

## Supporting Information

### Structure of silkworm *Bombyx mori* sucrose hydrolase and substrate specificity of GH13 subfamily 17 *exo- $\alpha$* -glucosidases

Takatsugu Miyazaki\*, and Enoch Y. Park

Green Chemistry Research Division, Research Institute of Green Science and Technology, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka, 422-8529, Japan.

\*To whom correspondence should be addressed: Takatsugu Miyazaki, Green Chemistry Research Division, Research Institute of Green Science and Technology, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka, 422-8529, Japan; Tel & Fax: +81-54-238-4886; E-mail: [miyazaki.takatsugu@shizuoka.ac.jp](mailto:miyazaki.takatsugu@shizuoka.ac.jp)

#### **This file includes:**

**Supplementary Table S1.** Enzymes structurally homologous to BmSUH.

**Supplementary Table S2.** Sequences of oligonucleotides used in this study.

**Supplementary Table S3.** GH13\_17 proteins used in the phylogenetic analysis in Figure 6.

**Supplementary Figure S1.** Effects of pH and temperature on the recombinant BmSUH hydrolysis of sucrose.

**Supplementary Figure S2.** Structures of *exo*-acting GH13 enzymes homologous to BmSUH.

**Supplementary Figure S3.** Metal binding sites of BmSUH.

**Supplementary Figure S4.** Dimeric structures of BmSUH and Cqm1.

**Supplementary Figure S5.** Sequence alignment of GH13\_17 enzymes.

**Supplementary Figure S6.** Sequence alignment of *exo*-acting enzymes belonging to GH13\_16, 17, 23, 29, 31, and 40 subfamilies.

**Supplementary Table S1. Enzymes structurally homologous to BmSUH.**

Enzyme	GH family	PDB code	Z score	RMSD (Å)	Sequence identity (%)
<i>Culex quinquefasciatus</i> maltase Cqml	GH13_17	6K5P	54.1	1.3	44
<i>Listeria monocytogenes</i> $\alpha$ -1,6-glucosidase	GH13_31	5DO8	47.5	2.4	33
<i>Pseudomonas mesoacidophila</i> sucrose isomerase	GH13_31	2PWH	47.2	2.2	32
<i>Bacillus subtilis</i> oligo- $\alpha$ -1,6-glucosidase	GH13_31	4M56	46.3	2.3	32
<i>Xanthomonas campestris</i> $\alpha$ -glucosyltransferase	GH13_23	6AAV	44.6	2.3	33
<i>Bacillus</i> sp. AHU2216 $\alpha$ -glucosidase	GH13_31	5ZCD	44.4	2.1	33
<i>Halomonas</i> sp. H11 $\alpha$ -glucosidase	GH13_23	3WY1	44.2	2.3	33
<i>Lactobacillus acidophilus</i> glucan $\alpha$ -1,6-glucosidase	GH13_31	4AIE	43.8	2.1	32
<i>Deinococcus radiodurans</i> trehalose synthase	GH13_16	4WF7	43.7	2.6	29
<i>Streptococcus mutans</i> dextran $\alpha$ -1,6-glucosidase	GH13_31	4WLC	43.5	2.2	32
<i>Klebsiella</i> sp. LX3 isomaltulose synthase	GH13_31	1M53	42.4	2.3	31
<i>Protaminobacter rubrum</i> sucrose isomerase	GH13_31	3GBE	41.8	2.3	32
<i>Mycolicibacterium smegmatis</i> trehalose synthase	GH13_16	5JY7	41.7	2.7	29
<i>Bacillus cereus</i> oligo- $\alpha$ -1,6-glucosidase	GH13_31	1UOK	41.4	2.2	33
<i>Erwinia rhapontici</i> isomaltulose synthase	GH13_31	4HP5	41.2	2.3	31
<i>Thermomonospora curvata</i> trehalose synthase	GH13_16	5H2T	41.1	2.7	30
<i>Bacillus licheniformis</i> trehalose-6-phosphate hydrolase	GH13_29	5BRP	41.0	2.6	32
<i>Thermobaculum terrenum</i> trehalose synthase	GH13_16	5X7U	40.7	2.5	28
<i>Geobacillus</i> sp. HTA-462 $\alpha$ -glucosidase	GH13_31	2ZE0	40.4	2.4	32
<i>Saccharomyces cerevisiae</i> oligo- $\alpha$ -1,6-glucosidase	GH13_40	3AXH	39.8	2.7	29
<i>Bacteroides thetaiotaomicron</i> $\alpha$ -amylase	GH13_36	3K8I	39.8	2.5	24
<i>Arthrobacter globiformis</i> cyclic maltosyl-maltose hydrolase	GH13_20	5ZXG	38.4	2.5	23
<i>Thermoactinomyces vulgaris</i> $\alpha$ -amylase TVAI	GH13_20	1JIB	38.1	2.5	25
<i>Thermotoga maritima</i> 4- $\alpha$ -glucanotransferase	GH13_NC <sup>a</sup>	1LWJ	37.1	2.6	24
<i>Anoxybacillus</i> sp. SK3-4 $\alpha$ -amylase	GH13_NC	5A2A	36.9	2.5	23
<i>Neisseria polysaccharea</i> amylosucrase	GH13_4	1MW1	36.8	2.5	21
<i>Geobacillus stearothermophilus</i> neopullulanase	GH13_20	1J0I	36.6	2.4	24
<i>Thermus</i> sp. IM6501 maltogenic $\alpha$ -amylase	GH13_20	1GVI	36.3	2.5	23
<i>Thermus thermophilus</i> HB8 TTHA1563	GH13_20	2Z1K	35.8	2.7	20
<i>Bacillus</i> sp. I-5 cyclomaltodextrinase	GH13_20	1EA9	35.4	2.5	22
<i>Thermoactinomyces vulgaris</i> $\alpha$ -amylase TVAI	GH13_21	2D0G	35.2	2.8	21
<i>Lactobacillus plantarum</i> $\alpha$ -amylase	GH13_NC	3DHU	34.2	2.4	20
<i>Pyrococcus furiosus</i> $\alpha$ -amylase	GH13_20	4AEF	33.9	2.5	24
<i>Streptomyces coelicolor</i> $\alpha$ -1,4-glucan: phosphate $\alpha$ -maltosyltransferase	GH13_3	3ZSS	33.7	2.8	16
<i>Thermococcus kodakarensis</i> amylopullulanase	GH13_20	5OT1	33.4	2.6	23
<i>Mycobacterium tuberculosis</i> $\alpha$ -1,4-glucan: phosphate $\alpha$ -maltosyltransferase	GH13_3	4U33	33.2	2.9	13
<i>Thermotoga petrophila</i> $\alpha$ -amylase	GH13_36	5M99	33.1	2.9	23
<i>Xanthomonas axonopodis</i> pv. <i>glycines</i> sucrose hydrolase	GH13_4	3CZL	31.3	2.6	19
<i>Bifidobacterium adolescentis</i> sucrose phosphorylase	GH13_18	6FME	29.7	2.5	18
<i>Xanthomonas campestris</i> pv. <i>campestris</i> sucrose hydrolase	GH13_4	2WPG	27.7	2.7	20

<sup>a</sup> Not classified into any subfamilies according to the CAZy database.

**Supplementary Table S2. Sequences of oligonucleotides used in this study.**

Oligonucleotide	Sequence (5'→3') <sup>a</sup>
BmSUH_F	TTTT <u>CATATG</u> TCAAATCAAATGCGCCAAC
BmSUH_R	TTTT <u>CTCGAG</u> TTATATAGGTTGGGCTTTCA
Q191V_F	GGTCAGTGTGTTTCGGAGGTTTCGGCCTGGGA
Q191V_R	TCCGAACACACTGACCCAATTAGAAGGAGG
D247N_F	TCAGACTAAATGCTCTACCGTATTTGATTG
D247N_R	TAGAGCATTTAGTCTGAAGCCGTCAGCTCC
Y251H_F	TCTACCGCATTTGATTGAAGCTGATCCAGC
Y251H_R	ATCAAATGCGGTAGAGCATCTAGTCTGAAG
E322Q_F	TGTTCTCACAAGGGTATGCCAACGTGTCTA
E322Q_R	ATACCCTTGTGAGAACCACTCTGGTATC
E440A_F	GGACATAGCAGCCTGTAACCGGGGAGACC
E440A_R	TACAGGCTGCTATGTCCACAGTGTCTTCC

<sup>a</sup> Restriction sites are underlined.

**Supplementary Table S3. GH13\_17 proteins used in the phylogenetic analysis in Figure 6.**

Name <sup>a</sup>	Species	GenBank accession number
APS1	<i>Acyrtosiphon pisum</i>	ABB55878.1
AAEL000392-PA	<i>Aedes aegypti</i>	EAT48590.1
AAEL000667-PA	<i>Aedes aegypti</i>	EAT48296.1
AAEL003434-PA	<i>Aedes aegypti</i>	EAT45278.1
AAEL009524-PA	<i>Aedes aegypti</i>	EAT38606.1
AAEL010532-PA	<i>Aedes aegypti</i>	EAT37483.1
AAEL010536-PA	<i>Aedes aegypti</i>	EAT37482.1
<b>Aam1 (AAEL010537-PA)</b>	<i>Aedes aegypti</i>	EAT37485.1
AAEL014710-PA	<i>Aedes aegypti</i>	EAT33036.1
AtSUH	<i>Amyelois transitella</i>	XP_013192965.1
AtMal1	<i>Amyelois transitella</i>	XP_013201068.1
AtMal2	<i>Amyelois transitella</i>	XP_013201277.1
AtMal3_1	<i>Amyelois transitella</i>	XP_013201279.1
AtMal3_2	<i>Amyelois transitella</i>	XP_013201280.1
AtMal4	<i>Amyelois transitella</i>	XP_013201278.1
alpha-amylase	<i>Anopheles albimanus</i>	ADD81256.1
alpha-amylase	<i>Anopheles darlingi</i>	ETN61169.1
alpha-amylase	<i>Anopheles darlingi</i>	ETN63249.1
alpha-glucosidase	<i>Anopheles darlingi</i>	ETN63250.1
alpha-glucosidase	<i>Anopheles darlingi</i>	ETN65874.1
alpha-amylase	<i>Anopheles darlingi</i>	ETN65875.1
alpha-amylase	<i>Anopheles darlingi</i>	ETN67705.1
Agm1	<i>Anopheles gambiae</i>	CAA60857.1
Agm2	<i>Anopheles gambiae</i>	CAA60858.1
Agm3	<i>Anopheles gambiae</i>	ABW98683.1
alpha-amylase	<i>Antheraea pernyi</i>	AJR29308.1
AciHBG-I	<i>Apis cerana indica</i>	AGE11865.1
<b>AciHBG-II</b>	<i>Apis cerana indica</i>	AFR59339.1
AciHBG-III	<i>Apis cerana indica</i>	ABO27432.1
<b>JBgI</b>	<i>Apis cerana japonica</i>	BAF44218.1
JBgII	<i>Apis cerana japonica</i>	ACN63343.1
JBgIII	<i>Apis cerana japonica</i>	ACQ45697.1
AdMal	<i>Apis dorsata</i>	ADB11049.1
<b>HBG-I</b>	<i>Apis mellifera</i>	BAE86926.1
<b>HBG-II</b>	<i>Apis mellifera</i>	BAE86927.1
<b>HBG-III</b>	<i>Apis mellifera</i>	BAA11466.1
BdMalA1	<i>Bactrocera dorsalis</i>	AIT38256.1
alpha-glucosidase	<i>Bombus ignitus</i>	BAI44030.1
alpha-glucosidase	<i>Bombus diversus diversus</i>	BAI44029.1
BmaSUH	<i>Bombyx mandarina</i>	XP_028038901.1
BmaSUH2	<i>Bombyx mandarina</i>	XP_028038902.1
BmaMal3	<i>Bombyx mandarina</i>	XP_028034157.1

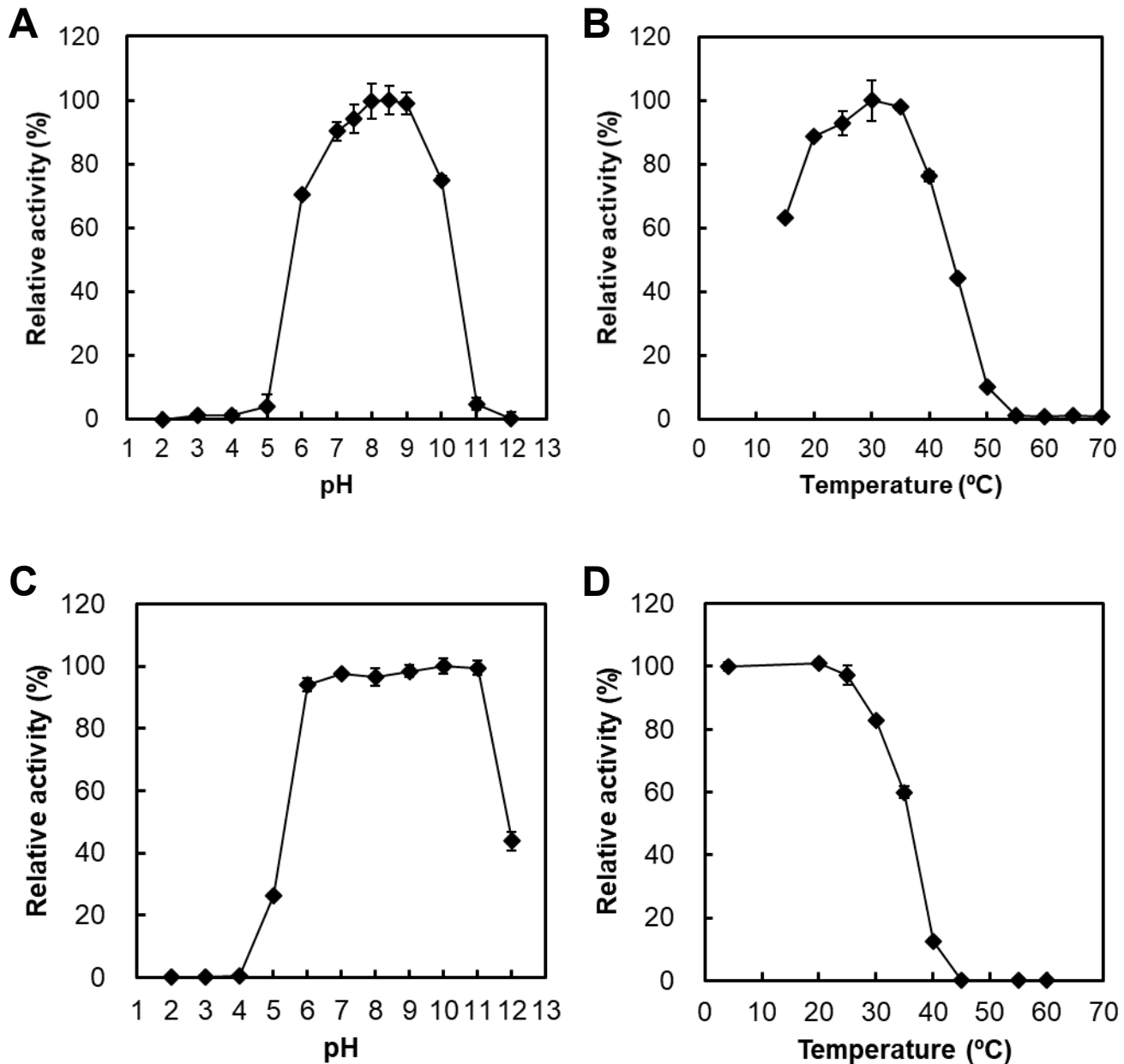
BmaMal4	<i>Bombyx mandarina</i>	XP_028034158.1
<b>BmSUH</b>	<i>Bombyx mori</i>	BAP18683.1
BmSUH2	<i>Bombyx mori</i>	BBE00602.1
BmMal2	<i>Bombyx mori</i>	XP_012549946.2
BmMal3	<i>Bombyx mori</i>	BBD49474.1
BmMal4	<i>Bombyx mori</i>	BBD13387.1
alpha-glucosidase	<i>Camponotus japonicus</i>	BAO48210.1
Cpm1	<i>Culex pipiens</i>	AAL05443.1
<b>Cqm1</b>	<i>Culex quinquefasciatus</i>	ASO96882.1
maltase	<i>Culicoides nubeculosus</i>	ACM40914.1
Cul o 1 allergen	<i>Culicoides obsoletus</i>	AGI16774.1
maltase	<i>Culicoides sonorensis</i>	AAU06480.1
venom polypeptide	<i>Dolopus genitalis</i>	AYV99560.1
DpSUH1	<i>Danaus plexippus plexippus</i>	OWR49826.1
DpSUH2	<i>Danaus plexippus plexippus</i>	OWR49828.1
DpMal1	<i>Danaus plexippus plexippus</i>	OWR49357.1
DpMal2	<i>Danaus plexippus plexippus</i>	OWR49126.1
DpMal3	<i>Danaus plexippus plexippus</i>	OWR49124.1
DpMal4	<i>Danaus plexippus plexippus</i>	OWR49125.1
DbMalA1	<i>Drosophila busckii</i>	ALC40924.1
DbMalA2	<i>Drosophila busckii</i>	ALC40926.1
DbMalA3	<i>Drosophila busckii</i>	ALC40927.1
DbMalA4	<i>Drosophila busckii</i>	ALC40928.1
DbMalA5	<i>Drosophila busckii</i>	ALC40929.1
DbMalA6	<i>Drosophila busckii</i>	ALC40930.1
DbMalA7	<i>Drosophila busckii</i>	ALC40931.1
DbMalA8	<i>Drosophila busckii</i>	ALC40932.1
DbMalB1	<i>Drosophila busckii</i>	ALC39348.1
DmMalA1	<i>Drosophila melanogaster</i>	AAF59089.3
DmMalA2	<i>Drosophila melanogaster</i>	AAF59088.1
DmMalA3	<i>Drosophila melanogaster</i>	AAM50308.1
DmMalA4	<i>Drosophila melanogaster</i>	AAF59086.1
DmMalA5	<i>Drosophila melanogaster</i>	AAV36953.1
DmMalA6	<i>Drosophila melanogaster</i>	AAM68846.2
DmMalA7	<i>Drosophila melanogaster</i>	AAF59084.2
DmMalA8	<i>Drosophila melanogaster</i>	AAF59083.2
DmMalB1	<i>Drosophila melanogaster</i>	AAF53127.1
DmMalB2	<i>Drosophila melanogaster</i>	AAF53128.4
DvMalB1	<i>Drosophila virilis</i>	AAB82327.1
DvMalB2	<i>Drosophila virilis</i>	AAB82328.1
maltase 1-like protein	<i>Habrobracon hebetor</i>	AXY94681.1
HaSUH1	<i>Helicoverpa armigera</i>	XP_021191332.1
HaSUH2	<i>Helicoverpa armigera</i>	XP_021191328.1
HaMal1_1	<i>Helicoverpa armigera</i>	XP_021188630.1
HaMal1_2	<i>Helicoverpa armigera</i>	XP_021188631.1

HaMal2	<i>Helicoverpa armigera</i>	XP_021193832.1
HaMal3	<i>Helicoverpa armigera</i>	XP_021193831.1
HaMal4	<i>Helicoverpa armigera</i>	XP_021193829.1
MsSUH	<i>Manduca sexta</i>	XP_030027415.1
MsSUH2	<i>Manduca sexta</i>	XP_030027440.1
MsMal1	<i>Manduca sexta</i>	XP_030034371.1
MsMal2	<i>Manduca sexta</i>	XP_030035809.1
MsMal3	<i>Manduca sexta</i>	XP_030031300.1
MsMal4	<i>Manduca sexta</i>	XP_030035812.1
MdMalA1	<i>Musca domestica</i>	AFP60283.1
MdMalA2	<i>Musca domestica</i>	AFP59562.1
MdMalA4	<i>Musca domestica</i>	AFP62365.1
MdMalA7	<i>Musca domestica</i>	AFP62166.1
MdMalB1	<i>Musca domestica</i>	AFP61790.1
alpha-amylase	<i>Musca domestica</i>	AFP62021.1
sucrase	<i>Myzus persicae</i>	AKM95031.1
NIMalA1	<i>Nilaparvata lugens</i>	AQW43009.1
NIMalA2	<i>Nilaparvata lugens</i>	ANJ04656.1
NISUH	<i>Nilaparvata lugens</i>	AEO89448.1
PmSUH1	<i>Papilio machaon</i>	XP_014360940.1
PmSUH2	<i>Papilio machaon</i>	XP_014360941.1
PmMal1	<i>Papilio machaon</i>	XP_014371748.1
PmMal2	<i>Papilio machaon</i>	XP_014368521.1
PmMal3_1	<i>Papilio machaon</i>	XP_014368517.1
PmMal3_2	<i>Papilio machaon</i>	XP_014368519.1
PmMal4	<i>Papilio machaon</i>	XP_014368518.1
<b>PpSUH1</b>	<i>Papilio polytes</i>	XP_013147169.1
<b>PpSUH2</b>	<i>Papilio polytes</i>	XP_013147171.1
PpMal1	<i>Papilio polytes</i>	XP_013137112.1
PpMal3_1	<i>Papilio polytes</i>	XP_013136071.1
PpMal3_2	<i>Papilio polytes</i>	XP_013136068.1
PpMal4	<i>Papilio polytes</i>	XP_013136069.1
PxSUH1	<i>Papilio xuthus</i>	XP_013161261.1
PxSUH2	<i>Papilio xuthus</i>	XP_013161263.1
PxMal1	<i>Papilio xuthus</i>	XP_013177984.1
PxMal2	<i>Papilio xuthus</i>	XP_013168325.1
PxMal3_1	<i>Papilio xuthus</i>	XP_013168326.1
PxMal3_2	<i>Papilio xuthus</i>	XP_013168328.1
PxMal4	<i>Papilio xuthus</i>	XP_013168327.1
alpha-glucosidase	<i>Periplaneta americana</i>	AIA09343.1
alpha-amylase	<i>Pristhesancus plagipennis</i>	ATU82842.1
alpha-amylase	<i>Riptortus pedestris</i>	BAN20159.1
predicted protein	<i>Riptortus pedestris</i>	BAN20192.1
<b>ScSUH</b>	<i>Samia cynthia ricini</i>	BAP18685.1
ScMal1	<i>Samia cynthia ricini</i>	BBD13389.1

ScMal2	<i>Samia cynthia ricini</i>	BBD49476.1
SISUH1	<i>Spodoptera litura</i>	XP_022821249.1
SISUH2	<i>Spodoptera litura</i>	XP_022821250.1
SIMal1_1	<i>Spodoptera litura</i>	XP_022816388.1
SIMal1_2	<i>Spodoptera litura</i>	XP_022816674.1
SIMal3	<i>Spodoptera litura</i>	XP_022820326.1
SIMal4	<i>Spodoptera litura</i>	XP_022820325.1
<b>TvSUH</b>	<i>Trilocha varians</i>	BAP18684.1
TvMal1	<i>Trilocha varians</i>	BBD13388.1
TvMal2	<i>Trilocha varians</i>	BBD49475.1

---

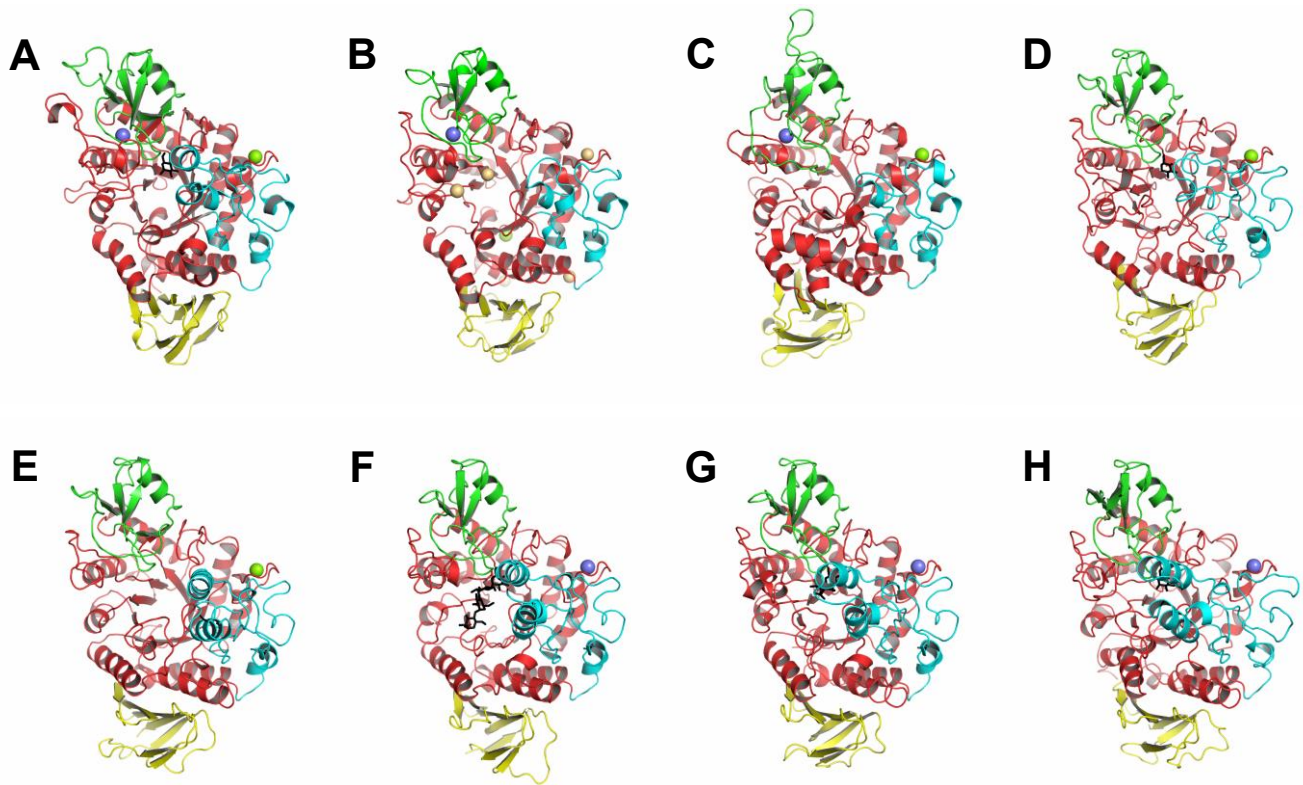
<sup>a</sup> Enzymatically characterized proteins (except for partially purified enzymes) are highlighted in bold.



**Supplementary Figure S1. Effects of pH and temperature on the recombinant BmSUH hydrolysis of sucrose.**

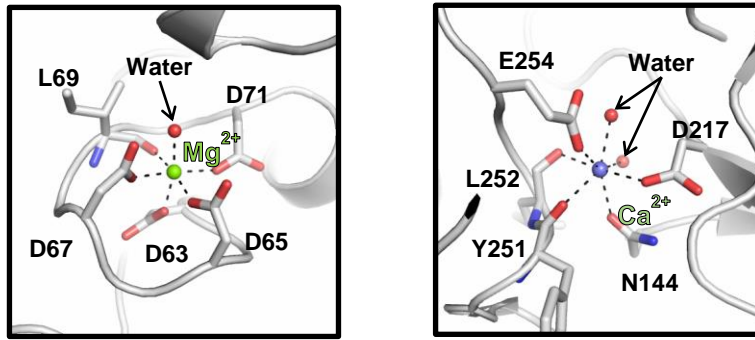
pH dependence (A), temperature dependence (B), pH stability (C), and thermostability of BmSUH. pH dependence was measured at 30 °C using Britton-Robinson buffer (pH 2.0–12), and temperature dependence was measured at 15 °C–70 °C using 50 mM HEPES-NaOH buffer (pH 8.0). pH stability was measured at 30 °C using enzymes after incubation in buffer with varying pH for 24 h at 4 °C. Thermostability was measured at 30 °C using enzymes pre-incubated at varying temperature for 30 min. All experiments were performed in triplicate.





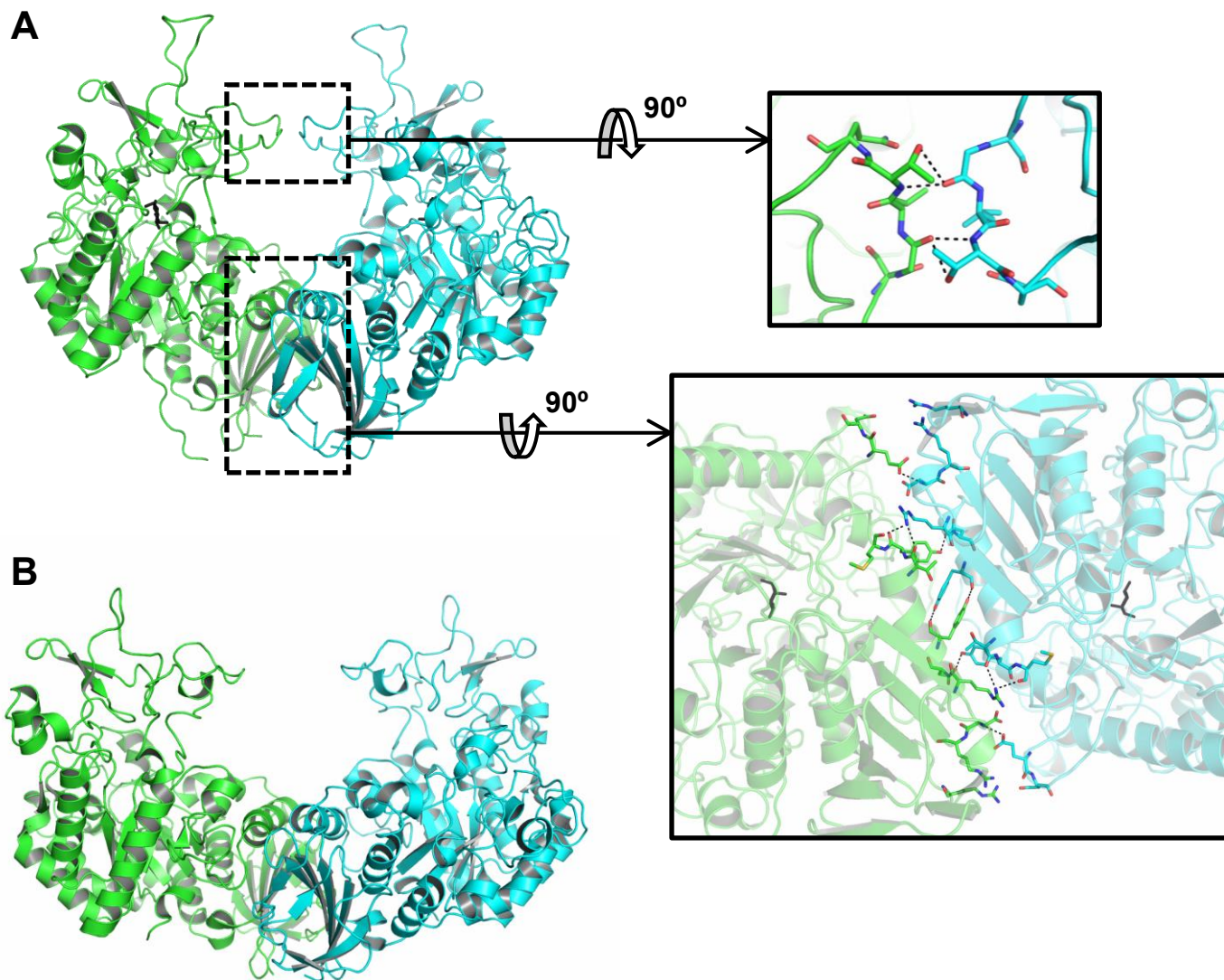
**Supplementary Figure S2. Structures of *exo*-acting GH13 enzymes homologous to BmSUH.**

(A) *B. mori* GH13\_17 sucrose hydrolase BmSUH (this study), (B) *C. quinquefasciatus* GH13\_17 maltase Cqm1, (C) *D. radiodurans* GH13\_16 trehalose synthase, (D) *Halomonas* sp. H11 GH13\_23  $\alpha$ -glucosidase HaG, (E) *B. licheniformis* GH13\_29 trehalose-6-phosphate hydrolase, (F) *Bacillus* sp. AHU2216 GH13\_31  $\alpha$ -glucosidase, (G) *P. mesoacidophila* GH13\_31 sucrose isomerase, (H) *S. cerevisiae* GH13\_40 oligo- $\alpha$ -1,6-glucosidase. Colors and symbols used are as follows: domain A, red; domain B, green; domain B', cyan; domain C, yellow; calcium ion, slate blue sphere; magnesium ion, light green sphere; cadmium ion, light brown sphere; ligands in the active centers, black stick.



**Supplementary Figure S3. Metal binding sites of BmSUH.**

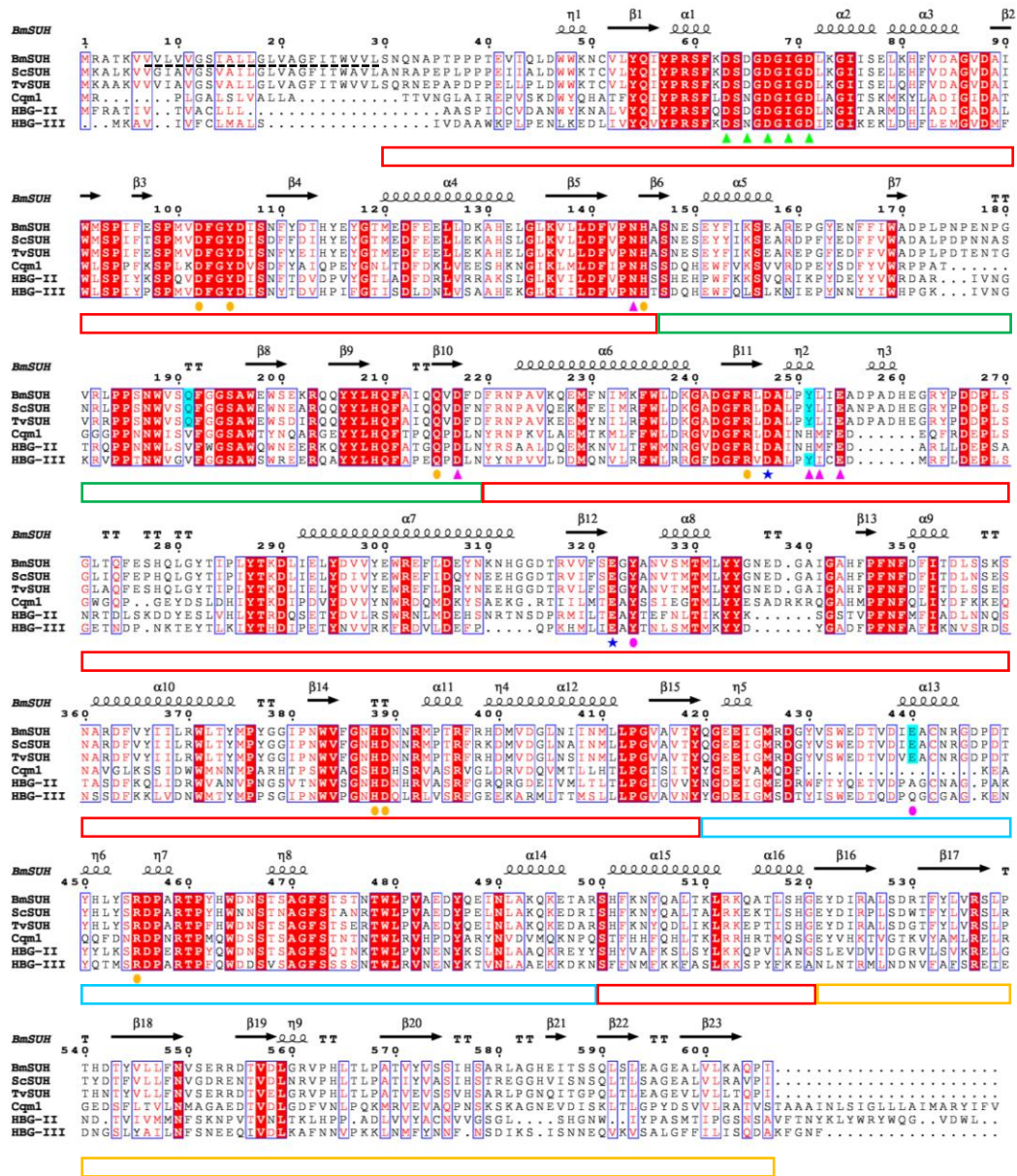
The amino acid residues that coordinate magnesium ion (*light green sphere*) in site I (*left panel*) and calcium ion (*slate blue sphere*) in site II (*right panel*) are shown in *stick models*.



**Supplementary Figure S4. Dimeric structures of BmSUH (A) and Cqm1 (B).**

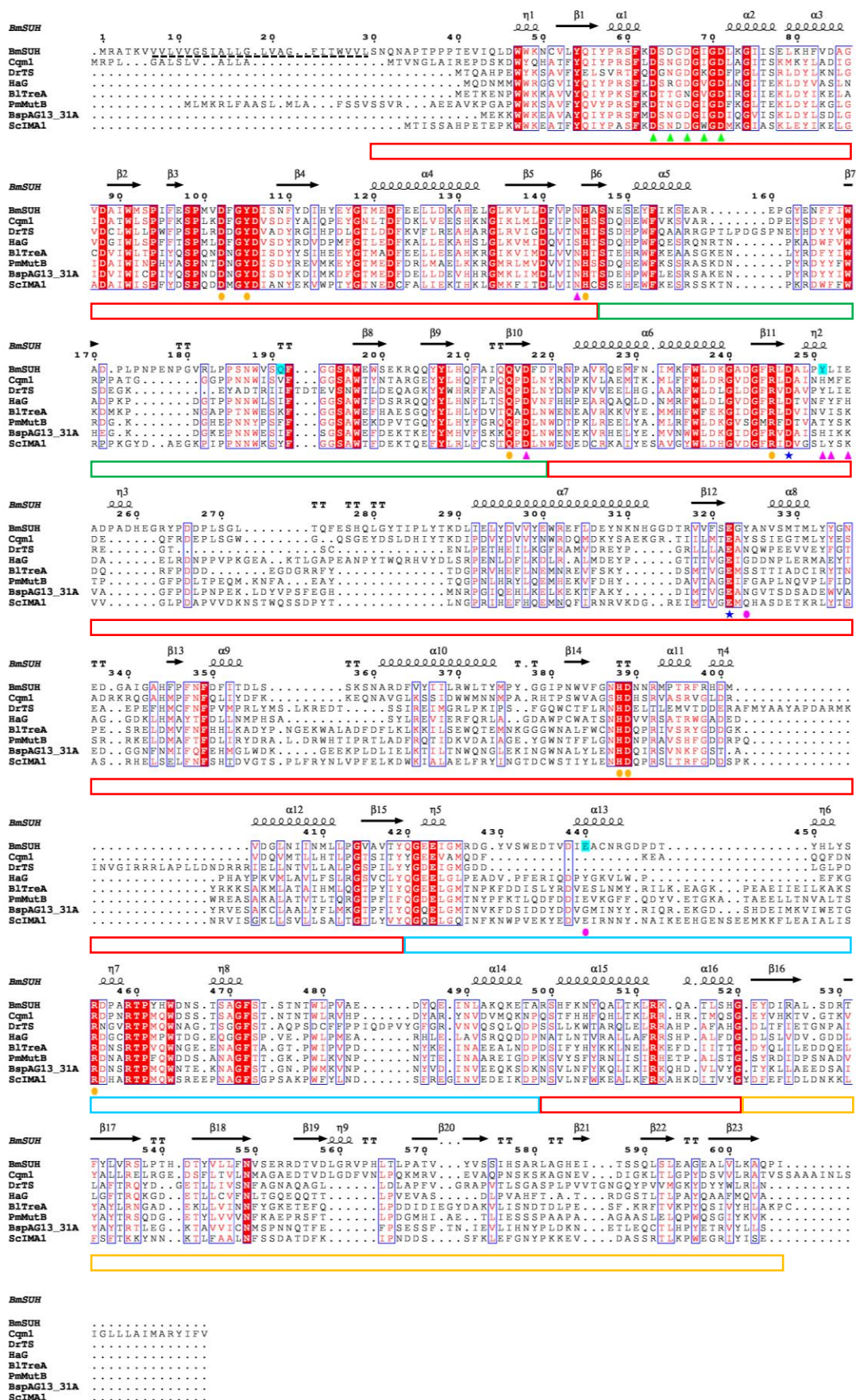
Two monomers are shown in *green* and *cyan*. In (A), amino acid residues located at the dimer interfaces are shown as *stick* models and hydrogen bonds are shown as *dashed lines*.





### Supplementary Figure S5. Sequence alignment of GH13\_17 enzymes.

The sequences of GH13\_17 enzymes were aligned using Clustal Omega [x] and the figure generated by ESPrnt 3.0. Sequences used are BmSUH (GenBank BAP18683.1), *Samia cynthia ricini* sucrose hydrolase (ScSUH, BAP18685.1), *Trilocha varians* sucrose hydrolase (TvSUH, BAP18684.1), Cqm1 (ASO96882.1), honeybee (*Apis mellifera*)  $\alpha$ -glucosidase II (HBG-II, BAE86927.1), and honeybee  $\alpha$ -glucosidase III (HBG-III, BAA11466.1). Secondary structures and domain architecture of BmSUH are described above and under sequences, respectively. Colors for the domains are as follows: domain A, red; domain B, green; domain B', cyan; domain C, yellow. The predicted transmembrane region is underlined with a dashed line. Identical residues are shown in white with a red background and conservative changes in red with a white background. The amino acid residues mutated in this study (Gln191, Tyr251, and Glu440) and their conserved residues are highlighted in cyan. Catalytic residues are marked with blue stars, and residues that interact with glucose and fructose residues of substrates are orange and magenta circles, respectively, under sequences. Green and magenta triangles indicate the magnesium- and calcium-binding residues of BmSUH, respectively.



**Supplementary Figure S6. Sequence alignment of *exo*-acting enzymes belonging to GH13\_16, 17, 23, 29, 31, and 40 subfamilies.**

The alignment was prepared and described in the same manner as Figure S5. Sequences used are BmSUH (GenBank BAP18683.1), Cqm1 (ASO96882.1), *D. radiodurans* GH13\_16 trehalose synthase (DrTS, AAF11586.1), *Halomonas* sp. H11 GH13\_23  $\alpha$ -glucosidase (HaG, BAL49684.1), *B. licheniformis* GH13\_29

trehalose-6-phosphate hydrolase (BTreA, AAU39732.1), *P. mesoacidophila* GH13\_31 sucrose isomerase (PmMutB, ABC33903.1), *Bacillus* sp. AHU2216 GH13\_31  $\alpha$ -glucosidase (BspAG13\_31A, BBB94188.1), and *S. cerevisiae* GH13\_40 oligo- $\alpha$ -1,6-glucosidase (ScIMA1, BAA07818.1).