

# SCIENTIFIC REPORTS



OPEN

## Crystal structure of a raw-starch-degrading bacterial $\alpha$ -amylase belonging to subfamily 37 of the glycoside hydrolase family GH13

Yanhong Liu<sup>1</sup>, Jigang Yu<sup>1</sup>, Fudong Li<sup>2</sup>, Hui Peng<sup>1</sup>, Xuecheng Zhang<sup>1</sup>, Yazhong Xiao<sup>1</sup> & Chao He<sup>1</sup>

Received: 18 October 2016

Accepted: 02 February 2017

Published: 17 March 2017

Subfamily 37 of the glycoside hydrolase family GH13 was recently established on the basis of the discovery of a novel  $\alpha$ -amylase, designated AmyP, from a marine metagenomic library. AmyP exhibits raw-starch-degrading activity and consists of an N-terminal catalytic domain and a C-terminal starch-binding domain. To understand this newest subfamily, we determined the crystal structure of the catalytic domain of AmyP, named AmyP <sub>$\Delta$ SBD</sub>, complexed with maltose, and the crystal structure of the E221Q mutant AmyP <sub>$\Delta$ SBD</sub> complexed with maltotriose. Glu221 is one of the three conserved catalytic residues, and AmyP is inactivated by the E221Q mutation. Domain B of AmyP <sub>$\Delta$ SBD</sub> forms a loop that protrudes from domain A, stabilizes the conformation of the active site and increases the thermostability of the enzyme. A new calcium ion is situated adjacent to the -3 subsite binding loop and may be responsible for the increased thermostability of the enzyme after the addition of calcium. Moreover, Tyr36 participates in both stacking and hydrogen bonding interactions with the sugar motif at subsite -3. This work provides the first insights into the structure of  $\alpha$ -amylases belonging to subfamily 37 of GH13 and may contribute to the rational design of  $\alpha$ -amylase mutants with enhanced performance in biotechnological applications.

$\alpha$ -Amylases (EC 3.2.1.1) are endo-hydrolases that act on  $\alpha$ -(1,4)-glucosidic linkages in starch and other related oligo- and polysaccharides, thus causing the release of malto-oligosaccharides and glucose in the  $\alpha$ -anomeric form.  $\alpha$ -Amylases are essential for the conversion of starch into oligosaccharides and are critical for many organisms that use starch as a primary source of energy.  $\alpha$ -Amylases also have potential applications in a number of industrial processes, such as processes in the food, fermentation, and pharmaceutical industries.

According to the Carbohydrate-Active enZymes (CAZy) database, glycoside hydrolase family 13 (GH13) is the main  $\alpha$ -amylase family that, which, together with GH70 and GH77, forms clan GH-H, in which the catalytic domain adopts a  $(\beta/\alpha)_8$ -barrel. GH13  $\alpha$ -amylases use a reaction mechanism that retains anomeric configuration, and share 4–7 conserved sequence regions and catalytic machinery<sup>1</sup>. The observation that some GH13 members exhibit high mutual sequence similarities has led to the current division of the GH13 family into 42 subfamilies. The GH13\_37 subfamily is the most recently established  $\alpha$ -amylase subfamily<sup>2</sup>. It was created on the basis of the isolation and phylogenetic analysis of a novel  $\alpha$ -amylase, designated AmyP, from a marine metagenomic library. AmyP shares 85.6% sequence identity with the glycosidase A0A0B9G5J0 from the bacterium *Photobacterium gaetbulicola*. AmyP is 639 amino acids in length and has a theoretical molecular weight of approximately 70 kD. It has been proposed to contain a conserved catalytic domain at the N terminus<sup>2</sup> and a starch binding domain at the C terminus<sup>3</sup>. AmyP exhibits high substrate specificity for starch over pullulan or cyclomaltoextrins, efficiently hydrolyzing starch into glucose, maltose and maltotriose<sup>2</sup>. Currently, the GH13\_37 subfamily contains, in addition to AmyP, approximately 56 hypothetical members, most of which are found in marine bacteria.

In general, GH13  $\alpha$ -amylases are three-domain proteins, consisting of domains A, B and C. Of special interest is that subfamily GH13\_37 members may lack domain B because the loop connecting strand  $\beta$ 3 to helix  $\alpha$ 3 of

<sup>1</sup>Anhui Provincial Engineering Technology Research Center of Microorganisms and Biocatalysis and School of Life Sciences, Anhui University, Hefei, Anhui 230601, China. <sup>2</sup>Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230026, China. Correspondence and requests for materials should be addressed to C.H. (email: chaohe@ahu.edu.cn)

their catalytic ( $\beta/\alpha$ )<sub>8</sub>-barrel appears to be too short to form a B domain typical of the GH13 family. For several  $\alpha$ -amylases, it has been reported that the metal and substrate binding sites are located in domain B and its interface with domain A and that these regions affect the activity and stability of the enzyme<sup>4–8</sup>. The specific role of domain B in the GH13\_37 subfamily will be clearer when the three-dimensional structure of the novel  $\alpha$ -amylase AmyP has been solved and described in detail.

AmyP is a raw-starch-degrading enzyme (RSDE), thus making it of potential interest in the food processing and green-fuels communities. AmyP exhibits the unique and remarkable ability to preferentially and very rapidly digest raw rice starch. The specific activity of AmyP at 40 °C against 4% raw rice starch reaches  $118.5 \pm 0.6$  U/mg, which is much higher than its activity against other raw starches<sup>9</sup>. Maximal hydrolysis by AmyP occurs within 4 h for 10 mg/ml and within 1 h for 80 mg/ml raw rice starch, thus indicating a very rapid rate of hydrolysis.

Although much work on the production, purification, characterization, and application of microbial RSDEs has been conducted in recent years<sup>10</sup>, studies focusing on the mechanism of action of RSDEs are rare. Most bacterial and fungal RSDEs possess starch-binding domains (SBDs), usually at the C terminus, that mediate attachment to starch granules<sup>11,12</sup> and in some cases disrupt the surface of the granules<sup>13</sup>. It has previously been shown that the C-terminal domain of AmyP has an affinity for raw rice starch granules, that is similar to that of full-length AmyP, whereas the binding of the catalytic domain of AmyP, designated AmyP <sub>$\Delta$ SBD</sub>, to raw starch granules is so weak that it is difficult to quantify. This result suggests that the C-terminal domain of AmyP can be unambiguously defined as an SBD and that the binding of AmyP to raw starch depends primarily on its SBD. However, some amylolytic enzymes that lack an SBD domain are capable of binding and digesting raw starch. These enzymes include the barley  $\alpha$ -amylase isozyme AMY1<sup>14–17</sup>, yeast glucoamylase from *Saccharomyces fibuliger*<sup>18</sup>, and human salivary  $\alpha$ -amylase<sup>19</sup>, which have well-defined structures and possess one or more secondary carbohydrate-binding sites called surface-binding sites (SBSs) on the catalytic domain or an intimately associated domain rather than on an SBD. The SBSs usually contain aromatic residues. Recently, the bacterial  $\alpha$ -amylases from *Anoxybacillus* species (ASKA and ADTA)<sup>20</sup>, *Geobacillus thermoleovorans* (GTA)<sup>21</sup>, Pizzo  $\alpha$ -amylase<sup>22</sup>, and GtamyII<sup>23</sup>), and *Bacillus aquimaris* (BaqA<sup>24</sup>) have been proposed to constitute a novel GH13 subfamily<sup>25</sup> of raw-starch-digesting  $\alpha$ -amylases, that lack SBDs but exhibit the ability to degrade raw starch. However, enzymes in this subfamily possess an extended C-terminal region that is rich in aromatic residues, and the actual role of this region remains to be investigated. Likewise, AmyP <sub>$\Delta$ SBD</sub> retains raw-starch digesting ability and has a specific activity of  $55.5 \pm 1.4$  U/mg for 4% raw rice starch<sup>3</sup>. Additionally, another GH13\_37  $\alpha$ -amylase that does not possess a distinct SBD, AmyASS from *Aeromonas salmonicida*, is able to degrade raw rice starch<sup>26</sup>. Thus, the knowledge of the three-dimensional structure of AmyP may improve understanding of the mechanism of action of RSDEs.

Here, we present the first crystal structures of  $\alpha$ -amylase from the GH13\_37 subfamily. This work provides further insights into the structure of  $\alpha$ -amylases that can directly degrade raw starch granules.

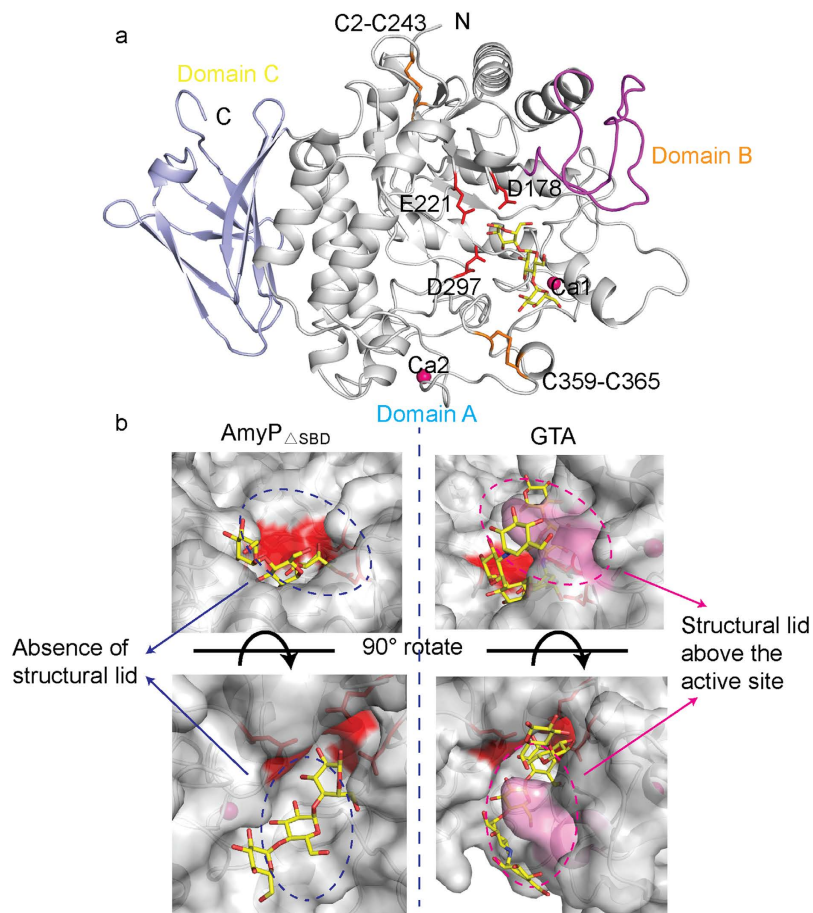
## Results

**Overall structure.** The catalytic residues of AmyP were initially identified by comparing the amino acid sequence of AmyP with sequences in the Protein Data Bank. Substitution of the glutamic acid at position 221 of AmyP with glutamine (E221Q mutation) resulted in nearly complete inactivation of the enzyme. This inactivation was confirmed by enzyme activity assays using 10 mg/ml of soluble starch as substrate. The E221Q mutant displayed 0.125% of the activity of the wild-type (WT) enzyme in these assays. The crystal structure of the E221Q catalytic domain mutant AmyP <sub>$\Delta$ SBD</sub> in complex with maltotriose was obtained through co-crystallization of the E221Q mutant of full-length AmyP with  $\gamma$ -cyclodextrin ( $\gamma$ -CD) (AmyP-E221Q/ $\gamma$ -CD). This structure was determined to a resolution of 2.55 Å by molecular replacement by using the crystal structure of the *Bacillus stearothermophilus* TRS40 neopullulanase (pdb accession code 1J0H<sup>27</sup>) as a search model. The crystal structure of WT AmyP <sub>$\Delta$ SBD</sub> complexed with maltose was obtained through co-crystallization of full-length WT AmyP with  $\beta$ -cyclodextrin ( $\beta$ -CD), an unpreferred substrate of AmyP (AmyP/ $\beta$ -CD). This structure was determined to a resolution of 1.95 Å by using the AmyP-E221Q/ $\gamma$ -CD structure as an initial model.

Both crystal structures contain four enzyme molecules in the asymmetric unit. Surprisingly, each molecule includes residues 1 to 496, corresponding to AmyP <sub>$\Delta$ SBD</sub>, the theoretical molecular weight of approximately 55 kD. Nevertheless, no electron density is invisible for residues 497–639, which correspond to the C-terminal SBD domain. Peptide sequencing of protein bands from the dissolved crystal sample by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and liquid chromatography-mass spectrometry (LC-MS) showed that partial degradation of the SBD domain of the AmyP-E221Q/ $\gamma$ -CD crystals had occurred (Fig. S1a). The far-UV circular dichroism (CD) spectrum showed that the complete SBD domain has some secondary structure, and is not totally disordered (Fig. S1b). These results suggested that the SBD region of recombinant AmyP expressed in *Escherichia coli* cells might be poorly organized, thus resulting in its absence in the crystal structure.

The AmyP <sub>$\Delta$ SBD</sub> is composed of A, B, and C domains and shares structural features with other GH13  $\alpha$ -amylases (Fig. 1a). Domain A (residues 1–125 and 155–410) forms a ( $\beta/\alpha$ )<sub>8</sub> barrel that contains the catalytic site. Domain B (residues 126–154) forms a loop that protrudes from domain A and has a non-regular secondary structure. Domain C (residues 411–498) consists of an antiparallel eight-stranded  $\beta$ -sheet. In addition, four cysteine residues in domain A are involved in two disulfide bridges (Cys2–Cys243 and Cys359–Cys365) that are conserved in the GH13\_37 subfamily (Fig. S2). The locations of the disulfide bridges in AmyP <sub>$\Delta$ SBD</sub> are not identical to the locations of the disulfide bridges in chloride-dependent  $\alpha$ -amylases<sup>28,29</sup>.

A search on the Dali server revealed that AmyP <sub>$\Delta$ SBD</sub> is structurally most similar to the *B. stearothermophilus* neopullulanase excluding domain N (pdb accession code 1J0H<sup>27</sup>, z-score 41.1 and rmsd of 2.0 Å for 279 Ca atoms, with 28% sequence identity), *G. thermoleovorans*  $\alpha$ -amylase GTA lacking transmembrane regions (pdb accession code 4e2o<sup>21</sup>, z-score 39.6 and rmsd of 2.2 Å for 288 Ca atoms, with 26% sequence identity), and *Thermus* maltogenic amylase excluding domain N (pdb accession code 1GVI<sup>30</sup>, z-score 39.5 and rmsd of 2.2 Å for 288 Ca atoms, with 27% sequence identity). The most divergent portion of the structure lies within the B domain, which

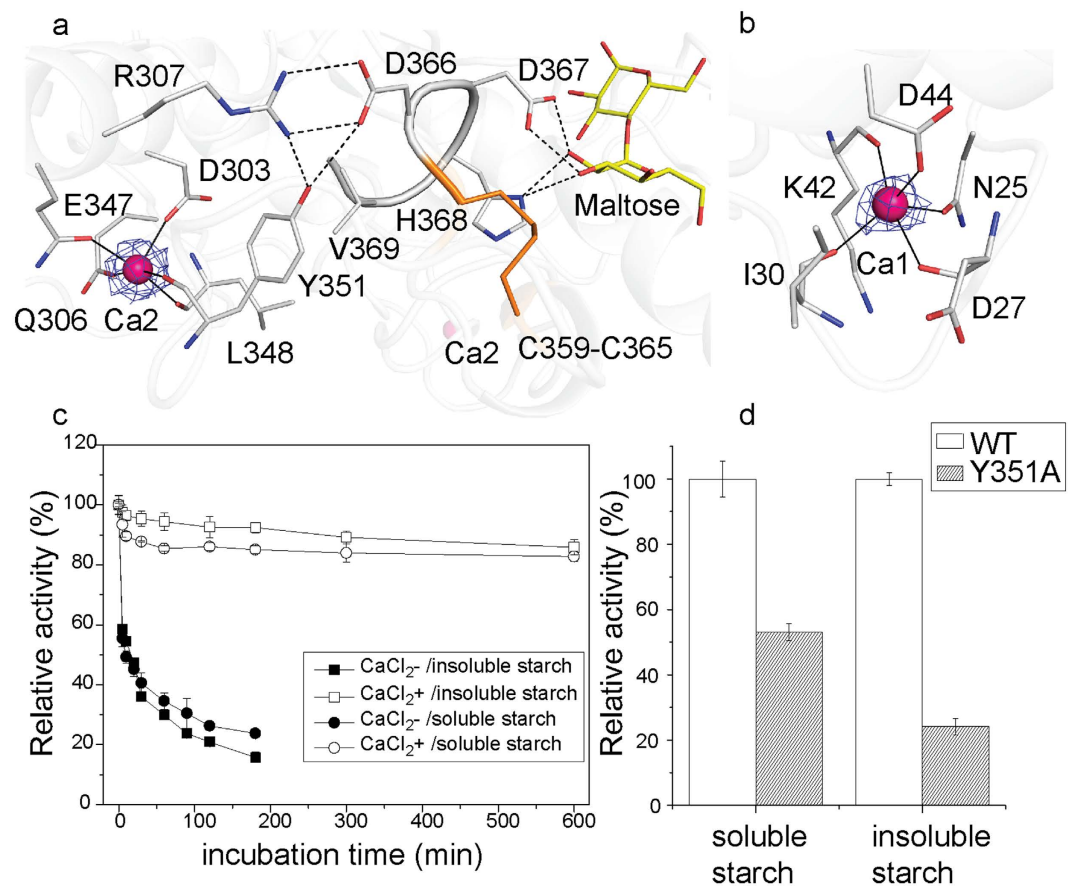


**Figure 1. Overall structure of AmyP<sub>ΔSBD</sub>.** (a) AmyP<sub>ΔSBD</sub> consists of domains A (shown in gray), B (shown in magenta) and C (shown in blue). Disulfide bridges are shown in orange. Calcium ions are shown as hot pink spheres. Catalytic residues are shown as red sticks. (b) Structural comparison of active sites of AmyP<sub>ΔSBD</sub> and GTA (pdb accession code 4e2o) shows the absence of a protruding lid above the active site for AmyP<sub>ΔSBD</sub>. The lid in GTA (shown in violet) is involved in the stabilization of the substrate. The catalytic residues are shown in red.

is adjacent to the active site. The B domain of AmyP<sub>ΔSBD</sub> is indeed considerably shorter than the B domains of  $\alpha$ -amylases belonging to other GH13 subfamilies. It is composed of only loop elements, does not bind metal ions and points away from the catalytic site. Structural comparison of AmyP<sub>ΔSBD</sub> and GTA showed that the B domain loop of AmyP<sub>ΔSBD</sub> lacks a region corresponding to residues 181–184 of GTA, which form a small structural lid above the active site (Fig. 1B). Owing to the small size of domain B in AmyP<sub>ΔSBD</sub>, a much wider and more open substrate binding pocket is created, which may facilitate access to starch granule surfaces and increase the rate of hydrolysis of raw starch.

**Ca<sup>2+</sup> ion binding to AmyP<sub>ΔSBD</sub>.** We identified two Ca<sup>2+</sup> ions in the AmyP<sub>ΔSBD</sub> structure on the basis of high electron densities at their positions and on coordination geometry (Fig. 2a,b). These Ca<sup>2+</sup> ions, which we designated Ca1 and Ca2, are located in the periphery of domain A (Fig. 1a). Ca1 lies at the back of the active center, whereas Ca2 lies in the vicinity of the -3 subsite binding loop. Their positions differ from that of the conserved Ca<sup>2+</sup> at the junction of domains A and B that is present in most GH13  $\alpha$ -amylases.

Ca1 is coordinated by the main-chain oxygens of Asp27, Ile30 and Lys42, and the side-chain oxygens of Asn25 and Asp44, with an average coordinating distance of 2.4 Å (Fig. 2b). This calcium-binding site resembles those in the structures of *B. stearothermophilus* neopullulanase (a GH13\_20 family member, pdb accession code 1J0H), *Halothermothrix orenii*  $\alpha$ -amylase AmyA (a GH13\_36 family member, pdb accession code 1wza<sup>31</sup>) and truncated *Anoxybacillus*  $\alpha$ -amylase TASKA (a member of a new GH13 subfamily, pdb accession code 5a2b<sup>20</sup>) (Fig. S3). In particular, the two side-chain carbonyl groups of Asp44 and Asn25 of AmyP<sub>ΔSBD</sub>, both of which function in calcium coordination, are strictly conserved among the four proteins (Figs S2 and S3). Ca2 is coordinated by the main-chain oxygens of Leu348 and Tyr351 and the side-chain oxygens of Asp303, Gln306 and Glu347, with an average coordinating distance of 2.3 Å (Fig. 2a). This calcium-binding site is unique among GH13  $\alpha$ -amylases. Notably, all residues in AmyP<sub>ΔSBD</sub> that interact with Ca<sup>2+</sup> through their side chains are conserved among GH13\_37 subfamily members (Fig. S2), thus indicating that calcium binding is highly conserved within this subfamily.



**Figure 2. Calcium ion binding sites in AmyP<sub>ΔSBD</sub>.** (a) Stick representation of the Ca2 binding site and the residues surrounding Ca2. Ca2 is coordinated by the main-chain oxygens of Leu348 and Tyr351 and the side-chain oxygens of Asp303, Gln306 and Glu347 in domain A. Close-up view showing the 2*Fo*–*Fc* electron density pattern (2σ contoured) of Ca2. Ca2 may help maintain the conformations of Arg307 and Tyr351, thereby stabilizing the substrate binding loop formed by Asp366–Val369. Hydrogen bonds are shown as dashed lines. (b) Stick representation of the Ca1 binding site. Ca1 is bound to the main-chain oxygens of Asp27, Ile30 and Lys42, and the side chain oxygens of Asn25 and Asp44 from domain A. Close-up view showing the 2*Fo*–*Fc* electron density pattern (2σ contoured) of Ca1. (c) Thermal inactivation of WT AmyP<sub>ΔSBD</sub> incubated in the absence or in the presence of 10 mM CaCl<sub>2</sub> at 40 °C and pH 6.5 for the times indicated. The enzyme activity against soluble and insoluble rice starch was then assayed in the standard buffer. (d) Specific activities of the Y351A variant relative to WT AmyP<sub>ΔSBD</sub> with soluble and insoluble rice starch as substrates.

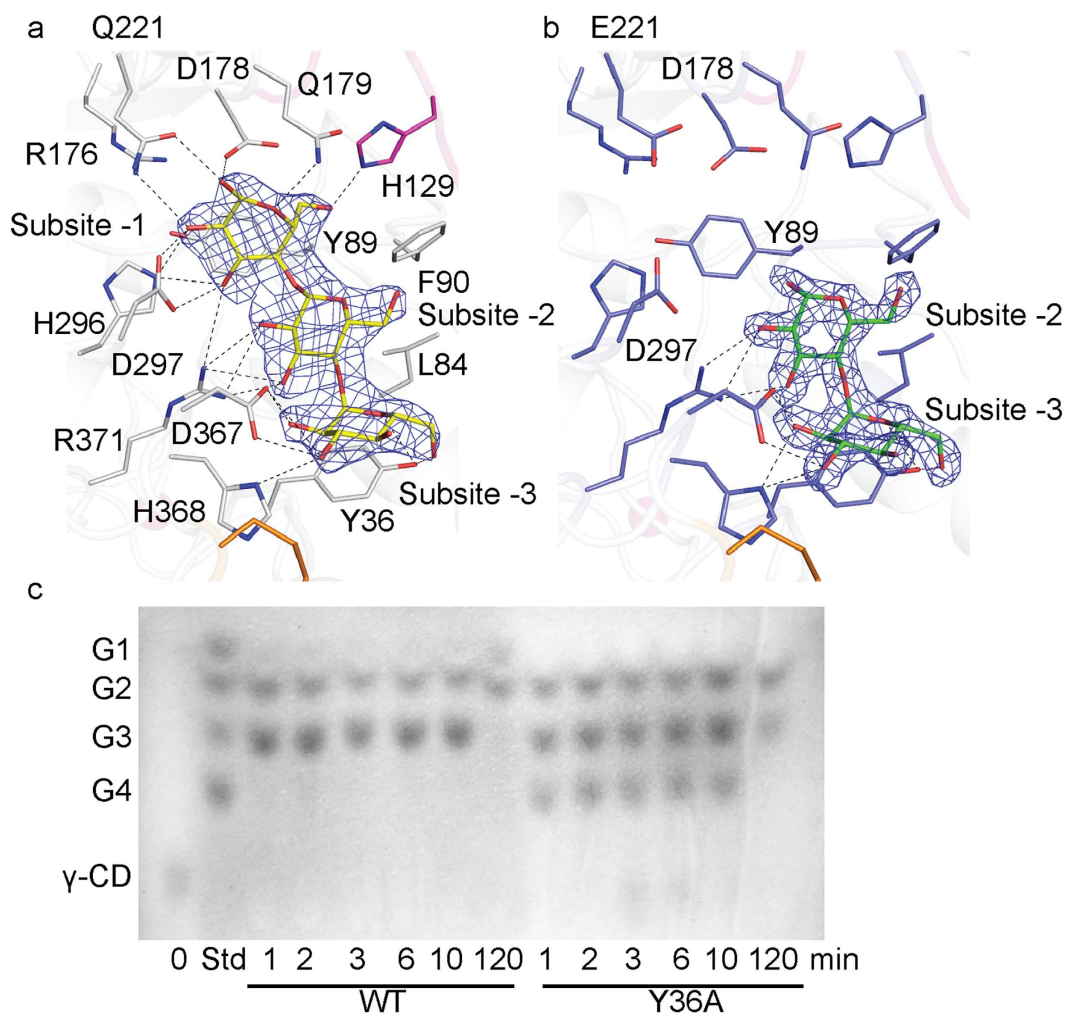
To probe the effect of Ca<sup>2+</sup> ions on enzyme stability, thermal inactivation of the hydrolytic activity of AmyP<sub>ΔSBD</sub> against both soluble and insoluble rice starch was examined in the presence and absence of added Ca<sup>2+</sup>. AmyP<sub>ΔSBD</sub> retained more than 80% of its maximal activity after incubation at 40 °C for 10 h in the presence of 10 mM CaCl<sub>2</sub>, in contrast, the enzyme lost most of its activity after incubation at 40 °C for 3 h in the absence of CaCl<sub>2</sub> (Fig. 2c). These results suggested that Ca<sup>2+</sup> ions bound to AmyP<sub>ΔSBD</sub> substantially contribute to the enzyme's thermostability, as has been found for other GH13 α-amylases. The stabilizing effect of bound Ca<sup>2+</sup> was ascribed mainly to Ca2, which is adjacent to the -3 subsite. Ca2 helps to maintain the conformations of nearby Arg307 and Tyr351 (Fig. 2a), which interact with their facing residues Asp366 and Val369 through hydrogen bonding and hydrophobic interactions. Asp366 and Val369 are situated in a -3 subsite loop that contains the substrate-binding residues Asp367 and His368. The importance of Tyr351 in stabilizing this -3 subsite loop is supported by the Y351A mutant's significantly decreased specific activity against both soluble and insoluble rice starch (Fig. 2d) and a greatly increased *K<sub>m</sub>* for the γ-CD substrate, as compared with the WT enzyme (Table 1). Furthermore, the overall conformation of the Y351A mutant is identical to that of the WT enzyme, as indicated by far-UV CD and tryptophan intrinsic fluorescence spectra (Fig. S4). Together, these results suggest that Ca2 promotes localized stabilization of a -3 subsite loop and that this stabilization helps to preserve the catalytic activity of the enzyme.

**The active site.** The structure of AmyP-E221Q/γ-CD contains electron density representing a maltotriose molecule bound to the active site and occupying subsites -1, -2 and -3 (Fig. 3a), whereas the structure of AmyP/β-CD contains electron density representing a maltose molecule at subsites -2 and -3 (Fig. 3b). These observations suggested that the E221Q mutant possesses activity for γ-CD and the WT enzyme possesses activity



Enzyme	$K_m$ (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_m$ ( $s^{-1} mM^{-1}$ )
WT	$0.12 \pm 0.005$	$39.58 \pm 0.004$	323.2
N148Y	$0.10 \pm 0.009$	$26.11 \pm 0.01$	273.7
N148A	$0.20 \pm 0.03$	$22.59 \pm 0.007$	112.8
Y351A	$2.11 \pm 0.09$	$25.02 \pm 0.006$	11.9
K131A	$1.29 \pm 0.13$	$7.47 \pm 0.004$	5.8

**Table 1.** Kinetic parameters of the WT and mutant AmyP $_{\Delta SBD}$  enzymes for the  $\gamma$ -CD substrate.



**Figure 3.** Active site of AmyP $_{\Delta SBD}$ . (a) Electron density of maltotriose bound to the active site of the E221Q mutant of AmyP $_{\Delta SBD}$  from the  $2Fo-Fc$  map contoured at  $2\sigma$  in the AmyP-E221Q/ $\gamma$ -CD structure. Glucose residues in the oligosaccharide are numbered from the non-reducing end. His129 is from domain B (magenta). Hydrogen bonds are shown as dashed lines. (b) Electron density of maltose bound to the active site of WT AmyP $_{\Delta SBD}$  from the  $2Fo-Fc$  map contoured at  $2\sigma$  in the AmyP/ $\beta$ -CD structure. Glucose residues in the oligosaccharide are numbered from the non-reducing end. (c) Time-course analysis of the hydrolysis products of the  $\gamma$ -CD substrate produced by WT AmyP $_{\Delta SBD}$  and its Y36A mutant, respectively, using TLC. Std, standards.

for  $\beta$ -CD at the high concentration in the crystallization drop, thus producing maltotriose from  $\gamma$ -CD and maltose from  $\beta$ -CD, respectively.

Structural comparison of the active sites of AmyP $_{\Delta SBD}$  and some other homologs in the DALI database indicated the presence of a well-conserved catalytic triad, Asp178, Glu221 and Asp297, located at the center of the barrel (Fig. 1a). Stick superposition of the active site of AmyP $_{\Delta SBD}$ -E221Q complexed with maltotriose on the structures of GTA (Fig. S5a) and TASKA (Fig. S5b) indicated the remainder of the strictly conserved residues for subsites -1 and -2, including residues participating in hydrogen bonding and aromatic stacking interactions. Owing to its shortened B domain, AmyP $_{\Delta SBD}$  lacks residues corresponding to Trp170, Tyr182, Arg220 and His221 for subsites +1 and +2 in GTA. Trp223 of AmyP $_{\Delta SBD}$  adopts a conformation identical to that of the

corresponding residue in GTA, but differing from that of the corresponding residue in TASKA, which has been proposed to undergo a conformational change after substrate binding<sup>20</sup>. Additionally, in AmyP<sub>ΔSBD</sub>, Gln179 provides a hydrogen bond for subsite -1, and Leu84 provides hydrophobic interactions for subsite -2 (Figs 3a and S5).

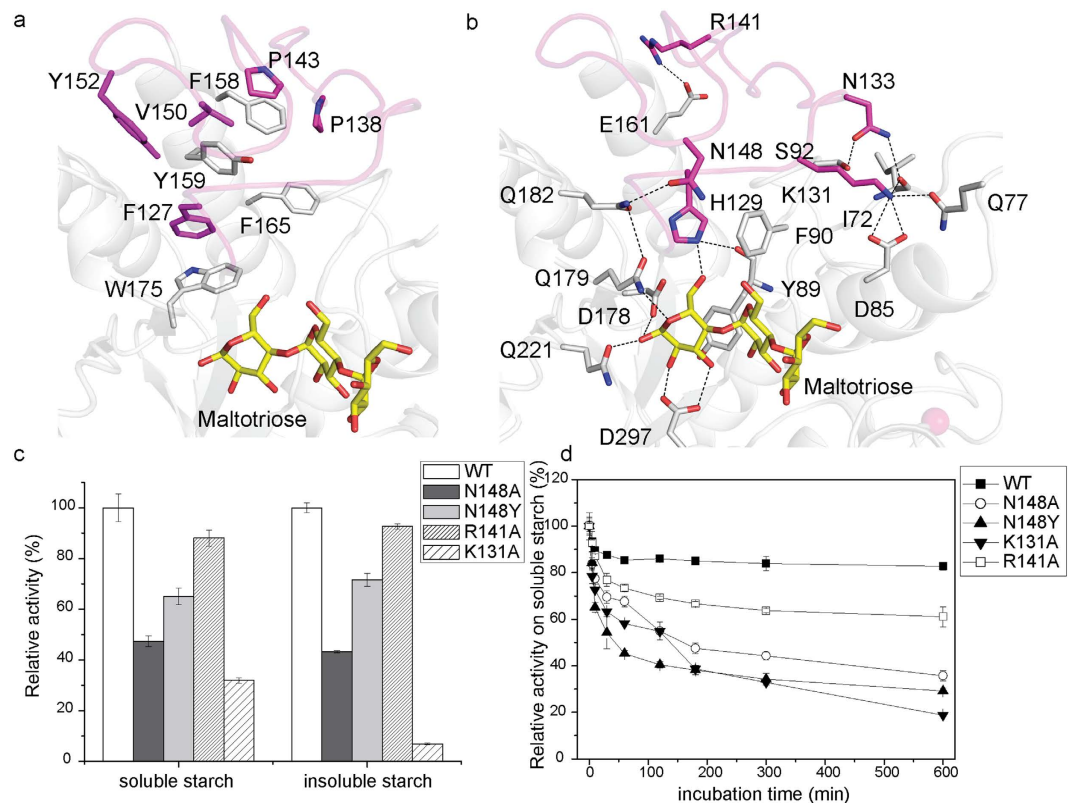
Interestingly, the -3 subsite of AmyP<sub>ΔSBD</sub>, which is formed by the hydrogen-bond-forming residues Tyr36, Asp367 and His368, is novel. Tyr36 also stacks against the sugar residue at subsite -3 (Figs 3a,b and S5). This tyrosine and other substrate binding residues of AmyP<sub>ΔSBD</sub> are all conserved in the GH13\_37 subfamily (Fig. S2). Thin-layer chromatography (TLC) of the reaction products showed that the Y36A mutant, compared with the WT, produced a lower ratio of maltotriose (G3) to maltose (G2) during the initial 10 min of hydrolysis with  $\gamma$ -CD as substrate (Fig. 3c). Additionally, the Y36A mutant produced some maltotetrose (G4), whereas the WT enzyme appeared to produce none of this product during the initial stage of hydrolysis (Fig. 3c). This finding may have occurred because the Y36A mutant, owing to sharply decreased interaction at subsite -3, was able to cleave G2/maltohexaose (G6) and G3/maltopentaose (G5) from  $\gamma$ -CD, and then G2/G4 and G2/G3 from G6 and G5, respectively. A new but weak substrate-binding site at subsite -4 may be created by the reorientation of nearby residues, such as Trp80, in the Y36A mutant (Fig. S5). This -4 subsite may temporarily protect G4 from rapid digestion. In contrast, the WT enzyme mainly produced G3/G5 from  $\gamma$ -CD, and then G3/G2 from G5, owing to the strong binding force at the -3 subsite. Thus, G3 is the major product of the WT enzyme during the early stage of hydrolysis. This finding confirms the critical role of Tyr36 in subsite -3 binding.

The catalytic and substrate-binding residues in the two structures of AmyP<sub>ΔSBD</sub> superimposed well. The maltose bound to the active pocket of the WT enzyme also superimposed well at subsites -2 and -3 with the maltotriose bound to the active site of the E221Q mutant (Fig. S6). The observed binding of maltose in subsites -2 and -3 indicated that the affinity of subsite -3 for a glucose residue is greater than that of subsite -1. It has previously been shown that AmyP degrades raw starch, releasing G1, G2 and G3, with complete loss of G4 over time<sup>9</sup>. It seems likely that a strong -3 subsite would protect G3 from rapid digestion, whereas a weak -1 subsite would limit the formation of glucans smaller than G3. To confirm this speculation, the hydrolysis products obtained from 10 mg/ml raw rice starch after incubation with the WT and Y36A mutant AmyP<sub>ΔSBD</sub> enzymes for various times were analyzed by high-performance liquid chromatography (HPLC). These results (Fig. S7) indicated more efficient production of G3 by the WT enzyme and more efficient production of G2 by the Y36A mutant during the early stage of hydrolysis of insoluble starch, which were generally consistent with the results obtained using  $\gamma$ -CD as substrate. The AmyP<sub>ΔSBD</sub>-E221Q/G3 and AmyP<sub>ΔSBD</sub>/G2 structures can therefore explain the nature of the limit digest of G1, G2 and G3 by the AmyP enzyme.

**Role of a short domain B.** In most GH13  $\alpha$ -amylases, the conserved calcium-binding site located between domains A and B is thought to significantly contribute to the protein's structural stability and to regulate its thermostability and enzymatic activity<sup>7,8,32</sup>. However, the B domain of AmyP is the smallest among those of GH13 subfamilies of  $\alpha$ -amylases, and it does not bind Ca<sup>2+</sup> or other metal ions. The crystal structures of AmyP<sub>ΔSBD</sub> show that several hydrophobic residues that are strictly conserved in the GH13\_37 subfamily (Fig. S2) cluster at the interface between domains A and B (Fig. 4a). The hydrophobic interactions mediated by these residues are likely to play critical roles in stabilizing the B domain and its interface with domain A. Hydrogen bonds are also involved in the inter-domain interactions within the enzyme (Fig. 4b).

As in other GH13  $\alpha$ -amylases, the small B domain of AmyP<sub>ΔSBD</sub> is situated adjacent to the active center of the enzyme. To determine the role of the B domain in enzyme catalysis, we generated four site-directed mutants of the B domain, K131A, N148A, N148Y and R141A, and compared their activities and stabilities to those of the WT enzyme (Fig. 4c). The far-UV CD and tryptophan intrinsic fluorescence spectra of all four mutants were essentially identical to those of the WT enzyme (Fig. S8), thus indicating that these mutations did not cause significant changes in the protein's overall conformation. The K131A mutation had the most profound effect on the enzyme's activity against soluble and insoluble starch substrates. Kinetic studies revealed that the  $K_m$  value of the K131A mutant for the  $\gamma$ -CD substrate was ten-fold higher than that of the WT enzyme (Table 1), thus indicating a decreased binding affinity for  $\gamma$ -CD. This mutant also showed a considerable decrease in turnover number ( $k_{cat}$ ) and catalytic efficiency ( $k_{cat}/K_m$ ). The K131A mutant showed markedly decreased activity against soluble starch at pH below 6.5, whereas the WT enzyme and the other mutants retained maximal hydrolytic activity in the pH range of 5.0 to 6.5 (Fig. S9). This phenomenon may have been because the replacement of Lys131 with an alanine significantly decreased the rigidity of the -1 subsite binding loop (residues 128–133) in domain B (Fig. 3a) because of the absence of the four hydrogen bonds normally formed by Lys131 with domain A (Fig. 4b); this change probably perturbs the proper conformation of the substrate-binding residue His129. It was noted that the N148A mutant displayed a greater decrease in activity against soluble and insoluble starch substrates than the N148Y mutant (Fig. 4c). Moreover, for the degradation of  $\gamma$ -CD, the N148A variant, compared with that of the N148Y mutant, exhibited a larger decrease in  $k_{cat}/K_m$  (Table 1). These results suggest that Asn148 lines one side of the entrance to the active site. Mutations at this residue disrupt the network of hydrogen bonds formed by Asn148, Gln182, Gln179 and the sugar residue at subsite -1 (Fig. 4b) and thus decrease the enzyme's catalytic efficiency. However, the substitution of Asn148 by Tyr might introduce extra stacking interactions of this residue with nearby residues such as His129, Phe90 and/or substrate sugar rings, thus disorganizing the architecture of the active-site residues to a lesser extent than occurs in the N148A mutation. The R141A mutant and WT AmyP<sub>ΔSBD</sub> displayed similar levels of activity against soluble and insoluble rice starch (Fig. 4c), possibly because Arg141, which is located opposite the catalytic region, had only a slight effect on catalysis.

Thermal inactivation studies showed that the activity of WT AmyP<sub>ΔSBD</sub> against both soluble (Fig. 4d) and insoluble starch substrates was more thermostable than that of the four mutants (Fig. S10), thus demonstrating the importance of the rigidity of domain B in the thermostability of the enzyme. Together, these results suggest that, despite its small size, the B domain of AmyP<sub>ΔSBD</sub> has a function similar to those of the B domains of



**Figure 4. The role of short domain B.** (a) Stick representation of hydrophobic residues clustered at the interface between domains A and B. (b) Stick representation of hydrogen-bond-forming residues between domains A and B. (c) Specific activities of the N148A, N148Y, R141A, and K131A variants of AmyP $_{\Delta\text{SBD}}$  relative to that of WT with soluble and insoluble rice starch as substrates. (d) Thermal inactivation of WT AmyP $_{\Delta\text{SBD}}$  and its N148A, N148Y, K131A, and R141A variants. The enzymes were incubated in the presence of 10 mM CaCl $_2$  at 40°C and pH 6.5 for the times indicated; enzyme activity against soluble rice starch was then assayed in the standard buffer.

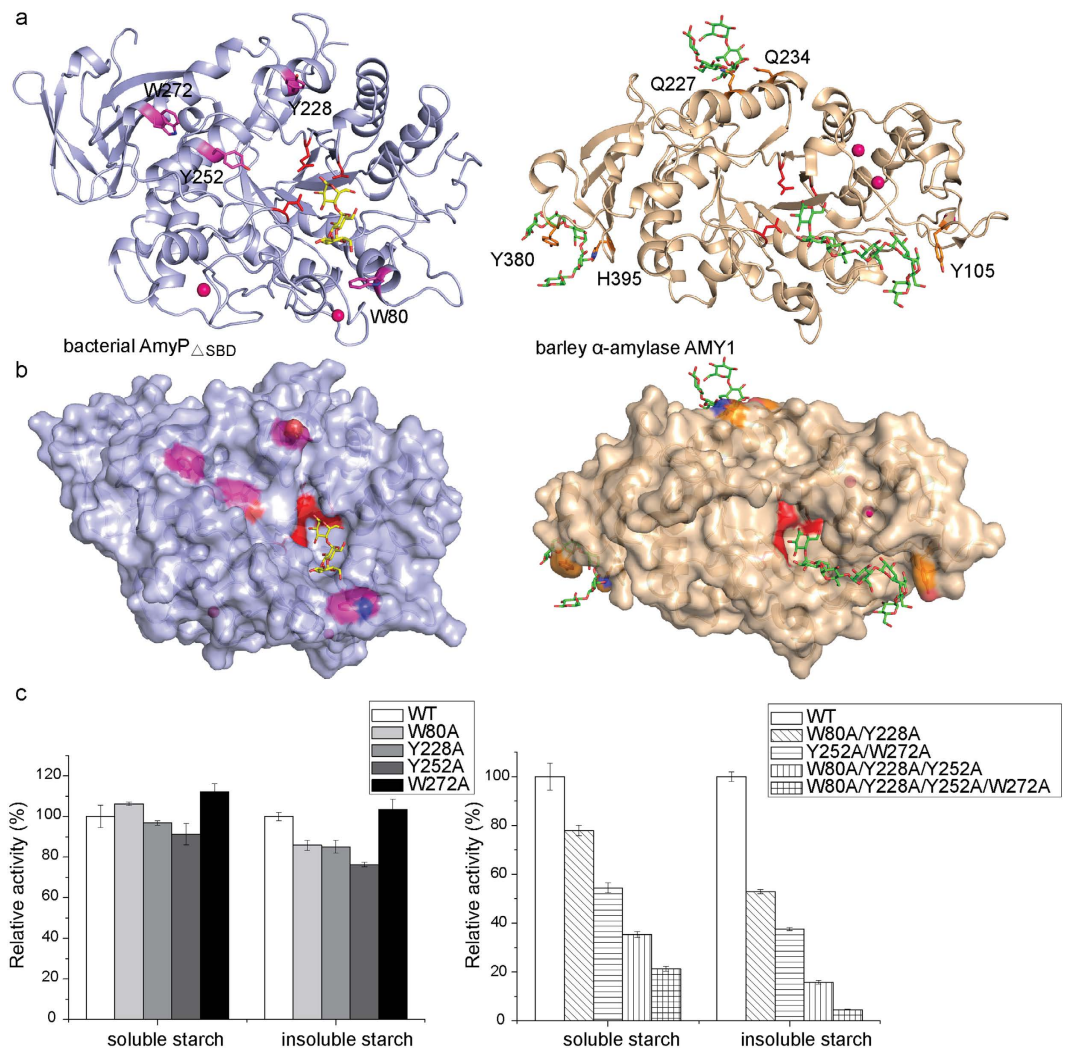
other GH13  $\alpha$ -amylases, that is, to reinforce the conformation of the active site and to enhance the enzyme's thermostability.

## Discussion

The crystal structures of AmyP $_{\Delta\text{SBD}}$  presented in this work provide the first insights into the overall structure, Ca $^{2+}$  binding sites, and substrate-binding subsites of  $\alpha$ -amylases belonging to the new subfamily GH13\_37. A key feature of these structures is the very short length of its B domain, a trait shared by GH13\_37 subfamily members (Fig. S2). To our knowledge, the B domain of AmyP $_{\Delta\text{SBD}}$  is the smallest B domain in presently known GH13  $\alpha$ -amylases. This B domain is formed exclusively by loop elements, does not bind metal ions, and helps to form a wide-open entrance to the catalytic cleft. In the case of AmyP $_{\Delta\text{SBD}}$ , hydrophobic interactions and hydrogen bonds play important roles in the association of domains A and B. Like B domains of other GH13  $\alpha$ -amylases, the B domain of AmyP $_{\Delta\text{SBD}}$ , which is located in proximity to the catalytic region, contributes to the thermostability and activity of the enzyme. Although a common calcium-binding site at the interface between the A and B domains is not present in AmyP $_{\Delta\text{SBD}}$ , two calcium-binding sites are found, one of which is unique and adjacent to the -3 subsite binding loop in domain A. The calcium binding in AmyP $_{\Delta\text{SBD}}$  significantly enhances the enzyme's thermostability. In this work, the resulting glucan in the crystal structures, however, was not the cyclodextrin, but a degradation product, G2 in the case of the WT enzyme and G3 in the case of the active-site mutant. The observation that G2 binds in subsites -2 and -3 indicates tighter binding in these subsites than in subsite -1, which is consistent with the production of a limit digest of G1, G2 and G3 by the enzyme. The novel Tyr36 of AmyP participates in both stacking and hydrogen-bonding interactions with the sugar residue at subsite -3. The importance of Tyr36 in the -3 subsite binding was confirmed by the significantly increased ratio of G2 to G3 produced by the Y36A mutant from both  $\gamma$ -CD and insoluble starch substrates during the early stage of hydrolysis. The information presented here on the new GH13 subfamily may aid in engineering the thermostability, calcium requirements and product composition of  $\alpha$ -amylases for industrial applications.

Here, we described the crystal structure of a bacterial RSDE. RSDEs may exhibit similar structural properties that are related to their ability to digest raw starch. The structural comparison of AmyP $_{\Delta\text{SBD}}$  and GTA reveals that AmyP $_{\Delta\text{SBD}}$  lacks a structural lid above the active site, owing to the short length of its B domain. GTA may also be a RSDE, because it has an identical protein sequence to that of Pizzo extracellular  $\alpha$ -amylase, which can





**Figure 5. The roles of surface aromatic residues.** (a) Stick representation of the surface aromatic residues (magenta) of AmyP $\Delta$ SBD (blue on the left) and the surface aromatic residues (orange) of barley  $\alpha$ -amylase AMY1 (pdb accession code 1RP8, light orange on the right). Saccharide bound in the active site cleft of AmyP $\Delta$ SBD is shown as yellow sticks. Saccharides bound in the active site cleft and at surface sites of AMY1 are represented by green sticks. In both structures, the catalytic residues are shown in red and calcium ions are shown as hot pink spheres. (b) Surface representation of AmyP $\Delta$ SBD (left) and barley  $\alpha$ -amylase AMY1 (right). (c) Comparison of the specific activities of WT AmyP $\Delta$ SBD, its W80A, Y228A, Y252A, and W272A single point mutants (left) and its W80A/Y228A, Y252A/W272A, W80A/Y228A/Y252A, and W80A/Y228A/Y252A/W272A multiple point variants (right) using soluble rice starch and insoluble rice starch, respectively, as substrates.

adsorb onto and hydrolyze raw starch<sup>21,22</sup>. However, GTA does not possess a C-terminal domain rich in aromatic residues, whereas Pizzo  $\alpha$ -amylase has been studied in the presence of this extended region. It has been reported that the B domain of GTA is one of the smallest B domains found in  $\alpha$ -amylases<sup>21</sup>. In contrast, the domain B of AmyA has an elongated architecture, which creates a structural lid above the active site over a wider area and probably restricts substrate access to a larger extent (Figs S2 and S11). Structural comparison of AmyP $\Delta$ SBD and the well-studied RSDE barley  $\alpha$ -amylase AMY1 (pdb accession code 1RP8<sup>16</sup>) shows that both of their B domains point away from the catalytic site, forming a much wider cleft allowing the substrate to enter (Fig. 5). Therefore, enhanced accessibility of the catalytic cleft may be necessary for the enzyme to degrade raw starch granules.

Aromatic residues are central in protein-carbohydrate interactions. Because AmyP $\Delta$ SBD has several aromatic residues exposed on its surface at a certain distance from the active site, we hypothesized that these residues might enable stacking interactions with glucose motifs and enhance the enzyme's activity against starch. To test this hypothesis, four surface aromatic residues (Fig. 5) were mutated to alanine to generate eight mutants: four single mutants (W80A, Y228A, Y252A and W272A), two double mutants (W80A/W228A, Y252A/W272A), one triple mutant (W80A/Y228A/Y252A) and one quadruple mutant (W80A/Y228A/Y252A/W272A). The aromatic side chains of Tyr228/Trp272 and Trp80/Tyr252 are perpendicular and parallel to the protein surface, respectively (Fig. 5). In addition, Trp80, Tyr252 and Trp272 are conserved in the GH13\_37 subfamily (Fig. S2). The far-UV CD and tryptophan intrinsic fluorescence spectra of the mutant proteins indicated that the overall conformation



of AmyP $_{\Delta\text{SBD}}$  was still retained in all mutants (Fig. S12). It appeared that a single mutation did not reduce the enzyme's specific activity against soluble or insoluble rice starch to any significant extent. Nevertheless, multiple point mutants at surface aromatic residues resulted in pronounced decreases in starch hydrolytic activity. The quadruple mutant was the most significantly altered; its activities against soluble and insoluble rice starch were approximately 20% and 5%, respectively, of those of the WT enzyme. These results indicated that these exposed aromatic residues may together facilitate the attachment of the enzyme to glucan substrates; the resulting absence of saccharide-binding ability at these secondary binding sites in the quadruple mutant resulted in a significant reduction in the enzyme activity. A similar synergistic effect has been observed in other RSDEs such as AMY1<sup>15,33</sup> and human salivary  $\alpha$ -amylase<sup>19</sup>. However, this feature may not be as obvious in AmyP $_{\Delta\text{SBD}}$  as it is in AMY1, because saccharide binding at SBSs is directly seen in the crystal structures of AMY1<sup>16,17</sup>, but not in those of AmyP $_{\Delta\text{SBD}}$ . This feature of AmyP $_{\Delta\text{SBD}}$  is consistent with its weak binding to raw starch granules. In general, RSDEs often possess extra starch-binding regions formed by surface aromatic residues that enhance their enzymatic activity.

The attachment of AmyP to starch granules is mainly ascribed to its C-terminal SBD domain. This SBD belongs to a new carbohydrate-binding-module family, CBM69. Further biochemical work is required to determine the structure-function relationship of this novel SBD. Two cysteine residues flank the N and C termini of the SBD and both of these residues are highly conserved in the GH13\_37 subfamily. These two cysteines may be involved in a disulfide bridge that enhances the conformational stability of the SBD. Improvement of the *E. coli* expression of the SBD domain of AmyP to facilitate correct protein folding is currently being investigated by our group.

## Methods

**Cloning, expression and purification.** The pET32a-amyP plasmid encoding full-length AmyP (GenBank accession number HM572234) was used as a PCR template. DNA fragments corresponding to full-length AmyP (residues 1–639), the SBD-truncated mutant AmyP $_{\Delta\text{SBD}}$  (residues 1–496) and the SBD (residues 491–639) were amplified by PCR and cloned into a modified pET-28a(+) vector in which the thrombin protease site was substituted for the tobacco etch virus (TEV) cleavage site. All of the mutants were generated with a Mutant Best kit (Takara). All clones were verified by DNA sequencing.

Recombinant AmyP and AmyP $_{\Delta\text{SBD}}$  were transformed into *E. coli* BL21 (DE3) cells. The transformed cells were grown at 37 °C to an OD<sub>600</sub> of approximately 0.8 before induction with 0.2 mM isopropyl- $\beta$ -D-thiogalactopyranoside. The culture was incubated at 16 °C for an additional 24 h before the cells were harvested and lysed by sonication.

Full-length AmyP and its E221Q mutant derivative were purified for crystallization using a HisTrap column (GE Healthcare), treated with TEV to remove the 6  $\times$  His tag at the N-terminus and was further purified on a HisTrap column and then subjected to size-exclusion chromatography on a HiLoad 16/60 Superdex 200 column (GE Healthcare) and anion-exchange chromatography on a MonoQ HR 5/5 column (GE Healthcare). The purified proteins were dialyzed against buffer A (20 mM Tris-HCl [pH 8.0], 0.5 M NaCl) and concentrated. AmyP $_{\Delta\text{SBD}}$  and all mutant proteins used in enzyme activity assays, as well as the SBD of AmyP used in CD spectrum examination, were purified using a HisTrap column, and subsequent on a HiLoad 16/60 Superdex 200 column equilibrated with buffer A.

**Crystallization, data collection and structure determination.** All crystals were grown using the sitting drop vapor diffusion method at 10 °C. Crystal screen I and II kits (Hampton Research) were used to identify initial crystallization conditions. The conditions were further optimized to obtain good diffraction-quality crystals. Crystals of the E221Q mutant of full-length AmyP co-crystallized with  $\gamma$ -CD (AmyP-E221Q/ $\gamma$ -CD) were obtained by mixing equal volumes of a 5 mg/ml protein sample pre-incubated with 10 mM  $\gamma$ -CD and a reservoir solution containing 0.1 M sodium citrate [pH 6.0], 22.5% (v/v) 2-propanol, and 22.5% (w/v) polyethylene glycol (PEG) 4000. Crystals of WT full-length AmyP co-crystallized with  $\beta$ -CD (AmyP/ $\beta$ -CD) were grown using a 7 mg/ml protein solution pre-incubated with 4 mM  $\beta$ -CD and then diluted 1:1 with reservoir solution containing 0.1 M Tris-HCl [pH 8.5], 0.2 M sodium acetate, and 30% (w/v) PEG 4000.

All crystals used for diffraction data collection were quickly transferred to a cryoprotectant solution consisting of 25% glycerol (v/v) in reservoir solution, and were flash-frozen in liquid nitrogen. The data sets of AmyP-E221Q/ $\gamma$ -CD and AmyP/ $\beta$ -CD were collected on beam line 17U (BL17U) at the Shanghai Synchrotron Radiation Facility (SSRF).

The data sets were indexed, integrated and scaled with HKL2000<sup>34</sup>. The AmyP-E221Q/ $\gamma$ -CD structure was solved by molecular replacement with Phaser<sup>35</sup> in the CCP4<sup>36</sup> software suite by using the *B. stearothermophilus* neopullulanase (pdb accession code 1J0H) as a search model, autobuilt with ArpWarp<sup>37</sup> and refined and manually built using Refmac5<sup>38</sup> and COOT<sup>39</sup>. The AmyP/ $\beta$ -CD structure was then solved using AmyP-E221Q/ $\gamma$ -CD as a model. The crystal diffraction data and refinement statistics are listed in Table 2.

**Enzyme activity assays.** The starch-hydrolyzing activities of the WT and E221Q mutant AmyP enzymes were measured at 40 °C for 5 min in buffer B (20 mM HEPES [pH 6.5], 150 mM NaCl, 10 mM CaCl<sub>2</sub>) using 10 mg/ml soluble starch as substrate, which was freshly prepared by dissolving soluble starch powder in buffer B with brief heating and then cooling to room temperature. The enzyme reactions were initiated by the addition of 0.05 ml of 5  $\mu$ g/ml WT or 1 mg/ml E221Q mutant AmyP enzyme to 0.55 ml of pre-heated soluble starch solution. The starch-hydrolyzing activities of the WT and mutant AmyP $_{\Delta\text{SBD}}$  enzymes were measured at 40 °C for 5 min in buffer B using 10 mg/ml soluble starch and 10 mg/ml insoluble rice starch as substrates. Raw rice starch granules were washed three times with ultrapure water followed by buffer B, suspended in buffer B with shaking, and used immediately. Reactions were initiated by the addition of 0.05 ml of 2  $\mu$ g/ml WT or 2–10  $\mu$ g/ml mutant AmyP $_{\Delta\text{SBD}}$

Structure	AmyP/ $\beta$ -CD	AmyP-E221Q/ $\gamma$ -CD
Wavelength (Å)	0.9793	0.9793
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub>
Cell parameters		
a, b, c (Å)	79.87, 130.53, 217.53	79.87, 130.53, 217.53
$\alpha$ , $\beta$ , $\gamma$ (°)	90.00, 90.00, 90.00	90.00, 93.14, 90.00
Resolution (Å)	40.00–1.95 (2.02–1.95) <sup>a</sup>	40.00–2.55 (2.64–2.55)
$R_{\text{merge}}$ (%)	10.9(47.4)	12.1(64.2)
I/ $\sigma$ I	25.5(4.0)	11.78(2.0)
Completeness (%)	87.6(85.0)	99.7(99.9)
Redundancy	5.0(5.2)	3.5(3.5)
Refinement		
No.reflections (overall)	144717	95368
No. reflections (test)	7292	4716
$R_{\text{work}}/R_{\text{free}}$ (%)	16.35/21.23	17.02/22.00
No. non-H atoms		
Protein	15550	15498
Sugar	92	136
Ca1/Ca2	4/4	4/4
Water	2177	431
B factors (Å <sup>2</sup> )		
Protein	13.96	37.20
Sugar	14.57	33.49
Ca1/Ca2	28.31/10.89	59.98/57.93
Water	22.05	36.34
R.m.s. deviations		
Bond lengths (Å)	0.006	0.008
Bond angles (°)	0.828	0.933
Rampage plot% residues		
Favored	97.28	96.32
Allowed	2.52	3.33
Outliers	0.20	0.35

**Table 2. Data collection and refinement statistics for AmyP/ $\beta$ -CD and AmyP-E221Q/ $\gamma$ -CD.** <sup>a</sup>Values in parentheses are for the highest-resolution shell.

enzyme to 0.55 ml pre-heated soluble starch solution, or by the addition of 0.05 ml of 20  $\mu$ g/ml WT or 20–100  $\mu$ g/ml mutant AmyP <sub>$\Delta$ SBD</sub> enzyme to 0.55 ml raw rice starch solution. The reducing sugars produced in the reaction were quantified using the 3,5-dinitrosalicylic acid (DNS) method<sup>40</sup>. The absorbance at 540 nm was measured and the amount of reducing sugar present was determined on the basis of a standard curve using glucose. The enzyme concentration was determined using the Bradford method. All experiments were performed in triplicate. The average and standard deviation values were calculated; the specific activity of the WT enzyme was taken as 100% for determining relative activities.

**Enzyme kinetics.** Reactions for assessing kinetics were conducted at 40 °C in buffer B using  $\gamma$ -CD as substrate at nine different concentrations. Reactions were initiated by the addition of 0.05 ml of 1  $\mu$ g/ml WT or 1–5  $\mu$ g/ml mutant AmyP <sub>$\Delta$ SBD</sub> enzyme to 0.45 ml  $\gamma$ -CD solution. Reducing sugar quantities were determined using the copper-bicinchoninate (CuBic) method<sup>41</sup>. The absorbance at 570 nm was measured; and the corresponding amount of reducing sugar was determined via a standard curve using glucose. The initial velocity of the reaction was determined by plotting absorbance versus time for each substrate concentration. The kinetic parameters  $k_{\text{cat}}$ ,  $K_{\text{m}}$ , and  $k_{\text{cat}}/K_{\text{m}}$  were calculated by fitting initial velocities versus substrate concentrations to the Michaelis-Menten equation using nonlinear regression with Origin.

**Enzyme thermal inactivation.** The effect of Ca<sup>2+</sup> ions on the thermal inactivation of the enzyme was evaluated by incubating WT AmyP <sub>$\Delta$ SBD</sub> at 40 °C in 20 mM HEPES [pH 6.5] and 150 mM NaCl with or without 10 mM CaCl<sub>2</sub>. Comparison of the thermal inactivation of WT and mutant AmyP <sub>$\Delta$ SBD</sub> was conducted by incubating the enzymes at 40 °C in buffer B; the residual activity was then determined at intervals with the DNS method using 10 mg/ml soluble starch and 10 mg/ml insoluble starch as substrates.

**TLC of hydrolysis products.** Hydrolysis products were detected by TLC after incubation of the WT and Y36A mutant AmyP <sub>$\Delta$ SBD</sub> enzymes with the  $\gamma$ -CD substrate for various times. Activity assays were performed at 40 °C in buffer C (20 mM HEPES [pH 6.5], 50 mM NaCl, 10 mM CaCl<sub>2</sub>). Each reaction contained 0.1 ml of

3 mg/ml enzyme and 0.9 ml of 10 mg/ml  $\gamma$ -CD solution. Three microliters of each reaction was blotted onto a 20 cm  $\times$  20 cm, 500- $\mu$ m-thick silica gel 60 A<sup>o</sup> plate (Qingdao Haiyang Chemical Co. Ltd., China). The plate was dried and placed in a chamber containing a mixture of isopropanol:ethyl acetate:water (3:1:1, v/v/v). The plate was dried, and the contents were visualized by dipping the plate into a solution containing 0.3% N-(1-naphthyl)-ethylenediamine and 5% H<sub>2</sub>SO<sub>4</sub> in methanol and heating for 10 min at 110 °C as described by Shim et al.<sup>42</sup>. Standards (1  $\mu$ l each of a 10 mg/ml solution) were also blotted on the plate for comparison and determination of the reaction products.

**HPLC of hydrolysis products.** The hydrolysis products against raw rice starch after incubation with the WT and Y36A mutant AmyP <sub>$\Delta$ SBD</sub> enzymes for various times were identified with HPLC. Activity assays were performed at 40 °C in buffer C. Each reaction was initiated by the addition of 0.1 ml of 1 mg/ml enzyme to 0.9 ml of 10 mg/ml raw rice starch solution and was terminated by heating at 100 °C for 10 min. The mixtures were then centrifuged at 10,000  $\times$  g for 10 min. Each supernatant was diluted 10-fold with ultrapure water, and 10  $\mu$ l of the diluted supernatant was applied to an HPLC system (Agilent 1260 Series). The column (TSK-GEL amide-80 column (4.6  $\times$  250 mm; Tosoh, Tokyo, Japan) was equilibrated with 55% acetonitrile. The components were separated in 55% acetonitrile at a flow rate of 1 ml/min at room temperature. The yields of sugar molecules were analyzed using an evaporative light scattering detector incorporated into the HPLC system. Glucose, maltose, maltotriose and maltotetraose (20  $\mu$ l each of a 5 mM solution) were used as standards.

**Far-UV CD spectra.** Far-UV CD spectra were recorded at 25 °C with a MOS-500 spectropolarimeter, at wavelengths between 200 and 250 nm, using a 0.1-cm path length cell and approximately 0.2 mg/ml protein sample in buffer D (0.15 M NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub> [pH 7.5], 50 mM NaCl). A buffer only reference was subtracted from each curve.

**Intrinsic fluorescence spectra.** Intrinsic fluorescence measurements of proteins were recorded at 25 °C with a F-4500 spectrofluorophotometer in a 0.1-cm path length quartz cell using approximately 0.2 mg/ml protein sample in buffer A. Fluorescence emission spectra were recorded from 300 nm to 400 nm using a maximum excitation wavelength of 288 nm.

## References

- Janecek, S., Svensson, B. & MacGregor, E. A. alpha-Amylase: an enzyme specificity found in various families of glycoside hydrolases. *Cellular and molecular life sciences: CMLS* **71**, 1149–1170, doi: 10.1007/s00018-013-1388-z (2014).
- Liu, Y. et al. Identification and phylogenetic characterization of a new subfamily of alpha-amylase enzymes from marine microorganisms. *Mar Biotechnol (NY)* **14**, 253–260, doi: 10.1007/s10126-011-9414-3 (2012).
- Peng, H. et al. A starch-binding domain identified in alpha-amylase (AmyP) represents a new family of carbohydrate-binding modules that contribute to enzymatic hydrolysis of soluble starch. *FEBS letters* **588**, 1161–1167, doi: 10.1016/j.febslet.2014.02.050 (2014).
- Khemakhem, B. et al. The importance of an extra loop in the B-domain of an alpha-amylase from *B. stearothermophilus* US100. *Biochemical and biophysical research communications* **385**, 78–83, doi: 10.1016/j.bbrc.2009.04.137 (2009).
- Machius, M., Declerck, N., Huber, R. & Wiegand, G. Activation of *Bacillus licheniformis* alpha-amylase through a disorder-to-order transition of the substrate-binding site mediated by a calcium-sodium-calcium metal triad. *Structure* **6**, 281–292 (1998).
- Igarashi, K. et al. Improved thermostability of a *Bacillus* alpha-amylase by deletion of an arginine-glycine residue is caused by enhanced calcium binding. *Biochemical and biophysical research communications* **248**, 372–377, doi: 10.1006/bbrc.1998.8970 (1998).
- Priyadharshini, R. & Gunasekaran, P. Site-directed mutagenesis of the calcium-binding site of alpha-amylase of *Bacillus licheniformis*. *Biotechnology letters* **29**, 1493–1499, doi: 10.1007/s10529-007-9428-0 (2007).
- Ghollasi, M., Khajeh, K., Naderi-Manesh, H. & Ghasemi, A. Engineering of a *Bacillus* alpha-amylase with improved thermostability and calcium independency. *Applied biochemistry and biotechnology* **162**, 444–459, doi: 10.1007/s12010-009-8879-2 (2010).
- Lei, Y. et al. Preferential and rapid degradation of raw rice starch by an alpha-amylase of glycoside hydrolase subfamily GH13\_37. *Applied microbiology and biotechnology* **94**, 1577–1584, doi: 10.1007/s00253-012-4114-0 (2012).
- Sun, H. et al. Recent advances in microbial raw starch degrading enzymes. *Applied biochemistry and biotechnology* **160**, 988–1003, doi: 10.1007/s12010-009-8579-y (2010).
- Machovic, M. & Janecek, S. Starch-binding domains in the post-genome era. *Cellular and molecular life sciences: CMLS* **63**, 2710–2724, doi: 10.1007/s00018-006-6246-9 (2006).
- Rodriguez-Sanoja, R., Oviedo, N. & Sanchez, S. Microbial starch-binding domain. *Current opinion in microbiology* **8**, 260–267, doi: 10.1016/j.mib.2005.04.013 (2005).
- Southall, S. M., Simpson, P. J., Gilbert, H. J., Williamson, G. & Williamson, M. P. The starch-binding domain from glucoamylase disrupts the structure of starch. *FEBS letters* **447**, 58–60 (1999).
- Robert, X. et al. The structure of barley alpha-amylase isozyme 1 reveals a novel role of domain C in substrate recognition and binding: a pair of sugar tongs. *Structure* **11**, 973–984 (2003).
- Nielsen, M. M. et al. Two secondary carbohydrate binding sites on the surface of barley alpha-amylase 1 have distinct functions and display synergy in hydrolysis of starch granules. *Biochemistry* **48**, 7686–7697, doi: 10.1021/bi900795a (2009).
- Robert, X., Haser, R., Mori, H., Svensson, B. & Aghajari, N. Oligosaccharide binding to barley alpha-amylase 1. *The Journal of biological chemistry* **280**, 32968–32978, doi: 10.1074/jbc.M505515200 (2005).
- Bozonnet, S. et al. The 'pair of sugar tongs' site on the non-catalytic domain C of barley alpha-amylase participates in substrate binding and activity. *The FEBS journal* **274**, 5055–5067, doi: 10.1111/j.1742-4658.2007.06024.x (2007).
- Sevcik, J. et al. Structure of the complex of a yeast glucoamylase with acarbose reveals the presence of a raw starch binding site on the catalytic domain. *The FEBS journal* **273**, 2161–2171, doi: 10.1111/j.1742-4658.2006.05230.x (2006).
- Raganath, C. et al. Probing the role of aromatic residues at the secondary saccharide-binding sites of human salivary alpha-amylase in substrate hydrolysis and bacterial binding. *Journal of molecular biology* **384**, 1232–1248, doi: 10.1016/j.jmb.2008.09.089 (2008).
- Chai, K. P. et al. Crystal structure of *Anoxybacillus* alpha-amylase provides insights into maltose binding of a new glycosyl hydrolase subclass. *Scientific reports* **6**, 23126, doi: 10.1038/srep23126 (2016).
- Mok, S. C., Teh, A. H., Saito, J. A., Najimudin, N. & Alam, M. Crystal structure of a compact alpha-amylase from *Geobacillus thermoleovorans*. *Enzyme and microbial technology* **53**, 46–54, doi: 10.1016/j.enzmictec.2013.03.009 (2013).
- Finore, I. et al. Purification, biochemical characterization and gene sequencing of a thermostable raw starch digesting  $\alpha$ -amylase from *Geobacillus thermoleovorans* subsp. stromboliensis subsp. nov. *World Journal of Microbiology and Biotechnology* **27**, 2425–2433, doi: 10.1007/s11274-011-0715-5 (2011).

23. Mehta, D. & Satyanarayana, T. Domain C of thermostable alpha-amylase of *Geobacillus thermoleovorans* mediates raw starch adsorption. *Applied microbiology and biotechnology* **98**, 4503–4519, doi: 10.1007/s00253-013-5459-8 (2014).
24. Puspasari, F. *et al.* Raw starch-degrading alpha-amylase from *Bacillus aquimaris* MKSC 6.2: isolation and expression of the gene, bioinformatics and biochemical characterization of the recombinant enzyme. *Journal of applied microbiology* **114**, 108–120, doi: 10.1111/jam.12025 (2013).
25. Ranjani, V. *et al.* Protein engineering of selected residues from conserved sequence regions of a novel *Anoxybacillus* alpha-amylase. *Scientific reports* **4**, 5850, doi: 10.1038/srep05850 (2014).
26. Peng, H. *et al.* Identification and characterization of a novel raw-starch-degrading a-amylase (AmyASS) from the marine fish pathogen *Aeromonas salmonicida* ssp *salmonicida*. *J Mol Catal B-Enzym* **119**, 71–77, doi: 10.1016/j.molcatb.2015.06.005 (2015).
27. Hondoh, H., Kuriki, T. & Matsuura, Y. Three-dimensional structure and substrate binding of *Bacillus stearothermophilus* neopullulanase. *Journal of molecular biology* **326**, 177–188 (2003).
28. Siddiqui, K. S. *et al.* Role of disulfide bridges in the activity and stability of a cold-active alpha-amylase. *Journal of bacteriology* **187**, 6206–6212, doi: 10.1128/JB.187.17.6206-6212.2005 (2005).
29. Aghajari, N., Feller, G., Gerday, C. & Haser, R. Crystal structures of the psychrophilic alpha-amylase from *Alteromonas haloplanctis* in its native form and complexed with an inhibitor. *Protein science: a publication of the Protein Society* **7**, 564–572, doi: 10.1002/pro.5560070304 (1998).
30. Lee, H. S. *et al.* Cyclomaltodextrinase, neopullulanase, and maltogenic amylase are nearly indistinguishable from each other. *The Journal of biological chemistry* **277**, 21891–21897, doi: 10.1074/jbc.M201623200 (2002).
31. Sivakumar, N., Li, N., Tang, J. W., Patel, B. K. & Swaminathan, K. Crystal structure of AmyA lacks acidic surface and provide insights into protein stability at poly-extreme condition. *FEBS letters* **580**, 2646–2652, doi: 10.1016/j.febslet.2006.04.017 (2006).
32. Liu, Y., Shen, W., Shi, G. Y. & Wang, Z. X. Role of the calcium-binding residues Asp231, Asp233, and Asp438 in alpha-amylase of *Bacillus amyloliquefaciens* as revealed by mutational analysis. *Current microbiology* **60**, 162–166, doi: 10.1007/s00284-009-9517-5 (2010).
33. Cockburn, D. *et al.* Surface binding sites in amylase have distinct roles in recognition of starch structure motifs and degradation. *International journal of biological macromolecules* **75**, 338–345, doi: 10.1016/j.ijbiomac.2015.01.054 (2015).
34. Otwinowski, Z. & Minor, W. In *Methods in Enzymology* Vol. Volume 276, 307–326 (Academic Press, 1997).
35. McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of applied crystallography* **40**, 658–674, doi: 10.1107/S0021889807021206 (2007).
36. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* **67**, 235–242, doi: 10.1107/S0907444910045749 (2011).
37. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* **3**, 1171–1179, doi: 10.1038/nprot.2008.91 (2008).
38. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D* **67**, 355–367, doi: 10.1107/S090744491001314 (2011).
39. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126–2132, doi: 10.1107/S0907444904019158 (2004).
40. Miller, G. L. Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Analytical Chemistry* **31**, 426–428, doi: 10.1021/ac60147a030 (1959).
41. McIntyre, A. P., Mukerjee, R. & Robyt, J. F. Reducing values: dinitrosalicylate gives over-oxidation and invalid results whereas copper bichinchonate gives no over-oxidation and valid results. *Carbohydrate research* **380**, 118–123, doi: 10.1016/j.carres.2013.06.008 (2013).
42. Shim, J. H. *et al.* Role of maltogenic amylase and pullulanase in maltodextrin and glycogen metabolism of *Bacillus subtilis* 168. *Journal of bacteriology* **191**, 4835–4844, doi: 10.1128/JB.00176-09 (2009).

## Acknowledgements

We thank Dr. Yiyang Jiang, Hongyu Bao and the staff at BL17U of Shanghai Synchrotron Radiation Facilities for their kind help in the collection and processing of the X-ray data. This work was supported by grants from the National Natural Science Foundation of China (No. 31400643), the National Natural Science Foundation of China (No. 31470207), the Natural Science Foundation of Anhui Province (1408085QC49), the Key Research Program of the Education Department of Anhui Province (KJ2014A008), and the Introduction Project of Academic and Technology Leaders in Anhui University (01001770).

## Author Contributions

C.H. and Y.L. conceived and designed the experiments. Y.L., J.Y., F.L., and C.H. performed the experiments. C.H., Y.L., F.L., and X.Z. analyzed the data. H.P, Y.X., and C.H. contributed reagents/materials/analysis tools. C.H., Y.L., and F.L. prepared figures and tables. C.H. and Y.L. wrote the paper.

## Additional Information

**Accession codes:** The atomic coordinates and structure factors for the E221Q mutant of full-length AmyP co-crystallized with  $\gamma$ -CD (AmyP-E221Q/ $\gamma$ -CD) and full-length AmyP co-crystallized with  $\beta$ -CD (AmyP/ $\beta$ -CD) have been deposited in the Protein Data Bank under accession codes 5H05 and 5H06, respectively.

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing Interests:** The authors declare no competing financial interests.

**How to cite this article:** Liu, Y. *et al.* Crystal structure of a raw-starch-degrading bacterial  $\alpha$ -amylase belonging to subfamily 37 of the glycoside hydrolase family GH13. *Sci. Rep.* **7**, 44067; doi: 10.1038/srep44067 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>