

SUPPLEMENTARY INFORMATION

The complex structure of bile salt hydrolase from *Lactobacillus salivarius* reveals the structural basis of substrate specificity

Fuzhou Xu^{1,2}, Xiao-Jian Hu^{3*}, Warispreet Singh⁴, Wenjing Geng¹, Irina G. Tikhonova^{4*},
and Jun Lin^{1*}

¹Department of Animal Science, The University of Tennessee, Knoxville, TN 37996. USA;

²Institute of Animal Science and Veterinary Medicine, Beijing Academy of Agriculture and

Forestry Sciences, Beijing, China; ³ State Key Laboratory of Genetic Engineering, School of

Life Sciences, Collaborative Innovative Centre of Genetics and Development,, Fudan University,

Shanghai 200438, China, ⁴ Molecular Therapeutics, School of Pharmacy, Medical Biology

Centre, Queen's University Belfast, BT9 7BL, Northern Ireland, United Kingdom

*Corresponding author:

Dr. Xiao-Jian Hu, State Key Laboratory of Genetic Engineering, School of Life Sciences, Collaborative Innovative Center of Genetics and Development, Fudan University, Shanghai, China; Email: xjhu@fudan.edu.cn

Dr. Irina G. Tikhonova, Molecular Therapeutics, School of Pharmacy, Medical Biology Centre, Queen's University Belfast, BT9 7BL, Northern Ireland, United Kingdom, Email: i.tikhonova@qub.ac.uk

Dr. Jun Lin, Department of Animal Science, The University of Tennessee, 2506 River Drive. Knoxville, TN, USA; Email: jlin6@utk.edu

Figure S1. The active *ls*BSH-GCA and *ls*BSH-CA complex structure. The wall-eye stereo presentation of the polar interaction around the oxidative cysteine sulfonic acid OCS-2 in chain F (A), substrate GCA in chain F (B) and product CA in chain G (C). The refined density map around OCS-2, GCA, CA and ligand-bound water molecules were shown. The 2FoFc map contoured at 1.0 σ (blue) and FoFc map contoured at $\pm 3.0 \sigma$ (green and red) were shown. For clarity residues 128-139 are hid in (B) and (C). The polar contacts were plotted as red dot line and the length below 3.20 Å were shown.

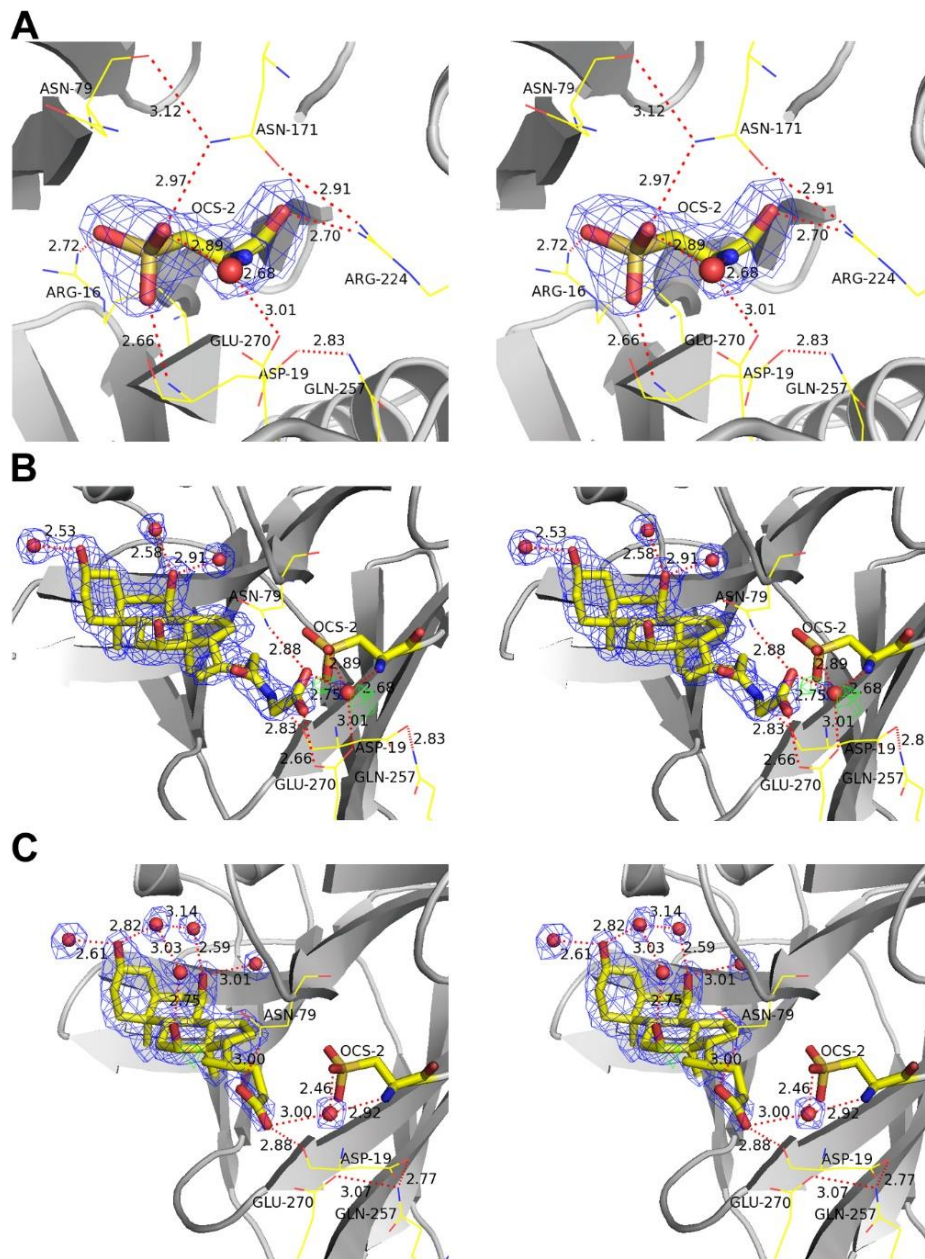


Figure S2. Sequence Alignment of BSH isotypes. The first four sequences of BSH with available crystal structures are shown with the PDB code. GenBank access code are given for all BSH sequences. Substrate preference: to glycol-conjugated bile acids, GC or tauro-conjugated bile acids, TC is shown from Dong *et al*, *Protein Science*, 2018, Vol. 27, pages 1742-1754. Loops 1-4 lining the active sites are highlighted in yellow. Residues involved in the catalytic reaction and substrate binding are in cyan and pink. Residues in pink contribute to substrate specificity. The red asterisk indicates the residues mutated in this study.

										Substrate	GenBank code	
5Y7P_L_salivarius	*			20		40		60		67	GC	AFP87505.1
2BJF_C_Perfringens		*								70	TC	AAC43454.1
4WL3_E_Feacalis					*					66	TC	EET97240.1
2HEZ_B_Longum										69	GC	AAF67801.1
BSHA_L_acidophilus_NCFM										67	GC	AAV42751.1
BSHB_L_acidophilus_NCFM										67	TC/GC	AAV42923.1
L_fermentum_NCDO394										69	GC	AEZ06356.1
BSH α _L_johnsonii100_100										68	TC/GC	AAG22541.1
BSHA_L_johnsoni_PF01										68	TC	EGP12224.1
BSHB_L_johnsoni_PF01										69	TC	EF536029.1
BSHC_L_johnsoni_PF01										67	GC	EGP12391.1
L_plantarum_80										67	GC	AAB24746.1
L_plantaru_WCFS1										67	GC	CAD65617.1
L_plantarum_80										67	TC	ACL98197.1
L_plantarum_80										67	TC/GC	ACL98205.1
L_salivarius_JCM1046										67	TC	ACL98194.1
L_salivarius_UCC118										67	GC	ACL98201.1
L_salivarius_B-30514										67	GC	AFP87505.1
BSH β _L_johnsonii100_100										69	TC/GC	AAC34381.1
B_animalis_subsp.lactis_Bi30										68	GC	AEK27050.1
B_animalis_subsp.Lactis_KL612										68	GC	AAS98803.1
B_bifidum_ATTCC_11863										70	GC	AAR39435.1
	2			16	19	24				58	65	

loop_3

80 * 100 * 120 * 140

5Y7P_L_salivarius : INEDGLGMAGLNFPGNAYYSDALENDKDNITPFEFIPWILGQCSDVNEARNLVEKINLINLSFS-EQLPL : 136
 2BJF_C_Perfringens : MNEKGLGCAGLNFVYVYSKEDIKGTNIPVYNFLWVLANFSSVEEVKEATKNANIVDIPIS-ENIPN : 139
 4WL3_E_Feacalis : INEKGLGMAGLNFSGYADYKKI-EEGKENVSPFEFIPWVLCQSTVDEAKKLLKNLNLVNINFS-DELPL : 134
 2HEZ_B_Longum : ANEHGLAIAGLNFPGYASFFVHEPVEGTENVATFEFPLWVARNFDSVDEVEETLRNVTLVSVQIVP--GQQE : 137
 BSHA_L_acidophilus_NCFM : INEKGLGIAGLNFQGNHYFVK-IEGKKNIASFEFLMPYLLSNCENTDDVKEITLDNANILNISFS-ANYP : 135
 BSHB_L_acidophilus_NCFM : TNEKGLGMAGLNFPGNATYEE-KENKDNIASFEFIPWILGQCSTISEVKDLSRINIADLNFS-EKMQA : 135
 L_fermentum_NCD0394 : VNEFGLMAQKLTFKNGARLVDRHPDKVQLAAEFELIFYLLGHFKSVADVAHLDQIELMNDVNADVPFGY : 139
 BSHα_L_johnsonii100_100 : SNEEGLGIAGLNFDPGCHYFPE-NAEKNNVTPFELIPYLLSQCTTVAEVKDALKDVSLVNINFS-EKLPL : 136
 BSHA_L_johnsoni_Pf01 : SNEEGLGIAGLNFDPGCHYFPE-VSGKNNVTPFELIPYLLSQYTTVAEVKEATKSVNLVKINFS-EKLQL : 136
 BSHB_L_johnsoni_Pf01 : YNEEGLGIAGLNFPHFAKFDGPIIDKINLASYEIMLVWVQNFTHVSEVKEATKNVNLVNEAIN-TSFAV : 138
 BSHC_L_johnsoni_Pf01 : INEKGLGIAGLNFDPGKYFAV-DESKNNVTSFELIPYLLSQSCETIEDVKKLLSETNITDESFS-KDLPV : 135
 L_plantarum_80 : MNEKGLCIAGLNFAGYADYKDY-DADKVNITPFEFIPWLLGQFSSVREVKKNIQKLNLVNINFS-EQLPL : 135
 L_plantaru_WCFs1 : MNEKGLCIAGLNFAGYADYKDY-DADKVNITPFEFIPWLLGQFSSVREVKKNIQKLNLVNINFS-EQLPL : 135
 L_plantarum_80 : INEDGLGMAGLNFPGNAYYSDALENDKDNITPFEFIPWILGQCSDVNEARNLVERINLINLSFS-EQLPL : 136
 L_plantarum_80 : ANEEGLGMAGLNFDPGPAHFPPV-EEGKDNVSPFEFIPYILGQCKNVAEAKELKSNLNVNINFS-DQLQL : 135
 L_salivarius_JCM1046 : INEDGLGMAGLNFPGNAYYSDALENDKDNITPFEFIPWILRQCSDVNEARNLVERINLINLSFS-EQLPL : 136
 L_salivarius_UCC118 : INEDGLGMAGLNFPGNAYYSDALENDKDNITPFEFIPWILGQCSDVNEARNLVEKINLINLSFS-EQLPL : 136
 L_salivarius_B-30514 : MNEKGLCIAGLNFAGYADYKDY-DADKVNITPFEFIPWLLGQFSSVREVKKNIQKLNLVNINFS-EQLPL : 135
 BSHβ_L_johnsonii100_100 : YNEEGLGIAGLNFPHFAKFDGPIIDKINLASYEIMLVWVQNFTHVSEVKEATKNVNLVNEAIN-TSFAV : 138
 B_animalis_subsp.lactis_Bi30 : MNDAGLAVAGLNFAGYCKYATEAVNFTTNVAAYEFPLWVTRNFTSVDDVQEAATKNVTIVGKPIN-DRFPV : 137
 B_animalis_subsp.Lactis_KL612 : MNDAGLAVAGLNFAGYCKYATEAVNFTTNVAAYEFPLWVTRNFTSVDDVQEAATKNVTIVGKPIN-DRFPV : 137
 B_bifidum_ATTCC_11863 : ANEHGLAIAGLNFPGYASFAHEPVEGTENVATFEFPLWVARNFDSVDEVEEATKNVTLVSVQVVP--GQQE : 138

79

134

* 160 *****180 * 200 *

5Y7P_L_salivarius : AGLHHLIADRE-KSIVVEVT-KSGVHIYDNPVGLTNNPEFNYQMYNLNKYRNLSISTPQNTFSDSVDLK : 204
 2BJF_C_Perfringens : TTHHWMISDITGKSIIVVEQT-KEKLNVDNNGVLTNSPTFDWHVANLNQYVGRYQVPEFKLGDQSLT : 208
 4WL3_E_Feacalis : SPLHHLIADRE-KSIVVEST-KEGLRVFDNPGVLTNNPTFDYQLFNLNRYRNLSTRTPKNNFSDQIELD : 202
 2HEZ_B_Longum : SLHHLWFIGDGK-RSIVVEQM-ADGMHVHDDVDVLTNQPTFDHFMENLRNYMCSVNEMAEPITWGWKASLT : 205
 BSHA_L_acidophilus_NCFM : ADLHHLIADRE-KSIVVEST-NSGLHVIYDNPVGLTNNPEFDQLIKLSDYADVTPHNPKNLTPVNDLN : 204
 BSHB_L_acidophilus_NCFM : SSLHHLIADKTGTSLVVETD-KDGMHIYDNPVGLTNNPQFPKQLFNLNRYRNLSPKMPKNNFSDKVNMA : 204
 L_fermentum_NCD0394 : SEQHVLSDPTGRGVVIEPS-EHPLKLDNPLGEMTNNPKFDHQLERLQDYLDFTPDFLNGTLAPNTFHV : 208
 BSHα_L_johnsonii100_100 : SPLHHLIADKTGSIIVVEST-LSGLHVIYDNPVGLTNNPEFPGQLRNLNRYRNLAPAQPKNTLVPVNDLN : 205
 BSHA_L_johnsoni_Pf01 : SPLHHLIADKTGSIIVVEST-LSGLHVIYDNPVGLTNNPEFPGQLRNLNRYRNLAPAQPKNTLVPVNDLN : 205
 BSHB_L_johnsoni_Pf01 : APHHLIADSD-EAIIIVEVSKQYGMKVFDKVGVLTSNPFNWHLTNLGNVYTGPHDATAQSWNGQKVA : 207
 BSHC_L_johnsoni_Pf01 : TTHHWMISDITGKSIIVVEST-ETGLHVIYDNPVGLTNNPVPFAQVETLANFASVSPAQPKNTLVPNADLN : 204
 L_plantarum_80 : SPLHHLIADKQ-ESIVIESV-KEGLKIYDNPVGLTNNPFDYQLFNLNRYRNLSTPQNSFSEKVDLD : 203
 L_plantaru_WCFs1 : SPLHHLIADKQ-ESIVIESV-KEGLKIYDNPVGLTNNPFDYQLFNLNRYRNLSTPQNSFSEKVDLD : 203
 L_plantarum_80 : AGLHHLIADRE-KSIVVEVT-KSGVHIYDNPVGLTNNPEFNYQMYNLNKYRNLSISTPQNTFSDSVDLK : 204
 L_plantarum_80 : SPLHHLIADKSGAATVVEST-ASGLHVIYDNPVGLTNNPEFDQLTNLNRYRNLSPANPANTLAPQATALA : 204
 L_salivarius_JCM1046 : AGLHHLIADRE-KSIVVEVT-KSGVHIYDNPVGLTNNPEFNYQMYNLNKYRNLSISTPQNTFSDSVDLK : 204
 L_salivarius_UCC118 : AGLHHLIADRE-KSIVVEVT-KSGVHIYDNPVGLTNNPEFNYQMYNLNKYRNLSINTPQNTFSDSVDLK : 196
 L_salivarius_B-30514 : SPLHHLIADKQ-ESIVIESV-KEGLKIYDNPVGLTNNPFDYQLFNLNRYRNLSTPQNSFSEKVDLD : 203
 BSHβ_L_johnsonii100_100 : APHHLIADSD-EAIIIVEVSKQYGMKVFDKVGVLTSNPFNWHLTNLGNVYTGPHDATAQSWNGQKVA : 207
 B_animalis_subsp.lactis_Bi30 : ATLHHLIADNT-RSIVVECT-EDGMHVYDDVDVLTNQPPFPQIEHLNRYRNLSPRTGKSVKWSSELE : 205
 B_animalis_subsp.Lactis_KL612 : ATLHHLIADNT-RSIVVECT-EDGMHVYDDVDVLTNQPPFPQIEHLNRYRNLSPRTGKSVKWSSELE : 205
 B_bifidum_ATTCC_11863 : SLHHLWFIGDGT-RSIVVEQM-ADGMHVHDDVDVLTNQPTFDHFMENLRNYMCSVNEMAEPITWGWKAELS : 206

171

loop_4

	220	*	240	*	260	*	*	*	280																											
5Y7P_L_salivarius	:	VDGTGFGGIGLPGDVSPESRFVRAAF	SKLNS	SKG	MTVEEDITQ	FFHILG	TV	EQIKGV	NKTESGKE	EYTVY	: 274																									
2BJF_C_Perfringens	:	ALGQGTGLVGLPGDFTPASRFIRVAF	FLR	DA	MIKNDKDSIDLIEFF	FHILNN	VAMVRG	STR	TVEEKS	DLTQY	: 278																									
4WL3_E_Feacalis	:	IYSRGMGGIGLPGDLSSVSRFVKATF	TKLNS	VSR	SSSEYESISQFF	FHILSS	VE	QKGL	CDVGDKEY	EYTIY	: 272																									
2HEZ_B_Longum	:	AWGACVSMHGIPGDVSSPSRFVRAAF	YTN	NAHY	PQQNDEAANVSRLE	FHTLGS	VQ	MV	DM	AKMGDQGF	ERTLF	: 275																								
BSHA_L_acidophilus_NCFM	:	LYSRGLGTHHLPDGMDS	SRFV	RAAF	VFLAHTPQ	GKNEVEN	VNTNY	FHIL	HSVE	QPKGL	DEVEDNRY	EYTM	: 274																							
BSHB_L_acidophilus_NCFM	:	GYSRGLGSHNLPDGMDS	SRFV	RAAF	FNKFN	APIAETEEEN	IDTY	FHIL	HSVE	QKGL	DEVGPN	SFEYTIY	: 274																							
L_fermentum_NCDO394	:	TTGKLSGKKTTPGAYTPKGFYVRAAF	YMKEL	ADQ	PASKDEALAT	TWHLL	DS	V	TP	KSKA	----	HRPT	F	SVY	: 274																					
BSHα_L_johnsonii100_100	:	LYSRGLGTHHLPDGMDS	SRFV	RAAF	VVRAHS	PQGN	NELSS	VNTNY	FHIL	HSVE	QPKGT	DE	V	GPN	SYEYTIY	: 275																				
BSHA_L_johnsoni_PF01	:	LYSRGLGTHHLPDGMDS	SRFV	RAAF	VVRAHS	PQGN	NELSS	VNTNY	FHIL	HSVE	QPKGT	DE	V	GPN	SYEYTIY	: 275																				
BSHB_L_johnsoni_PF01	:	PWGVGTGSLGLPGDSIPADRFVKAAY	LVN	VNY	PTAKGEKAN	VAKF	FNIL	KS	VAM	IKG	SV	VND	Q	GK	DEYTVY	: 277																				
BSHC_L_johnsoni_PF01	:	LYSRGLGTHHLPDGMDS	SRFV	RAAF	YVLAHS	PKGN	DEV	NVNTNY	FHIL	HSVE	QAKGT	DE	V	DN	VFEFTMY	: 274																				
L_plantarum_80	:	SYSRGMGGLGLPGDLSSMSRFVRAAF	TKLNS	LPM	QTESGS	VSQFF	FHIL	GS	VE	QKGL	CE	V	T	D	GKY	EYTIY	: 273																			
L_plantaru_WCFs1	:	SYSRGMGGLGLPGDLSSMSRFVRAAF	TKLNS	LPM	QTESGS	VSQFF	FHIL	GS	VE	QKGL	CE	V	T	D	GKY	EYTIY	: 273																			
L_plantarum_80	:	VDGTGFGGIGLPGDVSPESRFVRAAF	SKLNS	SKG	TTVEEDITQ	FFHILG	TV	EQIKGV	NKTESGKE	EYTVY							: 274																			
L_plantarum_80	:	SYSRAGSHFLPGDMDS	SRFV	KEVET	TLQH	APAG	ETE	VAN	VNTNY	FH	CL	H	A	E	QKGL	DE	V	GK	DQFEYTIY	: 274																
L_salivarius_JCM1046	:	VDGTGFGGIGLPGDVSPESRFVRAAF	SKLNS	SKG	TTVEEDITQ	FFHILG	TV	EQIKGV	NKTESGKE	EYTVY										: 274																
L_salivarius_UCC118	:	VDGTGFGGIGLPGDVSPESRFVRAAF	SKLNS	SKG	TTVEEDITQ	FFHILG	TV	EQIKGV	NKTESGKE	EYTVY										: 266																
L_salivarius_B-30514	:	SYSRGMGGLGLPGDLSSMSRFVRAAF	TKLNS	LSM	QTESGS	VSQFF	FHIL	GS	VE	QKGL	CE	V	T	D	GKY	EYTIY	: 273																			
BSHβ_L_johnsonii100_100	:	PWGVGTGSLGLPGDSIPADRFVKAAY	LVN	ANY	PTVKGEKAN	VAKF	FNIL	KS	VAM	IKG	SV	VND	Q	G	QSD	EYTVY	: 277																			
B_animalis_subsp.lactis_Bi30	:	TNQDSNSSQGLPGGYGSMARFVRAAF	YNN	TH	YPTQ	SGENAN	VN	R	L	E	K	T	L	S	T	A	A	V	E	G	T	A	I	S	A	N	E	F	E	K	T	L	F	: 275		
B_animalis_subsp.Lactis_KL612	:	TNQDSNSSQGLPGGYGSMARFVRAAF	YNN	TH	YPTQ	SGENAN	VN	R	L	E	K	T	L	S	T	A	A	V	E	G	T	A	I	S	A	N	E	F	E	K	T	L	F	: 275		
B_bifidum_ATTCC_11863	:	AWGACVSMHGIPGDVSSPSRFVRAAF	YTN	TH	YP	QQN	NEA	N	V	S	R	L	E	F	H	T	L	V	S	V	Q	M	V	D	G	M	S	K	M	G	N	G	Q	F	ERTLF	: 276

	*	300	*	320	*	340																								
5Y7P_L_salivarius	:	SNCYDLNKTLYYTTYENRQIVAVTLNK	-DKD	GNRLV	TY	PFERKQI	INKLNL	LER	HHHHH	: 333																				
2BJF_C_Perfringens	:	TSCMCLEKGIYYNTYENNQINAIDM	NKEN	LDGNE	LKTY	KYNK	TL	SIN	HVN	~~~~~	: 329																			
4WL3_E_Feacalis	:	SSCCNLEKGIYYRTYDNSQITAVDM	NKEN	LEKDS	LIVY	PM	V	ET	QQ	I	N	Y	A	N	L	E	H	H	H	H	~	: 331								
2HEZ_B_Longum	:	TSGYSSKTNYYMNTYDDPAIRSYAM	ADY	D	M	D	S	S	E	L	S	V	A	R	~~~~~	: 316														
BSHA_L_acidophilus_NCFM	:	TDCMNLDKGILYFTTYDNNRINAV	DM	H	K	A	D	L	S	E	D	L	I	C	Y	D	L	F	K	Q	D	I	E	Y	M	N	~~~~~	: 325		
BSHB_L_acidophilus_NCFM	:	SDGTNLDKGIYYTTYSNKQINVVDM	NK	E	D	L	S	N	L	I	T	Y	D	M	L	D	K	T	K	F	N	H	Q	N	~~~~~	: 325				
L_fermentum_NCDO394	:	RAATVAEDRTYEFQSYHQAVTSVKL	T	D	D	L	L	K	R	A	T	P	L	V	F	D	T	A	D	V	W	A	P	V	K	L	N	~~~~~	: 325	
BSHα_L_johnsonii100_100	:	SDGTNLETGTFYYTYENNQINAIE	L	N	K	E	N	L	N	G	D	E	L	D	Y	K	L	I	E	K	Q	T	I	N	Y	Q	N	~~~~~	: 326	
BSHA_L_johnsoni_PF01	:	SDGTNLETGTFYYTYENNQINAIE	L	N	K	E	N	L	N	G	D	E	L	D	Y	K	L	I	E	K	Q	T	I	N	Y	Q	N	~~~~~	: 326	
BSHB_L_johnsoni_PF01	:	TACYSSGSKTYICNFEDDFELKTY	K	L	D	D	H	T	M	N	S	T	S	L	V	T	Y	~~~~~	: 316											
BSHC_L_johnsoni_PF01	:	SDCMNLDKGILYFTTYDNNQINAV	D	M	N	E	N	L	D	T	S	D	L	I	T	Y	E	L	F	K	D	Q	A	I	K	F	E	N	~~~~~	: 325
L_plantarum_80	:	SSCCMDKGVYYRTYDNSQINSVNL	N	H	E	H	L	D	T	T	E	L	I	S	Y	P	L	R	S	E	A	Q	Y	Y	A	V	N	~~~~~	: 324	
L_plantaru_WCFs1	:	SSCCMDKGVYYRTYDNSQINSVNL	N	H	E	H	L	D	T	T	E	L	I	S	Y	P	L	R	S	E	A	Q	Y	Y	A	V	N	~~~~~	: 324	
L_plantarum_80	:	SNCYDLNKTLYYTTYENRQIVSVTL	NK	-DKN	GNL	V	V	Y	P	F	E	R	K	Q	I	I	N	K	L	N	~~~~~	: 324								
L_plantarum_80	:	SDGVNLTGTFYYTYENNQINAVK	M	H	A	E	D	M	E	G	Q	L	H	R	F	P	I	A	S	H	Q	S	I	N	M	Q	N	~~~~~	: 325	
L_salivarius_JCM1046	:	SNCYDLNKTLYYTTYENRQIVAVTL	N	E	-DKN	GNL	I	A	Y	P	F	E	R	K	Q	V	I	N	K	L	N	~~~~~	: 324							
L_salivarius_UCC118	:	SNCYDLNKTLYYTTYENRQIVAVTL	NK	-DKD	GNRLV	TY	PFERKQI	INKL	N	~~~~~	: 316																			
L_salivarius_B-30514	:	SSCCMDKGVYYRTYDNSQINSVNL	N	H	E	H	L	D	T	T	E	L	I	S	Y	P	L	R	S	E	A	Q	Y	Y	A	V	N	~~~~~	: 324	
BSHβ_L_johnsonii100_100	:	TACYSSGSKTYICNFEDDFELKTY	K	L	D	D	H	T	M	N	S	T	S	L	V	T	Y	~~~~~	: 316											
B_animalis_subsp.lactis_Bi30	:	SDCYSTATQTVLKKYDDMAVHSY	AV	K	D	F	D	A	S	S	N	L	Q	S	K	~~~~~	: 314													
B_animalis_subsp.Lactis_KL612	:	SDCYSTATQTVLKKYDDMAVHSY	AV	K	D	F	D	A	S	S	N	L	Q	S	K	~~~~~	: 314													
B_bifidum_ATTCC_11863	:	TSGYSGKTNYYMNTYEDPAIRSF	A	M	S	D	F	D	M	D	S	S	E	L	I	T	A	D	~~~~~	: 316										

Figure S3. Proposed hydrolysis mechanism of GCA in BSH. (a) The neutral form of the N-terminal Cys2; (b) the zwitterionic state of the N-terminal Cys2; in this state the nucleophilic thiolate anion attacks the carbonyl carbon of the amide bond of GCA; (c) the proposed transition state associated with the nucleophilic attack by the thiolate anion and the proton abstraction by the lone pair of the nitrogen of the amide bond; (d) the acylated enzyme substrate complex and the free glycine amino acid; (e) proton abstraction from the water molecule by the N-terminal primary amine; (f) nucleophilic attack of the hydroxyl group to the carbonyl carbon of the acylated enzyme substrate complex, (g) the proposed transition state associated with the nucleophilic attack by the hydroxyl anion and proton abstraction by the sulphur atom of Cys2, (i) the neutral form of the N-terminal Cys2 and cholic acid. The lone pair of the electrons are represented by the double dots; breaking and forming bonds are in dashed lines. The protein residues, the glycine fragment of GCA and the water molecule are shown in black, pink and blue, respectively.

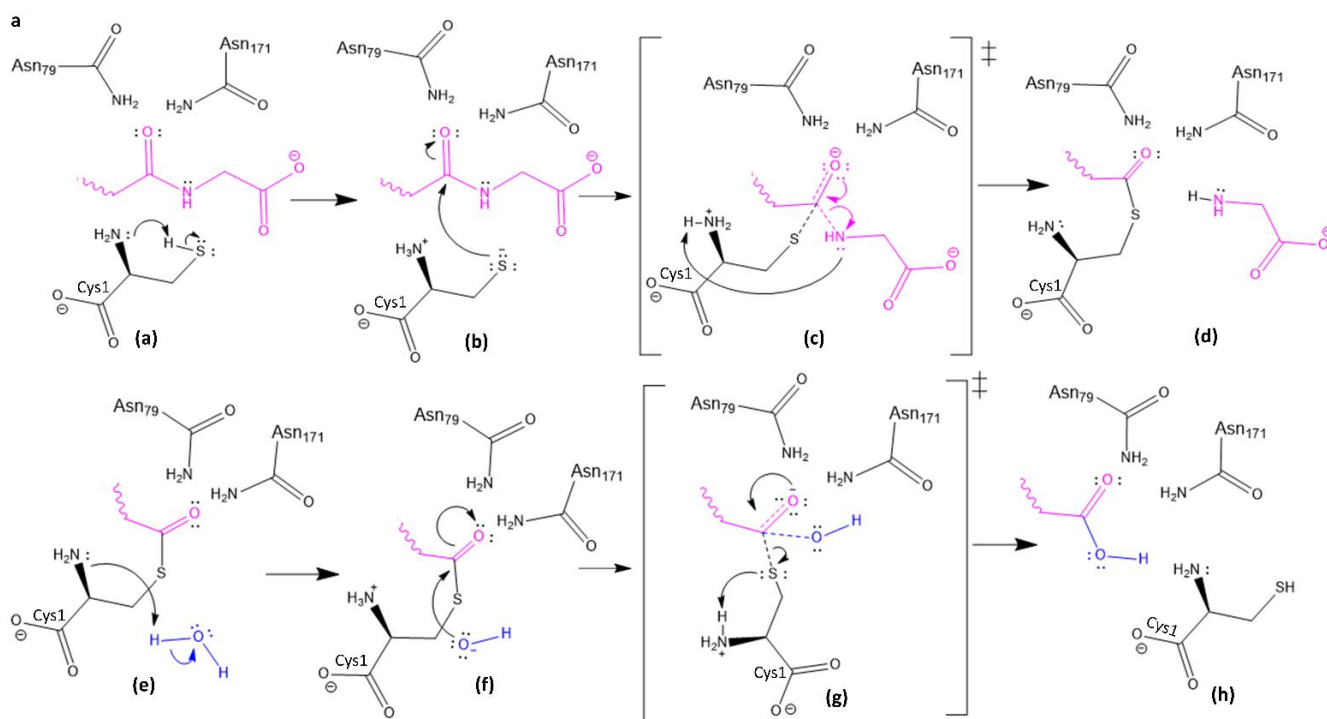


Figure S4. The active site of the *ls*BSH in complex with GCA. The snapshot of the most populated cluster obtained from the 50ns MD simulation. The GCA molecule is colored yellow and the active site residues are shown in grey stick representation. The loop region consisting of residues 164-171 are shown in red ribbon representation.

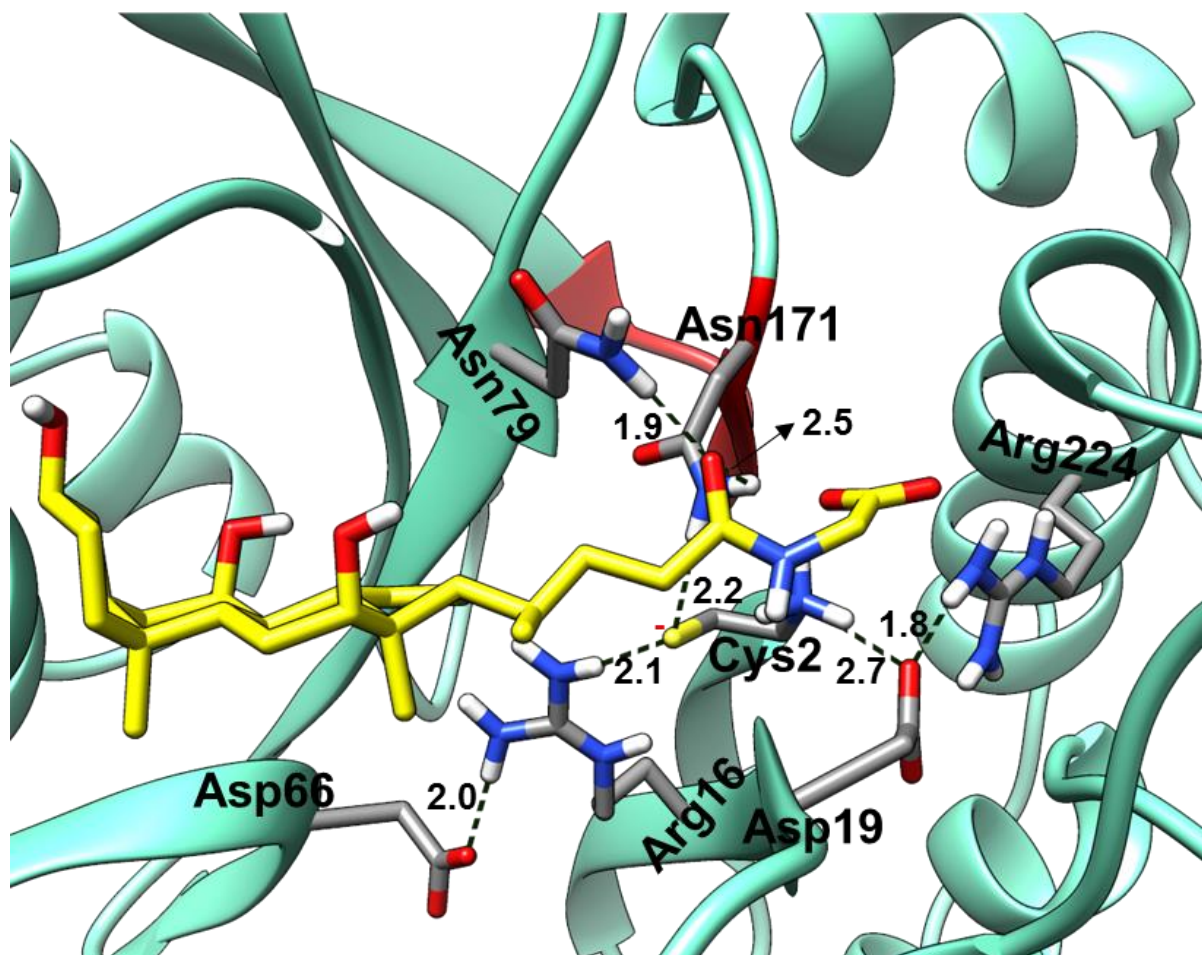


Figure S5. The surface representation of *ls*BSH bound to GCA obtained from the MD simulations. The hydrogen bonds formed by water molecules with the substrate are shown in the black dashed lines. The substrate binding cavity of the *ls*BSH is highlighted with the orange colour. The substrate GCA is shown in yellow colour.

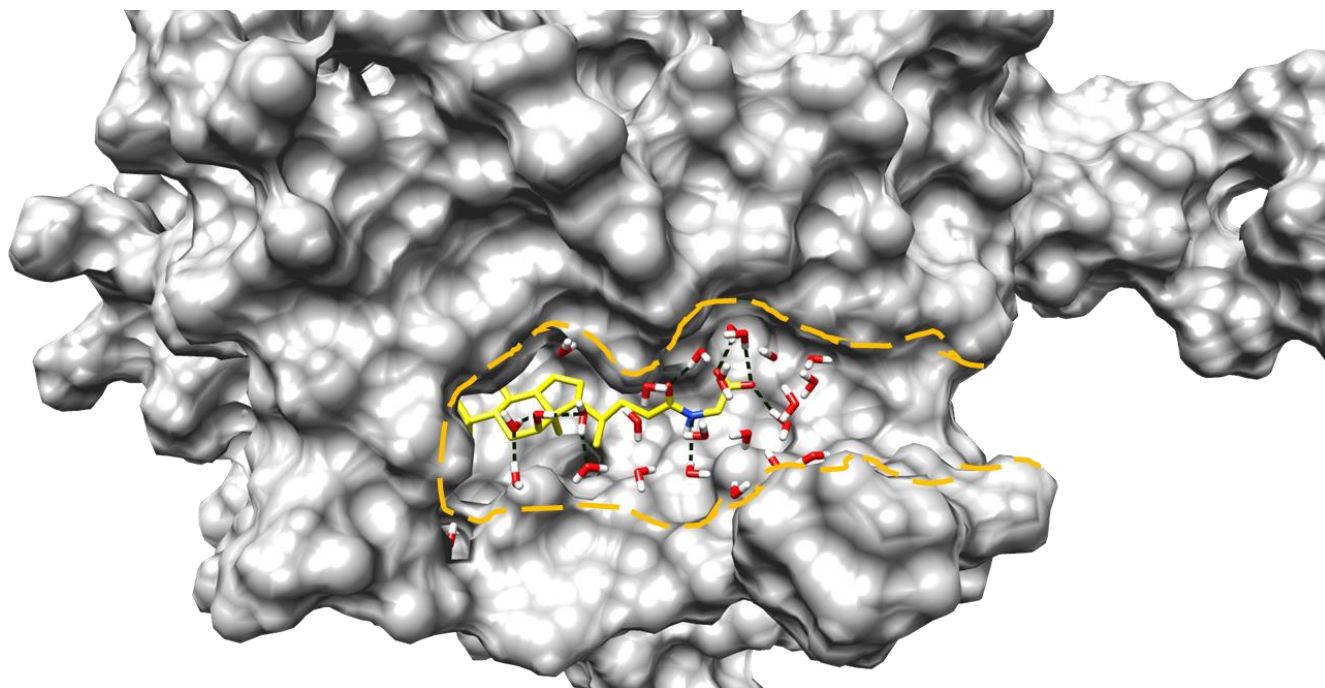


Figure S6. The substrate binding site of *ls*BSH in complex with GCA. The snapshot of the most populated cluster obtained from the 50ns MD simulation. GCA and the substrate stabilizing residues are shown in yellow and grey stick representation, respectively. The hydrophobic interactions mediated by the *ls*BSH residues with the substrate are shown in the orange dashed lines.

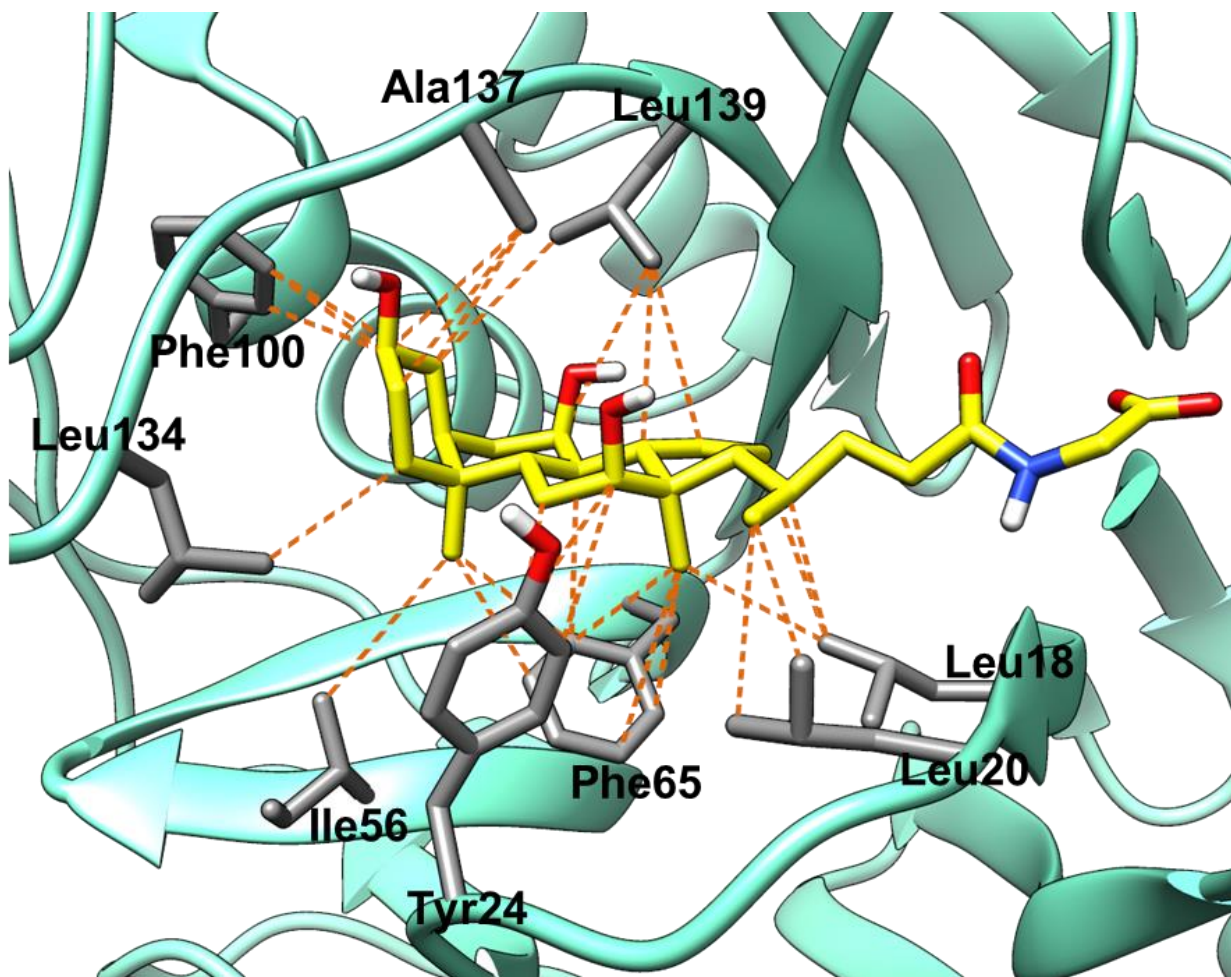


Table S1. Key bacterial plasmids and strains used in this study.

Plasmids or strains	Description	Source or Reference
Plasmids		
pBSH	pET21b containing <i>bsh</i> gene from <i>Lactobacillus salivarius</i> NRRL B-30514, Amp ^r	<u>11</u>
pBSH (C2S)	pBSH derivative with C2S mutation	This study
pBSH (Y24F)	pBSH derivative with Y24F mutation	This study
pBSH (F65A)	pBSH derivative with F65A mutation	This study
pBSH (N171A)	pBSH derivative with N171A mutation	This study
pBSH (164-171)	pBSH derivative with the 164 -171 aa motif deleted	This study
pBSH (Q257A)	pBSH derivative with Q257A mutation	This study
pBSH (E270A)	pBSH derivative with E270A mutation	This study
pBSH (Q257A E270A)	pBSH derivative with double mutations (Q257A and E270A)	This study
Strains		
<i>E.coli</i> BL21(DE3)	F ⁻ ompT gal dcm lon hsdSB(rB ⁻ mB ⁻) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	Novagen
<i>E.coli</i> XL1-Blue	EndA1 gyrA96(nal ^R) thi-1 recA1 relA1 lac glnV44 F' ⁺ [::Tn10 proAB ⁺ lacI ^q Δ(lacZ)M15] hsdR17(r _K ⁻ m _K ⁺)	Agilent Technologies
JL885	<i>E.coli</i> BL21(DE3) containing the plasmid pBSH	<u>11</u>
JL1189	<i>E.coli</i> BL21(DE3) containing pBSH (C2S)	This study
JL1191	<i>E.coli</i> BL21(DE3) containing pBSH (Y24F)	This study
JL1193	<i>E.coli</i> BL21(DE3) containing pBSH (F65A)	This study
JL1196	<i>E.coli</i> BL21(DE3) containing pBSH (N171A)	This study
JL1198	<i>E.coli</i> BL21(DE3) containing pBSH (164-171)	This study
JL1201	<i>E.coli</i> BL21(DE3) containing pBSH (Q257A)	This study
JL1203	<i>E.coli</i> BL21(DE3) containing pBSH (E270A)	This study
JL1205	<i>E.coli</i> BL21(DE3) containing pBSH (Q257A E270A)	This study

Figure S7. SDS-PAGE analysis of the purified wild-type (WT) *Is*BSH and its derivative with specific amino acid substitution mutagenesis or deletion.

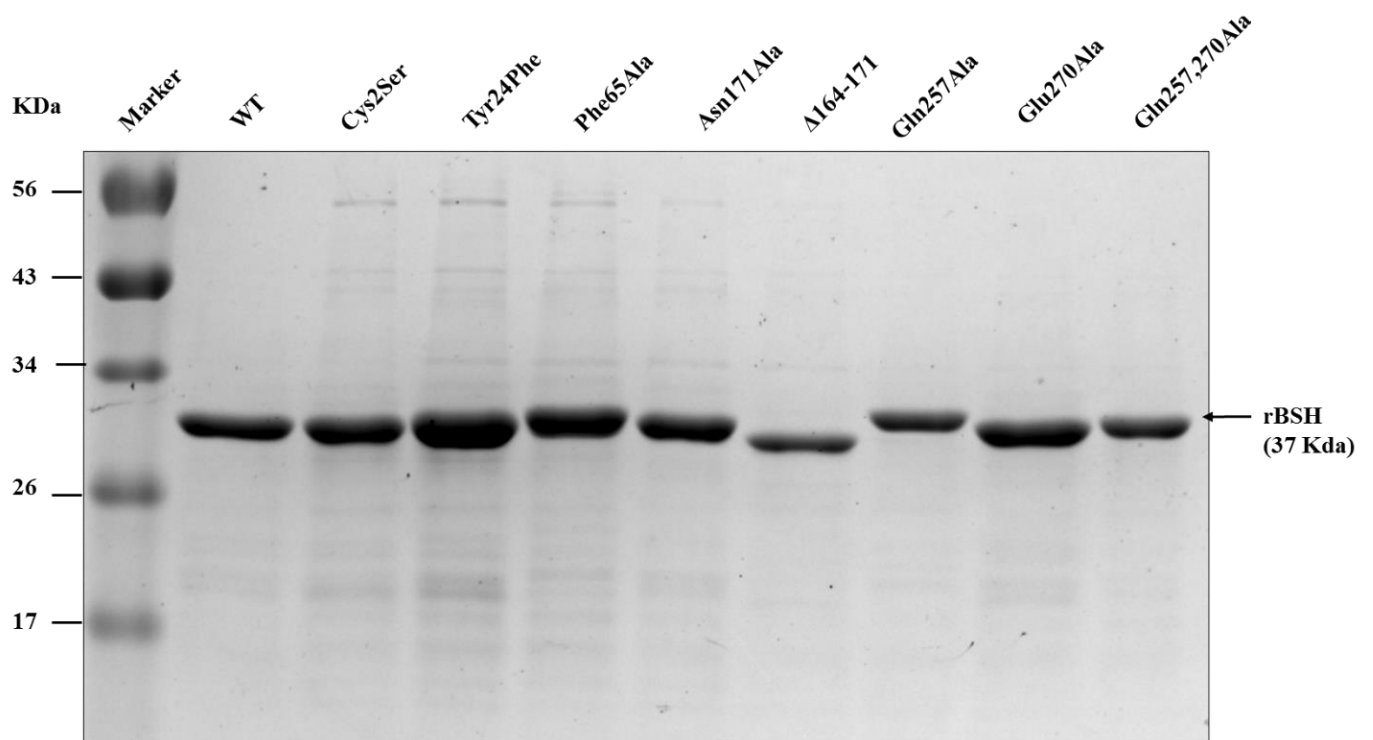


Table S2. The primers used for site-directed amino acid substitution mutagenesis in this study.

Primer Name	Sequence 5'- 3'^a	AA Mutation
BSHMT1-F	CCATTAAAGTAATTGCTGT AG ACATGGATCCCGACCCATTT	Cys2Ser
BSHMT1-R	AAATGGGTCGGGATCCATG TCT TACAGCAATTACTTTAAATGG	
BSHMT2-F	GATTACCTCCTCACC AA ATGAAAAATCTAAATCTAAGTTTCTTCCAA	Tyr24Phe
BSHMT2-R	TTGGAAGAACTTAGA TTT TAGATTTTTTCATTTGGTGAGGAGGTAATC	
BSHMT3-F	TCCATCCTCATTAATAGCATCT TGC ATACAATGGGTAATCGTTAGCGACAATTCC	Phe65Ala
BSHMT3-R	GGAATTGTCGCTAACGATTACCCATTGTAT GC AGATGCTATTAATGAGGATGGA	
BSHMT4-F	CAGATTGTACATCTGATAATTAATCTGGATT TGC AGTCAATATTCCAATTGG	Asn171Ala
BSHMT4-R	CCAATTGGAATATTGACT GCA AATCCAGAATTTAATTATCAGATGTACAATCTG	
BSHMT5-F	GTACATCTGATAATTAATCTGGATTATCATAAATATGTACTCCAGATTTAGTTACTTC	Δ 164-171
BSHMT5-R	GAAGTAACTAAATCTGGAGTACATATTTATGATAATCCAGAATTTAATTATCAGATGTAC	
BSHMT6_Q257A-F	CCATATACTAGGGACAGTAGA AGCA ATAAAGGGCGTTAATAAGACAG	Gln257Ala
BSHMT6_Q257A-R	CTGTCTTATTAACGCCCTTTAT TGCT TCTACTGTCCCTAGTATATGG	
BSHMT7_E270A-F	GACAGAATCAGGAAAAGA GC ATACACTGTATATTCTGAATTGC	Glu270Ala
BSHMT7_E270A-R	GCAATTCGAATATACAGTGTAT TGCT TCTTTCTGATTCTGTC	

^a Bolded letters indicate the nucleotides designed for aa substitution mutagenesis

Figure S8. Comparison of the *ls*BSH binding site in the wild type (A) and Tyr24Phe mutation (B). The water-mediated hydrogen bond between Tyr24 and the hydroxyl group of GCA is shown in black dashed lines. The hydrophobic interactions between Phe24 and GCDCA are shown in the orange dashed lines.

