SCIENTIFIC REPORTS

Received: 31 August 2016 Accepted: 31 January 2017 Published: 13 March 2017

OPEN A new group of glycoside hydrolase family 13 lpha-amylases with an aberrant catalytic triad

Fean D. Sarian^{1,2}, Štefan Janeček^{3,4}, Tjaard Pijning⁵, Ihsanawati¹, Zeily Nurachman¹, Ocky K. Radjasa⁶, Lubbert Dijkhuizen⁷, Dessy Natalia¹ & Marc J. E. C. van der Maarel²

 α -Amylases are glycoside hydrolase enzymes that act on the α (1 \rightarrow 4) glycosidic linkages in glycogen, starch, and related α -glucans, and are ubiquitously present in Nature. Most α -amylases have been classified in glycoside hydrolase family 13 with a typical $(\beta/\alpha)_8$ -barrel containing two aspartic acid and one glutamic acid residue that play an essential role in catalysis. An atypical α -amylase (BmaN1) with only two of the three invariant catalytic residues present was isolated from Bacillus megaterium strain NL3, a bacterial isolate from a sea anemone of Kakaban landlocked marine lake, Derawan Island, Indonesia. In BmaN1 the third residue, the aspartic acid that acts as the transition state stabilizer, was replaced by a histidine. Three-dimensional structure modeling of the BmaN1 amino acid sequence confirmed the aberrant catalytic triad. Glucose and maltose were found as products of the action of the novel α -amylase on soluble starch, demonstrating that it is active in spite of the peculiar catalytic triad. This novel BmaN1 α -amylase is part of a group of α -amylases that all have this atypical catalytic triad, consisting of aspartic acid, glutamic acid and histidine. Phylogenetic analysis showed that this group of α -amylases comprises a new subfamily of the glycoside hydrolase family 13.

 α -Amylases are ubiquitously present in nature. They act on the $\alpha(1 \rightarrow 4)$ glycosidic linkages in glycogen, starch, and related α -glucans and thereby play an important role in the digestion of starch in humans, plants and microorganisms¹⁻³. Most α -amylases belong to glycoside hydrolase (GH) family 13⁴, constituting 30 different reaction and product specificities including, glycoside hydrolases (EC 3.2.1.x), glucanotransferases (EC 2.4.1.x and EC 2.4.99.16), and isomerases (EC 5.4.99.x), all sharing a conserved structural scaffold^{4.5}. The crystal structure of Taka α -amylase A from Aspergillus oryzae (TAA), the first experimentally determined three-dimensional (3D) structure of α -amylase⁶, revealed that α -amylases have three characteristic domains: A, B, and C^{6,7}. The A domain containing the catalytic residues is the most conserved domain, with a typical $(\beta/\alpha)_{8^{-}}$ or TIM-barrel comprised of eight stranded parallel β -sheet surrounded by eight α -helices. Domain B is inserted between the third β -strand and the third α -helix of the $(\beta/\alpha)_8$ -barrel and varies in length and structure. The C domain folds into eight antiparallel β -strands, is connected to the A domain by loops and seems to be an independent domain with unknown function^{5,8}. Despite low similarity between the amino acid sequences of α -amylases from animals, plants, and microorganisms, the GH13 enzymes share seven highly conserved regions⁷ that are involved in the formation of the catalytic site. The α -amylase active site is located in an open cavity between the A and B domains, and contains the invariably carboxylic acid Asp206, Glu230 and Asp297 (TAA numbering) being essential for catalysis, acting as the nucleophile, and as the general acid/base and transition state stabilizer, respectively⁶.

¹Biochemistry Research Division, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Jl. Ganesha 10, Bandung, 40132, Indonesia. ²Aguatic Biotechnology and Bioproduct Engineering, Engineering and Technology institute Groningen (ENTEG), University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands. ³Laboratory of Protein Evolution, Institute of Molecular Biology, Slovak Academy of Sciences, Dubravska cesta 21, SK-84551 Bratislava, Slovakia. ⁴Department of Biology, Faculty of Natural Sciences, University of SS. Cyril and Mehtodius, Nam. J. Herdu 2, SK-91701 Trnava, Slovakia. ⁵X-Ray Crystallography, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands. ⁶Department of Marine Science, Faculty of Fishery and Marine Science, Diponegoro University, Semarang 50275, Central Java, Indonesia. ⁷Microbial Physiology, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands. Correspondence and requests for materials should be addressed to M.J.E.C.v.d.M. (email: m.j.e.c.van.der.maarel@ rug.nl)

Twenty years ago, several α -amylases and related enzymes composed of a $(\beta/\alpha)_7$ -barrel (an irregular TIM-barrel domain) were classified into the family GH57; more recently also the family GH119 was established^{9,10}. Both of these enzyme families are at present considerably smaller than GH13 and only few members have been characterized in detail¹¹. The first determined 3D structure of GH57 was that of the 4- α -glucanotransferase from *Thermococcus litoralis* (TLGT). X-ray crystallography supported by site-directed mutagenesis of TLGT revealed that it has two catalytic residues, Glu123 and Asp214, as the catalytic nucleophile and the general acid/base, respectively¹². No 3D structure is currently available for GH119 members. In addition to the structural differences between GH13 and GH57-GH119 family members, there are also distinctive conserved regions between these families⁹. The GH57 and GH119 families share five conserved sequence regions¹³.

Several microbial strains isolated from a unique land-locked marine lake located in Kakaban island, part of the Derawan Islands, East Kalimantan, Indonesia, were screened for the production of α -amylases. The lake originally was the lagoon of an atoll, formed by corals over a period of two million years. As a result of movements in the earth's crust the coral reef was raised above the sea level, trapping 5 km² of seawater within a 50 meter high ridge, effectively creating a land-locked marine lake¹⁴. One of the isolates, *Bacillus megaterium* NL3, contained an active GH13 α -amylase with only the general acid/base residue (Glu231) at the conserved position. Amino acid sequence alignments and 3D homology modeling showed that the nucleophile may be shifted one position downstream (Asp203) and that the transition state stabilizing residue is not the canonical Asp but a His residue (His294). Phylogenetic analysis clustered this new α -amylase and its homologs, which also possess the incomplete GH13 catalytic machinery, as a separate branch in family GH13, representing a novel subfamily.

Results and Discussion

Screening of Kakaban lake isolates producing extracellular amylases. Eight of twenty bacterial isolates from Kakaban landlocked marine lake tested positive for the hydrolysis of starch by producing a clear halo around their colonies on red-dyed amylopectin agar plates. Isolate NL3 showed the largest clearing zone, indicating a relatively high α -amylase activity and was selected for further study. 16S rDNA sequence analysis showed that strain NL3 was most closely related to *Bacillus megaterium*. This result was in agreement with biochemical and physiological properties (data not shown) and hence the selected isolate was designated as *B. megaterium* NL3. The culture medium of strain NL3 showed activity towards soluble starch. The 50–80% ammonium sulphate precipitate of the culture medium gave a single protein band with molecular mass of approximately 55 kDa on SDS-PAGE in combination with activity staining with soluble starch (data not shown).

Molecular identification of the NL3 amylase. Using degenerate α -amylase specific primers and inverse PCR, a DNA fragment of 2.3 kb was obtained from genomic DNA of strain NL3. Analysis of the nucleotide sequence of this fragment showed that an open reading frame of 1515 bp with clear α -amylase sequence similarity was present. This gene was designated as *bmaN1*. A putative ribosomal binding site (RBS) corresponding to the AGGAGG sequence located 12 nucleotides upstream of the start codon was predicted. A probable catabolite responsive element (CRE) was found together with possible -10 (TATAAT) and -35 (TTAACA) regions. The CRE sequence showed only one mismatch in the last position when compared to the consensus sequence (TGT/AAANCGNTNA/TCA)¹⁵. The BmaN1 polypeptide deduced is 505 amino acid residues in length with a clear putative signal peptide sequence of 23 residues preceding the mature enzyme, as predicted by SignalP 4.0 Server¹⁶. The molecular weight and *p*I of BmaN1 were predicted using ExPASy server (http://web.expasy.org/protparam) as 56934 Da and 9.05, respectively. The full-length DNA sequence of the putative α -amylase gene of *B. megaterium* NL3, *bmaN1*, has been deposited in the GenBank database¹⁷ under the accession no. AGT45938.

In silico analysis of BmaN1 and its homologues. The BLAST search using the BmaN1 amino acid sequence as a query resulted in retrieving of more than 30 highly similar sequences of putative α -amylases (Fig. 1) some of which have already been classified in the family GH13¹⁸. Although all of them possess variations in the three residues forming the family GH13 catalytic machinery, it is possible to divide them into two groups: (i) the first, larger group (Nos 1-27 in Fig. 1) with Lys202 and His294 in the positions of the catalytic nucleophile and transition state stabilizer, respectively (instead of normally occurring aspartates); and (ii) a second, smaller group (Nos 28-34 in Fig. 1) exhibiting substitutions in positions of the entire catalytic triad, but rather without an obvious regularity (Fig. 2). While the sequences of the members of the former group are almost identical to BmaN1, those of the latter one are slightly different (Fig. 2). Interestingly, there is a strictly conserved aspartic acid residue succeeding the "strange" lysine, which corresponds with the position of the catalytic nucleophile (Fig. 2). The sequences of both these groups, proposed here to constitute a novel GH13 subfamily xy around the α -amylase from *B. megaterium* BmaN1, are all highly similar to those of α -amylases around the α -amylase from B. aquimaris BaqA (Nos 35-39 in Fig. 1) suggested recently to define also a new and independent GH13 subfamily xx¹⁹. The main difference between the α -amylases around the BmaN1 and those around BaqA is that the BaqA α -amylase and the members of its subfamily possess the complete catalytic machinery (Fig. 2) characteristic for the α -amylase family GH13⁷. The other feature of interest is the presence of a tryptophan pair in both BmaN1 and BaqA (Fig. 2) between the CSR-V (loop 3) and CSR-II (strand β 4), located in the helix α 3 of the catalytic $(\beta/\alpha)_8$ -barrel¹⁹

In addition to the incomplete catalytic machinery mentioned above, the most striking differences of BmaN1 and its close homologues discriminating them from other well-established GH13 subfamilies with the α -amylase specificity (Fig. 2) are the presence of a glutamic acid instead of aspartate at the beginning of the CSR-I (strand β 3) and in addition the position of the histidine in the middle of the CSR-V (loop3), a position usually occupied by aspartic acid²⁰. With regard to alignment of the representative α -amylases studied here (Fig. 1), its substantial part covering almost the entire (β/α)₈-barrel including domain B clearly document a very close homology of both eventual GH13 subfamilies, i.e. BmaN1 and BaqA. All these sequences (Nos 1–39 in Fig. 1) go well

Bacillus megaterium ADA1096374 WP 000746111 504 255 xy Bacillus seguterium ADA1096374 WP 00074361 504 255 xy Bacillus flexus UP1000491A179 WP 00074361 504 255 xy Bacillus flexus UP100471190C WP 02549923.1 504 255 xy Bacillus angaterium UP100471190C WP 02549923.1 504 255 xy Bacillus angaterium UP1000573870 WP L4380050.1 504 255 xy Bacillus angaterium UP1000570870 WP L4380050.1 504 255 xy Bacillus angaterium UP100070704AE WP 06183655.1 504 255 xy Bacillus angaterium UP10007AED4DA WP 06183656.1 504 255 xy Bacillus angaterium UP10007AED4DA WP 06183656.1 504 255 xy Bacillus angaterium UP10004804200 WP 0518241451 504 255 xy Bacillus angaterium UP1000684141 <td< th=""><th>No.</th><th>Organism</th><th>UniProt^a</th><th>GenBank⁵</th><th>Length^c</th><th>β2-β8^d</th><th colspan="2">GH13°</th></td<>	No.	Organism	UniProt ^a	GenBank⁵	Length ^c	β2-β8 ^d	GH13°	
2 Bacillus megaterium AA0109G3F4 WP_060746441.1 504 255 xy Bacillus fixus ADA0L1LRE8 KNH167821.1 504 255 xy Bacillus fixus UP100071060C WP.02574982.1 504 255 xy Bacillus megaterium DSD2X3 ADEC7693.1 504 255 xy Bacillus speathata UP100071984AE KNE75282.1 504 255 xy Bacillus sp. Fool123 ADA00017284AE KNE75282.1 504 255 xy Bacillus sp. Fool123 ADA0080CR KRE5002.1 504 255 xy Bacillus argabatata UP100034C5414 WP_01673014.1 504 255 xy Bacillus argabatata UP100034C5414 WP_01673014.1 504 255 xy Bacillus argabatata UP100048C200 WP_01673014.1 504 255 xy Bacillus argabatata UP100048C200 WP_01573014.1 504 255 xy Bacillus argabatatai UP10005836707 <t< td=""><td>1</td><td>Bacillus megaterium BmaN1</td><td>T1SIF2</td><td>AGT45938.1</td><td>504</td><td>255</td><td>ху</td></t<>	1	Bacillus megaterium BmaN1	T1SIF2	AGT45938.1	504	255	ху	
3 Bacillus sp. 171065_106 UP1000491A179 VM_028407436.1 504 255 xy 5 Bacillus flexus UP1000471606C VM_025749923.1 504 255 xy 7 Bacillus flexus UP1000471606C VM_025749923.1 504 255 xy 7 Bacillus megaterium UP100077054AE KOU28073.1 504 255 xy 10 Bacillus megaterium UP100077054AE VM_05189559.1 504 255 xy 11 Bacillus megaterium AAA0A08CR8 KK1E10025.1 504 255 xy 12 Bacillus megaterium AAA0A09G072 KKD04449.1 504 255 xy 13 Bacillus megaterium AAA0049G072 KKD04449.1 504 255 xy 18 Bacillus megaterium AAA0049G072 KKD04449.1 504 255 xy 18 Bacillus angaterium AAA0045042144 504 255 xy 18 Bacillus angaterium AAA0A0450448	2	Bacillus megaterium	A0A109G3F4	WP_060746441.1	504	255	xy	
4 Bacillus fixus ADAOL 1LREB KNH187821 504 255 xy Bacillus megaterium DSD2X3 ADEC676931 504 255 xy Bacillus splatitim DSD2X3 ADAOCB42B KOL280731 504 255 xy Bacillus splatitim UPRIOMYTBARE KNF5282013001 504 255 xy Bacillus splatitim UPRIOMYTBARE KNF5282013001 504 255 xy Bacillus splatitim UPRIOMYTBARE KNF5282013001 504 255 xy Bacillus splatitim UPRIOMYTBARE KNF508211 504 255 xy Bacillus splatitim UPRIOMYTBARE KNF508211 504 255 xy Bacillus sylbhattai UPRIOMACC200 VP.05324115 504 255 xy Bacillus sylbhattai UPRIOMARC200 VP.05324174 504 255 xy Bacillus sylbhattai UPRIOMARC200 VP.05324174 504 255 xy Bacillus sylbhattai UPRIOMARC200	3	Bacillus sp. 171095_106	UPI000491A179	WP_028407436.1	504	255	xy	
b Bacillus Revuis UPI004/190C VVP_U2/37498/23.1 S04 255 xy Bacillus sp. Leaf75 ADA006H/23 KOU26073.1 S04 255 xy Bacillus sp. Solb531 ADA006H/23 KOU26073.1 S04 255 xy Bacillus sp. Solb531 ADA006H/23 KOU26073.1 S04 255 xy Bacillus sp. Fold230 ADA0009(CR6 KPE10025.1 S04 255 xy Bacillus sp. Fold230 ADA0009(CR6 KPE10025.1 S04 255 xy Bacillus sp. Fold230 ADA0009(CR6 KRD84049.1 S04 255 xy Bacillus aryabhatai UPI003745414 KDP3005.1 S04 255 xy Bacillus aryabhatai UPI004470200 VP_2047751650.1 S04 255 xy Bacillus aryabhatai UPI000480C200 VP_2042812402 504 255 xy Bacillus aryabhatai UPI00055340C VP_204292334.1 504 255 xy Bacillus aryabhatai UPI0005524.00	4	Bacillus flexus	A0A0L1LRE8	KNH18792.1	504	255	xy	
0 Bacillus sp. Leaff 3 ADACOA (JZS ADECINSS.1 SUP 225 xy Bacillus sp. Solf31 ADACOA (JZS KIPE (JSSS)1 SUP 225 xy Bacillus sp. Solf31 ADACOA (JZS KIPE (JSSS)1 SUP 225 xy Bacillus sp. Solf31 ADACOA (JZS KIPE (JSSS)1 SUP 225 xy Bacillus sp. Root (JZ) ADACOGICR6 KRE1002515 SUP 2255 xy Bacillus sp. Root (JZ) ADACOGICR6 KRE1002511 SUP 2255 xy Bacillus anguabratai UPICO034C5414 WP_015730141 SUP 255 xy Bacillus anguabratai UPICO044DC200 WP_0234267734,1 S04 255 xy Bacillus anguabratai UPICO044DC20 WP_033578026,1 S04 255 xy Bacillus anguabratai UPICO05514241 WP_047731650,1 S04 255 xy Bacillus anguabratai UPICO05524026 WP_0477931024,1 S04 255 xy Bacillus anguabratai	5	Bacillus flexus	UPI0004/1906C	WP_025749923.1	504	255	xy	
8 Bacillus sp. SoliS31 ADA0008UT.12 KRF52820.11 504 255 xy 10 Bacillus megatarium UPI000570870 WP_043808030.1 504 255 xy 11 Bacillus sp. FolZ39 ADA00090CR6 KRE10025.1 504 255 xy 12 Bacillus sp. FolZ-230 ADA00090CR6 KRE10025.1 504 255 xy 12 Bacillus sp. FolZ-230 ADA00090CR6 KRE10025.1 504 255 xy 12 Bacillus megatarium ADA00403C5414 WP_01673014.1 504 255 xy 15 Bacillus anyabhatiai UPI00048D200 WP_01673014.1 504 255 xy 18 Bacillus anyabhatiai UPI00048D2000 WP_034287134.1 504 255 xy 18 Bacillus anyabhatiai UPI0005538C70 WP_034280734.1 504 255 xy 28 Bacillus anyabhatiai UPI000552A00 WP_034280734.1 504 255 xy 28 Bacillus any	7	Bacillus sn Leaf75	A0A006I476	KOU26073 1	504	255	XY	
9 Bacillus anyabritatia UPI0005C75870 WP_043890630.11 504 255 xy 11 Bacillus sp. Root239 AAAQ09ICR6 KRE10025.1 504 255 xy 12 Bacillus sp. Root147 AAAQ09ICR6 KRE10025.1 504 255 xy 13 Bacillus sp. Root147 AAAQ09CR2 KRE10025.1 504 255 xy 13 Bacillus sp. Root147 AAAQ09C072 KRD84049.1 504 255 xy 15 Bacillus anyabhatai UPI00034C5414 WP_015736145.1 504 255 xy 16 Bacillus anyabhatai UPI0004ADC200 WP_024212402 504 255 xy 17 Bacillus anyabhatai UPI000532A0E WP_033578028.1 504 255 xy 18 Bacillus anyabhatai UPI0005542A17 WP_047751650.1 504 255 xy 18 Bacillus anyabhatai UPI00052420E WP_04779161651.1 504 255 xy 18 Bacillus anyabhata	8	Bacillus sp. Soil531	A0A0Q9UTJ2	KRF52620.1	504	255	xv	
10 Bacillus megatarium UPI00077D34AE WP_06189555.1 504 255 xy 12 Bacillus Sp. FkJT-21351 ADAQM0VCB1 KCPT3005.1 504 255 xy 12 Bacillus Sp. FkJT-21351 ADAQM0VCB1 KCPT3005.1 504 255 xy 13 Bacillus sp. Root147 ADAQ060072 KRD84494.1 504 255 xy 14 Bacillus anyabhatai UPI0007ABDADA WP_01673014.1 504 255 xy 15 Bacillus anyabhatai UPI00048DC200 WP_01573014.1 504 255 xy 16 Bacillus anyabhatai UPI00048DC200 WP_03428773.4 504 255 xy 18 Bacillus anyabhatai UPI0005314.4 WP_04739102.1 504 255 xy 18 Bacillus anyabhatai UPI00054A441 WP_0438102.1 504 255 xy 28 Bacillus anyabhatai UPI000552A0E WP_0438102.1 504 255 xy 28 Bacillus anyabhat	9	Bacillus aryabhattai	UPI0005C75870	WP 043980530.1	504	255	XV	
11 Bacillus sp. Rot/239 ADADQSICR6 KRE 10025.1 504 255 xy 13 Bacillus megaterium ADADBGAUA8 ALI2477.1 504 255 xy 15 Bacillus sp. Rot147 ADADBGAUA8 ALI2477.1 504 255 xy 15 Bacillus megaterium UPI00034C5414 WP_063248145.1 504 255 xy 16 Bacillus aryabhattai UPI00034C5414 WP_063248145.1 504 255 xy 17 Bacillus aryabhattai UPI000480C200 WP_02414202 504 255 xy 19 Bacillus aryabhattai UPI000480C200 WP_0323241.1 504 255 xy 18 Bacillus aryabhattai UPI000523400E WP_04332934.1 504 255 xy 18 Bacillus aryabhattai UPI000522400E WP_047831024.1 504 255 xy 28 Bacillus aryabhattai UPI00052441 264 255 xy 28 Bacillus aryabhattai UPI0003056	10	Bacillus megaterium	UPI00077D84AE	WP_061859559.1	504	255	xy	
12 Bacillus sp. FA/T-21351 ADADMOV/GB1 KOP73005.1 504 255 xy Bacillus sp. Root147 ADADQ3G072 KRD84949.1 504 255 xy Bacillus sp. Root147 ADADQ3G072 KRD84949.1 504 255 xy Bacillus aryabhattai UPI0007ABD4DA WP_016783014.1 504 255 xy Bacillus aryabhattai ADADHR840 WP_016783014.1 504 255 xy Bacillus aryabhattai ADADUSDUMA KNN97899.1 504 255 xy Bacillus aryabhattai UPI0004520.200 WP_03378026.1 504 255 xy Bacillus aryabhattai UPI000532A90 WP_03378026.1 504 255 xy Bacillus aryabhattai UPI00052C303 WP_047931024.1 504 255 xy Bacillus aryabhattai UPI00052C303 WP_047931024.1 504 255 xy Bacillus aryabhattai UPI00052C303 WP_045931024.1 504 255 xy Bacillus aryabhattai	11	Bacillus sp. Root239	A0A0Q9ICR6	KRE10025.1	504	255	xy	
13 Bacillus megaterium ADADBSALA8 ALI2477.1 504 255 xy Bacillus segaterium UPI00034C5414 VVP_063248145.1 504 255 xy Bacillus megaterium UPI00034C5414 VVP_063248145.1 504 255 xy Bacillus segaterium ADADH4R640 AKP75627.1 504 255 xy Bacillus aryabhattai UPI00048DC20M VVP_063248145.1 504 255 xy Bacillus aryabhattai UPI000528C7C VVP_0324267734.1 504 255 xy Bacillus aryabhattai UPI0005282409E VVP_042392334.1 504 255 xy Bacillus aryabhattai UPI000522809E VVP_042392823.1 504 255 xy Bacillus aryabhattai UPI0005252801 504 255 xy Bacillus aryabhattai UPI0005252801 504 255 xy Bacillus aryabhattai UPI000525263.1 504 255 xy Bacillus aryabhattai UPI00035086A VVP_041931024.1 504 <td>12</td> <td>Bacillus sp. FJAT-21351</td> <td>A0A0M0WGB1</td> <td>KOP73005.1</td> <td>504</td> <td>255</td> <td>xy</td>	12	Bacillus sp. FJAT-21351	A0A0M0WGB1	KOP73005.1	504	255	xy	
Instructure AUAQUSCU2 NRU64949.1 D04 255 XY Bacillus aryabhattai UPI0007ABD4DA VWP_016783014.1 504 255 XY Bacillus aryabhattai A0A0HAR640 VWP_016783014.1 504 255 XY Bacillus sp. 27892_107 UPI0004BC200 VWP_0281241202 504 255 XY Bacillus aryabhattai A0A0JSUGMS VMP_0147751659.1 504 255 XY Bacillus aryabhattai UPI000532A9E VWP_031378026.1 504 255 XY Bacillus aryabhattai UPI00052404441 WP_04393026.1 504 255 XY Bacillus aryabhattai UPI000524050 WP_047931024.1 504 255 XY Bacillus bataviensis KGEA75 EN/N7296.1 508 267 XY Bacillus bataviensis UPI00036040467 WP_01931024.1 504 225 XY Bacillus bataviensis UPI000525200 WP_01931024.1 504 255 XY Bacillus bataviensis UPI000261002044	13	Bacillus megaterium	A0A0B6AUA8	AJI24747.1	504	255	xy	
Instructure Definition OPTION/SECTION State St	14	Bacillus sp. Root147	AUAUQ9G072	KRD84949.1	504	255	xy	
Decking angeletismen Concord Ref 200 App 79637.1 Sol 255 xy Bacillus sp. 278922_107 UP1000480C200 WP_028412402 Sol 255 xy Bacillus sp. 278922_107 UP1000480C200 WP_028412402 Sol 255 xy Bacillus argabhattai UP10005538C7C WP_004751659.1 Sol 255 xy Bacillus argabhattai UP10005538C7C WP_047751659.1 Sol 255 xy Bacillus megaterium UP10005538C7C WP_04375026.1 Sol 255 xy Bacillus megaterium UP1000523A902 WP_043731024.1 Sol 255 xy Bacillus megaterium UP1000524C402 WP_043731024.1 Sol 255 xy Bacillus megaterium UP100052538C7C WP_04529528.1 Sol 255 xy Bacillus bateviensis K6EA75 EKN70286.1 Sol 255 xy Bacillus bateviensis UP1000380E8A WP_0191930978.1 Sol 267 xy Bacillus socennisectininis UP10025F2619 WP_019380978.1 Sol 267 xy Bacillus socennisectininis UP10025F2	16	Bacillus anyabhattai		WP_010703014.1	504	255	xy	
18 Bacillus sp. 278922, 107 UP1000480C200 WP_028412402 504 255 Wy 10 Bacillus aryabhattai UP10005481914 WP_0477516591 504 255 Xy 20 Bacillus aryabhattai UP10005588C7C WP_0477516591 504 255 Xy 21 Bacillus aryabhattai UP10005532807C WP_0473516591 504 255 Xy 23 Bacillus megaterium UP10005523209E WP_033378026.1 504 255 Xy 24 Bacillus megaterium UP10005E04C04C8 WP_047931024.1 504 255 Xy 25 Bacillus megaterium UP10005E04C04C8 WP_047931024.1 504 255 Xy 26 Bacillus megaterium UP10005E04C04C WP_0477031024.1 508 257 Xy 26 Bacillus metarionis G2RMI2 ARN91476.1 468 255 Xy 26 Bacillus metarionis UP1000360346A WP_01970641.1 509 265 Xy 27	17	Bacillus megaterium	A0A0H4R640	AKP75627.1	504	255	xv	
19 Beacilus aryabhattai AQAQISUQM8 KMispr899.1 504 255 y 20 Bacillus aryabhattai UP10006481914 WP_04267734.1 504 255 xy 21 Bacillus aryabhattai UP1000584A44 WP_043973024.1 504 255 xy 23 Bacillus argabattai UP1000584A444 WP_043931024.1 504 255 xy 24 Bacillus argabattai UP1000544C4CB WP_047931024.1 504 255 xy 25 Bacillus argabattai UP100054CC4CB WP_047931024.1 504 255 xy 26 Bacillus argabattai UP100054CC4CB WP_047931024.1 504 255 xy 28 Bacillus bataviensis K6EA75 EKN70296.1 509 262 xy 28 Bacillus bataviensis VP_0100306EA4 WP_01195200.1 462 255 xy 28 Bacillus argainaris UP10003076AA WP_011930074.1 507 264 xy 36 28	18	Bacillus sp. 278922 107	UPI00048DC200	WP 028412402	504	255	xv	
20 Bacillus aryabhattai UP10005481914 WP_047751659.1 504 255 xy 21 Bacillus megaterium UP10005581C7.0 WP_042892334.1 504 255 xy 22 Bacillus megaterium UP100055400E WP_04357026.1 504 255 xy 24 Bacillus megaterium D5DAE5 ADF37524.1 497 255 xy 25 Bacillus megaterium D5DAE5 ADF37524.1 402 255 xy 26 Bacillus megaterium D5DAE5 ADF37524.1 402 255 xy 27 Bacillus megaterium G2RNI2 AEN91476.1 468 255 xy 28 Bacillus ordini UP1000300344 WP_01870541.1 509 267 xy 38 Bacillus sp. 1NLA3E NNAV77 AGK2591.1 507 264 xy 34 Bacillus sp. 1NLA3E NNAV77 AGK2591.1 507 xx 36 Acoxybacillus sp. S13.4 11VW19 AF49455.1	19	Bacillus aryabhattai	A0A0J5UQM8	KMN97899.1	504	255	xy	
21 Bacillus sp. Aph1 UP1005538C7C WP_034267734.1 504 255 xy 23 Bacillus argabnattai UP1005532A90E WP_033578026.1 504 255 xy 24 Bacillus argabnattai UP100552A90E WP_047931024.1 504 255 xy 25 Bacillus megaterium UP10005EC3C03 WP_047931024.1 504 255 xy 26 Bacillus megaterium GZRMI2 AEN91476.1 468 255 xy 28 Bacillus bataviensis K6EA75 EKN70296.1 508 267 xy 29 Bacillus medaterium UP100380248A WP_019153200.1 492 255 xy 30 Bacillus methanolicus UP10038024F2619 WP_019153200.1 492 255 xy 31 Bacillus apos.1NLA3E UP1003267F2619 WP_019153200.1 492 257 xy 32 Bacillus apos.1LA3E UP1003267F2619 WP_019153200.1 492 267 xy 33 Bacillus apos.1LA3E UP10038024F2619 WP_019153200.1 492 267 xy	20	Bacillus aryabhattai	UPI00064B1914	WP_047751659.1	504	255	xy	
22 Bacillus argabhattai UP10005824041 WP_042992334.1 504 255 xy 24 Bacillus argabhattai UP1000582200E WP_033578026.1 504 255 xy 24 Bacillus megaterium UP1000582200E WP_045295263.1 504 255 xy 25 Bacillus argabhattai UP10005822003 WP_045295263.1 504 255 xy 27 Bacillus bataviensis K6EA75 EKN70296.1 508 267 xy 29 Bacillus bataviensis K6EA75 EKN70296.1 509 262 xy 30 Bacillus methanolicus ISE5X2 WP_0198380978.1 509 267 xy 32 Bacillus sp. 1NLA3E NDAVI77 AGK52891.1 507 284 xy 34 Bacillus sp. 5N3_4 IVWH9 AFI49455.1 505 270 xx 35 Bacillus sp. 5N3_4 IVWH9 AFI49455.1 505 270 xx 36 Geobacillus themoleovorans GRN74<	21	Bacillus sp. Aph1	UPI000553BC7C	WP_034267734.1	504	255	xy	
23 Bacillus megaterium DEPOLES 2X40E WP_0335/8028.1 504 255 xy 25 Bacillus megaterium DEDAE5 ADF37524.1 504 255 xy 25 Bacillus megaterium UP100054CC4CB WP_047931024.1 504 255 xy 27 Bacillus aryabhatia UP100054CC4CB WP_047931024.1 504 255 xy 28 Bacillus bataviensis K6EA75 EKN70296.1 508 267 xy 28 Bacillus bataviensis UP100038048A WP_019153290.1 492 255 xy 30 Bacillus methanolicus 13E5X2 EUE1893.1 511 270 xy 32 Bacillus aceanisediminis UP10028F2619 WP_019380978.1 507 244 xy 34 Bacillus aceanisediminis UP10028F2619 EV74976.1 509 267 xy 35 Bacillus acyuimaris G8IJA7 AER68125.1 512 269 xx 36 Anoxybacillus sp. SK3_4	22	Bacillus megaterium	UPI0005B4A441	WP_042992334.1	504	255	xy	
Abc Abc <td>23</td> <td>Bacillus aryabhattai</td> <td>UPI000532A90E</td> <td>WP_033578026.1</td> <td>504</td> <td>255</td> <td>xy</td>	23	Bacillus aryabhattai	UPI000532A90E	WP_033578026.1	504	255	xy	
Lobility any abhattai UP (D005EC3C03 WP_045295283.1 S04 205 Xy 27 Bacillus megaterium G2RMI2 AEN91476.1 468 255 Xy 28 Bacillus megaterium G2RMI2 AEN91476.1 508 255 Xy 29 Bacillus rotali UP (D005EC3C03 WP_01870296.1 508 257 Xy 29 Bacillus megaterium G2RMI2 AEN91476.1 509 262 Xy 20 Bacillus nettianolicus ISEX2 EUB (1893.1 501 255 Xy 31 Bacillus sp. 1NLA5E NDAW77 AGRS2891.1 509 267 Xy 32 Bacillus sp. 1NLA5E NDAW77 AGRS2891.1 509 267 Xy 34 Bacillus sp. 1NLA5E NDAW77 AGRS2891.1 505 270 Xx 35 Bacillus sp. 1D13 11 11/VVI9 AFI49455.1 505 270 Xx 36 Aooxybacillus sp. SK3.4 11/VVI9 AFI49455.1 <td>24</td> <td>Bacillus megaterium</td> <td></td> <td>MP 047031024.1</td> <td>497</td> <td>255</td> <td>xy</td>	24	Bacillus megaterium		MP 047031024.1	497	255	xy	
27 Bacillus megaterium G2RMI2 AEN91476.1 468 255 xy 28 Bacillus bataviensis K6EA75 EKN70286.1 508 267 xy 29 Bacillus braiviensis K6EA75 EKN70286.1 509 262 xy 30 Bacillus massilicsenegalensis UPI00330034BA WP_018153290.1 492 255 xy 31 Bacillus ceanisedminis UPI0032024EA EU01893.1 511 270 xy 32 Bacillus sp. TLA3E NDAW77 AGK52691.1 509 267 xy 33 Bacillus aquimaris G8IJA7 AER68125.1 512 269 xx 34 Bacillus aquimaris G8IJA7 AER68125.1 505 270 xx 35 Bacillus aquimaris G8IJA7 AER68125.1 512 269 xx 36 Geobacillus thermoleovarans G8IJA7 AER68125.1 513 270 xx 37 Anoxybacillus psp. D13_1 HVWI0 AFH	26	Bacillus arvabhattai	UPI0005EC3C03	WP_045295263.1	504	255	xv	
28 Bacillus bataviensis K6EA75 EKN70296.1 508 267 xy 29 Bacillus fordil UP100030034BA WP_01915290.1 492 255 xy 31 Bacillus matsiliosenegalensis UP100030034BA WP_019153290.1 492 255 xy 31 Bacillus methanolicus UP10002FF2619 WP_019380978.1 509 267 xy 32 Bacillus apuimaris GBIJA7 AER68125.1 512 269 xx 33 Bacillus aquimaris GBIJA7 AER68125.1 512 269 xx 34 Bacillus aquimaris GBIJA7 AER68125.1 512 269 xx 35 Bacillus aquimaris GBIJA7 AER68125.1 512 269 xx 36 Geobacillus thermoleovorans GBN704 AEV18110.1 511 270 xx 40 Aspergillus oryzae PO0693 AAA32128.1 498 261 1 41 Saccharomycopsis fibuligera OP0692	27	Bacillus megaterium	G2RMI2	AEN91476.1	468	255	xy	
29 Bacillus fordi UP1000380E64F WP_019705441.1 509 282 xy 30 Bacillus massiliosenegalensis UP100030034BA WP_019153290.1 492 255 xy 31 Bacillus sections (instantian instantian instantinstantian instantinstea instantian instantian insta	28	Bacillus bataviensis	K6EA75	EKN70296.1	508	267	XV	
30 Bacilius messiliosenegalensis UPI00030034BA WP_019153290.1 492 255 xy 31 Bacilius methanolicus I3E5X2 EU81893.1 511 270 xy 32 Bacilius sp. 1NLA3E N0AW77 AGK52691.1 509 267 xy 33 Bacilius sp. 2.4, 57C12 E5W0K9 EFV74976.1 509 267 xy 34 Bacilius sp. 2.4, 57C12 E5W0K9 EFV74976.1 505 270 xx 35 Bacilius sp. DT3_1 11VW10 AFI49455.1 505 270 xx 36 Anoxybacilius sp. SK3_4 11VW19 AFI49455.1 505 270 xx 37 Anoxybacilius sp. SK3_4 11VW19 AFI49455.1 505 270 xx 38 Geobacilius thermoleovorans I3014 AFK98071.1 513 270 xx 41 Saccharomycopsis fibuligera D4P4Y7 ADD80242.1 494 276 1 42 Bacilius amyloiquefaciens P00693 <td>29</td> <td>Bacillus fordii</td> <td>UPI000380E64F</td> <td>WP_018705441.1</td> <td>509</td> <td>262</td> <td>xy</td>	29	Bacillus fordii	UPI000380E64F	WP_018705441.1	509	262	xy	
31 Bacillus methanolicus I3E5X2 EUB1809.1 511 270 xy 33 Bacillus sp. 1NLA3E N0AV77 AGK52691.1 507 264 xy 34 Bacillus sp. 2_A_57_CT2 E5W0K9 EFV74976.1 509 267 xy 35 Bacillus sp. 2_A_57_CT2 E5W0K9 EFV74976.1 505 270 xx 36 Anoxybacillus sp. DT3_1 11VW10 AFI49456.1 505 270 xx 37 Anoxybacillus sp. SK3_4 11VW19 AFI49456.1 511 270 xx 39 Geobacillus thermoleovorans G8N704 AEV18110.1 511 270 xx 40 Aspergillus oryzae P0013 CAA31218.1 499 276 1 41 Saccharomycopsis fibuligera D4P4Y7 AD080242.1 494 276 1 42 Bacillus amyloiquefaciens P00692 AAA22191.1 614 332 5 43 Paracoccidiodes brasiliensis A7L832 ABS11196.1 635 331 5 44 Hordeum vulgare <td>30</td> <td>Bacillus massiliosenegalensis</td> <td>UPI00030034BA</td> <td>WP_019153290.1</td> <td>492</td> <td>255</td> <td>xy</td>	30	Bacillus massiliosenegalensis	UPI00030034BA	WP_019153290.1	492	255	xy	
32 Bacillus oceanisediminis UPI0002FF2619 WP_019380978.1 507 264 xy 34 Bacillus sp.1 507 264 xy 34 Bacillus sp.1 507 264 xy 35 Bacillus sp.1 507 264 xy 35 Bacillus sp.1 64 64 509 267 xy 36 Anoxybacillus sp.1 11 VW0 AF184855.1 505 270 xx 37 Anoxybacillus sp.SK3_4 11 VW19 AF149455.1 505 270 xx 39 Geobacillus thermoleovorans G8N704 AEV18110.1 511 270 xx 41 Saccharomycopsis fibuligera D4P4Y7 AD080242.1 494 276 1 42 Bacillus anyloiquefaciens P00693 AAA3229.1 438 293 6 44 Hordeum vulgare P00693 AAA3229.1 438 293 6 45 Malus domestica Q5BLY0	31	Bacillus methanolicus	13E5X2	EIJ81893.1	511	270	×y	
33 Bability Sp. 1NLA3E NUMW/r AGRS2991.1 SU/ SU/ Zef Xy 34 Bability sp. 2_6,57_C12 E5WQR9 EFV74976.1 509 267 xy 35 Bacillus squimaris G8IJA7 AER68125.1 512 269 xx 36 Anoxybacillus sp. 2_A_57_C12 E5WQR9 EFV74976.1 505 270 xx 37 Anoxybacillus sp. SN3_4 11VWI0 AFI49455.1 505 270 xx 38 Geobacillus thermoleovorans G8N704 AEV18110.1 513 270 xx 40 Aspergillus oryzae POC1B3 CAA31218.1 499 276 1 41 Saccharomycopsis fibuligera D4F4Y7 ADD80242.1 494 276 1 42 Bacillus amyloliquefaciens P00692 AAA22191.1 514 332 5 43 Paracoccidicides brasiliensis A7L832 ABS11190.1 535 331 5 44 Hordeum vulgare P0693 AAA322929.1 438 293 6 45 Mal	32	Bacillus oceanisediminis	UPI0002FF2619	WP_019380978.1	509	267	×y	
Boolinus aguimaris CBILLIN Control of the second s	33	Bacillus sp. 2 A 57 CT2		AGK52691.1 EE\/74976.1	507	264	xy	
Database Database Description AFL49456.1 505 270 xx 37 Anoxybacillus sp. DT3_1 11VW10 AFL49456.1 505 270 xx 37 Anoxybacillus sp. SK3_4 11VW10 AFL49456.1 505 270 xx 39 Geobacillus thermoleovorans G8N704 AEV18110.1 511 270 xx 40 Aspergillus oryzee P0C1B3 CAA31218.1 499 276 1 41 Saccharomycopsis fibuligera D4P4Y7 ADb80242.1 494 276 1 42 Bacillus amyloliquefaciens P00692 AAA22191.1 614 332 5 43 Paracoccidioides brasiliensis A7L832 ABS11196.1 635 331 5 44 Hordeum vulgare P00693 AAA3229.1 414 280 6 45 Malus domestica Q58LY0 AAX3323.1 414 280 6 46 Pyrococcus wosesi Q7LYT7 AAD54338.1	35	Bacillus aquimaris	GBLIAT	AER68125.1	512	269	NY NY	
37 Anoxybacillus sp. SK3_4 I1VWH9 AFI49455.1 505 270 xx 38 Geobacillus thermoleovorans I3QII4 AFK08971.1 513 270 xx 40 Aspergillus oryzae P0C1B3 CAA31218.1 499 276 1 41 Saccharomycopsis fibuligera D4P4Y7 AD80242.1 494 276 1 42 Bacillus amyloliquefaciens P00692 AAA22191.1 514 332 5 44 Hordeum vulgare P00693 AAA32929.1 438 293 6 45 Malus domestica Q5BLY0 AAX33234.1 414 280 6 46 Pyrococcus woesei Q7LYT7 AAD54338.1 460 276 7 47 Thermococcus hydrothermalis O93647 AAC97877.1 457 276 7 48 Drosophila melanogaster P08144 X04569.1 494 295 15 50 Bacillus haldourans A8QWW3 BAF93484.1	36	Anoxybacillus sp DT3 1		AE149456 1	505	270	Ŷ	
38 Geobacillus thermoleovorans I3QII4 AFK08971.1 513 270 xx 39 Geobacillus thermoleovorans G8N704 AEV18110.1 511 270 xx 40 Aspergillus oryzae POC1B3 CAA31218.1 499 276 1 41 Saccharomycopsis fibuligera D4P4Y7 ADD80242.1 494 276 1 42 Bacillus amyloiquefaciens P00692 AAA22191.1 514 332 5 43 Paracoccidioides brasiliensis ATL832 ABS11196.1 535 331 5 44 Hordeum vulgare P00693 AAA32229.1 438 293 6 45 Malus domestica Q5BLY0 AAX33234.1 414 280 6 46 Pyrococcus woesei Q7LYT7 AAD54338.1 460 276 7 47 Thermococcus hydrothermalis 093647 AAC97877.1 457 276 7 48 Drosophila melanogaster P03144 X04569.1<	37	Anoxybacillus sp. SK3 4	11VWH9	AFI49455.1	505	270	XX	
39 Geobacillus thermoleovorans G8N704 AEV/18110.1 511 270 xx 40 Aspergillus oryzae P0C1B3 CAA31218.1 499 276 1 41 Saccharomycopsis fibuligera D4P4Y7 ADD80242.1 494 276 1 42 Bacillus amyoliquefaciens P00692 AAA22191.1 514 332 5 43 Paracoccidicides brasiliensis A7L832 ABS11196.1 535 331 5 44 Hordeum vulgare P00692 AAA22191.1 418 280 6 45 Malus domestica Q5BLY0 AAX33234.1 414 280 6 46 Pyrococcus woesei Q7LYT7 AAD54338.1 460 276 7 47 Thermococcus hydrothermalis Q93847 AAC97877.1 457 276 7 48 Drosophila melanogaster P08144 X04569.1 471 291 15 50 Bacillus halodurans A8QWV3 BAF93484.1	38	Geobacillus thermoleovorans	13Q114	AFK08971.1	513	270	xx	
40 Aspergillus oryzae P0C1B3 CAA31218.1 499 276 1 41 Saccharomycopis fibuligera D4P4Y7 ADD80242.1 494 276 1 42 Bacillus amyloliquefaciens P06692 AAA22191.1 514 332 5 44 Hordeum vulgare P00693 AAA32929.1 438 293 6 45 Malus domestica Q5BLY0 AAX33234.1 414 280 6 46 Pyrococcus woesei Q7LYT7 AAD54338.1 460 276 7 47 Thermococcus hydrothermalis O93847 AAC97877.1 457 276 7 48 Drosophila melanogaster P08144 X04569.1 494 295 15 50 Bacillus halodurans A8QWV3 BAF93484.1 958 361 19 51 Escherichia coli P25718 CAA41740.1 676 352 19 52 Gallus gallus O98942 AAC60246.1 511	39	Geobacillus thermoleovorans	G8N704	AEV18110.1	511	270	XX	
41 Saccharomycopsis fibuligera D4P4Y7 AD080242.1 494 276 1 42 Bacillus amyloliquefaciens P00692 AAA22191.1 514 332 5 43 Paracoccidioides brasiliensis ATL832 ABS11196.1 535 331 5 44 Hordeum vulgare P00693 AAA32929.1 438 293 6 45 Malus domestica Q5BLY0 AAX33234.1 414 280 6 46 Pyrococcus woesei Q7LYT7 AAD54338.1 465 276 7 47 Thermococcus hydrothermalis O93647 AAC97877.1 457 276 7 49 Tenebrio molitor P56634 471 291 15 50 Bacillus halodurans A8QVW3 BAF93484.1 958 361 19 51 Escherichia coli P25718 CAA1740.1 676 352 19 52 Gallus gallus Q98842 AAC60246.1 512 307 24 53 Homo sapiens P04746 M18785.1	40	Aspergillus oryzae	P0C1B3	CAA31218.1	499	276	1	
42 Bacillus anyloliquefaciens Paracoccidioides brasiliensis P00692 ATL832 AAA22191.1 ABS11196.1 514 535 332 331 5 43 Hordeum vulgare Malus domestica Q5BLY0 AAA32229.1 438 414 280 6 44 Hordeum vulgare Malus domestica Q5BLY0 AAA33234.1 414 280 6 45 Malus domestica Q5BLY0 AAA33234.1 414 280 6 46 Pyrococcus woesei Q7LYT7 AAD54338.1 460 276 7 47 Thermococcus hydrothermalis 093647 AAC97877.1 457 276 7 48 Drosophila melanogaster P08144 X04569.1 494 295 15 49 Tenebrio molitor P56634 471 291 15 50 Bacillus haldoturans A8QWV3 BAF93484.1 958 361 19 51 Escherichia coli P25718 CAA41740.1 676 352 19 54 Aeromonas campestris	41	Saccharomycopsis fibuligera	D4P4Y7	ADD80242.1	494	276	1	
43 Paracoccidioides brasiliensis A7L832 ABS11196.1 535 331 5 44 Hordeum vulgare P00693 AAA32929.1 418 293 6 45 Malus domestica Q5BLYO AAX33234.1 414 280 6 46 Pyrococcus woesei Q7LYT7 AAD54338.1 460 276 7 47 Thermococcus hydrothermalis Q93647 AAC97877.1 457 276 7 48 Drosophila melanogaster P08144 X04569.1 494 295 15 50 Bacillus halodurans A8QWV3 BAF93484.1 958 361 19 51 Escherichia coli P25718 CAA41740.1 676 352 19 52 Gallus gallus Q98942 AAC60246.1 512 307 24 54 Aeromonas hydrophila P22630 AAA21936.1 464 307 27 54 Aeromonas hydrophila P22630 AAA21936.1 466 307 27 55 Xanthomonas campestris Q56791 AAA	42	Bacillus amyloliquefaciens	P00692	AAA22191.1	514	332	5	
44 Hordeum vulgare P00693 AAA32929.1 438 293 6 45 Malus domestica Q5BLY0 AAX33234.1 414 280 6 46 Pyrococcus woesei Q7LYT7 AAD54338.1 460 276 7 47 Thermococcus hydrothermalis O93647 AAC97877.1 457 276 7 48 Drosophila melanogaster P08144 X04569.1 494 295 15 50 Bacillus halodurans A8QWV3 BAF93484.1 958 361 19 51 Escherichia coli P25718 CAA41740.1 676 352 19 52 Gallus gallus O98942 AAC60246.1 512 307 24 54 Aeromonas hydrophila P22630 AAA21936.1 464 307 27 55 Xanthomonas campestris Q56791 AAA21936.1 475 307 27 56 Bacillus substilis P00691 CAB12096.2 659 269	43	Paracoccidioides brasiliensis	A7L832	ABS11196.1	535	331	5	
45 Malus domestica Q5BLY0 AAX33234.1 414 280 6 46 Pyrococcus woesei Q7LYT7 AAD54338.1 460 276 7 47 Thermococcus hydrothermalis O93647 AAC97877.1 457 276 7 48 Drosophila melanogaster P08144 X04569.1 494 295 15 49 Tenebrio molitor P56634 471 291 15 50 Bacillus halodurans A8QWV3 BAF93484.1 958 361 19 51 Escherichia coli P25718 CAA41740.1 676 352 19 52 Gallus gallus Q98942 AAC60246.1 512 307 24 53 Homo sapiens P04746 M18785.1 611 307 27 54 Aeromonas hydrophila P22630 AAA21936.1 475 307 27 55 Xanthomonas campestris Q56791 AAA25591.1 475 307 27 55 Salilus substilis P00691 CAB12098.2 659<	44	Hordeum vulgare	P00693	AAA32929.1	438	293	6	
46 Pyrococcus woesei QTLYT7 AAD54338.1 460 276 7 47 Thermococcus hydrothermalis O93647 AAC97877.1 457 276 7 48 Drosophila melanogaster P08144 X04569.1 494 295 15 49 Tenebrio molitor P56634 471 291 15 50 Bacillus halodurans A8QWV3 BAF93484.1 958 361 19 51 Escherichia coli P25718 CAA41740.1 676 352 19 52 Gallus gallus O98842 AAC60246.1 512 307 24 53 Homo sapiens P04746 M18785.1 511 307 27 54 Aeromonas hydrophila P22630 AAA21938.1 464 307 27 55 Xanthomonas campestris Q56791 AAA25791.1 475 307 27 56 Bacillus substilis P00691 CAB12098.2 659 269 28 57 Lactobacillus amylovorus Q48502 AAC45781.1	45	Malus domestica	Q5BLY0	AAX33234.1	414	280	6	
41 Interinductocus injutoinentations Op3647 AAC97677.1 457 276 7 48 Drosophila melanogaster P08144 X04569.1 494 295 15 48 Drosophila melanogaster P08144 X04569.1 494 295 15 50 Bacillus halodurans A8QWV3 BAF93484.1 958 361 19 51 Escherichia coli P25718 CAA41740.1 676 352 19 52 Gallus galus O98842 AAC60246.1 512 307 24 53 Homo sapiens P04746 M18785.1 511 307 24 54 Aeromonas hydrophila P22630 AAA21936.1 464 307 27 55 Xanthomonas campestris Q56791 AAA27591.1 475 307 27 56 Bacillus substilis P00691 CAB12098.2 659 269 28 57 Lactobacillus amylovorus Q48502 AAC45781.1 953 271 28 58 Streptomyces limosus P09794 <	46	Pyrococcus woesei	Q7LYT7	AAD54338.1	460	276	7	
48 Drosophila melanogaster P08144 X04569.1 494 295 15 49 Tenebrio molitor P56634 471 291 15 50 Bacillus halodurans A8QWV3 BAF93484.1 958 361 19 51 Escherichia coli P25718 CAA41740.1 676 352 19 52 Galius galius 098842 AAC60246.1 512 307 24 53 Homo sapiens P04746 M18785.1 511 307 27 54 Aeromonas hydrophila P22630 AAA21936.1 464 307 27 55 Xanthomonas campestris Q56791 AAA21936.1 464 307 27 56 Bacillus substilis P00691 CAB12098.2 659 269 28 57 Lactobacillus amylovorus Q48502 AAC45781.1 953 271 28 58 Streptomyces limosus P09794 AAA88554.1 605 286	47	memococcus nyurotnermails	093647	AAC97077.1	457	270	1	
49 Jenebno molitor P50634 4/1 291 15 50 Bacillus halodurans A8QWV3 BAF9348.1 958 361 19 51 Escherichia coli P25718 CAA41740.1 676 352 19 52 Gallus gallus O98942 AAC60246.1 512 307 24 53 Homo sapiens P04746 M18785.1 511 307 27 54 Aeromonas hydrophila P22630 AAA21936.1 464 307 27 55 Xanthomonas campestris O56791 AAA27591.1 475 307 27 56 Bacillus substilis P00691 CAB12098.2 659 269 28 57 Lactobacillus amylovorus Q48502 AAC45781.1 953 271 28 58 Streptomyces limosus P09794 AAA85541.1 605 286 32 60 Dictyoglomus thermophilum P14899 AC119039.1 499 281 36 61 Halothermothrix orenii Q8GPL8 AAN52525.1	48	Drosophila melanogaster	P08144	X04569.1	494	295	15	
50 Bacillus halodurans A8QWV3 BAFB3484.1 958 361 19 51 Escherichia coli P25718 CAA41740.1 676 352 19 52 Gallus gallus O98842 AAC60246.1 512 307 24 53 Homo sapiens P04746 M18785.1 511 307 24 54 Aeromonas hydrophila P22630 AAA21936.1 464 307 27 55 Xanthomonas campestris Q56791 AAA27591.1 475 307 27 56 Bacillus substilis P00691 CAB12098.2 659 269 28 57 Lactobacillus amylovorus Q48502 AAC45781.1 953 271 28 58 Streptomyces limosus P09794 AA88554.1 666 24 32 60 Dictyoglomus thermophilum P14899 ACI19039.1 499 281 36 61 Halothermothrix orenii Q8GPL8 AAN52525.1 515	49	Tenebrio molitor	P56634		471	291	15	
51 Escherichia coli P25718 CAA41740.1 676 352 19 52 Gallus gallus O98942 AAC60246.1 512 307 24 54 horno sapiens P04746 M18785.1 511 307 24 54 Aeromonas hydrophila P22630 AAA21936.1 464 307 27 55 Xanthomonas campestris Q56791 AAA27591.1 475 307 24 56 Bacillus substilis P00691 CAB12008.2 659 269 28 57 Lactobacillus amylovorus Q48502 AAA45781.1 953 271 28 58 Streptomyces limosus P09794 AAA88554.1 665 286 32 50 Dictyogiomus thermophilum P14899 AC119039.1 499 281 36 61 Halothermothrix orenii Q8GPL8 AAN52525.1 515 291 36 62 Photobacterium profundum Q6LIA8 CAC22972.1 687 284 37 63 Uncultured bacterium D9MZ14 <	50	Bacillus halodurans	A8QWV3	BAF93484.1	958	361	19	
52 Gallus gallus Q98942 AAC60246.1 512 307 24 53 Homo sapiens P04746 M18785.1 511 307 24 54 Aeromonas hydrophila P22630 AAA21936.1 464 307 27 55 Xarthomonas campestris Q56791 AAA27991.1 475 307 27 56 Bacillus substilis P00691 CAB12098.2 659 269 28 57 Lactobacillus amylovorus Q48502 AAA245781.1 953 271 28 58 Streptomyces limosus P09794 AAA88554.1 566 274 32 59 Thermonospora curvata P29750 CAA1881.1 605 286 32 60 Dictyoglomus thermophilum P14899 AC119039.1 499 281 36 61 Halothermothrix orenii Q8GPL8 AAN52525.1 515 291 36 62 Photobacterium profundum QBLIA8 CAG22972.1 687<	51	Escherichia coli	P25718	CAA41740.1	676	352	19	
53 Homo sapiens P04746 M18785.1 511 307 24 54 Aeromonas hydrophila P22630 AAA21936.1 464 307 27 55 Xanthomonas campestris Q56791 AAA27591.1 475 307 27 56 Bacillus substilis P00691 CAB12098.2 659 269 28 57 Lactobacillus amylovorus Q48502 AAC45781.1 953 271 28 58 Streptomyces limosus P09794 AA848554.1 605 286 32 60 Dictyoglomus thermophilum P14899 AC119039.1 499 281 36 61 Halothermothrix orenii Q8GPL8 AAN52525.1 515 291 36 62 Photobacterium profundum Q8LN8 CAG22972.1 687 284 37 63 Uncultured bacterium D9MZ14 ADK21254.1 639 284 37 64 Flavobacterium pp. No. 92 Q8KKQ0 CAD32957.1.	52	Gallus gallus	Q98942	AAC60246.1	512	307	24	
54 Aeromonas hydrophila P22630 AAA21936.1 464 307 27 55 Xanthomonas campestris Q56791 AAA27591.1 475 307 27 56 Bacillus substilis P00691 CAB12098.2 659 269 28 57 Lactobacillus amylovorus Q48502 AAC45781.1 953 271 28 58 Streptomyces limosus P09794 AAA88554.1 666 264 32 59 Thermonospora curvata P29750 CAA41881.1 605 286 32 60 Dictyoglomus thermophilum P14899 AC19039.1 499 281 36 61 Halothermothrix orenii Q8GPL8 AAN52525.1 515 291 36 62 Photobacterium profundum Q6LN8 CAC22972.1 687 284 37 63 Uncultured bacterium D9MZ14 ADK21254.1 639 284 37 64 Flavobacterium sp. No. 92 Q8KKQ0 CAD32957.1 </td <td>53</td> <td>Homo sapiens</td> <td>P04746</td> <td>M18785.1</td> <td>511</td> <td>307</td> <td>24</td>	53	Homo sapiens	P04746	M18785.1	511	307	24	
55 Xahthomonas campestris Q50/91 AAA2 (591.1) 475 307 27 56 Bacillus substilis P00691 CAB12098.2 659 269 28 57 Lactobacillus amylovorus Q48502 AAC45781.1 953 271 28 58 Streptomyces limosus P09794 AAA88554.1 566 274 32 59 Thermonospora curvata P29750 CAA41881.1 605 286 32 60 Dictyoglomus thermophilum P14899 ACI19039.1 499 281 36 61 Halothermothrix orenii Q8GPL8 CAG22972.1 567 291 36 62 Photobacterium profundum Q6LIA8 CAG22972.1 687 284 37 63 Uncultured bacterium D9MZ14 ADK21254.1 639 284 37 64 Flavobacterium sp. No. 92 Q8KKQ0 CAD32957.1 619 294 ??	54	Aeromonas hydrophila	P22630	AAA21936.1	464	307	27	
56 Bacillus substilies P00691 CAB12098.2 659 269 28 57 Lactobacillus amylovorus Q48502 AAC45781.1 953 271 28 58 Streptomyces limosus P09794 AAA88554.1 566 274 32 50 Dictyoglomus thermophilum P14899 AC119039.1 499 281 36 60 Dictyoglomus thermophilum P14899 AC119039.1 499 281 36 61 Halothermothrix orenii Q8CPL8 AAN52525.1 515 291 36 62 Photobacterium profundum Q6LIA8 CAG22972.1 687 284 37 64 Flavobacterium ps. No. 92 Q8KKG0 CAD32957.1. 619 294 ??	55	Xanthomonas campestris	Q56791	AAA27591.1	475	307	27	
57 Lactobacillus amylovorus Q48502 AAC45781.1 953 271 28 58 Streptomyces limosus P09794 AAA88554.1 566 274 32 59 Thermonospora curvata P29750 CAA41881.1 605 286 32 60 Dictyoglomus thermophilum P14899 AC119039.1 499 281 36 61 Halothermothrix orenii Q8GPL8 AAN52525.1 515 291 36 62 Photobacterium profundum Q6LIA8 CAG22972.1 687 284 37 63 Uncultured bacterium D9MZ14 ADK21254.1 639 284 37 64 Flavobacterium sp. No. 92 Q8KKG0 CAD32957.1 619 294 ??	56	Bacillus substilis	P00691	CAB12098.2	659	269	28	
58 Streptomyces limosus P09794 AAA88554.1 566 274 32 59 Thermonospora curvata P29750 CAA41881.1 605 286 32 60 Dictyoglomus thermophilum P14899 ACI19039.1 499 281 36 61 Halothermothrix orenii Q8GPL8 AAN52525.1 515 291 36 62 Photobacterium profundum Q6LIA8 CAG22972.1 687 284 37 63 Uncultured bacterium D9MZ14 ADK21254.1 619 294 ??	57	Lactobacillus amylovorus	Q48502	AAC45781.1	953	271	28	
59 Thermonospora curvata P29750 CAA41881.1 605 286 32 60 Dickyoglomus thermophilum P14899 AC19039.1 499 281 36 61 Halothermothrix orenii Q8GPL8 AAN52525.1 515 291 36 62 Photobacterium profundum Q6LIA8 CAG22972.1 687 284 37 64 Flavobacterium sp. No. 92 Q8KKG0 CAD32957.1. 619 294 ??	58	Streptomyces limosus	P09794	AAA88554.1	566	274	32	
60 Dictyoglomus thermophilum P14899 ACI19039.1 499 281 36 61 Halothermothrix orenii Q8GPL8 AANS2525.1 515 291 36 62 Photobacterium profundum Q6LIA8 CAG22972.1 687 284 37 64 Flavobacterium sp. No. 92 Q8KKG0 CAD32957.1. 619 294 ??	59	Thermonospora curvata	P29750	CAA41881.1	605	286	32	
61 Halothermothrix orenii Q8GPL8 AAN52525.1 515 291 36 62 Photobacterium profundum Q6LIA8 CAG22972.1 687 284 37 63 Uncultured bacterium D9MZ14 ADK21254.1 639 284 37 64 Flavobacterium sp. No. 92 Q8KKQ0 CAD32957.1. 619 294 ??	60	Dictyoglomus thermophilum	P14899	ACI19039.1	499	281	36	
62 Photobacterium profundum Q6LIA8 CAG22972.1 687 284 37 63 Uncultured bacterium D9MZ14 ADK21254.1 639 284 37 64 Flavobacterium sp. No. 92 Q8KKG0 CAD32957.1. 619 294 ??	61	Halothermothrix orenii	Q8GPL8	AAN52525.1	515	291	36	
63 Uncultured bacterium D9MZ14 ADK21254.1 639 284 37 64 Flavobacterium sp. No. 92 Q8KKG0 CAD32957.1. 619 294 ??	62	Photobacterium profundum	Q6LIA8	CAG22972.1	687	284	37	
64 Flavobacterium sp. No. 92 Q8KKG0 CAD32957.1. 619 294 ??	63	Uncultured bacterium	D9MZ14	ADK21254.1	639	284	37	
	64	Flavobacterium sp. No. 92	Q8KKG0	CAD32957.1.	619	294	??	

Figure 1. α -**Amylases used in the present study**^{*a*}. ^{*a*}The list involves: (i) the members of the newly proposed GH13 subfamily xy represented by the " α -amylase" from *Bacillus megaterium* BmaN1 (Nos 1–27) and its closely related homologues (Nos 28–34) - probably intermediates between BmaN1 and the α -amylase from *Bacillus aquimaris* BaqA - caught by BLAST; (ii) the members of the recently proposed GH13 subfamily xx¹⁹ represented by the BaqA (Nos 35–39); (iii) representatives of the individual GH13 subfamilies with the specificity of α -amylases - subfamilies GH13_1, 5, 6, 7, 15, 19, 24, 27, 28, 32, 36 and 37 (Nos 40–63); and (iv) the currently unassigned cyclomaltodextrinase (GH13_??; No. 64); for details, see the Materials and methods section. ^{*bc*}Accession numbers from the UniProt (UniParc) and GenBank sequence databases, respectively. ^{*d*}The length of the entire amino acid sequence of the protein. ^{*c*}The length of the polypeptide chain spanning the segment from the beginning of the CSR-VI (strand β 2) to the end of the CSR-VII (strand β 8). ^{*f*}The GH13 subfamily (if available).

together exhibiting their own pattern of the alignment in comparison to remaining α -amylases that represent well-established GH13 subfamilies (Nos 40–63 in Fig. 1). Of note is also the fact that the small group of putative α -amylases with irregular substitutions in catalytic positions (Nos 28–34 in Fig. 1) exhibits obviously a higher similarity to the BaqA subfamily than to that around BmaN1, especially in domain B (preceding the CSR-V in loop3) as well as in the segment preceding the CSR-IV at strand β 7 (Fig. 2). The cyclomaltodextrinase from *Flavobacterium* sp. No. 92²¹ was added to the comparison as an interesting example since it was recently found to possess the pair of adjacent tryptophan residues¹⁹, typical for both BmaN1 and BaqA (Fig. 2). This is of interest because the cyclomaltodextrinase belongs to GH13 members intermediate between subfamilies of

	β2	β3	100p3		B 4		β5		β/	βB
xy_Bacillus_megaterium_T1SIF2	G <mark>F</mark> TA <mark>VLL</mark> TP	EFPLTI	L PHL N	WW	GXXA K	DDQ	LLIG E	ING	FLDDV H	SVPIVFYIS
xy_Bacillus_megaterium_A0A109G3F4	GFTAVLLTP	EFPLTI	L PHLN	WW	GXXA K	D DQ	LLIG E	ING	FLDDV H	SVPIVFYGT
xy Bacillus flexus A0A0LILEE8	GETAVLLTP	EFPLTT	TPHIN	WW						SVPIVEIGT
xy Bacillus flexus UPI000471906C	GFTAVLLTP	EFPLTI	LPHLN	ww	GYYV K	DDO	LLIGE	ING	FLDDV H	SVPIVFYGT
xy_Bacillus_megaterium_D5DZX3	g <mark>f</mark> ta <mark>vll</mark> tp	EFPLTI	LPHLN	ww	GYY <mark>V</mark> K	DIDQ	LLIG E	ING	FLDDV H	S <mark>VPIVFY</mark> GT
xy_Bacillus_sp_Leaf72_A0A0Q6I4Z6	G <mark>F</mark> TA <mark>VLL</mark> TP	EFPLTI	LPHLN	WW	GYY <mark>V</mark> K	DIDQ	LLIG E	ING	FLDD <mark>V</mark> H	S <mark>VPIVFY</mark> GT
xy_Bacillus_sp_Soil531_A0A0Q9UTJ2	G <mark>F</mark> TAVLLTP	EFPLTI	LPHLN	WW	GYYV K	DIDQ	LLIG E	ING	FLDDV H	SVPIVFYGT
xy Bacillus aryabhattai UP10005C75870	GETAVLLTP GETAVLLTP	EFPLT1 FFDT TT	PHUN	WW	GYYV K		GE	NG		SUPIVEYGT
xy Bacillus sp Root239 A0A009ICR6	GETAVLLTP	EFPLTI	LPHLN	ww	GYYV	DIDO	LLIGE	ING	FLDDV H	SVPIVEYGT
xy Bacillus sp FJAT-21351 A0A0M0WGB1	G <mark>F</mark> TA <mark>VLL</mark> TP	EFPLTI	LPHLN	ww	GYYV K	DDQ	LLIG E	ING	FLDDV H	S <mark>VPIVFYG</mark> T
xy_Bacillus_megaterium_A0A0B6AUA8	G <mark>F</mark> TA <mark>VLL</mark> TP	EFPLTI	LPHLN	ww	GYY <mark>V</mark> K	DIDQ	LLIG E	ING	FLDDV H	S <mark>VPIVFY</mark> GT
xy_Bacillus_sp_Root147_A0A0Q9G072	GFTAVLL TP	EFPLTI	L PHLN	WW	GYYV K	D DQ	LLIG E	NG	FLDDV H	SVPIVFYGT
xy_Bacillus_megaterium_0P100034C5414	GETAVLLTP	EFPLTI FFDT TT	T PHUN	WW TUTUT	GY YV			NG		SUPTUFICI
xy Bacillus megaterium AOAOH4R640	GFTAVLLTP	EFPLTI	LPHLN	ww	GYYV K	DDO	LLIGE	ING	FLDDV H	SVPIVFYGT
xy_Bacillus_sp_278922_107_UPI00048DC200	G <mark>F</mark> TA <mark>VLL</mark> TP	EFPLTI	LPHLN	ww	GYY <mark>V</mark> K	DIDQ	LLIG E	ING	FLDDV H	S <mark>VPIVFY</mark> GT
xy_Bacillus_aryabhattai_A0A0J5UQM8	G <mark>F</mark> TA <mark>VLL</mark> TP	EFPLTI	LPHLN	WW	GYY <mark>V</mark> K	DIDQ	LLIG E	ING	FLDDV H	S <mark>VPIVFY</mark> GT
xy_Bacillus_aryabhattai_UPI00064B1914	GFTAVLL TP	EFPLTI	PHLN	WW	GYYV K	DDQ	LLIGE	NG	FLDDV H	SVPIVFYGT
xy Bacillus sp Aphi UPI000553BC/C	GETAVILTP CETAVILTP	FFPLTI	TPHIN	WW	GY YV			NG	FLDDV H	SUPTUPYOT
xy Bacillus arvabhattai UPI000532A90E	GETAVLLTP	EFPLTV	LPHLN	ww	GYYV	DDO	LLIGE	ING	FLDDV H	SVPIVEYGT
xy Bacillus megaterium D5DAE5	G <mark>F</mark> TA <mark>VLL</mark> TP	EFPLTI	LPHLN	ww	GYY <mark>V</mark> K	DIDQ	LLIG E	ING	FLDDV H	S <mark>VPIVFY</mark> GT
xy_Bacillus_megaterium_UPI00064CC4CB	G <mark>F</mark> TA <mark>VLL</mark> TP	EFPLTI	LPHLN	ww	GYY <mark>V</mark> K	DIDQ	LLIG E	ING	FLDDV H	S <mark>VPIVFYG</mark> T
xy_Bacillus_aryabhattai_UPI0005EC3C03	GFTAVLL TP	EFPLTI	PHEN	WW	GYYV K	D DQ	LLIG E	ISG	FLDDV H	SVPIVFYGT
xy_Bacillus_megaterium_G2RM12	GETAVLLTP	EFPLIT	PHUN	ww	GYIV	202	and G	NG	FLUDV H	SVELVEIGT
xy_Bacillus_bataviensis_K6EA75	GFTAIRL TP	DEVTNN	I PD I N	WW	GYSL P	QINH		PAE	FLDNE Y	GIPIFYYGT
xy_Bacillus_fordi1_0F1000380E64F	GETTICISS	DENVNH		TATA		SVEH		AGE		GTPEVYYGS
xy Bacillus methanolicus I3E5X2	GFTAICLTP	DEVANH	I PDIA	ww	GYRL H	AVNH	FLLGE	IWS	FMDNH D	GIPIVYYGS
xy Bacillus oceanisediminis UPI0002FF2619	G <mark>Y</mark> TT <mark>LM</mark> LTP	EFPFNN	LPELD	ww	GYQL N	LIQN	YLI <mark>A</mark> D	TVP	YLDNP H	G <mark>VPIVYY</mark> GS
xy_Bacillus_sp_1NLA3E_NOAW77	GFTA <mark>IC</mark> LSP	EFRANS	LPDLN	WW	GYKL D	HAAN	FLL <mark>G</mark> D	1BA	FMDNQ R	G <mark>IPIVYY</mark> GS
xy_Bacillus_sp_2_A_57_CT2_E5WQK9	GFTT <mark>L</mark> MLTP	DFPSNN	L PE I N	WW	GFQL N	TQN	YLI <mark>A</mark> G	TVP	YMDNP H	GVPIVYYGS
xx_Bacillus_aquimaris_G8IJA7	GFTS <mark>IWL</mark> TP	D F V V NH	LPDLN	ww	GYRL D	TVRH	YLLG E	VFD	FIDNH D	GIPIVYYGS
xx Anoxybacillus sp DT3 1 IIVWI0	GETA W TP	DEVVNH	PD A	WW	GY II D	T	FILG E	WH		GIPIMYYGT
xx_AnoxyDacillus_sp_SK5_4_IIVWH9	GETA INT TE	DEVANH		5757				WS		GTPTMYYGT
xx Geobacillus thermoleovorans G8N704	GFTAIWL TP	DFVANH	LPDLA	ww	GYRL D	TVRH	FLLG E	w s	FLDNH D	GIPIMYYGT
1 Aspergillus orvzae POC1B3	GETAIWITE	DVVANH	L PD LD	s	GLRI D	TVKE				G D VAGO
1_Saccharomycopsis_fibuligera_D4P4Y7	GFTA <mark>IWI</mark> SP	DIVTNH	L PDLR	DF	GLRI D	SAKH	YS <mark>V</mark> G E	VEQ	FVENH D	G <mark>IPVIYY</mark> GQ
5 Bacillus amyloliquefaciens P00692	GITAVWI PP	DVVLNH	YADVD	WY	GFRI D	AA	ETVA B	WQ	FVENH D	GYPOVFYGD
5 Paracoccidioides brasiliensis A7L832	G <mark>V</mark> TS <mark>IWL</mark> PP	D <mark>AVL</mark> NH	FADLD	WL	G <mark>LRF</mark> D	AAKH	FFVA E	YWK	FVMNH D	GYPCLFYGD
6 Hordeum vulgare P00693	GVT <mark>HVWL</mark> PP	DIVINH	APDD	WL	AW II	FARG	ZAVA E	VWD	FVDNH D	GIPCIFYDH
6_Malus_domestica_Q5BLY0	G <mark>F</mark> TSA <mark>WL</mark> PP	DIVI <mark>N</mark> H	VPNID	WL	DFRF D	FARG	F <mark>SV</mark> G E	YWD	FLDNH D	G <mark>I</mark> PTVFYDH
7 Pyrococcus woesei Q7LYT7	GISA <mark>IWL</mark> PP	DVVINH	FPDIC	AY	G <mark>WRE</mark> D	YVKG	WAVG	YWD	EVANH D	GOPVIFYRD
7 Thermococcus hydrothermalis 093647	G <mark>I</mark> SA <mark>IWI</mark> PP	DIVI <mark>NH</mark>	YPDIC	AY	AWRE D	Y <mark>VK</mark> G	WAVG E	YWD	FVANH D	GQPAIFYRD
15_Drosophila_melanogaster_P08144	G <mark>Y</mark> AG <mark>V</mark> Q <mark>V</mark> SP	DVVFNH	LEDLN	HL	GFRV D	AAKH	YIVQ E	VID	FVDNH D	GTPRVMSSF
15_Tenebrio_molitor_P56634	G <mark>F</mark> GG <mark>V</mark> Q <mark>I</mark> SP	DAVINH	LRDLN	HМ	GFRV D	AAKH	FIYQ E	VID	FVDNH D	GTTRIMSS
19_Bacillus_halodurans_A8QWV3	G <mark>I</mark> NA <mark>IWI</mark> TA	DVV MNH	L PD FR	AW	GFRV D	TAKH	WMVG E	<mark>∨w</mark> G	YLSQH D	GG <mark>VQVFYGD</mark>
19_Escherichia_coli_P25718	G <mark>V</mark> NA <mark>LWI</mark> SA	D <mark>VV</mark> MNH	1PDIK	Q <mark>W</mark>	GFRV D	TAKH	WMTG E	AWG	YLSSH D	GAVQIFYGD
24_Gallus_gallus_Q98942	G <mark>F</mark> GG <mark>V</mark> Q <mark>V</mark> SP	D <mark>AVV</mark> NH	LIDIA	HL	GFRI D	AAKH	FIYQ E	VID	FVDNH D	GFT <mark>RV</mark> MSS <mark>Y</mark>
24_Homo_sapiens_P04746	G <mark>F</mark> GG <mark>V</mark> Q <mark>V</mark> S₽	DAVINH	LLDLA	HL	GFRL D	AS	FIYQ 🖻	VID	FVDNH D	GFTRVMSSY
27_Aeromonas_hydrophila_P22630	G <mark>yk</mark> QvliSP	DVVLNH	LPDLD	A	GFRV D	A <mark>VK</mark> H	H <mark>VF</mark> G E	VI T	FAITH D	GSPLVYSDH
27_Xanthomonas_campestris_Q56791	G <mark>yrkvlv</mark> ap	DVVFNH	LPDLL	A	GFRV D	AAK	YVFG E	VI T	FAVTH D	GVPMVYTDN
28 Bacillus subtilis P00691	GYTA <mark>I Q</mark> TSP	DAVINH	LYDWN	RA	GFRF D	AAKH	FQYG E	Ω	WVESH D	STPLFFSRP
28_Lactobacillus_amylovorus_Q48502	GYTAVQTSP	DATOND	FYDWN	81	GPRY D	AATH	EQXC E	[™] 2	WESHD	SVELFEDRE
32_Streptomyces_limosus_P09794	GYGYVQVSP	DSVINH	ADLD		GFRI D	AA	HGAG E	AVQ	FVDNH D	GSPDVHSGY
32_Inermonospora_curvata_P29750	GEGAVQVSP	DAVINH	AD	2	GE L D	AA	<u>19</u> 0	<u> A</u>	EVVNH D	GTESVMSSY
36_Dictyoglomus_thermophilum_P14899	N TALWIMP		MPDIN	FW	GFRL D	AA	YLVG E	WP		GIPFIYYGE
30_naiothermothrix_orenii_Q8GPL8	GVNG WLMP		rus on	New York	GE ST	CAM			TNE	GNEELYIGE
37 Photobacterium profundum Q6LIA8	CMNAVWL TP	DGVFGH		YW	G	QAYQ		WN		GPITLYYGD
22 Warehastering on NoO2 OPW20		DOVE OF								
TT FIGYODACERIIUM SD NO92 USANGU	1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1000		1 1 1 1 1 1 1 1 1	NP1 191 10	11-111-2	- CICINIAI	

CSR-VI

CSR-I

CSR-V

CSR-II

CSR-III

CSR-IV

CSR-VII

Figure 2. Sequence alignment of CSRs of studied family GH13 enzymes with focus on the novel α -amylase subfamily. The two consecutive tryptophans characteristic for the novel α -amylase subfamily are also shown. Colour code for the selected residues: W, yellow; F, Y - blue; V, L, I - green; D, E - red; R, K - cyan; H - brown; C - magenta; G, P - black. The positions of the three catalytic residues are boxed and signified by asterisks under the alignment. The label of the protein source consists of the name of the organisms and the UniProt (UniParc) accession number. The number at the beginning of the protein source label indicates the number of known (already established) GH13 subfamily. For the newly proposed BmaN1 GH13 subfamily, the label "xy" is used; similarly ("xx") for until now non-defined subfamily around the BaqA. The alignment of all 64 enzymes spanning the sequence segment from the beginning of the strand β 2 (CSR-VI) to the end of the strand β 8 (CSR-VII) is shown in Fig. S2.

oligo-1,6-glucosidases and neopullulanases²² that are closely related to α -amylases from the subfamily GH13_36 that, however, do not possess the tryptophan pair (Fig. 2).

A topological alignment of BmaN1 and the putative α -amylases of *B. megaterium* DSM319, *Bacillus* sp. 278922, *B. flexus*, *B. aryabhattai*, *B. megaterium* WSH-002, and GTA was made (Fig. S1). Almost all β -strands and α -helices of the TIM barrel in domain A and the Greek key motif in domain C are conserved in these α -amylases. A model of the 3D structure of BmaN1 was generated by the PHYRE server²³ and visualized by the MacPymol software²⁴ (Fig. 3). The BmaN1 protein displayed 40% homology (100% confidence, 85% sequence coverage) with the X-ray crystal structure of *Geobacillus thermoleovorans* α -amylase (GTA, PDB code: 4E2O)²⁵ which was used as a template for the modeling. The comparison between the model and the GTA crystal structure revealed that the global topology is almost the same (Fig. 3). The BmaN1 protein model folds into three distinct domains: a central A domain of 366 residues harboring a (β/α)₈ barrel, with an irregular loop domain of 37 residues (domain B) connecting the third β -sheets strand and the third α -helix of the barrel. The C domain of 79 residues has an eight-stranded anti-parallel β -sandwich-like fold (Fig. 3). The (β/α)₈ barrel is quite similar to the (β/α)₈-barrel found in maltogenic amylase from *Pseudomonas saccharophila* and *Bacillus stearothermophilus*²⁶, in that there is an additional helix between A α 6 and A β 7, which is a three-turn helix lying nearly parallel with the A α 6 strand.

Superposition of a carbose-bound GTA with the BmaN1 model demonstrated that of the three catalytic residues found in GH13 α -amylases, only residue Glu231 of BmaN1 superimposes with the corresponding residue in



Figure 3. Structural comparison of BmaN1 of *B. megaterium* NL3 and GTA of *Geobacillus thermoleovorans.* (A) BmaN1 model structure, (B) Structure of BmaN1 (orange) superimposed on GTA (grey) structure, (C) Active-site region in a superposition of BmaN1 with GTA including the acarbose bound in its subsites -2 to +2 (white carbon atoms). Active-center residues of BmaN1 (orange) and GTA (grey) are given as stick models and labeled in orange (BmaN1) and black (GTA).

GTA (Glu246), and presumably is the general acid/base in BmaN1 (Fig. 3). As already concluded from sequence alignments, two of the three catalytic residues are not conserved in BmaN1. Lys202 replaces the catalytic aspartate (Asp217 of GTA); however, Asp203 directly downstream of the lysine is positioned nearby and has its carboxylic acid side chain pointing into the presumed substrate-binding groove (Fig. 3). Furthermore, at the position corresponding to the nucleophile, His294 replaces the transition-state stabilizing aspartate residue (Asp314) found in α -amylases. The absence of any one of the catalytic residues normally causes partial or complete loss of hydrolysis activity²⁷. Remarkably, the mutant α -amylase from *Xanthomonas campestris* truncated from the C-terminal part of domain B and thus lacking any of the three conserved catalytic residues, still exhibited starch-hydrolyzing activity²⁸, but that observation has never been supported by other examples.

Phylogeny of BmaN1 and other α **-amylases.** The evolutionary relatedness of the α -amylase from *B. megaterium* BmaN1, representing all its homologues with lysine and histidine in positions of the catalytic nucleophile and transition state stabilizer, respectively (Fig. 2; Nos 1–27 in Fig. 1), to members of the recently proposed GH13 subfamily around the BaqA (Nos 35–39 in Fig. 1) as well as to representatives of remaining well-established GH13 subfamilies with α -amylase specificity (Nos 40–63 in Fig. 1), is shown in the evolutionary tree (Fig. 4). It is clear that both subfamilies BmaN1 and BaqA are most closely related to each other among all family GH13 α -amylases. Furthermore, a small group of putative α -amylases with irregular substitutions in catalytic positions (Nos 28–34 in Fig. 1) may be considered as an intermediary connection between both BmaN1 and BaqA subfamilies since, despite the lack of complete family GH13 catalytic machinery (similar to BmaN1), they cluster together with representatives of the BaqA subfamily (Fig. 4). One of the most convincing sequence-structural features characteristic for all these α -amylases is the presence of the pair of adjacent tryptophan residues in helix α 3 of the catalytic (β/α)₈-barrel (Fig. S2)¹⁹. Interestingly, the *Flavobacterium* sp. No. 92 cyclomaltodextrinase (with the tryptophan pair) is positioned in the evolutionary tree between the subfamilies of BmaN1 and BaqA and all other remaining GH13 families with the α -amylase specificity (Fig. 4).

BmaN1 encodes an active exo-acting α -amylase. The gene encoding BmaN1 was cloned in vector pMM1525 and this recombinant plasmid was transformed to *B. megaterium* MS941. A transformant with clear α -amylase activity, as detected on starch plates by iodine staining, was selected and grown in liquid medium. The culture medium was saturated with 50–80% concentrations of ammonium sulphate to purify the BmaN1 α -amylase enzyme. The molecular weight of the partially purified BmaN1 was estimated to be 55kDa as judged



Figure 4. Evolutionary tree of studied family GH13 enzymes with focus on the novel α -amylase subfamily around the BmaN1. The label of the protein source consists of the name of the organisms and the UniProt (UniParc) accession number preceded by GH13 subfamily indication. The tree is based on the alignment shown in Fig. S2.



Figure 5. Incubation of BmaN1 (open squares) with soluble starch; open circles: control (empty vector). 1% (w/v) soluble starch and 12.5 µg/mL of the BmaN1 protein were incubated for various time at 55 °C. Each data point represents the means of triplicate experiments.

from activity staining after protein renaturing on SDS-PAGE gels (Fig. S3). In contrast, no band was observed in the culture supernatant of *B. megaterium* MS941 carrying pMM1525 without any insert (Fig. S3). Amylolytic activity of BmaN1 was measured spectrophotometrically by incubating it with soluble starch and measuring the increase in the amount of reducing sugars released over a period of 40 min (Fig. 5). A clear increase in reducing ends was observed, resulting in an activity of 8.4 U/mg of protein. BmaN1 was found to be most active at 55 °C and pH 6.0. The main end products formed from soluble starch were glucose and maltose, indicating an exo-acting mode of action. Minor amounts of longer chain maltooligosaccharides were also found (Fig. 6). This





mode of action is very similar to that of the amylase from *Bacillus* sp. IMD 435 that releases glucose and maltose as the major products on hydrolysis of both soluble starch and raw corn starch²⁹.

The results presented above indicate that the substitution of an aspartate residue by a histidine, a positively charged amino acid, still gives (some) amylase activity. The reaction mechanism of BmaN1 may be essentially different from the general catalytic reaction mechanism of α -amylases. Further experiments are needed to demonstrate whether the His residue indeed is one of the catalytic residues of α -amylases.

Methods

Materials. All chemicals used were reagent grade and were obtained from either Fermentas (Maryland, USA) or Difco Laboratories (New Jersey, USA).

Bacterial strains, plasmids, and growth conditions. Twenty microbial strains (gift of Prof. Ocky Karna Radjasa of Diponegoro University, Indonesia) that had been isolated from Kakaban landlocked marine lake (Derawan Islands, East Kalimantan, Indonesia) were screened for extracellular α -amylase activity. The isolates were cultured in marine broth (MB) medium containing 0.25% (w/v) yeast extract, and 0.5% (w/v) peptone in filtered sea water (Seaworld, Ancol, Jakarta, Indonesia) at 30 °C. *B. megaterium* MS941 (MoBiTec, Germany) and *Escherichia coli* TOP10 were grown at 37 °C in LB medium (1% (w/v) Bacto-tryptone, 1% (w/v) NaCl and 0.5% (w/v) yeast extract). Ampicillin and tetracyclin were used at concentrations of 100µg/ml and 12µg/ml, respectively. The medium was autoclaved at 120 °C for 30 min prior to adding the antibiotics. Plasmid pGEM-T (Promega, USA) was used for PCR product cloning, whereas pMM1525 (MoBiTec, Germany) was used as expression vector.

Screening of α **-amylase producing bacteria.** Bacterial isolates were inoculated on MB agar plates supplemented with 1.0% (v/v) red-dyed amylopectin³⁰ and then incubated at 30 °C for 24 h. The appearance of a clear zone against a red background was indicative for the production of α -amylase activity. The positive isolates were then subjected to a second screening round using MB agar plates containing 1.0% (w/v) potato or wheat starch. A clearing zone around the bacterial colony indicated that the starch was hydrolyzed and thus that the isolate produced extracellular amylase activity.

Bacterial identification. The isolate showing the largest clearing zone on starch-agar plate was identified by 16SrDNA sequencing. Chromosomal DNA was isolated using Wizard Genomic DNA Purification (Promega). The 16S rDNA gene was amplified by PCR using universal primers UniB1 and BactF1 (Supplementary Table 1). The resulted 1.4 kb fragment was sequenced using the dideoxy-chain termination method (Macrogen, South Korea). The bacterial isolate was identified by aligning the 16s rDNA sequences with other known bacteria using NCBI BLASTn (http://www.ncbi.nlm.nih.gov). 16S rDNA gene sequence was submitted to GenBank.

Cloning of the α **-amylase-encoding gene and plasmid construction.** Two degenerate primers (Table S1) were designed based on the amino acid sequences of the well-conserved regions (region VI-VII) of α -amylases from several Bacilli. The first α -amylase gene fragment was amplified by polymerase chain reactions (PCR), using chromosomal DNA from *B. megaterium* NL3 as a template and the two degenerate primers. The PCR products were inserted into pGEM-T vector (Promega) and transformed into *E. coli* TOP10. Plasmid DNA of the transformed *E. coli* TOP10 was isolated and the nucleotide sequence of the inserted DNA was determined using the dideoxy-chain termination method (Macrogen). The resulting nucleotide sequence was used to design a set of primers, NL3_SP8-invF1 and NL3_SP8-invR1 (Table S1), to amplify parts of α -amylase gene beyond the conserved region. The chromosomal DNA was partially digested with *EcoRV* and then self-ligated using T4 DNA ligase (Fermentas). Inverse PCR (iPCR) was performed with Dream *Taq* polymerase (Fermentas) and the primers listed in Supplementary Table 1 using the self-ligated DNA fragment as a template. Analysis of sequence data and sequence similarity searches was performed using the BLAST program of the National Center for Biotechnology Information (NCBI). Primers pMM-NL3-F and pET/MM-NL3-R (Table S1) were used to amplify the complete open reading frame of the α -amylase gene which was designated as *bmaN1*.

Transformation of *B. megaterium.* The recombinant plasmid containing the α -amylase gene, pMM1525-bmaN1, was transformed into the expression host, *B. megaterium* MS941. The transformation procedure was essentially conducted as described by Puyet *et al.* with some modifications³¹. A 0.5 ml protoplast suspension was added to a tube containing 5.0 µg DNA and 1.5 ml PEG-P (40% (w/v) PEG6000 in 1x SMMP) for each transformation and incubated for 2 min at room temperature. SMMP medium contains 3.5% (w/v) AB3 (Antibiotic medium no. 3, Difco), 1 M sucrose, 40 mM disodium maleic acid and 40 mM MgCl₂ (pH adjusted to 6.5 before autoclaving for 12 min) and prepared freshly before use. To the mixture, 5.0 ml SMMP was added and mixed by rolling the tube carefully. Cells were harvested by centrifugation at 2,700 × g for 10 min at room temperature and the supernatant was poured off immediately. The pellets were resuspended with 0.5 ml SMMP and incubated at 37 °C for 90 min with gentle shaking or rolling of tubes (max. 100 rpm). Then, 50 to 200 µl of cells were added into top agar and mixed gently by rolling the tube. The mixture was poured on a pre-warmed plate of LB containing 12 µg/ml of tetracycline and incubated at 37 °C overnight.

Expression and partial purification of recombinant α **-amylase.** α -Amylase was produced by growing the *B. megaterium* harboring recombinant plasmids in 20 ml of LB medium supplemented with 12 µg/ml tetracycline at 37 °C with shaking. The overnight culture was transferred into fresh media and incubated until the 546-nm absorbance reached 0.8–1.0. Subsequently, expression was induced by adding 1% (v/v) xylose, and the culture was incubated at 18 °C with constant shaking at 150 rpm for 24 h. Cells were removed by centrifugation (6000 × g, 10 min) and the resulted supernatant was subjected to ammonium sulphate precipitation at a saturation value up to 80%. The precipitate was dissolved and dialyzed against 50 mM maleate buffer pH 6.0 at 4 °C. This partially purified α -amylase was used for further studies.

Gel electrophoresis and activity staining. SDS-PAGE was carried out as described by Laemmli³² and gels were then stained with Coomassie Brilliant Blue (Bio-Rad). For α -amylase activity test, the protein samples were separated by SDS-PAGE containing 1% soluble starch. After electrophoresis, SDS was removed by washing the gel with water followed by 10 min incubation at room temperature. This was repeated twice. The gels were then immersed in the enzyme reaction buffer (50 mM maleate buffer pH 6.0) for 4.0 h at 55 °C and then stained with KI/I₂ solution for 10 min and followed by rinsing with water. The α -amylase activity was detected as a clear zone against a purple background.

Enzyme assay. The amylase activity assay was conducted using the 3,5-dinitrosalicylic acid (DNS) method described by Miller (1955) with a slight modification³³. Briefly, the assay was performed by adding $25 \,\mu$ l of enzyme sample into $25 \,\mu$ l of 1% (w/v) soluble starch (Fermentas, USA) in 50 mM of the appropriate buffer and then incubated at 55 °C for 10 min. To the reaction mixture, 50 μ l of DNS reagent was added. The absorbance at 500 nm was measured and then the amount of reducing sugar-end was calculated using a glucose standard curve. One unit of α -amylase activity was defined as the amount of enzyme that releases 1 μ mol of reducing sugar per min under the assay conditions. The protein concentration was determined using the Lowry method and bovine serum albumin as the standard.

Analysis of sugars. The starch hydrolysis products were analyzed by high-performance liquid chromatography (HPLC). Aliquots of $100 \,\mu$ l of enzyme solution were incubated at 55 °C in the presence of 1% (w/v) soluble starch, maleate buffer 50 mM. After specific time intervals, samples were withdrawn and hydrolysis was stopped by incubation at 90 °C for 10 min. After centrifugation at $12000 \times g$ for 10 min at 4 °C, the products were then analyzed by HPLC (Aminex[®] HPX-87H system). The separated hydrolysis products were identified by calculating based on peak areas compared to standard glucose, maltose, and purified maltooligosaccharide (Sigma).

Bioinformatics. The sequences eventually forming the new GH13 subfamily xy were collected based on protein BLAST³⁴ searches against the non-redundant database using the entire amino acid sequence of *Bacillus megaterium* NL3 α -amylase BmaN1 (UniProt accession No.: T1SIF2) as well as on previous bioinformatics analyses when the BLAST was performed with the *Bacillus aquimaris* α -amylase BaqA^{19,35}. The main criterion applied for the selection was the lack of at least one residue from the catalytic triad characteristic for the α -amylase family GH13. In addition to BmaN1 and its closely related homologues, five experimentally characterized α -amylases from the recently proposed GH13 subfamily around the *B. aquimaris* α -amylase BaqA^{19,35} were added. These α -amylases – BaqA from *B. aquimaris*³⁵, ASKA and ADTA from *Anoxybacillus* sp.^{36,37}, GTA and GTA-II from *Geobacillus thermoleovorans*^{25,38,39} – exhibit a high degree of sequence similarity with BmaN1, but possess the

complete GH13 catalytic machinery¹⁹. The entire set was finally completed by two selected representatives from well-established GH13 subfamilies with the α -amylase specificity, i.e. 1, 5, 6, 7, 15, 19, 24, 27, 28, 32, 36 and 37¹⁹ including also the related but until now unclassified cyclomaltodextrinase from *Flavobacterium* sp. No. 192²¹ so that the final number of studied enzymes and hypothetical proteins was 64 (Fig. 1).

All 64 GH13 sequences were retrieved from GenBank¹⁷ and UniProt⁴⁰ sequence databases and the set was aligned using the program Clustal-Omega⁴¹ available at the European Bioinformatics Institute's web-site (http:// www.ebi.ac.uk/). A subtle manual tuning was done in order to maximize similarities, especially with regard to aligning the individual CSRs. The boundaries of the CSRs were defined based on previous bioinformatics studies^{19,20}. The evolutionary tree was constructed based on the final alignment of the sequence segment corresponding to a 255-residue long region of BmaN1 α -amylase spanning almost the entire catalytic (β/α)₈-barrel domain including the domain B from the beginning of the CSR-VI (strand β 2; starting with Gly79) to the end of the CSR-VII (strand β 8; ending with Ser333). The tree was calculated as a Phylip-tree type using the neighbour-joining clustering and the bootstrapping procedure - the number of bootstrap trials used was 1,000⁴² implemented in the Clustal-X package⁴³, and then displayed with the program iTOL⁴⁴.

The 3D structure of BmaN1 was predicted by QuickPhyre structure program server (http://www.sbg.bio.ic.ac. uk/phyre)²³. Structural modeling of the BmaN1 was performed based on the crystal structures of α -amylase of *G. thermoleovorans* [PDB access code: 4E20]. The generated BmaN1 structures were displayed and drawn by MacPymol.

References

- Van der Maarel, M. J. E. C., van der Veen, B., Uitdehaag, J. C. M., Leemhuis, H. & Dijkhuizen, L. Properties and applications of starch-converting enzymes of the α-amylase family. J. Biotechnol. 94, 137–155 (2002).
- 2. Butterworth, P. J., Warren, F. J. & Ellis, P. R. Human α-amylase and starch digestion: An interesting marriage. *Starch Stärke* 63, 395–405 (2011).
- 3. Zeeman, S. C., Kossmann, J. & Smith, A. M. Starch: its metabolism, evolution and biotechnological modification in plants. *Ann. Rev. Plant Biol.* **61**, 209–234 (2010).
- Cantarel, B. L. et al. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res. 37, D233–238 (2009).
- MacGregor, E. A., Janeček, Š. & Svensson, B. Relationship of sequence and structure to specificity in the α-amylase family of enzymes. Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol. 1546, 1–20 (2001).
- Matsuura, Y., Kusunoki, M., Harada, W. & Kakudo, M. Structure and possible catalytic residues of Taka-amylase A. J. Biochem. 95, 697–702 (1984).
- Janeček, Š., Svensson, B. & MacGregor, E. A. α-Amylase: an enzyme specificity found in various families of glycoside hydrolases. Cell. Mol. Life Sci. 71, 1149–1170 (2014).
- 8. MacGregor, E. A. α-Amylase structure and activity. J. Protein Chem. 7, 399–415 (1988).
- Janeček, Š. & Kuchtová, A. In silico identification of catalytic residues and domain fold of the family GH119 sharing the catalytic machinery with the α-amylase family GH57. FEBS Lett. 586, 3360–3366 (2012).
- Blesák, K. & Janeček, Š. Two potentially novel amylolytic enzyme specificities in the prokaryotic glycoside hydrolase α-amylase family GH57. *Microbiology* 159, 2584–2593 (2013).
- 11. Henrissat, B. & Davies, G. Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* 7, 637–644 (1997).
- 12. Imamura, H. *et al.* Crystal structures of 4-α-glucanotransferase from *Thermococcus litoralis* and its complex with an inhibitor. *J. Biol. Chem.* **278**, 19378–19386 (2003).
- Zona, R., Chang-Pi-Hin, F., O'Donohue, M. J. & Janeček, Š. Bioinformatics of the glycoside hydrolase family 57 and identification of catalytic residues in amylopullulanase from *Thermococcus hydrothermalis. Eur. J. Biochem.* 271, 2863–2872 (2004).
- 14. Radjasa, O. K., Limantara, L. & Sabdono, A. Antibacterial activity of a pigment producing-bacterium associated with *Halimeda sp.* from island-locked marine lake kakaban, Indonesia. J. Coast. Dev. **12**, 100–104 (2019).
- Hueck, C. J. & Hillen, W. Catabolite repression in *Bacillus subtilis*: a global regulatory mechanism for the Gram-positive bacteria? *Mol. Microbiol.* 15, 395–401 (1995).
- Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Meth 8, 785–786 (2011).
- 17. Benson, D. A. et al. GenBank. Nucleic Acids Res. 43, D30-5 (2015).
- 18. Lombard, V. et al. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, D490-D495 (2013).
- Janeček, Š., Kuchtová, A. & Petrovičová, S. A novel GH13 subfamily of α-amylases with a pair of tryptophans in the helix α3 of the catalytic TIM-barrel, the LPDlx signature in the conserved sequence region V and a conserved aromatic motif at the C-terminus. *Biologia* 70, 1284–1294 (2015).
- 20. Janeček, Š. How many conserved sequence regions are there in the α -amylase family? *Biologia* 57 (Suppl. 11), 29–41 (2002).
- 21. Fritzsche, H. B., Schwede, T. & Schulz, G. E. Covalent and three-dimensional structure of the cyclodextrinase from *Flavobacterium sp.* no. 92. *Eur. J. Biochem.* **270**, 2332–2341 (2003).
- Majzlová, K., Pukajová, Z. & Janeček, Š. Tracing the evolution of the α-amylase subfamily GH13_36 covering the amylolytic enzymes intermediate between oligo-1,6-glucosidases and neopullulanases. *Carbohydr. Res.* 367, 48–57 (2013).
- 23. Kelley, L. A. et al. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10, 845-858 (2015).
 - 24. Schrodinger LLC. The PyMOL molecular graphics system, Version 1.3r1 (2010).
 - Mok, S.-C. et al. Crystal structure of a compact α-amylase from Geobacillus thermoleovorans. Enzyme Microb. Technol. 53, 46–54 (2013).
 - Jespersen, H. M., MacGregor, E. A., Sierks, M. R. & Svensson, B. Comparison of the domain-level organization of starch hydrolases and related enzymes. *Biochem. J.* 280, 51–55 (1991).
 - Nielsen, J. E. & Borchert, T. V. Protein engineering of bacterial α-amylases. Biochim. Biophys. Acta Protein Struct. Mol. Enzymol. 1543, 253–274 (2000).
 - 28. Ke, T. *et al.* A mutant α-amylase with only part of the catalytic domain and its structural implication. *Biotechnol. Lett.* **29**, 117–122 (2007).
 - Hamilton, L. M., Kelly, C. T. & Fogarty, W. M. Production and properties of the raw starch-digesting α-amylase of *Bacillus sp.* IMD 435. *Process Biochem.* 35, 27–31 (1999).
 - Jørgensen, S., Vorgias, C. E. & Antranikian, G. Cloning, sequencing, characterization, and expression of an extracellular α-amylase from the hyperthermophilic archaeon *Pyrococcus furiosus* in *Escherichia coli* and *Bacillus subtilis. J. Biol. Chem.* 272, 16335–16342 (1997).

- Puyet, A. et al. A simple medium for rapid regeneration of Bacillus subtilis protoplasts transformed with plasmid DNA. FEMS Microbiol. Lett. 40, 1–5 (1987).
- 32. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685 (1970).
- 33. Miller, G. L. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal. Chem. 31, 426–428 (1959).
- 34. Altschul, S. F. et al. Basic local alignment search tool. J. Mol. Biol. 215, 403-10 (1990).
- 35. Puspasari, F. *et al.* Raw starch-degrading α-amylase from *Bacillus aquimaris* MKSC 6.2: isolation and expression of the gene, bioinformatics and biochemical characterization of the recombinant enzyme. *J. Appl. Microbiol.* **114**, 108–120 (2013).
- 36. Chai, Y. Y., Rahman, R. N., Illias, R. M. & Goh, K. M. Cloning and characterization of two new thermostable and alkalitolerant α-amylases from the *Anoxybacillus* species that produce high levels of maltose. *J. Ind. Microbiol. Biotechnol.* **39**, 731–741 (2012).
- 37. Chai, K. P. *et al.* Crystal structure of *Anoxybacillus* α -amylase provides insights into maltose binding of a new glycosyl hydrolase subclass. *Sci. Rep.* **6**, 23126 (2016).
- Finore, I. et al. Purification, biochemical characterization and gene sequencing of a thermostable raw starch digesting α-amylase from *Geobacillus thermoleovorans* subsp. stromboliensis subsp. nov. World J. Microbiol. Biotechnol. 27, 2425–2433 (2011).
- Mehta, D. & Satyanarayana, T. Domain C of thermostable α-amylase of *Geobacillus thermoleovorans* mediates raw starch adsorption. Appl. Microbiol. Biotechnol. 98, 4503–4519 (2014).
- 40. UniProt Consortium. UniProt: a hub for protein information. Nucleic Acids Res. 43, D204-12 (2015).
- 41. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539 (2011).
- 42. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution (N. Y). 39, 783-791 (1985).
- 43. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947-2948 (2007).
- 44. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23, 127–128 (2007).

Acknowledgements

This research was funded by Asahi Glass Foundation, Institut Teknologi Bandung. We thank Prof. F. Meinhardt, University of Munster, Germany, for *B. megaterium* MS941 strain and pMM1525 plasmid. We are grateful to Dr. Lily Eurwilaichitr and Dr. Sutipa Tanapongpipat for allowing us to do the preliminary experiments on BIOTEC, NSTDA, Thailand. ŠJ thanks the Slovak Grant Agency VEGA – the project No. 2/0150/14 – for financial support.

Author Contributions

E.D.S., I., Z.N., D.N., and O.K.R. planned the experiments on the isolation and characterization of the strain and the isolation of the amylase gene and analysed the results; F.D.S., D.N. and M.M. planned the biochemical analysis of the amylase and analysed the results; F.D.S. performed all these experiments; F.D.S., D.N., S.J., T.P., L.D. and M.M. wrote the manuscript; F.D.S. and T.P. performed the 3D modeling and interpreted the results; F.D.S. and S.J. performed the phylogenetic analysis and interpreted the results; all authors have seen the final version of the manuscript and agree with the content.

Additional Information

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: Sarian, F. D. *et al.* A new group of glycoside hydrolase family 13 α -amylases with an aberrant catalytic triad. *Sci. Rep.* 7, 44230; doi: 10.1038/srep44230 (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/

© The Author(s) 2017