B. E., and Jones, A. D. (2013) Profiling of soluble neutral oligosaccharides from treated biomass using solid phase extraction and LC-TOF MS. *Carbohydr. Polym.* **94**, 791–799
34. Gessler, K., Usón, I., Takaha, T., Krauss, N., Smith, S. M., Okada, S., Sheldrick, G. M., and Saenger, W. (1999) V-amylose at atomic resolution: X-ray structure of a cycloamylose with 26 glucose residues (cyclomaltohexaicosaose). *Proc. Natl. Acad. Sci. U. S. A.* **96**, 4246–4251