

## Supplementary Information for

### Conformational rearrangements enable iterative backbone *N*-methylation in RiPP biosynthesis

Fredarla S. Miller,<sup>1†</sup> Kathryn K. Crone,<sup>1†</sup> Matthew R. Jensen,<sup>2,3</sup> Sudipta Shaw,<sup>1</sup> William R. Harcombe,<sup>2,4</sup> Mikael H. Elias,<sup>1,2\*</sup> and Michael F. Freeman<sup>1,2\*</sup>

<sup>1</sup> Department of Biochemistry, Molecular Biology, and Biophysics, the <sup>2</sup> BioTechnology Institute, and the <sup>4</sup> Department of Ecology, Evolution and Behavior, University of Minnesota-Twin Cities, St. Paul, Minnesota, USA

<sup>3</sup> Present address: Science Department, Concordia University-St. Paul, St. Paul, Minnesota, USA

† These authors contributed equally to this work: F.S. Miller and K.K. Crone

\* Correspondence and requests for materials should be addressed to M.H. Elias (email: [mhelias@umn.edu](mailto:mhelias@umn.edu)) or M.F. Freeman ([mffreema@umn.edu](mailto:mffreema@umn.edu))

## Table of Contents

Supplementary Table 1. Split borosins encoded in NCBI representative genomes from the genus <i>Shewanella</i> .	3
Supplementary Table 2. All primers and plasmids used in this manuscript.	4
Supplementary Table 3. All genes used in this manuscript.	5
Supplementary Table 4. Parameter values used in our base kinetics model.	8
Supplementary Table 5. Crystallographic data and statistics.	9
Supplementary Figure 1. Phylogenic tree of borosin methyltransferase domains including putative split borosin domains encoded in bacteria.	10
Supplementary Figure 2. Geneious sequence alignment of previously identified fungal borosin and putative bacterial split borosin methyltransferase domains.	11
Supplementary Figure 3. Gene clusters of the two split borosins analyzed in this study.	15
Supplementary Figure 4. Mass spectrometric analysis of split borosin coexpressions.	16
Supplementary Figure 5. Global structural comparisons of borosin $\alpha$ -N-methylating enzymes.	27
Supplementary Figure 6. Oligomeric state of his <sub>6</sub> -SonM and his <sub>6</sub> -SonA proteins by size exclusion chromatography.	28
Supplementary Figure 7. Structural overlay of SonM—SonA-2Me—SAH.	29
Supplementary Figure 8. Structural overlay of the BBD with other protein domains.	30
Supplementary Figure 9. Thermal motion B-factors for SonA and SonM.	31
Supplementary Figure 10. SonM—SonA-2Me—SAH active site coordination.	32
Supplementary Figure 11. SonM interactions with the SonA core peptide in SonM—SonA-2Me—SAH.	33
Supplementary Figure 12. Superposition of SonM—SonA-2Me complexes.	34
Supplementary Figure 13. SAH makes extensive contacts in the SonM—SonA-2Me—SAH complex.	35
Supplementary Figure 14. SonM – fitted Michaelis-Menten kinetic curves.	36
Supplementary Figure 15. Active site tyrosine mutant structures.	38
Supplementary Figure 16. SonM in vitro reactions analyzed by LC-MS/MS and compared to kinetic model simulations.	39
Supplementary Figure 17. SonM – fitted Michaelis-Menten competitive inhibition kinetic curves.	40
Supplementary Figure 18. Superposition of SonM—SonA-2Me and SonM—SonA-BBD—( $\pm$ )SAM complexes.	41
Supplementary Figure 19. Bottom lock configurations of the SonM—SonA-BBD—( $\pm$ )SAM complex.	42
Supplementary Figure 20. Top lock configurations between borosin $\alpha$ -N-methyltransferases.	43
Supplementary Figure 21. Active site configuration differences among split borosin complexes.	44
Supplementary Figure 22. Interaction networks in different configurations of SonM and mutant complexes.	45
Supplementary Figure 23. Structural differences in the heterodimers of SonM-R67A—SonA-0Me—SAH.	46
Supplementary Figure 24. Structural differences in the BBD of SonM-R67A—SonA-0Me—SAH and <i>apo</i> SonM—SonA-2Me—SAH.	47
Supplementary Figure 25. Mass spectrometric analysis of SonM mutant in vitro reactions.	48
References	55

**Supplementary Table 1. Split borosins encoded in NCBI representative genomes from the genus *Shewanella*.**

	Representative species	Genome assembly	Gene cluster	Putative split borosin methyltransferase Protein ID	Putative split borosin precursor Protein ID	Syntenic with <i>S. oneidensis</i> MR-1 cluster (Y/N)
1	<i>Shewanella aestuarii</i> strain PN3F2	GCA_011765625.1_ASM1176562v1	1	WP_167676051.1	WP_167676053.1	Y
2	<i>Shewanella algae</i> strain MARS 14	GCA_000947195.1	1	WP_044735032.1	WP_028781613.1	N
3	<i>Shewanella algidipiscicola</i> strain LMG 23746	GCF_003605125.1_ASM360512v1	1	WP_119977437.1	WP_110456456.1	Y
4	<i>Shewanella amazonensis</i> SB2B	NC_008700	1	WP_011760496.1	WP_011760495.1	Y
			2	WP_011758902.1	WP_011758901.1	N
5	<i>Shewanella atlantica</i> strain HAW-EB5	GCF_003966265.1_ASM396626v1	1	WP_126503733.1	WP_126503731.1	Y
			2	WP_126507131.1	WP_126507132.1	N
6	<i>Shewanella baltica</i> OS678	GCF_000178875.2_ASM17887v2	1	WP_006085112.1	WP_006085113.1	Y
7	<i>Shewanella benthica</i> KT99	GCF_000172075.1_ASM17207v1	1	WP_005497745.1	WP_005497743.1	Y
8	<i>Shewanella bicestrii</i> strain JAB-1	GCF_002216875.1_ASM221687v1	1	WP_089067321.1	WP_086904596.1	Y
9	<i>Shewanella canadensis</i> strain HAW-EB2	GCF_003966225.1_ASM396622v1	1	WP_126520857.1	WP_126520690.1	Y
			2	WP_126518905.1	WP_126518907.1	N
10	<i>Shewanella carassii</i> strain 08MAS2251	GCA_002777975.1_ASM277797v1	-	-	-	-
11	<i>Shewanella chilikensis</i>	GCF_003217175.1_ASM321717v1	-	-	-	-
12	<i>Shewanella colwelliana</i> ATCC 39565	GCF_000518705.1_ASM51870v1	1	WP_051413204.1	WP_028765344.1	Y
13	<i>Shewanella corallii</i> strain A687	GCF_003353085.1_ASM335308v1	1	WP_115137829.1	WP_115137830.1	Y
14	<i>Shewanella decolorationis</i> S12	GCF_000485795.1_SheDec2.0	1	WP_039978563.1	WP_006085113.1 WP_023268149.1	Y
15	<i>Shewanella denitrificans</i> OS217	NC_007954	1	WP_041405716.1	WP_041405717.1	Y
			2	WP_011495401.1	WP_011495402.1 WP_041405677.1	N
			3	WP_011495261.1	WP_011495262.1 WP_157599827.1 WP_041405665.1	N
16	<i>Shewanella donghaensis</i> strain LT17	NZ_CP041783	1	WP_144213469.1	WP_144213471.1	Y
			2	WP_144214053.1	WP_144206800.1 WP_144206802.1	N
17	<i>Shewanella fidelis</i> ATCC BAA-318	GCF_000518605.1_ASM51860v1	1	WP_028768597.1	WP_028768598.1 WP_028768599.1	Y
			2	WP_037410318.1	WP_028766355.1	N
18	<i>Shewanella fodinae</i> strain 74A	GCF_004342405.1_ASM434240v1	1	WP_133038992.1	WP_133038991.1	Y
			2	WP_133039724.1	-	N
19	<i>Shewanella frigidimarina</i> NCIMB 400	NC_008345	1	WP_049763522.1	WP_011636573.1	Y
20	<i>Shewanella hafniensis</i> isolate NILIPAHB1	GCA_902728295.1_NILIPAHB_1	1	CAA7314325.1	CAA7314324.1	Y
21	<i>Shewanella halifaxensis</i> HAW-EB4	NC_010334	1	WP_012276456.1	WP_012276457.1	Y
22	<i>Shewanella hanedai</i> strain JCM 20706	GCF_007197645.1_ASM719764v1	1	WP_143563155.1	WP_143563156.1	Y
23	<i>Shewanella indica</i>	GCF_002836975.1_ASM283697v1	-	-	-	-
24	<i>Shewanella japonica</i> strain KCTC 22435	NZ_CP020472	1	WP_080916319.1	WP_080916317.1	Y
			2	WP_080915084.1	WP_080915085.1	N
25	<i>Shewanella khirikhana</i> strain TH2012	GCA_003957745.1_ASM395774v1	1	AZQ11249.1	AZQ11248.1	Y
			2	AZQ09459.1	AZQ09458.1	N
26	<i>Shewanella litoralis</i> strain JCM 32306	GCF_009828585.1_ASM982858v1	1	WP_160053732.1	WP_160053733.1	Y
27	<i>Shewanella livingstonensis</i> strain LMG 19866	NZ_CP034015	1	<i>disrupted</i>	WP_124730162.1	Y
28	<i>Shewanella loihica</i> PV-4	NC_009092	1	WP_011865074.1	WP_011865075.1	Y
29	<i>Shewanella mangrovi</i> strain YQH10	GCF_000753795.1_ASM75379v1	1	WP_037443456.1	WP_037443455.1	N
30	<i>Shewanella marina</i> JCM 15074	GCF_000614975.1_ASM61497v1	-	-	-	-
31	<i>Shewanella marisflavi</i> strain EP1	NZ_CP022272	1	WP_088905093.1	WP_088905092.1	Y
32	<i>Shewanella maritima</i> strain D4-2	NZ_CP036200	1	WP_130597927.1	WP_130597926.1	Y
			2	WP_130598048.1	WP_130598049.1	N
33	<i>Shewanella morhuae</i> strain ATCC BAA-1205	GCF_900156405.1_IMG-taxon_2681812898	1	WP_076497071.1	WP_006085113.1 WP_076496897.1	Y
34	<i>Shewanella oneidensis</i> MR-1	GCF_000146165.2_ASM14616v2	1	WP_011071665.1	WP_011071666.1	Y
35	<i>Shewanella pealeana</i> ATCC 700345	NC_009901	1	WP_012154539.1	WP_012154541.1	Y
36	<i>Shewanella piezotolerans</i> WP3	NC_011566	-	-	-	-
37	<i>Shewanella polaris</i> strain SM1901	NZ_CP041036	1	WP_140234809.1	WP_140234808.1	Y
38	<i>Shewanella psychrophila</i> strain WP2	NZ_CP014782	1	WP_077753941.1	WP_077753940.1	Y
			2	WP_077754459.1	WP_077754460.1 WP_077754461.1	N
39	<i>Shewanella putrefaciens</i> CN-32	NC_009438	1	WP_011918927.1	WP_011918928.1	Y
40	<i>Shewanella sediminis</i> HAW-EB3	NC_009831	1	WP_041421581.1	WP_012141759.1	Y
			2	WP_012143903.1	WP_012143904.1	N
41	<i>Shewanella vesiculosa</i> LMG 24424	GCF_003797885.1_ASM379788v1	1	WP_124016834.1	WP_124016835.1	Y
42	<i>Shewanella violacea</i> DSS12	NC_014012	1	WP_013050640.1	WP_013050641.1	Y
43	<i>Shewanella waksmanii</i> ATCC BAA-643	GCF_000518805.1_ASM51880v1	1	WP_028771625.1	WP_028771626.1	Y
			2	WP_051484558.1	WP_028773238.1 WP_028773237.1	N
44	<i>Shewanella woodyi</i> ATCC 51908	NC_010506	1	WP_041418034.1	WP_012324159.1	Y
45	<i>Shewanella xiamenensis</i> strain T17	GCF_001723195.1_ASM172319v1	1	WP_039978563.1	WP_006085113.1 WP_037425540.1	Y

**Supplementary Table 2. All primers and plasmids used in this manuscript.**

<b>Primers used in this study</b>	
<b>Name</b>	<b>Sequence (5'-3')</b>
prmMRJ036_fw	ACTTTAAGAAGGAGATATACCATGGGATCACTCGTCTGTG
prmMRJ043_rev	GATGATGATGATGATGCATGTTTTCTCCTTATTGTTAATAATGATTCAATAAC
prmMRJ044_fw	AGGAGAAAACATGCATCATCATCATCACATGTCTGGATTATCGGATTTT TTTAC
prmMRJ045_rev	CGAGTGC GGCCGCAAGCTTGTGCGACTTAATCACCATTACCATGTG
T7_fw	TAATACGACTCACTATAGGG
T7_rev	GCTAGTTATTGCTCAGCGG
prFM1175	TTTAAGAAGGAGATATACATGCATCATCATCATCAT
prFM1176	AGTGC GGCCGCAAGCTTGTTAATCACCATTACCATG
prFM1177	TAAGAAGGAGATATACATGCATCATCATCATCATCACAGCAGCATGGGATCA CTCGTC
prFM1178	AGTGC GGCCGCAAGCTTGTATCCCAAATCTTCGGG
prFM1191	GAAGTTAAAAATAAACGAGACACCTACGA
prFM1192	GAAGTTAAAAATGCCCGAGACACCTAC
prFM1193	ACCATTTTGC GCATAAAACTGCTG
prFM1194	CGAGACACCTTCGAGCAAATGGTC
prFM1195	GCGATTTTAACTTCACCATTTTGC G
prFM1212	GCAGCAGTTTTTTGCGCAAAA
prFM1213	AAATTGATGACATTGGGGTTGAGC
prFM1214	TGTGCACTCTTCGGTCATCC
prFM1215	CACGGTTTTTTTACCCGCTCTC
prKKC1010	GAGCTCGAATTCGGATCTTAACCACTTAACGT
prmMRJ_066_fwd	ATATAACATATGCAGGAGACCACCG
prmMRJ_067_rev	TTATATGGATCCTTAACGACGCGCCG
prmMRJ_068_fwd	ATATAACATATGCCGGCGGC
prmMRJ_069_rev	TTATATGGATCCTTACGCACCGCTCGG
<b>Plasmids used in this study</b>	
<b>ID</b>	<b>Description</b>
pMF1181	SonM-gRBS-His-SonA_pET28b
pMF1235	His-SonA_pET28b
pMF1236	His-SonM_pET28b
pMF1230	His-ADE (JW_3640 ASKA collection)
pMF1231	His-SAHN (JW_0155 ASKA collection)
pMF1256	SonM-R67A-gRBS_His-SonA_pET28b
pMF1257	His-SonM-R67A_pET28b
pMF1258	SonM-R67K-gRBS_His-SonA_pET28b
pMF1259	His-SonM-R67K_pET28b
pMF1260	SonM-Y71F-gRBS_His-SonA_pET28b
pMF1261	His-SonM-Y71F_pET28b
pMF1263	SonM-Y58F-gRBS_His-SonA_pET28b
pMF1264	His-SonM Y58F_pET28b
pMF1265	SonM-Y58F-Y71F-gRBS_His-SonA_pET28b
pMF1266	His-SonM Y58F + Y71F_pET28b
pMF1267	SonM-Y93F-gRBS_His-SonA_pET28b
pMF1268	His-SonM Y93F_pET28b
pMF1269	His-SonA_helicalbundle_pET28b
pMF1283	SonM-gRBS-His-SonA_helical bundle_pET28b
pMF1197	pET28b-His-StrA

**Supplementary Table 3. All genes used in this manuscript.**

UniProt ID	Name	Description
P31441	<i>ade</i>	Adenine deaminase
<p>ATGAATAATTCTATTAACCATAAATTTTCATCACATTAGCCGGGCTGAATACCAGGAATTGTTAG            CCGTTTCCCGTGGCGACGCTGTTGCCGATTATATTATTGATAATGTCTCTATTCTCGACCTGAT            CAATGGCGGAGAAATTTCCGGCCCAATTGTGATTAAGGACGTTACATTGCCGGTGTGGCGC            AGAATACACTGATGCTCCGGCTTTGCAGCGGATTGATGCTCGCGGCAACGGCGGTGCCAG            GTTTTATTGATGCTCACCTGCATATTGAATCCAGCATGATGACGCCGGTCACTTTTGAACCG            CTACCCTGCCGCGCGCCTGACGACCGTTATTTGCGACCCTCATGAAATCGTCAACGTGATG            GGCGAAGCCGGATTCCGCTGGTTTGCCGCTGTGCCGAACAGGCAAGGCAAAACCAGTACTT            ACAGGTCAGCTCTTGCGTACCCGCCCTGGAAGGCTGCGATGTTAACGGTGCCAGTTTTACCC            TTGAACAGATGCTCGCCTGGCGGGACCATCCGCAGGTTACCGGCCCTGCAGAAATGATGGAC            TACCCTGGCGTAATTAGCGGGCAGAATGCGCTGCTCGATAAACTGGATGCATTTCCGCCACT            GACGCTGGACGGTCACTGCCCGGGTTGGGTGGTAAAGAACTTAACGCCTATATTACTGCGG            GTATTGAAACTGCCACGAAAGTTATCAGCTGGAAGAAGGACGCCGGAAATTACAACCTCGGCA            TGTCGTTGATGATCCGCGAAGGGTCCGCTGCCCGCAATCTCAACGCGCTGGCACCGTTGATC            AACGAATTTAACAGCCCGCAATGCATGCTCTGTACCGATGACCGTAACCCGTGGGAGATCGC            CCATGAAGGACACATCGATGCCTTAATTCGCCGCTGATCGAACAACACAATGTGCCGCTGCA            TGTGGCATATCGCGTCGCCAGCTGGTCGACGGCGGCCACTTTGGTCTGAATCACCTCGGCT            TACTGGCACCCGGCAAGCAGGCCGATATCGTCTGTTGAGCGATGCGCGTAAGGTCACGGTG            CAGCAGGTAAGGCGAGCCGATTGATGCGCAAACCTTACAGGCGGAAGAGTCGG            CGAGACTGGACAATCCGCTCCGCCATATGGCAACACCATTGCCCGCCAGCCAGTTTTCCGCC            AGCGACTTTGCCCTGCAATTTACGCCCGGAAAACGCTATCGGGTCATTGACGTCATCCATAAC            GAATTGATTACGCACTCCCACTCCAGCGTCTACAGCGAAAATGGTTTTGATCGCGATGATGTG            AGCTTTATTGCCGTAATTGAGCGTTACGGGCAACGGCTGGCTCCGGCTTGTGGTTTTGCTTGG            CGGCTTTGGACTGAATGAAGGTGCGCTGGCTGCGACGGTCAGCCATGACAGCCATAATATTG            TGGTGATCGGTGCGAGTGCCGAAGAGATGGCGCTGGCGGTCAATCAGGTGATTCAGGATGG            CGGCGGGCTGTGCGTGGTACGTAACGGCCAGGTACAAAGTCATCTGCCGTTACCCATTGCCG            GGCTGATGAGCACCGACACGGCGCAGTCGCTGGCGGAACAAATTGACGCCTTAAAAGCCGC            CGCCCGTGAATGCGGTCCGTTACCCGATGAGCCGTTTATTCAGATGGCGTTTCTTTCTGCC            AGTGATCCCCGCGCTAAAACAAACCAGTCAGGGGCTATTTGATGGCGAGAAGTTTGCCTTAC            TAGCTGGAAGTCACGGAATAA</p>		
P0AF12	<i>sahn</i>	S-adenosylhomocysteine nucleosidase
<p>ATGAAAATCGGCATCATTGGTGCAATGGAAGAAGAAGTTACGCTGCTGCGTGACAAAATCGAA            AACCGTCAAACATCAGTCTCGGCGGTTGCGAAATCTATACCGGCCAACTGAATGGAACCGAG            GTTGCCTTCTGAAATCGGGCATCGGTAAAGTCGCTGCGGCGCTGGGTGCCACTTTGCTGTT            GGAACACTGCAAGCCAGATGTGATTATTAACACCGGTTCTGCCGGTGGCCTGGCACCAACGT            TGAAAGTGGGCGATATCGTTGTCTCGGACGAAGCAGTTATCACGACGCGGATGTCACGGCA            TTTGGTTATGAATACGGTCAGTTACCAGGCTGTCCGGCAGGCTTTAAAGCTGACGATAAACTG            ATCGCTGCCGCTGAGGCCTGCATTGCCGAACTGAATCTTAACGCTGTACGTGGCCTGATTGTT            AGCGGCGACGTTTTCATCAACGGTTCTGTTGGTCTGGCGAAAATCCGCCACAACCTCCCACA            GGCCATTGCTGTAGAGATGGAAGCGACGGCAATCGCCCATGTCTGCCACAATTTCAACGTCC            CGTTTGTGTCGTACGCGCCATCTCCGACGTGGCCGATCAACAGTCTCATCTTAGCTTCGATG            AGTTCCTGGCTGTTGCCGCTAAACAGTCCAGCCTGATGTTGAGTCACTGGTGCAGAACTTG            CACATGGCTAA</p>		
Q8EGW3	<i>sonM</i> (SO1478)	Borosin methyltransferase
<p>ATGGGATCACTCGTCTGTGTGGGCACTGGGTTACAGCTCGCGGGGCAAATTAGCGTATTAAG            CCGCAGCTATATTGAACATGCCGATATTGATTTTCACTCTTACCTGACGGTTTCTCGCAGCGT            TGGTTGACGAAGCTCAACCCCAATGTCATCAATTTGCAGCAGTTTTATGCGCAAAATGGTGAA            GTTAAAAATCGCCGAGACACCTACGAGCAAATGGTCAATGCCATTCTAGATGCGGTGAGAGC            GGGTAAAAAACCGTGTGTGCACTCTACGGTCATCCGGGGGATTTGCTGTGTATCCCATAT            GGCATAACTCGGGCGAAGGCCGAAGGGTTTTCGGCAAAGATGGAGCCGGGGATTTCCGGCC            GAAGCTTGCCTGTGGGCCGACTTAGGGATTGACCCCGGCAACTCGGGGCATCAAAGTTTTGA            AGCTAGCCAGTTTATGTTTTTCAACCATGTGCCCGATCCCACTACCCACTTATTACTCTGGCAA            ATCGCCATTGCAGGCCAACATACTTAACCCAATTTTACATACCTCGAGTGATAGGTTGCAGATC            CTCGTGGAGCAGTTGAATCAATGGTATCCCTCGACCATGAGGTGGTCATATACGAAGCGGC            CAATTTGCCAATCCAAGCCCCGCGTATCGAGCGTTTACCTTTAGCGAATTTACCCCAAGCACA</p>		

CTTAATGCCGATTAGTACGTTGTTAATTCCGCCAGCAAAAAAGCTGGAGTACAACCTATGCTATT TTGGCTAAGTTAGGGATCGGTCCCGAAGATTTGGGATAA		
Q8EGW2	<i>sonA</i> (SO1479)	Borosin RiPP precursor
ATGTCTGGATTATCGGATTTTTTTACCCAGTTAGGCCAAGATGCGCAGTTAATGGAAGACTATA AACAGAATCCTGAGGCGGTGATGCGTGCCACGGATTAAGTATGAACAAATTAACGCTGTAA TGAAGTGGGATATGGAAAAGCTCAAACGTTAAGTGGTGATAGTATCAATCTTACCTTGT TATTCACATGGTAATGGTGATTA		
n/a	<i>his<sub>6</sub>-sonM</i>	Hexahistidine tagged borosin precursor for heterologous expression
ATGCATCATCATCATCACAGCAGCATGGGATCACTCGTCTGTGTGGCACTGGGTTACAG CTCGCGGGGCAAATTAGCGTATTAAGCCGCAGCTATATTGAACATGCCGATATTGTATTTTAC TCTTACCTGACGGTTTCTCGCAGCGTTGGTTGACGAAGCTCAACCCCAATGTCATCAATTTGC AGCAGTTTTATGCGCAAAATGGTGAAGTTAAAAATCGCCGAGACACCTACGAGCAAATGGTCA ATGCCATTCTAGATGCGGTGAGAGCGGGTAAAAAACCGTGTGTGCACTCTACGGTCATCCG GGGGTATTTGCCTGTGTATCCCATATGGCGATAACTCGGGCGAAGGCCGAAGGGTTTTCGGC AAAGATGGAGCCGGGGATTTCCGGCCGAAGCTTGCCTGTGGGCCGACTTAGGGATTGACCC GGCAACTCGGGGCATCAAAGTTTTGAAGCTAGCCAGTTTATGTTTTTCAACCATGTGCCCGAT CCCACTACCCACTTATTACTCTGGCAAATCGCCATTGCAGGCCGAACATACCTTAACCCCAATTC ATACCTCGAGTATAGTTGCGATCCTCGTGAGCAGTTGAATCAATGGTATCCCCTCGACC CTAGGGTGGTCATATACGAAGCGGCCAATTTGCCAATCCAAGCCCCGCGTATCGAGCGTTTAC CTTTAGCGAATTTACCCCAAGCACACTTAATGCCGATTAGTACGTTGTTAATTCGCCAGCAAA AAAGCTGGAGTACAACCTATGCTATTTTGGCTAAGTTAGGGATCGGTCCCGAAGATTTGGGATA A		
n/a	<i>his<sub>6</sub>-sonA</i>	Hexahistidine tagged borosin methyltransferase for heterologous expression
ATGCATCATCATCATCACATGTCTGGATTATCGGATTTTTTTACCCAGTTAGGCCAAGATG CGCAGTTAATGGAAGACTATAAACAGAATCCTGAGGCGGTGATGCGTGCCACGGATTAAGT GATGAACAAATTAACGCTGTAATGACTGGGGATATGGAAAAGCTCAAACGTTAAGTGGTGAT AGTAGCTATCAATCTTACCTTGTATTTTACATGGTAATGGTGATTA		
n/a	<i>his<sub>6</sub>-sonA-BBD</i>	Hexahistidine tagged SonA helical bundle/BBD (SonA- BBD)
ATGCATCATCATCATCACATGTCTGGATTATCGGATTTTTTTACCCAGTTAGGCCAAGATG CGCAGTTAATGGAAGACTATAAACAGAATCCTGAGGCGGTGATGCGTGCCACGGATTAAGT GATGAACAAATTAACGCTGTAATGACTGGGGATATGGAAAAGCTCAAACGTTAAGTGGTTAA		
n/a	<i>sspA<sub>NRRLS118</sub></i>	Codon optimized borosin RiPP precursor
ATGCCGGCGGCGGTGGTTGACTTCATGGAGGAACTGGTGACCCAGCCGCGTCAACACG CGTACCGTCTAGCGCGGAGGCGTATGTTGCGGATAGCGCGCTGACCGCTAGCGAGCGTGA AGCGGTGGTTAGCGGTGACGTGGATCGTATGCGTGCGGTTCTGGCCGAGCACAGCGGCGTG AAAGAGGAGTGCCACGCGGTTCTGGTGGTTATCATTTTTGACCCGGATGAAGTTCCGAGCGG TGCCTAA		
n/a	<i>sspM<sub>NRRLS118</sub></i>	Codon optimized borosin methyltransferase
ATGCAGGAGACCACCGGTAACGCGCAACTGGTGGTTGTGGGTACCGGTTTCCGTGCGATTGG TGACCTGACCGTTGAAGCGCGTGCGTGCCGGAACAGGCGGACAAGGTTCTGTGCCTGATCG GTGATCCGCTGGTGACCCGTCACATTGAGAACTGAACGCGAGCGTTGAAACCTGGATGTT CATTATGCGGTGGGCAAGCCGCGTAGCGCGAGCTATGAGGACATGGTGGAAACACATTATGAG CGAACTGCACCGTGATCAATTCGTTTGCCTGGCGCTGTACGGTCACCCGGGCGTTTTTGCCT ATACCGGTCATGAGGCGATCCGTCGTGCGCGTGAGGAAGGCATCGCGGCGCGTATGCTGCC GGCGTGACGCGCGGAAGACTGGCTGTTTGCCTGATCTGGGTCTGGACCCGGGCGAGCGTGGC TGCCAGAGCTTCGAAGCGACCGACTTTCTGATCCGTCACCGTGTGTTTGTATCCGACCGGCT GCTGATTCTGTGGCAAGTTGGTGTGATCGGCATGATTGATCGTATCCGGGTTATGATGCGCG TCCGGGCGTTACCACCCTGACCGATGCGCTGGTTGCGAGCTACGGTAGCGGCCACCCGGTT ACCGTGTACGAGGCGAGCCGATGTTACCGCGGAACCGGTACCACCACCGTGCCGCTGG CGGAGCTGCCGGACACCCCGTGAAGCGGCGAGCACCTGGTTGTGCCGCGCTGCCGC CGCGTCCGGTGGATCGTGAACCTGCTGGCGCGTCTGGCGGCGCGTCTGTTAA		
n/a	<i>his<sub>6</sub>-sspA<sub>NRRLS118</sub></i>	Hexahistidine tagged borosin RiPP precursor: <i>SspA<sub>NRRLS118</sub></i>
ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCGCTGGTGCCGCGCGGCAGCCATAT GCCGGCGGCGGTGGTTGACTTCATGGAGGAACTGGTGACCCAGCCGCGTCAACACGCG TACCGTCTAGCGCGGAGGCGTATGTTGCGGATAGCGCGCTGACCGCTAGCGAGCGTGAAG		

CGGTGGTTAGCGGTGACGTGGATCGTATGCGTGCGGTTCTGGCCGAGCACAGCGGCGTGAA  
AGAGGAGTGCCACGCGGTTCTGGTGGTTATCATTTTTGACCCGGATGAAGTTCCGAGCGGTG  
CGTAA

n/a	<i>his<sub>6</sub>-SUMO- sspM<sub>NRRLS118</sub></i>	Hexahistidine and SUMO tagged borosin methyltransferase: SspM <sub>NRRLS118</sub>
-----	----------------------------------------------------------	--------------------------------------------------------------------------------------

ATGGGTAGCCACCACCACCACCATCATCATAGCAGCGGTTTAGTTCCTCGTGGTTCAGCTAGC  
CACATCAACCTGAAGGTGAAAGGCCAGGATGGCAACGAGGTGTTCTTCCGCATTAACGCTC  
AACCCAGCTGAAGAAGCTGATGAACGCGTACTGCGATCGTCAGAGCGTGGATATGACCGCAA  
TTGCGTTCCTGTTTCGATGGTCGTCGTTTACGTGCAGAACAAACCCCGGACGAACCTGGAATGG  
AAGATGGCGACGAGATTGATGCCATGCTGCATCAGACCGGTGGCCATATGCAGGAGACCACC  
GGTAACGCGCAACTGGTGGTTGTGGGTACCGGTTTCCGTGCGATTGGTGACCTGACCGTTGA  
AGCGCGTGCGTGCCCTGGAACAGGCGGACAAGGTTCTGTGCCTGATCGGTGATCCGCTGGTG  
ACCCGTCACATTGAGAACTGAACGCGAGCGTTGAAACCCTGGATGTTTCATTATGCGGTGGG  
CAAGCCGCGTAGCGCGAGCTATGAGGACATGGTGGAAACACATTATGAGCGAACTGCACCGTG  
ATCAATTCGTTTGCCTGGCGCTGTACGGTCACCCGGGCGTTTTTGCCTATACCGGTCATGAGG  
CGATCCGTCGTGCGCGTGAGGAAGGCATCGCGGCGCGTATGCTGCCGGCGTGACGCGCGG  
AAGACTGGCTGTTTGCGGATCTGGGTCTGGACCCGGGCGAGCGTGGCTGCCAGAGCTTCGA  
AGCGACCGACTTTTCTGATCCGTCACCGTGTGTTTATCCGACCGGCCTGCTGATTCTGTGGCA  
AGTTGGTGTGATCGGCATGATTGATCGTGATCCGGTTATGATGCGCGTCCGGGCGTTACCA  
CCCTGACCGATGCGCTGTTGCGAGCTACGGTAGCGGCCACCCGTTACCGTGTACGAGGC  
GAGCCCGTATGTTACCGCGGAACCGCGTACCACCACCGTGCCGCTGGCGGAGCTGCCGGAC  
ACCCCGCTGAGCGCGGCGAGCACCCCTGGTTGTGCCGCGCTGCCGCCGCGTCCGGTGGATC  
GTGAACTGCTGGCGCGTCTGGCGGCGCGTCTGTTAA

**Supplementary Table 4. Parameter values used in our base kinetics model.**

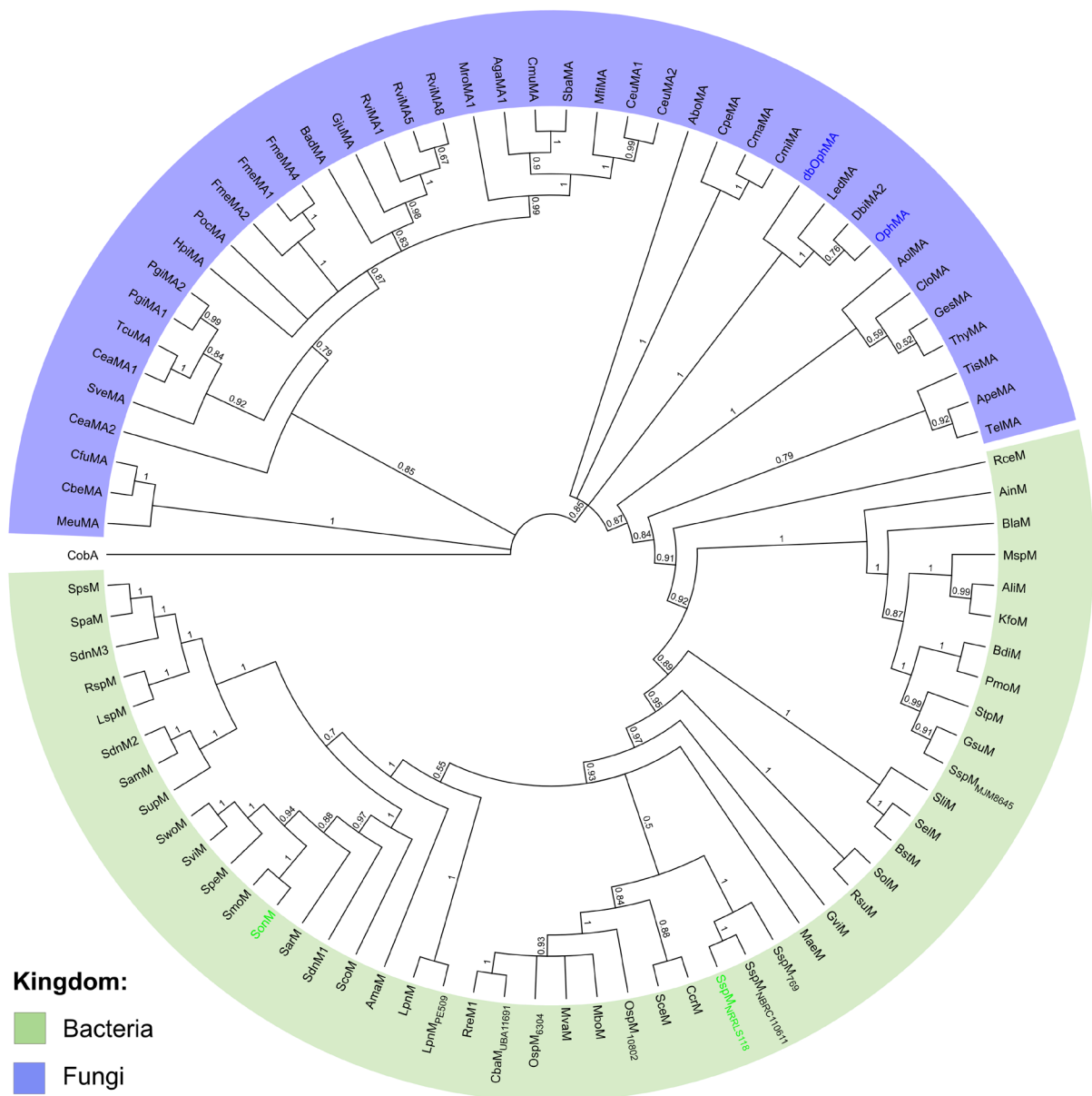
<b>Variables</b>	<b>Value (M)</b>
$E$	0
$S$	7.57E-05
$ES$	4.30E-06
$EP$	7.00E-07
$P_1$	1.03E-05
$P_2$	9.00E-06
<b>Parameters</b>	<b>Value (s<sup>-1</sup>)</b>
$k_1$	1.21E+03
$k_2$	8.70E-03
$k_3$	1.21E+03
$k_4$	8.70E-02
$k_{1r}$	1.21E-02
$k_{2r}$	0
$k_{3r}$	1.21E-03
$k_{4r}$	0



Supplementary Table 5. Crystallographic data and statistics.

DATA COLLECTION						
	SonM— SonA- 2Me—SAH	SonM— SonA-2Me	SonM- Y58F— SonA-2Me	SonM- Y93F— SonA-2Me	SonM-R67A— SonA-0Me— SAH	SonM—SonA- BBD—(±)SAM
<b>PDB ID</b>	7LTE	7LTC	7LTF	7LTH	7LTS	7LTR
<b>Resolution (Å)</b>	2.0	2.0	2.2	2.1	2.32	1.75
<b>Diffraction source</b>	APS Argonne 23ID-B					
<b>Wavelength (Å)</b>	1.033167				0.991840	
<b>Detector</b>	EIGER 16M					
<b>Rotation range per image (°)</b>	0.4	0.2	0.4	0.4	0.5	0.5
<b>Total rotation range (°)</b>	180	250	260	300	150	225
<b>Space group</b>	P2 <sub>1</sub>	P2 <sub>1</sub>	P2 <sub>1</sub>	P2 <sub>1</sub>	P422	P2 <sub>1</sub>
<b>Unit-cell parameters (Å)</b>	a = 52.54, b = 108.75, c = 59.14; α = γ = 90.000, β = 94.0	a = 52.51, b = 108.62, c = 59.02; α = γ = 90.000, β = 94.0	a = 52.37, b = 108.50, c = 58.95; α = γ = 90.000, β = 94.1	a = 52.34, b = 108.71, c = 58.96; α = γ = 90.000, β = 93.9	a = b = 80.95, c = 236.47; α = γ = β = 90.0	a = 72.24, b = 112.86, c = 85.87; α = γ = 90.000, β = 97.8
<b>Resolution range (Å)</b>	2.0 (2.1-2.0)	2.0 (2.1-2.0)	2.2 (2.3-2.2)	2.1 (2.2-2.1)	2.32 (2.45-2.32)	1.75 (1.85-1.75)
<b>N° of reflections (last bin)</b>	134820 (19119)	197428 (27550)	152302 (17833)	205844 (27677)	370988 (53390)	579688 (85609)
<b>N° of unique reflections (last bin)</b>	42332 (6106)	43871 (6015)	33041 (4031)	38054 (4943)	35110 (5161)	134823 (20453)
<b>Completeness (%) (last bin)</b>	94.6 (95.8)	98.4 (99.3)	99.1 (97.8)	99.2 (99.5)	99.9 (99.9)	98.3 (97.4)
<b>Redundancy</b>	3.18 (3.27)	4.50(4.58)	4.61 (4.42)	5.41(5.60)	10.56 (10.34)	4.30 (4.19)
<b><math>\langle I/\sigma(I) \rangle</math></b>	14.49 (8.41)	15.53 (8.45)	15.48 (9.42)	17.50 (12.27)	19.16 (2.33)	24.88 (2.56)
<b>R<sub>meas</sub> (%)</b>	6.2 (14.1)	6.1 (17.3)	6.3 (13.0)	5.8 (8.8)	8.9 (172.5)	3.5 (70.0)
<b>CC<sub>1/2</sub></b>	99.7 (99.1)	99.8 (98.8)	99.7 (99.4)	99.8 (99.7)	99.9 (84.6)	99.9 (82.2)
REFINEMENT STATISTICS						
<b>R<sub>free</sub>/R<sub>work</sub></b>	24.44 / 21.37	23.31 / 20.55	23.76 / 22.55	22.92 / 18.55	23.76 / 19.14	21.05 / 17.96
<b>N° of total model atoms</b>	6266	5711	5678	6181	5295	10920
<b>N° of water molecules</b>	1044	543	518	1007	69	993
<b>Ramachandran core and allowed (%)</b>	99.6	99.5	99.4	99.3	99.3	99.3
<b>Ramachandran generously allowed (%)</b>	0.0	0.2	0.3	0.2	0.0	0.2
<b>Ramachandran outliers (%)</b>	0.3	0.3	0.3	0.5	0.7	0.5
<i>Rmsd from ideal</i>						
<b>Bond lengths (Å)</b>	0.012	0.006	0.004	0.012	0.002	0.002
<b>Bond angles (°)</b>	1.40	1.26	1.40	1.65	1.16	1.12

**Supplementary Figure 1. Phylogenetic tree of borosin methyltransferase domains including putative split borosin domains encoded in bacteria.** Bayesian posterior probability values support tree branching using the sequence alignment (Supplementary Fig. 2) along with the methyltransferase CobA from *Bacillus megaterium* as the outgroup. The tree was constructed using the MrBayes plugin in Geneious 2019.2 using the following parameters: [Rate Matrix (fixed): wag; Rate Variation: propinv; Gamma Categories: 4; Chain Length: 1,110,000; Subsampling Freq: 1,000; Heated Chains: 4; Burn-in Length: 10,000; Heated Chain Temp: 0.2; Random Seed: 25,028; Unconstrained Branch Lengths: GammaDir (1, 0.1, 1, 1)]. Protein names are listed in general agreement with suggested RiPP nomenclature.<sup>1</sup> Proteins are named in an XxxM/A format that signifies the first letter of the encoding organism's genus followed by two lowercase letters from the species. The terminal letters MA denote a fused methyltransferase and precursor, while the individual letters denote separate methyltransferases (M) or precursors (A). Strain specific identifiers are added when species names are unavailable or when strain-specific genetic differences are present. Previous structurally defined borosin precursors are highlighted in blue; split borosin pathways interrogated in this study are highlighted in green.



**Supplementary Figure 2. Geneious sequence alignment of previously identified fungal borosin and putative bacterial split borosin methyltransferase domains. A cutoff of 90% amino acid identity was used to remove near-duplicate sequences. Methyltransferase domain sequences correspond to Gly10-Ala242 of SonM. Key active site residues are marked with an asterisk (\*). Previous structurally defined borosin precursors are highlighted in blue; split borosins interrogated in this study are highlighted in green.**

	1	10	20	30	40	50	60
AboMA	GKLV	LVGSGI	GSII	G.QF.	T.L.	S.AVAH	IEADRVFV
Agama1	GTLL	IAGSGI	ASIG	G.HI.	T.L.	ETLSYI	QEQADKVVYV
AclMA	GKLI	LVGTGVR	SLIC	Q.L.	T.L.	EALDEI	IRADVIVYV
ApeMA	GKLV	MVGSIG	KSTIS	H.M.	T.L.	ETVSHI	EQADKVFYV
BadMA	GSLT	IAGSGI	ASVA	H.II.	T.L.	ETLSHI	IRADKVFYV
CbeMA	GELV	VVGTGI	ASIR	Q.M.	T.V.	EALDYI	IQRADKVFYV
CeaMA1	GSLI	IAGSGI	SSVA	H.H.	T.L.	ETVSHI	EQADKVFYV
CeaMA2	GTLI	IAGSGI	ASIR	H.II.	T.L.	ETLSYI	IEADKVIYV
CeuMA1	GSLT	IAGSGI	ASIG	H.H.	T.L.	ETLSYI	EQADKVVYV
CeuMA2	GSLT	IAGSGI	ASVA	H.H.	T.L.	EVLSTY	IQEQADKVIYV
CfuMA	GELV	VVGTGI	ASLR	Q.L.	T.V.	EALDYI	IQRADKVFYV
CloMA	GSLT	IAGSGI	FRSII	Q.F.	T.T.	EALMHY	IEAAEKLYV
CmaMA	GQLT	IAGSGI	ASIN	H.M.	T.L.	QAVACI	IEADVVYV
CmiMA	QQLT	IAGSGI	ASIS	H.L.	T.L.	QAVSAI	ENADIVYV
CmuMA	GTLI	IAGSGI	ASIG	H.II.	T.L.	ETLSHI	IQADKVIYV
CpeMA	GSLT	IAGAGV	STIG	H.H.	T.L.	QTVSAI	ENADIVYV
dbOphMA	GSLT	IAGSGI	ASIG	Q.II.	T.L.	QALSHI	IEAAKVFYV
DbiMA2	GSLI	VVGTGI	ESIG	Q.M.	T.L.	QALSYI	IEAAKVFYV
FmeMA1	GSLT	IAGTGI	ASIK	H.H.	T.L.	ETLSYI	IEAAEKVVYV
FmeMA2	GSLT	IAGSGI	ASIK	H.H.	T.L.	ETVSHI	EQADKVVYV
FmeMA4	GSLT	IAGTGI	ASIK	H.H.	T.L.	ETLSYI	IEAAEKVVYV
GesMA	GGLV	VVGSIG	RSVS	Q.L.	T.L.	EAVMHY	IEADTVLYV
GjuMA	GSLT	IAGSGI	ASV	G.HI.	T.L.	ETLAIY	IKFSKVFYV
HpiMA	GSLT	IAGSGI	ASIR	H.M.	T.L.	ETLSAI	IKSADKVVYV
LedMA	GSLT	IAGTGI	ESIG	Q.M.	T.L.	QTLSTY	IEAADKVFYV
MfiMA	GSLT	IAGSGI	ASIR	H.II.	T.L.	ETLSHI	IRADKVVYV
MroMA1	GSLT	IAGSGI	ASIG	H.H.	T.L.	ETLALY	IEADKVIYV
MeuMA	GELV	VVGTGI	ASLR	Q.M.	T.V.	EALDYI	IQRADKVFYV
OphMA	GSLT	IAGTGI	ESIG	Q.M.	T.L.	QALSYI	IEAAKVFYV
PgiMA1	GSLT	IAGSGI	ASV	H.M.	T.L.	ETLAIY	IEADKVFYV
PgiMA2	GSLT	IAGSGI	ASVA	H.II.	T.L.	ETVAYL	IEADKVFYV
PochMA	GTLV	IAGSGI	ASIA	H.II.	T.L.	ETLSHI	IKESDRVYV
RviMA1	GTLT	IAGSGI	ASVA	H.H.	T.L.	ETLSYI	IKESKIFLYV
RviMA5	GTLT	IAGSGI	ACVA	H.H.	T.L.	ETLSYI	IKESKIFLYV
RviMA8	GTLT	IAGSGI	ASIA	H.H.	T.L.	ETLSYI	IKESKIFLYV
SbaMA	GTLT	IAGSGI	ASIA	H.H.	T.L.	ETLSYI	IKESKIFLYV
SveMA	GSLT	IAGTGI	ATLA	H.M.	T.L.	ETVSHI	IEADKVVYV
TcuMA	GSLI	IAGSGI	ASVA	H.F.	T.L.	ETVSHI	EQADKVFYV
TelMA	GRLV	MVGSIG	KSTIA	H.L.	T.L.	EALGHY	IEADKVFYV
ThyMA	GSLF	IAGSGI	RSIA	Q.L.	T.L.	EAMHII	ENADKVFYV
TisMA	GKLV	IAGSGI	RSIS	Q.F.	T.L.	EAVAHY	IEADKVFYV
SonM	GSLV	CVGTGL	QLAG	Q.I.	S.VL	SR.SYI	IEADIVFSL
AnaM	GSLV	CVGTGM	MVGA	H.L.	S.PI	CQ.SHI	EQADVVYV
LspM	GSLI	ACVGMGI	TLGSH	L.TP	S.R.	SHIEQAD	VVFAALSD
RspM	GSLI	CVGLGM	TLGSH	L.GP	A.R.	SHIEQAD	VVFAALSD
SpaM	GSLI	ACVGMGI	TLGAH	L.CP	S.K.	SYIEQAD	VVFAALSD
SpsM	GSLI	ACVGMGI	TLGAH	L.CP	A.K.	SYIEQAD	VVFAALSD
SdnM3	GSLI	CVGMGM	TMGSH	L.TP	S.R.	NYIEQAD	VVFAALSD
SamM	GNI	CVGTGI	L.IGGH	L.SP	A.Q.	NLIEQAD	VVFAALSD
SdnM2	GSLT	CVGMGM	MLGGH	L.SP	A.H.	SHIEQAD	VVFAALSD
SupM	GSLT	CVGMGM	MLGAH	L.SP	S.S.	LSHIEQAD	VVFAALSD
SarM	GNI	CVGTGL	QLAG	Q.I.	C.A.	L.S.	LSYIEHAD
SmoM	GSLV	CVGTGL	QLAG	Q.I.	N.V.	MRSYI	IEHADIVFSL
SpeM	GTLV	CVGTGL	NLAG	Q.I.	S.VL	SK.SYI	ENADVVFSI
SviM	GSLV	CVGTGL	NLAG	Q.I.	S.VL	SK.SYI	ENADVVFSI
SwoM	GSLV	CVGTGL	NLAG	Q.I.	S.VL	SK.SYI	ENADVVFSI
SdnM1	GSLV	VGTGL	QLAG	Q.I.	S.VL	SR.SYI	ENADRVFSL
ScoM	GSLV	CVGTGL	KMAG	H.I.	S.VI	SR.SYI	EHADKIFTL
CcrM	GSLT	VVGTGIR	RLVT	Q.L.	T.P.	EAAARI	IRDAEKVFYV
CbaM <sub>HBA11691</sub>	GCLT	IAGTGI	QFV	G.QV.	T.L.	AAKAWI	EQADKVLV
MboM	GSLT	IAGTGI	QLV	H.H.	T.L.	GAKAWI	EQADKVLV
MvaM	GSLT	VVGTGI	QLV	H.H.	T.L.	AAKAWI	EQADKVLV
OspM <sub>10802</sub>	GSLT	VVGTGI	QLV	H.H.	T.L.	AARSWI	EQADKVLV
OspM <sub>304</sub>	GSLT	IAGTGI	Q	LGGH	L.T.	AAQAWI	EQADKVLV
RreM <sub>1</sub>	GCLT	IAGTGI	QFV	G.QV.	T.L.	AAKAWI	EQADKVLV
SceM	GSLV	VVGTGI	QFV	G.QV.	T.L.	AAQAWI	EQADKVLV
SolM	GSLV	VAGSGI	KGIA	H.L.	T.Q.	EAAAGW	EQADHVVYV
RsuM	.....	.....	.....	.....	.....	ETKGM	IQSDVVYV
SspM <sub>769</sub>	AELI	VVGTG	YRAV	G.D.L.	T.L.	EAKACI	EQADTVLCL
SspM <sub>HARC110611</sub>	AQLV	VVGTG	YRAI	G.D.L.	T.V.	EARACI	EQADKVLCL
SspM <sub>HRRLS118</sub>	AQLV	VVGTG	YRAI	G.D.L.	T.V.	EARACI	EQADKVLCL
GviM	GSLI	VVGTGIR	A	G.H.L.	S.Q.	EAI	SAVRTADAVAYCIAE
MaeM	GSLV	VVGTGLH	V	G.Q.L.	T.L.	ESRANI	IEADHVVYV
SelM	GRLV	IAGSGI	KAVS	H.F.	T.L.	EAAQHI	IRNADIVLYA
SliM	GKLT	IAGSGI	KSTIA	H.F.	T.L.	ESQAHY	EQADIVLYA
AliM	GELT	IIGSGI	ETMG	F.T.	I.GD	E.LI	RGADAVFPCVAD
MspM	GELT	IIGSGI	ETV	G.F.	S.L.	GDQELI	RSADHVFYV
KfoM	GELI	IIGSGI	ETI	G.F.	S.L.	GDQELI	IRADKVFYV
BdiM	GSLI	IAGSGI	EASG	F.F.	I.RA	DE.A	IRAADHVFYV
PmoM	GTLI	IAGSGI	EASG	F.F.	S.R.	SDEARI	LVADHVFYV
GsuM	GELT	IAGSGI	ETV	G.F.	T.SA	DE.IR	IRADKVFYV
StpM	GDLT	IIGSGI	EAVG	F.F.	T.SV	DE.SL	IKRADHVFYV
SspM <sub>MUM645</sub>	GELT	IAGSGI	EAA	G.F.	I.TRA	DE.RL	IRADHVFYV
AinM	GELT	IIGSGI	GYM	G.F.	T.R.	DAFYI	TDADHVFYV
RceM	GSLT	VVGTGL	RALS	H.M.	T.L.	EAI	SHIRNADRVFV
BlaM	GDLI	ILGAGI	ASV	G.F.	T.M.	DAEMYI	IRADSVFV
LpnM <sub>LPE509</sub>	FKLI	IAGTGI	KFLS	H.L.	T.I.	EVKSAI	ETSCVFL
LpnM	.....	.....	.....	.....	.....	.....	.....
BstM	GRLV	IAGSGI	KAVS	H.F.	T.L.	EAAQAWI	EQADIVLYA

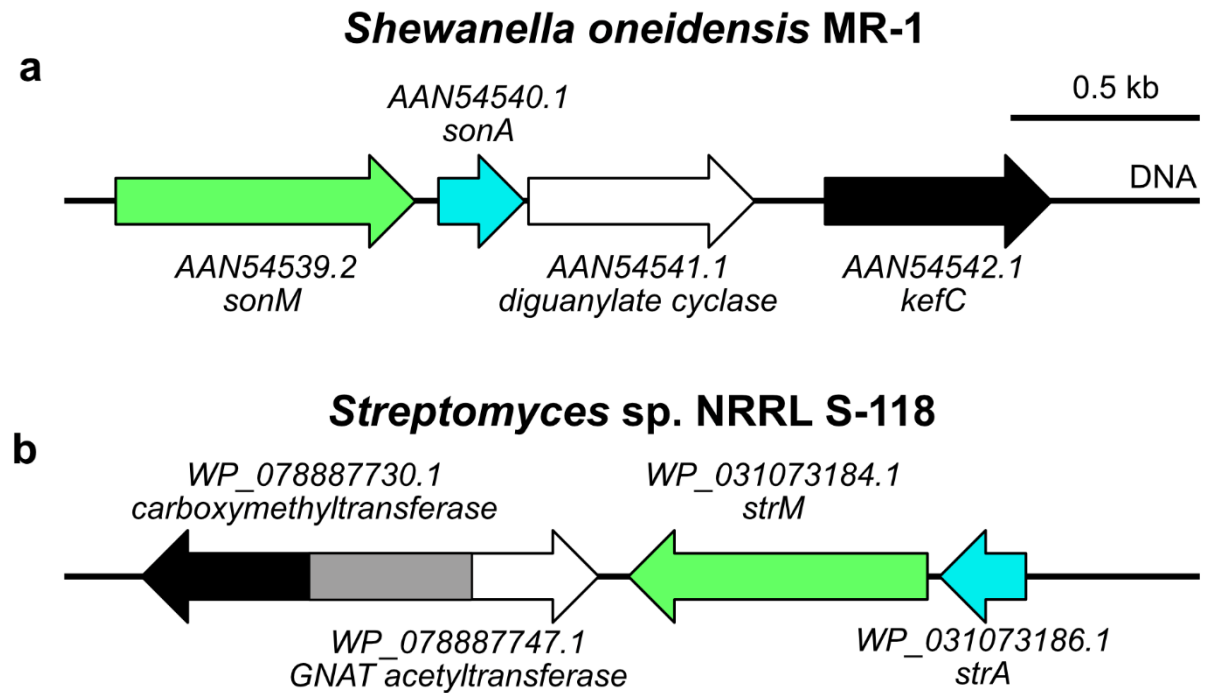
	70	80	90	100	110	120	130				
Aboma	.P.RMDTYIQMAE	EVMLRE	LRKKG	YSVVGVYI	GHPGVFV	.T.FSHRAL	S.IARDEGYS	AKMLPGVSAED	NLFADIGID	SR	
AgaMA1	.I.RYETVYVQMAE	EVMLRDV	RRAD	YNVVGVFY	GHPGVFV	.S.FSHRAL	.AIARDEGYS	AKMLPGVSAED	YMFSDLGDF	AV	
AolMA	IP.EADITYIQIAE	EVMLAATR	KDQ	RRVVGAFF	GHPGLFM	.S.FNRRAL	.AIAQAEQY	TAKILPGVSD	CLLADLGVD	PSF	
ApeMA	.N.RYITVYVQMAE	EVMLCLQA	ARRD	GFSSV	GVFYGHP	.S.FSHRAL	.GIAKREGIE	YAKMLPGVSAED	CLFADLGVD	SF	
BadMA	.A.RYDSYVQMAE	EVMLQDVR	CGG	KDVLGIFY	GHPGVFV	.S.FSHRAL	.AIARSEGY	KAKMLPGVSAED	YLFADLEFD	SV	
CbeMA	.N.RYITVYVQMAE	EVMLSSVR	KKG	KLTVAVFY	GHPGVFV	.T.FSHRAL	.YIARDEGYS	KAQMLPGVSAED	CLYADLGID	PAS	
CeaMA1	.H.RYQTYVEMAE	EVMLREVR	AG	HSVFGIFY	GHPGVL	.T.FAHRAL	.TIARDEGYS	EARMLPGVSD	YMFADLELD	GGQ	
CeaMA2	.I.RYDITYVQMS	EVMLRDVR	AG	FDVLGIFY	GHPGVFV	.S.FTORAM	.SIALDEGYS	QARMLPGVSAED	YLFADLRVD	CKM	
CeuMA1	.I.RFETIYIQMS	EVMLRDVR	AG	HSVVLGIFY	GHPGVFV	.S.FSHRAL	.AIALDEGYS	KARMLPGVSAED	YMFSDIGFD	PAL	
CeuMA2	.M.RSETIYVQMS	EVMLRDVR	SG	YNVLAIFY	GHPGVFV	.C.FTHRAL	.SIAARSEGY	TAKMLPGVSAED	YMFSDIGFD	PAV	
CfuMA	.P.RNATYIQMAE	ETILASV	RKKG	NMTVAVFY	GHPGVFV	.T.FSHRAL	.YIARDEGYS	KAKMLPGVSAED	CLYADLDID	PAS	
CloMA	.P.RYETIYIQM	TEAMLR	SVRD	GKATV	VLYGHP	.H.FSHRAL	.AIARSEGY	DWMLLPGVSD	YLFADLGID	PSN	
CmaMA	.E.RDITYIQMAE	EVMLNHVR	AG	KNVGVFY	GHPGVFV	.C.FTHRAL	.YIARNEGY	RAVMLPGVSAED	CLYADLGID	PSN	
CmiMA	.Q.RDITYIQMAE	EVMLIRVR	CGG	QNVGVFY	GHPGVFV	.C.FTHRAL	.YIARSEGY	KARMLPGVSAED	CLFADLGID	SS	
CmuMA	.M.RYDITYVQMS	EVMLRDVR	SG	HNVLGIFY	GHPGVFV	.S.FTHRAL	.AIARDEGYS	TAKMLPGVSAED	YMFSDLGDF	PAF	
CpeMA	.P.RLEIYTHQMV	EVMLSKVR	SG	QDVLGVFY	GHPGV	.NT.FAAQAF	.KIAARDEGYS	TAKMLPGIT	INDALLAD	VVAD	PAL
dbOphMA	.H.RMDTYIQMAE	EVMLKEVR	NG	LDVVGVFY	GHPGVFV	.S.FSHRAL	.AIARSEGY	KARMLPGVSAED	CLFADLRID	PSH	
DbiMA2	.S.RMDTYIQMAE	EVMLKEVR	NG	LDVVGVFY	GHPGVFV	.N.FSHRAL	.AIARSEGY	QARMLPGVSAED	CLFADLRID	PSN	
FmeMA1	.H.RYDSYVQMAE	EVMLDVR	AG	HSVVLGIFY	GHPGVFV	.S.FSHRAL	.AIARDEGYS	KARMLPGVSAED	YMFADIGFD	PAV	
FmeMA2	.P.RYDSYVQMS	EVMLRDVR	AG	HSVVLGIFY	GHPGVFV	.S.FSHRAL	.AIARDEGYS	KARMLPGVSAED	YMFADIGFD	PAV	
FmeMA4	.L.RYDSYVQMAE	EVMLRDVR	AG	NSVVLGIFY	GHPGVFV	.S.FSHRAL	.AVARDEGYS	KARMLPGVSAED	YMFADIGFD	PAV	
GesMA	.E.RPDIFYVQMAE	EVMLREVR	KG	INVVAVFY	GHPGIFV	.H.FSHRAL	.AIARDEGYS	AKMLPGVSAED	CLFADLEFD	SF	
GjuMA	.S.RYDSYVQMS	EVMLRDVR	NG	LDVLGVFY	GHPGVFV	.S.FSHRAL	.AIARDEGYS	NAKMLAGVSAED	CLFADLEFD	PAS	
HpiMA	.S.RYDITYVQMS	EVMLREVR	AG	HNVLGVFY	GHPGVFV	.S.FSHRAL	.AIARDEGYS	KARMLAGVSAED	YMFADLGDF	AAA	
LedMA	.S.RMDTYIQMS	EVMLREVR	KG	LDVVGVFY	GHPGVFV	.N.FSHRAL	.AIARSEGY	KARMLPGVSAED	CLYADLRID	PSN	
MfiMA	.N.RYDSYVQMS	EVMLNDVR	AG	YNVLGVFY	GHPGVFV	.S.FSHRAL	.AIARDEGYS	RVNMLPGVSAED	YMFSDIGFD	PAI	
MroMA1	.K.RYDSYVQMS	EVMLREVR	AG	HNVLGVFY	GHPGVFV	.A.FSHRAL	.AIARDEGYS	KARMLPGVSAED	YMFADLGDF	PSN	
MeuMA	.P.RNASYVQMAE	ELMVSV	RD	NLTVAVFY	GHPGVFV	.F.FTHRAL	.HIAARDEGYS	KARMLPGVSAED	CLYADLRID	DGT	
OphMA	.S.RLNTYIQM	SELMVREVR	KG	LDVVGVFY	GHPGVFV	.N.FSHRAL	.AIARDEGYS	RAKMLPGVSAED	CLFADLRID	PSN	
PgiMA1	.V.RYDITYVQMAE	TMLNAVR	REG	QKVLGIFY	GHPGVFV	.S.FSHRAL	.SIAARDEGYS	QARMLPGVSAED	YMFADLGDF	PAV	
PgiMA2	.S.RYDITYVQMAE	TMLNSVR	AG	HKVLGIFY	GHPGVFV	.S.FSHRAL	.AIARDEGYS	KARMLPGVSAED	YMFADLEFD	PAI	
PocMA	.V.RYDSYVQMS	EVMLRDVR	AG	HTVLGVFY	GHPGVFV	.S.FSHRAL	.AIARDEGYS	KARMLPGVSAED	YLFADLGDF	PAI	
RviMA1	.G.RHDSYIQMS	EVMLKAVR	AG	HDVLGVFY	GHPGVFV	.S.FSHRAL	.AVARDEGYS	KARMLPGVSAED	YMFADLEFD	PSL	
RviMA5	.S.RYDSYIQMS	EVMLKAVR	AG	HDVLGVFY	GHPGVFV	.S.PSYRAL	.AVARDEGYS	KARMLPGVSAED	YLFADLEFD	PCF	
RviMA8	.S.RYDSYIQMS	EVMLKAVR	AG	HSVVLGIFY	GHPGVFV	.S.FSHRAL	.AVARDEGYS	KARMLPGVSAED	YMFADLEFD	PSQ	
SbaMA	.M.RYDSYVQMS	EVMLRDVR	CGG	YNVLGVFY	GHPGVFV	.S.FSHRAL	.AIARDEGYS	IAKMLPGVSAED	YMFSDIGFD	PAV	
SveMA	.Y.RYDSYVQMAE	EVMLNAVR	REG	CNVLGVFY	GHPGIFV	.S.FSHRAL	.AIARDEGYS	EARMLPGVSAED	YMFADLGID	PAL	
TcuMA	.P.RYDSYVQMAE	EVMLNAVR	REG	CNVLGVFY	GHPGIFV	.S.FSHRAL	.AIARDEGYS	EARMLPGVSAED	YMFADLGID	PAL	
TelMA	.P.RYHTYVQMAE	RMLREVR	AG	FYVGVFY	GHPGLFV	.N.FSHRAL	.AIARDEGYS	QARMLPGVSAED	CLFADVGD	PSN	
ThyMA	.L.RNETIYIQMAE	EVMLREVR	SG	LRVGVFY	GHPGNFV	.S.FTHRAL	.AIARDEGYS	KARMLPGVSAED	CLFADLRID	PSY	
TisMA	.P.RHQTYIQMAE	EVMLQEV	RKKG	FSVGVFY	GHPGVFV	.N.FAHRAL	.SIAARDEGYS	EATMLPGVSAED	CLYADLRID	PSR	
SonM	.N.RRDITYVQMV	NAILDVAVR	AG	KTKVCLAL	YGHPGVFA	.C.VSHMAL	.TRAKAEGYS	AKMEPGVSAED	CLWADLGID	PSN	
Amam	.D.RHDITYAQMS	AMLEREVR	KG	EKVVGAFY	GHPGVFA	.KVP.HDAT	.AIARDEGYS	RAKMLPGVSAED	CLYADLGID	PSA	
LspM	.S.RMRTYREVM	QALMIAEVR	AG	KSVCAIFY	GHPGIFA	.P.WSTHKVV	.ELARAEQYS	KALMEAGVSAED	CLYADLGID	PSR	
RspM	.S.RMLTYRQMV	FAITAEVR	AG	KRVCAIFY	GHPGVFA	.P.HKST	.ETARDEGYS	SAHMEPGVSAED	CLYADLGID	PSR	
SpaM	.S.RLLITYNEM	VDAIMTEVR	AG	KKVVGAFY	GHPGVFA	.P.HKST	.AMAKAEGYS	AKMIPGVSAED	CLYADLGID	PSR	
SpsM	.S.RHITYREM	VDAIMTEVR	AG	KKVVGAFY	GHPGVFA	.P.HKST	.ELAKAEGYS	DAVMVPGVSAED	CLYADLGID	PSR	
SdnM3	.S.RLLSYQNM	INAVLREVR	AG	KNVVGAFY	GHPGVFA	.M.VTHKLL	.AQSKKEGYS	YCHMEAGVSAED	CLYADLGID	PSR	
SamM	.S.RNLTYNEM	VDAIMTEVR	AG	KKVVGAFY	GHPGVFA	.P.HKST	.ATARAEQYS	AKMIPGVSAED	CLYADLGID	PSR	
SdnM2	.S.RNISYGM	VQVMLAEVR	AG	KKVVGAFY	GHPGIFV	.K.STHEATAK	.AKKEGYS	AKMIPGVSAED	CLYADLGID	PSR	
SupM	.S.RLLSYREM	VDAIMTEVR	AG	KKVVGAFY	GHPGVFA	.P.HKST	.IARNEGYS	AKMIPGVSAED	CLYADLGID	PSR	
SarM	.N.RRDITYAEM	VDAIMTEVR	AG	KLVVCAIFY	GHPGVFA	.C.VAHWSIKQ	.ARSEGYS	AKMIPGVSAED	CLWADLGID	PSH	
SmoM	.N.RRDITYEM	VDAIMTEVR	AG	KKTVCALY	GHPGVFA	.C.VSHLAI	.AQAKAEGYS	AKMEPGVSAED	CLWADLGID	PSN	
SpeM	.S.RRDITYDQM	VDAIMTEVR	AG	KVVVCAIFY	GHPGVFA	.C.VSHFAIAQ	.ARDEGYS	AKMIPGVSAED	CLWADLGID	PSN	
SviM	.S.RRDITYDQM	VDAIMTEVR	AG	KVVVCAIFY	GHPGVFA	.C.VSHFAIAQ	.IARDEGYS	AKMIPGVSAED	CLWADLGID	PSN	
SwoM	.N.RRDITYDQM	VDAIMTEVR	AG	KVVVCAIFY	GHPGVFA	.C.VSHFAIAQ	.ALDEGYS	AKMIPGVSAED	CLWADLGID	PSN	
SdnM1	.N.RRHITYKQ	MVEAMLVAVR	AG	EQVVCALY	GHPGVFA	.C.VSHLAI	.AQARDEGYS	AKMIPGVSAED	CLWADLGID	PSN	
ScoM	.S.RRDITYHQM	VDAIMTEVR	AG	ERVVCAIFY	GHPGVFA	.C.VHMAIKQ	.ARDEGYS	AKMIPGVSAED	CLWADMGID	PSD	
CcrM	.D.RLSTYIE	IVERALLAA	ARRG	VRVCLVMY	GHPGVLV	.T.FAHDAT	.SARAEGYS	AKMIPGVSAED	CLFADLGVD	PSA	
CbaM	.D.RKETYIQMV	EVMLIAEVR	AG	LNVCVAFY	GHPGVFA	.D.SAHQAL	.RRARDEGYS	AKMIPGVSAED	CLFADLRID	PSN	
MboM	.QWRDITYQEM	TERILTDVR	AG	LNVCVAFY	GHPGVFA	.N.FAQAL	.KQARDEGYS	AKMIPGVSAED	CLFADLRID	PSN	
MvaM	.R.RRETYREM	VDAIMTEVR	AG	LNVCVAFY	GHPGVFA	.Y.PTHEAT	.KQARDEGYS	AKMIPGVSAED	CLFADLRID	PSN	
OspM	.R.RRKTYDEM	VGRILVREVR	AG	LNVCVAFY	GHPGVFA	.D.FAHEAT	.RQARDEGYS	AKMIPGVSAED	CLFADLRID	PSN	
OspM	.R.RRQTYGKM	VDRMFAVR	AG	GNTCAIFY	GHPGVFA	.D.FAHGAT	.AQARDEGYS	AKMIPGVSAED	CLFADLRID	PSN	
RreM1	.W.RRQTYGEM	VDAIMTEVR	AG	LDVCAIFY	GHPGVFA	.N.PTREAT	.RQARDEGYS	AKMIPGVSAED	CLYADLRID	PSN	
SceM	.TP.RRSITYAEM	VERILAEVR	AG	LNVCVAFY	GHPGVFA	.S.FAHAAV	.RQARDEGYS	AKMIPGVSAED	CLFADLRID	PSN	
SolM	.P.RIRTYREM	TDAVREVR	AG	LNVCVAFY	GHPGVFA	.H.PGHAAV	.RTVREGY	AKMIPGVSAED	CLFADLRID	PSN	
RsuM	.P.RLDITYIDM	SDEMLKYVR	AG	KSVVCAIFY	GHPGVFA	.Y.FSHRAL	.RIARDEGYS	AKMIPGVSAED	CLYADLRID	PSN	
SspM	.H.RSESYEM	VQVMLAEVR	AG	LRVCCALY	GHPGVFA	.Y.FGHEAV	.RRARLEGI	AKMIPGVSAED	CLFADLRID	PSN	
SspM	.P.RYASIEEM	VQHILSEVR	AG	QFVCCALY	GHPGVFA	.Y.AGHEAV	.RRARLEGI	AKMIPGVSAED	CLFADLRID	PSN	
SspM	.P.RSASIEDM	VQHILSEVR	AG	QFVCCALY	GHPGVFA	.Y.TGHEAV	.RRARLEGI	AKMIPGVSAED	CLFADLRID	PSN	
GviM	.P.RDQTYHQM	VDAIMTEVR	AG	LQVAVFY	GHPGVFA	.Y.PSHES	.RQARDEGYS	AKMIPGVSAED	CLFADLRID	PSN	
MaeM	.N.RLLSYEEM	IAHITNLV	HNN	LEVCAIFY	GHPGIFV	.Y.PSHESI	.QRLREGYS	AKMIPGVSAED	CLYADLRID	PSN	
SelM	.S.RLLITYSQM	TERILAEVR	AG	KYVCAIFY	GHPGVFA	.T.PSHNAL	.ELARSEGY	HAVMIPGVSAED	CLYADLRID	PSN	
SliM	.N.RIITYTQM	TERVMMEL	RS	KYVCAIFY	GHPGVFA	.T.PSHNAL	.ELARSEGY	HAVMIPGVSAED	CLYADLRID	PSN	
AliM	.V.RYTYTQM	SEAMLEHVR	AG	KVVAVFY	GHPGIFV	.L.STHRAL	.MIARRECH	KAVMIPGVSAED	CLCADLRID	PSH	
MspM	.V.RYTYTQM	SEAMLEHVR	AG	KVVAVFY	GHPGIFV	.L.STHRAL	.LIARRECH	KAVMIPGVSAED	CLCADLRID	PSH	
KfoM	.V.RYTYTQM	TEAQLYVR	AG	LKVAVFY	GHPGIFV	.L.STHRAL	.KIARRECH	KAVMIPGVSAED	CLCADLRID	PSH	
BdiM	.R.RYTYTQM	TEAQLYVR	AG	LKVAVFY	GHPGIFV	.L.STHRAL	.KIARRECH	KAVMIPGVSAED	CLCADLRID	PSH	
PmoM	.P.RYTYTQM	TEAQLYVR	AG	LKVAVFY	GHPGIFV	.L.STHRAL	.KIARRECH	KAVMIPGVSAED	CLCADLRID	PSH	
GsuM	.I.RYTYTQM	TEAQLYVR	AG	LKVAVFY	GHPGIFV	.L.STHRAL	.KIARRECH	KAVMIPGVSAED	CLCADLRID	PSH	
StpM	.L.RYTYTQM	TEAQLYVR	AG	LKVAVFY	GHPGIFV	.L.STHRAL	.KIARRECH	KAVMIPGVSAED	CLCADLRID	PSH	
SspM	.D.RYTYTQM	TEAQLYVR	AG	LKVAVFY	GHPGIFV	.L.STHRAL	.KIARRECH	KAVMIPGVSAED	CLCADLRID	PSH	
AinM	.P.RYTYTQM	TEAQLYVR	AG	LKVAVFY	GHPGIFV	.L.STHRAL	.KIARRECH	KAVMIPGVSAED	CLCADLRID	PSH	
RceM	.P.RKQTYVQMS	EVMLREVR	AG	SAVAVFY	GHPGIFV	.F.FARRIL	.SIAARKEGYS	RAVMIPGVSAED	CLCADLRID	PSH	
BlaM	.P.RYNTYMQM	TEAQLYVR	AG	RTVVGIFY	GHPGIFA	.L.STHRAL	.MIARRECH	KAVMIPGVSAED	CLCADLRID	PSH	
LpnM	.Q.RSESYDK	IANELLSTI	QKN	DHVVCLY	GHPGIFV	.S.VVEK	.IKKII	SEEDII	IQIMPISAM	D	
LpnM	.D.RQISYDK	IRDKILTE	LE	DDVCLY	GHPGIFV	.S.VVEK	.IKKII	SEEDII	IQIMPISAM	D	
LpnM	.Q.RQISYDK	IRDKILTE	LE	DDVCLY	GHPGIFV	.S.VVEK	.IKKII	SEEDII	IQIMPISAM	D	
BstM	.A.RLITYIQM	TEAQLYVR	AG	KYVCAIFY	GHPGVFA	.T.PSHNAL	.AIARDEGYS	HAVMIPGVSAED	CLYADLRID	PSN	

140 150 160 170 180 190

Aboma P. GCLTYEATDILLR...NR TLV.PSSHVLVFOVGC.IG.LSDF...R.FK.G.FD.NIN.FD.VL.LDRLEQVYGP.DHA  
 AgaM1 P. GCMTEATAMLNH...NK KLD.PSIHNLIWQVGA.VG.IDTM...V.F.D...NRK.FH.VL.VDRLEEDFVGG.DHR  
 AolM1 I. CCLTCEARDFMIH...DH.LGLTSRHHVIMYEVGV.LC.EYG...D.D.S...KTDY.FE.YF.VNRLLEIYCN.EHS  
 ApeMA T. GCOHYEATDILLR...DRPIS.PYSHLIVWQVGV.VG.DTGF...N.FG.G.FTCTK.FQ.VL.VDRLEEVYGS.DHR  
 BadMA H. CCAATFEATELLLR...EKPLN.TTMHNLIIWQVGA.VG.VDDMV...F.T...NSK.LH.VL.VDRLEKDFVGG.EHQ  
 CbeMA S. GCSMYEASFLLLE.P.DR.LD.SRHHLIIWQVGC.VG.KEAMV...F.D...NKE.IY.KL.ADYLEAEYGP.DHP  
 CeaMA1 H. GCMTHI EATDILLR.D.RR.LD.PSVHNIILQPSR.VG.SATL...E.K.E.AS.K.FQ.VL.VDRLEVRDFVGG.DHK  
 CeaMA2 F. CCAAYEATELLYR.K.RR.LN.PTMQNIWQVQKRF.T.IIKL...TSPD.T...QNSK.FG.VL.VDRLEEDYGP.DHK  
 CeuMA1 P. GCTIQEATHELLLH...NK KLD.PSMHNIWQVGV.VG.ADTM...N.F.D...NRQ.FH.VL.VDRLEEDFVGG.SHK  
 CeuMA2 P. GCMTEATSLLIY...NK QLD.PSVHNIWQVGS.VG.VDNMV...F.D...NKQ.FH.VL.VDRLEEDFVGG.IHK  
 CfuMA P. GCSMYEASFLLLE.P.DR.LD.SRHHLIIWQVGC.VG.KEAMV...F.D...NKE.IY.KL.ADYLEAEYGP.KHP  
 CloMA P. CTOIV EATEILLK...ERPLL.TSSHVILVWQVGC.VG.NFTF...N.FS.G.IKNDK.FD.VL.VDRLEIQEYGP.DHP  
 CmaMA V. GCOHYEATDMLV...NRPLN.SSSHVLVWQVGI.VG.KADFK...FAYD.P.KENHE.FG.KL.VL.VDRLELEVYGP.DHT  
 CmiMA V. GCVTYEATDILLV...KR PIN.PASHVLVWQVGI.VG.KSNFK...FDYT.S.DENIH.FT.KL.VL.VDRLEAEYGP.EHS  
 CmuMA P. GCMTEATDILLV...RKL.D.PSVHNIWQVGV.VG.VDTMV...F.D...NAN.FY.VL.VDRLEEDLGP.DHK  
 CpeMA G. GAMAYEATDFLNN...NRVLH.PEMNVIWQVGV.VG.NKHF...N.FM.E.MRSSL.LD.KL.VL.VDRLEETVGG.EKE  
 dbOphMA P. GCMTYEASDFLIR...ERPVN.IHSHVLVWQVGC.VG.VADF...N.SG.G.FKNTK.FD.VL.VDRLEQYGA.DHP  
 DbiMA2 F. GCLTYEASDFLIR...ERPVN.VHSHLIIWQVGC.VG.IADF...N.FS.G.FD.NSK.FT.VL.VDRLEQYGA.DHT  
 FmeMA1 H. GCVSYEATELLVR...DKPLL.PSSHNIWQVGA.VG.ANAMV...F.D...NGK.FN.VL.VDRLEQVYGP.DHK  
 FmeMA2 H. GCSMYEATELLVR...DKPLL.TSTHNIWQVGV.VG.AEAMV...F.D...NAK.FN.VL.VDRLEQVYGP.DHK  
 FmeMA4 H. GCVSYEATDILLR...DKPLL.PSSHNIWQVGA.VG.ANAMV...F.D...NGK.FN.VL.VDRLEEDFVGG.NHK  
 GesMA P. CCAQV EASDIIVYR.A.RPLA.TSCHVVIWQVGC.VG.HWKY...N.FT.A.FENCK.FD.VL.VDRLEQVYGP.DHP  
 GjuMA F. GCMTEATDILLR...NRPLN.PATHNIWQVGV.VG.VIDM...T.F.N.NN.K.FP.VL.VDRLEKDFVGG.NHT  
 HpiMA H. GCVTYEATEMLLR...KKQLN.PATHNIWQVGV.VG.VSNMI...F.D...NAR.FH.VL.VDRLEEDFVGG.DHQ  
 LedMA P. GCLTYEASDFLIR...ERPPTN.IYSHVILVWQVGC.VG.IADF...N.FT.G.FENSK.FG.VL.VDRLEKQYGA.EHP  
 MfiMA P. GCTIQEASTIFLLD.KR.LD.PTVHNIWQVGC.VG.VGTMA...F.D...NRQ.FH.VL.VDRLEKDFVGG.EHK  
 MroMA1 Q. GCMTEATDILLR...NK KLD.PSVHNIWQVGS.VG.VDTMV...F.D...NGK.FH.VL.VDRLEKDFVGG.DHK  
 MeuMA T. GCSMYEATYLLN.EPDR.LD.PRNHNIWQVGC.VG.K.STM...V.F.DN.S.E.IH.VL.VDRLEKTYGP.EYP  
 OphMA P. CCLTYEASDFLIR...DRPVS.IHSHVLVWQVGC.VG.IADF...N.FT.C.FD.NNK.FG.VL.VDRLEQYGA.EHP  
 PgiMA1 H. GCCAYEATQLLL...EVS.LD.TAMSNLIWQVGV.VG.VSKI...D.F.E...NSK.VK.VL.VDRLEKDFVGG.DHH  
 PgiMA2 H. GCCAYEATHILLK...NIPLD.TSINNIWQVGV.VG.VTKI...D.F.E...NSK.FK.VL.VDRLEKDFVGG.DHK  
 PocMA H. GCTSYEATDILLR...NKPLN.ASTHNIWQVGV.VG.VDTMV...F.D...NAK.FH.VL.VDRLEKDFVGG.SHT  
 RviMA1 S. GCKTCEATEILLR...DKPLD.PSIQNIWQVGS.VG.VDME...F.E.KS.K.FQ.VL.VDRLEKDFVGG.GHK  
 RviMA5 P. GCMTEATELLLR...DRSLD.PSIHNIWQVGS.VG.VDME...F.E.KS.K.LN.VL.VDRLEEDFVGG.DHK  
 RviMA8 S. ICNTYEATELLLR...DRPLD.PAIQNIWQVGS.VG.VDME...F.E.KS.K.FH.VL.VDRLEEDFVGG.DHK  
 SbaMA P. GCMTEATEGLLV...KKKLD.PSIHNIWQVGS.VG.VDTM...N.F.D...NR.E.FH.VL.VDRLEEDFVGG.DHK  
 SveMA P. GCVTYEATNPLR...NKPLN.PATHNIWQVGA.VG.ITAMD...F.E.NS.K.FS.VL.VDRLEEDLGP.NHK  
 TcuMA F. GCOHYEATELITR...NRPLN.TSVHNIWQVGV.VG.VSTL...EYQ.FS...K.FQ.VL.VDRLEEDFVGG.EHK  
 TelMA S. GCOHYEATDILLR...NRPIN.TGSHVILVWQVGV.VG.DSGFH...PQ.G.FKNTK.LH.VL.VDRLEEDFVGG.GHR  
 ThyMA P. GLOHYEATDVLV...NRPLQ.TTSHVILVWQVGV.VG.IC.SGF...F.Y.S.IENDK.FD.VL.VDRLEEDFVGG.NHP  
 TisMA P. GCOHYEATDVLV...KRPIA.KDCHVILVWQVGA.VG.DLGF...N.FK.G.FKNTK.FE.VL.VDRLEEDFVGG.DHS  
 SonM S. GHQSYEASQFMFF.N.HVP.D.PTTHLIIWQVAI.AG.EHTLT...QFH.T.S.SDR.LQ.VL.VDRLEEDFVGG.DHE  
 AmaM L. GCOHYEANQFLLY...KRREVD.TAAYVLVWQVGV.VG.DFSSA...V.FT.S.SSLO.RK.KL.VL.VDRLEEDFVGG.DHS  
 LspM V. GCOHYEASQFLLY...ERNLD.PTSHVILVWQVGV.VG.DRSLG...RFR.T.PDAYR.E.L.VL.VDRLEEDFVGG.DHE  
 RspM F. GCOHYEASQFLLY.Q.RR.ID.PTAYVLVWQVGV.VG.DQSL...ARFSTGP.AYR.Q.VL.VDRLEEDFVGG.DHE  
 SpaM F. GCOHYEASQFMFF...KRQYD.PSCY.LIIWQVGL.AG.DRSLA...KFSTG.AAHR.Q.VL.VDRLEEDFVGG.DHE  
 SpsM F. GCOHYEASQFMFF.K.RR.FD.PSSYVLVWQVGL.AG.DKSMA...KFATG.AAHR.Q.VL.VDRLEEDFVGG.DHE  
 SdnM3 L. GCOHYEASQFMFF...RRPID.PSALVILWQVGL.AG.DQSLA...RF.A.TAES.FRAVL.VL.VDRLEEDFVGG.DHE  
 SamM T. GCOHYEATQLLM...QRRID.PSALVILWQVGL.AG.DLTGL...IRPT.G.SAER.FE.LL.VL.VDRLEEDFVGG.DHE  
 SdnM2 V. GCOHYEATQFLM...HRQLD.PSALVILWQVGL.AG.DLTYG...IKPT.G.RAER.Q.VL.VDRLEEDFVGG.DHE  
 SupM N. GCOHYEATQFLM...HRRVD.PTATLIIWQVGL.AG.DLDMG...LTIT.D.AKNR.Q.VL.VDRLEEDFVGG.DHE  
 SarM S. GHQSYEATQFLM...HHPD.PTTHLIIWQVAI.AG.EHTLT...QFS.S.TKDK.LQ.VL.VDRLEEDFVGG.DHE  
 SmoM S. GHQSYEATQFMFF...NHVPD.PTTHLIIWQVAI.AG.EHTLT...QFH.T.TDR.LQ.VL.VDRLEEDFVGG.DHE  
 SpeM S. GHQSYEATQFMFF...KHTPD.PTTHLIIWQVAI.AG.EHTLT...TEFH.T.SSDR.LQ.VL.VDRLEEDFVGG.DHE  
 SviM S. GHQSYEATQFMFF...KHTPD.PTTHLIIWQVAI.AG.EHTLT...TEFH.T.SSDR.LQ.VL.VDRLEEDFVGG.DHE  
 SwoM S. GHQSYEATQFMFF...KHTPD.PTTHLIIWQVAI.AG.EHTLT...TEFH.T.SSDR.LQ.VL.VDRLEEDFVGG.DHE  
 SdnM1 C. GHQSYEATQFLM...HRPID.PSALVILWQVGL.AG.DHTL...TQFS.T.TSRR.LQ.VL.VDRLEEDFVGG.DHE  
 ScoM Y. GHQSYEATQFMFF...KRTPD.PAAHVLVWQVGL.AG.EHTLT...TQFS.T.TSER.LE.VL.VDRLEEDFVGG.DHE  
 CcrM R. GCOHYEATDVLV...RRRFD.PRSALVILWQVGL.AG.VRAH...T.A.T.LP.NLRGLR.VL.VDRLEEDFVGG.DHE  
 CbaM<sub>BA11691</sub> F. GCOHYEATDVLV...RRRFD.PRSALVILWQVGL.AG.VRAH...T.A.T.LP.NLRGLR.VL.VDRLEEDFVGG.DHE  
 MboM N. GCOHYEATDVLV...RRRFD.PRSALVILWQVGL.AG.VRAH...T.A.T.LP.NLRGLR.VL.VDRLEEDFVGG.DHE  
 MvaM N. GCOHYEATDVLV...RRRFD.PRSALVILWQVGL.AG.VRAH...T.A.T.LP.NLRGLR.VL.VDRLEEDFVGG.DHE  
 OspM<sub>10802</sub> N. GCOHYEATDVLV...RRRFD.PRSALVILWQVGL.AG.VRAH...T.A.T.LP.NLRGLR.VL.VDRLEEDFVGG.DHE  
 OspM<sub>6304</sub> G. GCOHYEATDVLV...RRRFD.PRSALVILWQVGL.AG.VRAH...T.A.T.LP.NLRGLR.VL.VDRLEEDFVGG.DHE  
 RreM1 Y. GCOHYEATDVLV...RRRFD.PRSALVILWQVGL.AG.VRAH...T.A.T.LP.NLRGLR.VL.VDRLEEDFVGG.DHE  
 SceM D. GLOHYEATDVLV...RAVD.ARTPLVLVWQVGV.VG.CAGVF...DPSD.A.ARIRLGE.VL.VDRLEEDFVGG.DHE  
 SolM P. GCOHYEATDVLV...RPLL.AESHVILVWQVGV.VG.DLGF...N.FA.G.FTNKE.LG.VL.VDRLEEDFVGG.DHE  
 RsuM V. GCOHYEATDVLV...RQLL.TDEHVVIVWQVGV.VG.DLGF...N.FS.G.YDNRR.LN.VL.VDRLEEDFVGG.DHE  
 SspM<sub>69</sub> C. GCOHYEATDVLV...RRVFD.ATSLVILVWQVGL.VG.MTDRD...PDF.D.ARVGA.A.LL.VDRLEEDFVGG.DHE  
 SspM<sub>NRRC110611</sub> Q. GCOHYEATDVLV...RRRFD.PTSALVILWQVGL.VG.MPDAH...RQF.D.PAPGA.S.CL.VL.VDRLEEDFVGG.DHE  
 SspM<sub>NRRL118</sub> R. GCOHYEATDVLV...RVFD.PTGLVILVWQVGV.VG.MTDRD...PGY.D.ARVGV.T.TL.VL.VDRLEEDFVGG.DHE  
 GviM A. GCOHYEATDVLV...RRHLD.TACGVILVWQVGV.VG.HGDYQ...GS.G.Y.DLRYP.TL.VL.VDRLEEDFVGG.DHE  
 MaeM Q. GCOHYEATDVLV...RIPD.PHSY.LIIWQVGL.VG.SYTF...S.S.T.G.IYDRRQV.VL.VDRLEEDFVGG.DHE  
 SelM P. GLOHYEATDVLV...RHID.TSANVILVWQVGV.VG.DLGF...KFG.G.YQNDK.LD.VL.VDRLEEDFVGG.DHE  
 SliM P. GLOHYEATDVLV...RSIN.TEINAVIWQVGV.VG.DVGF...KFG.G.YDNRR.LN.VL.VDRLEEDFVGG.DHE  
 AliM P. GCOHYEATDVLV...RIPD.TSLHVVIVWQVGL.VG.EMGF...RR.R.G.YINNRR.FS.VL.VDRLEEDFVGG.DHE  
 MspM P. GLOHYEATDVLV...RKP.D.TSLHVVIVWQVGL.VG.EMGF...RR.R.G.YINNRR.FS.VL.VDRLEEDFVGG.DHE  
 KfoM P. GLOHYEATDVLV...RNL.D.TSLHVVIVWQVGL.VG.ELGF...RR.Q.G.YLNNRR.FS.YF.VL.VDRLEEDFVGG.DHE  
 BdiM P. GCOHYEATDVLV...RQP.D.TGLHVVIVWQVGL.VG.ELGF...RR.Q.G.YLNNRR.FS.VL.VDRLEEDFVGG.DHE  
 PmoM P. GCOHYEATDVLV...RQP.D.TGLHVVIVWQVGL.VG.ELGF...RR.Q.G.YLNNRR.FS.VL.VDRLEEDFVGG.DHE  
 GsuM P. GCOHYEATDVLV...RHL.D.PELHVVIVWQVGL.VG.DLGF...RR.E.G.SLNSG.FA.VL.VDRLEEDFVGG.DHE  
 StpM P. GCOHYEATDVLV...RQID.TGLHVVIVWQVGL.VG.ELGF...RR.Q.G.YLNNRR.FS.VL.VDRLEEDFVGG.DHE  
 SspM<sub>KJM9645</sub> P. GCOHYEATDVLV...RID.PGLHVVIVWQVGL.VG.ELGF...RR.Q.G.YLNNRR.FS.VL.VDRLEEDFVGG.DHE  
 AinM P. GCOHYEATDVLV...RAID.PGSHVVIVWQVGL.VG.EMGF...RR.R.G.FHNNRR.FD.VL.VDRLEEDFVGG.DHE  
 RceM N. GCOHYEATDVLV...RPII.TSGHVVIVWQVGL.VG.DSAF...S.FTAG.FRHAKR.A.VL.VDRLEEDFVGG.DHE  
 BlaM P. GCOHYEATDVLV...KQP.D.TTGHVVIVWQVGL.VG.IYDF...RR.R.G.FINNNRR.FN.VL.VDRLEEDFVGG.DHE  
 LpnM<sub>PS509</sub> M. GLOHYEATDVLV...DENFS.TSHVILVWQVGL.VG.ELGVNNEINLD.RKQKAIT.VL.VDRLEEDFVGG.DHE  
 LpnM F. GCOHYEATDVLV...DKIID.PTVHLCIWQVGM.VG.NRSV...P.XP.N.Q.KSNE.LE.VL.VDRLEEDFVGG.DHE  
 BstM P. GCOHYEATDVLV...RKVD.TTANVILVWQVGV.VG.DLGF...K.FG.G.YKNDK.FD.VL.VDRLEEDFVGG.DHE

	200	210	220	230	240
AboMA	V.IH.YMAAV.L.P	.QSTTTIDRYT	IKELRDPV.IKK.R.I..TAI	STFYLP	KA
Agama1	V.VN.YIGAV.L.P	.QSTTVMDEFT	IGDLRKED.VVK.Q.F..TTV	STFYVP	RT
AclMA	L.LH.YTAAI.S.P	.LMQPVINTLT	IGDLRKPE.VRK.Q.I..TSA	STLYFP	KE
ApeMA	L.LH.YFAST.L.S	.HGPAHLEPLR	SDLRKPE.VEK.R.M..NGI	STFYVP	IG
BadMA	V.VH.YIGAV.L.P	.GSRVTMDTFT	VADLCKDD.VVK.Q.F..NPS	STLYFP	RS
CbeMA	V.IA.YLAAI.Q.P	.FHDSKMDKMT	VQDLRDQDKVQNIP.IT.AG.	TTLYVP	KK
CeaMA1	V.IH.YSGAV.L.P	.QSSSAMVVFV	ENLRNEQ.LAN.Q.I..RST	SILYIP	RD
CeaMA2	V.VH.YIGAV.L.P	.QATTVIQPYT	ISELRKPE.VAS.Q.I..RAC	STFYIP	RD
CeuMA1	V.VH.YIGAV.M.P	.QSTTIMDEFS	IADLRKEE.VVK.Q.F..TTW	STFYIP	RD
CeuMA2	V.IH.YVGAI.M.P	.QSATVMDEYT	ISDLRKED.VVK.K.F..TTT	STLYIP	RE
CfuMA	A.IA.YLAAI.Q.P	.FNDSKMDHMT	VEDLRDPEKVR SIP.IN.AG.	TTLYVP	KK
CloMA	L.VN.YQAAI.S.P	.LSEASIGRHIV	SDLRKAE.VQE.S.V..TGA	STFYIP	KT
CmaMA	V.VH.YIAPL.F.P	.TEEPVMEFRT	IGQLKKE.NSD.K.I..ATI	STFYLP	KA
CmiMA	V.VH.YIAPL.F.P	.TEDPIAEEYT	IAQLRLEP.IRD.K.I..HTI	STFYVP	KT
CmuMA	V.VH.YIGAV.L.P	.QSTAVIDEFT	VAGLRKEE.VVK.Q.I..TTV	STFYLP	RT
CpeMA	I.IH.YIAPM.L.P	.IDKPMQKMT	VSDLRKPE.YKA.K.I..VPS	STFYIP	NE
dbOphMA	V.VH.YMASI.L.P	.YEDPVTDKFT	VSQFRDPQ.IAK.R.I..CGI	STFYIP	KE
DbiMA2	V.VH.YIAAM.M.P	.HQDPVTDKFT	IGQLREPE.IAK.R.V..GGV	STFYIP	KA
FmeMA1	V.VH.YIGAV.L.P	.QSTSTIEAYT	TSDFRKGD.VVE.K.F..STT	STLYVP	SV
FmeMA2	V.VH.YIGAI.L.P	.QADPTVEAYI	VADLRKED.VVK.Q.F..NAI	STLYIP	RV
FmeMA4	V.VH.YIGAV.L.P	.QSTSKVEQYV	VADLRKDY.VVK.T.F..TTT	STLYVP	CV
GesMA	I.VS.YMAAV.S.P	.LEDPVINRHT	ISDLYKAD.VKK.E.I..TPN	CTLYIP	KD
GjuMA	V.IH.YVGRV.I.F	.QSVSKLLEIFT	IADLRKEE.VMN.H.F..DAI	STLYVP	RD
HpiMA	V.VH.YIGAV.L.P	.LSVKTMEITY	IADLRKED.VVA.Q.F..NPT	STLYIP	RD
LedMA	V.VH.YIAAM.L.P	.HEDPVTDQWT	IGQLREPE.FYK.R.V..GGV	STFYIP	KE
MfiMA	V.VH.YIGAV.L.P	.QSATVKDEFK	IADLRKDD.VVK.Q.I..STI	STFYIP	RQ
MroMA1	I.OH.YIGAI.L.P	.QSVTVKDAFA	IRDLRKEE.VLK.Q.F..TTT	STFYIP	RA
MeuMA	I.IA.YLAAV.R.P	.FNDPQIDKLMV	KDLRDLE.KLK.A.IPFNAA	TTLYIP	KT
OphMA	V.VH.YIAAM.M.P	.HQDPVTDKYT	VAAQLREPE.IAK.R.V..GCV	STFYIP	KA
PgiMA1	V.VH.YIGAV.L.P	.QSATVQDVVK	ISDLRKEE.IVA.Q.F..NSC	SILYVP	LT
PgiMA2	V.VH.YIGAV.L.P	.QSATVKEVYT	ISDLRKEE.VAT.Q.F..NAC	STLYVP	RK
PocMA	V.VH.YIGAV.L.P	.QSLTMDKLT	IADLRKDA.VVK.Q.F..NPT	STFYIP	RD
RviMA1	V.VH.YIGAV.L.P	.QSTTMDTFT	IADLRKED.VAK.Q.F..GTI	STLYVP	RD
RviMA5	V.VH.YIGAV.L.P	.QSTTMDTFA	VSDLRKED.VAK.Q.F..GTI	STLYIP	RD
RviMA8	V.VH.YIGAV.L.P	.QSTTMDIFT	ISDLRKEN.VAK.Q.F..GTI	STLYIP	RD
SbaMA	V.VH.YICAV.L.P	.QSTTVMDEFT	IADLRKEE.VVK.Q.I..TTT	STFYLP	RS
SveMA	V.VH.YVGAV.L.P	.QSATIMEYTY	IAELRKPE.VIK.R.IS.TTS	STFYIP	RD
TcuMA	V.VH.YVGAI.RMT	.QAQSAMVYVS	QETRNPA.VANF..I..NSG	STLYVP	RL
TelMA	L.VH.YIAPS.M.A	.TVEPTIDFLT	LGALKKSR.NAR.R.V..TGI	STFYIP	KH
ThyMA	V.VN.YVAAV.S.P	.LAEPTIQRHT	ISELPKDS.VKA.S.I..SGV	STFYIP	KE
TisMA	V.VH.YIASQ.L.T	.FAAPIRDRYA	IQDLVKPE.VAK.R.I..TGI	STFYLP	KD
SonM	V.VI.YEAAN.L.P	.IQAPRIERLPL	LANL..PQ..AHL..M..PI	STLLIP	PA
AnaM	A.IL.YEAAT.L.P	.IDEFRATTIA	IGEL..HN..HE..I..NQH	STLVIP	VV
LspM	V.VV.YEAPT.L.P	.IESAHIQRMPL	KDF..VN..AQF..T..TGO	ETLVLP	PA
RspM	I.TT.YRSAT.L.P	.IQQPRIFRMAL	IGDT..PQ..AD..I..GVP	TVVVP	PA
SpaM	V.IL.YEAAV.L.P	.IDNIRIETLS	LSEEL..VT..AEIH..M..H..	TTLVIP	PS
SpsM	V.IH.YEAAV.L.P	.IDTVRKQTLT	LAEL..AD..ADIY..M..H..	TTLVIP	PS
SdnM3	V.IH.YEAKT.L.A	.FNNPRIDTLP	LGEL..AK..ATIH..L..H..	TTLVIP	PS
SamM	C.LL.YEAAT.T.A	.LHSTRIERLPL	LSAL..PQ..ARLE..L..H..	TTLVLP	PS
SdnM2	C.LL.YEAAT.M.P	.LQSGRIERLPL	LNHL..P..LAQIA..L..H..	TTLVLP	PS
SupM	C.WL.YEAAT.T.A	.LASPRREAIK	LREL..VE..ATLA..L..H..	TTLVIP	PA
SarM	V.II.YEAAN.L.P	.FQSPVIERIP	LSLL..PE..ATL..T..TTI	STLLIP	PA
SmoM	V.VV.YEAAN.L.P	.IQAPRVDYLL	SEEL..PK..AHL..S..SPI	STLLIP	PA
SpeM	I.VL.YEAPN.L.P	.TQSPRIFERVAL	KQFT..P..FAQL..T..TTL	TLVLP	PS
SviM	I.IH.YEAPN.L.P	.IQQARIDKLP	LKNL..P..FARL..T..TAI	TTLVLP	PS
SwoM	I.TL.YEAPT.L.P	.IHQARVDKLL	LRDL..P..FARL..T..TSI	STLLIP	PS
SdnM1	V.VI.YEAAN.L.P	.LQQPRIERLAL	KKAL..PQ..ARL..T..TVI	STLLIP	PS
ScoM	V.VI.YEAAN.L.P	.VQQPRIDWLE	LQAL..PG..TRL..T..TAA	STLLIP	PS
CcrM	A.IV.YTAAE.Y.P	.GCRAIIEPST	IAGL..PS..APVQ..A..T..	ATLYVP	PL
ChaM <sub>UBA11691</sub>	A.IV.YEAAV.YHP	.VCEPSIQRIP	LEKL..PE..AKV..T..TES	STLYMP	EN
MboM	V.IV.YEAAV.Y.P	.VCKPVIQRIPL	SKL..PQ..SSV..T..TEV	STLYIP	PQ
MvaM	V.IV.YEAAV.Y.P	.VCEPVIQRIPL	SKL..AE..I..RV..T..TLV	STLYVP	PQ
OspM <sub>10802</sub>	V.IV.YEAAV.Y.P	.VCSPTIERMRL	CCEL..PN..ASV..T..TPV	STLYVP	PK
OspM <sub>6304</sub>	V.VI.YEAAV.Y.P	.VCQPVTINPT	LDST..P..IAKV..T..TTV	STLYVP	PL
RreM1	V.IV.YEAAI.Y.P	.VCQPIIEPMP	LSKL..PE..ALV..T..SEV	STLYVP	PQ
SceM1	V.IH.YEAAA.H.P	.LASPRIDRVLS	LSAL..SA..APV..T..DDV	STLYVP	PA
SolM	I.TH.YEASQ.Y.P	.VCAPHIRITLP	LRER..QE..PL..T..TGI	STLYLP	PL
RsuM	V.VH.YVGSQ.Y.P	.MCPAIVQRMPL	KDLR.T..ASV..T..TGL	STLYIP	PQ
SspM <sub>769</sub>	V.VV.YEASP.Y.P	.VTSPRIISVPL	LAKL..AD..TQL..T..TNK	STLVVP	PL
SspM <sub>NR3110611</sub>	V.VV.YEASP.Y.P	.AAEPRATTIVP	LAKL..PD..TSL..T..SAA	STLVVP	PL
SspM <sub>NR3110611</sub>	V.VV.YEASP.Y.P	.TAEPRATTIVP	LAET..PD..TPL..T..SAA	STLVVP	PL
GviM	V.VV.YQAAH.F.A	.MCDPIVECVAI	EAEML..K..ASI..T..TAV	STLYIP	PL
MaeM	V.IH.YKASV.I.P	.LCSHEVQRFAL	CDL..VT..AK..T..NGS	MTLIIP	PC
SelM	A.IN.YVANM.F.S	.GEBPIDRHR	IADYRDPE.VKR.T.V..SGI	STFFLP	AK
SliM	I.VYN.YVASM.F.S	.MAKPKKDKFK	LSDFRDPS.IAK.E.V..TGI	STFFIP	AV
AliM	V.VH.YIASR.Y.P	.TIPPTIEVYPL	LSALHDPO.IQT.R.V..TGV	STFYVP	PK
MspM	V.VH.YIASR.Y.P	.TIPPLIETYP	LSRLHEPD.I.QLR.V..TGI	STFYLP	PR
KfoM	V.VH.YIGSR.Y.P	.TIPPLIEIYHL	LNEMHDPE..TQLK.I..TGL	STFYIP	PR
EdiM	V.VN.YVGSR.Y.P	.GV.DPLIDRQT	LASLRDPS.VQNW..V..TGI	STFYLP	PK
PmoM	V.IN.YVGSR.Y.P	.GI.DPLIDRQT	LASLRDP..LAQ.SWV..TGI	STFYLP	PR
GsuM	V.VN.YIGSR.Y.P	.GADPVRDRHT	ISSLRNPA.VQST..I..TGI	STFYIP	PA
StpM	I.VN.YIGSR.Y.P	.GI.DPTKQHTV	SSLRVPE.AQST..V..TGI	STFYIP	PK
SspM <sub>NR3110611</sub>	V.VN.YIGAR.Y.P	.GTAPVTDTRHT	VAGLRDRA.VQRT..V..TSA	STFYLP	PA
AinM	I.VH.YIASR.Y.P	.TLPPIVIERY	LDELQNPV.VHAL..I..TGI	STFYVP	PA
RceM	S.VL.YLAAT.Y.P	.GLDGAQAVR	PLGAYRDPK.VLA.S.V..PPA	GTLYIP	AK
BlaM	I.VH.YVAA.S.P	.TLAPLIQNF	LRLELYPD.IQA.K.V..QSM	STFYLP	PK
IpnM <sub>LP8509</sub>	I.VYL.YEASQ.Y.P	.SIE..FELVSI	.DL..OK.LTEIN.IS.R.L	ATLYIP	PV
IpnM	I.SIL.YEASM.Y.P	.GV.EPIIHKFF	LYDI..ED..QNI..T..GTL	STLYIP	PL
BstM	A.LN.YVANM.F.S	.GPPQIDRHV	LDYRDPE.VKA.K.V..SGI	STFFIP	AK

**Supplementary Figure 3. Gene clusters of the two split borosins analyzed in this study.** The putative split borosin gene clusters of (a) *Shewanella oneidensis* MR-1 and (b) *Streptomyces* sp. NRRL S-118 are depicted with genes as arrows. Protein IDs are given along with the gene name or proposed enzyme function. Putative split borosin precursor genes are colored in cyan, while  $\alpha$ -N-methyltransferases are colored green.



**Supplementary Figure 4. Mass spectrometric analysis of split borosin coexpressions.** (a-c) HPLC-MS/MS spectra from AspN-digested SonA peptides after incubation at 30°C with wt SonM, saturating [SAM], and the other enzymes and kit reagents used in the kinetics assay (see materials and methods). The amino acid sequence above each spectra depicts the *N*-methylated residues that could be confirmed by MS/MS fragmentation (solid orange circles) or are inferred *N*-methylated since the position is not completely defined by MS/MS (unfilled orange circles). Lowercase 'c' denotes cysteine derivatized by iodoacetimide. Observed MS/MS fragmentation masses are listed above (b-ions) and below (y-ions) the amino acid sequence. The gray lines within the sequence mark the sites of fragmentation. Masses of methylation-containing ions are denoted in brackets, where 'Me' stands for methylation. The ppm difference from the observed masses to the theoretical expected masses are labeled in parentheses. A 10.0-ppm mass cutoff for annotated HPLC-MS/MS peaks was used. The protein, time of *in vitro* reaction, parent ion information and HPLC retention time (RT) are listed in the upper right corner of the LC-MS/MS spectra. Note, panel c depicts the same raw data as seen in Fig. 1d. (d-i) HPLC-MS/MS spectra from GluC-digested SspA<sub>NRRLS118</sub> peptides after a 15 min incubation at room temperature with SspM<sub>NRRLS118</sub>, saturating [SAM], and *S*-adenosylhomocysteine nucleosidase in 50 mM HEPES buffer. (j) Relative abundance of SspA<sub>NRRLS118</sub> peptides from panels (d-h). The methylation state is indicated over each graph (0-4) in an orange box with the relative abundance (%) of the methylated species directly below. Relative abundance (intensity %) was determined by integrating under each peak from the MS1 extracted ion chromatogram. Each peak is plotted over its retention time (x-axis).

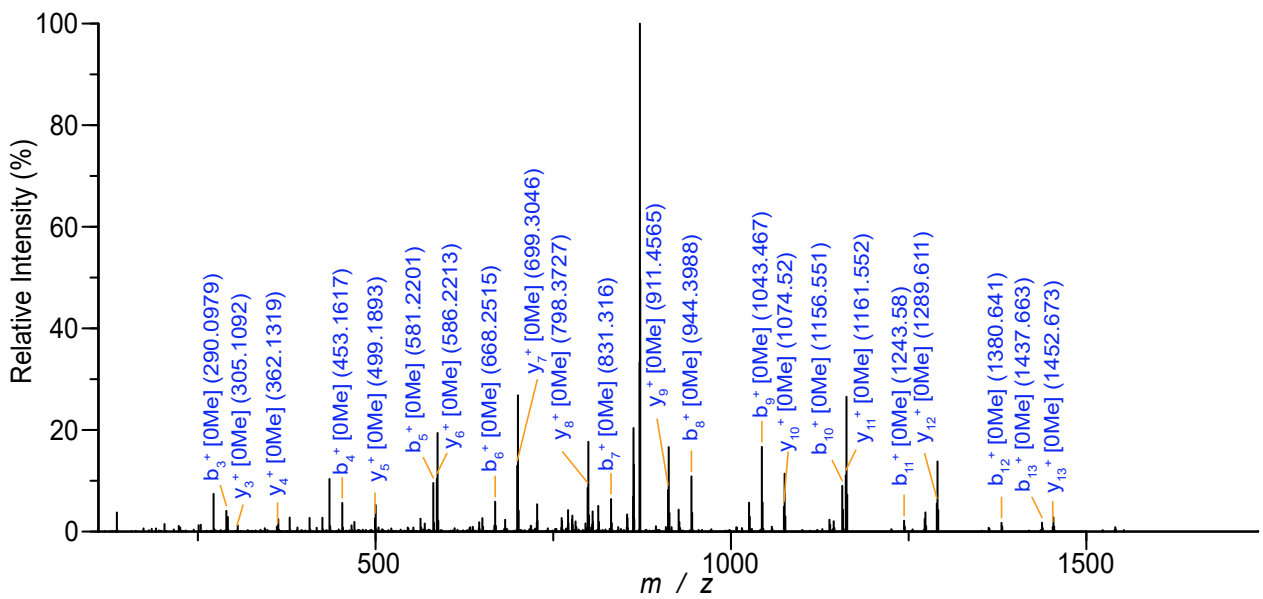


**a**

- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SonA (with SonM, at 24.5 min rxn)  
Parent ion: [M+OMe+2H]<sup>2+</sup> (871.39)  
RT: 11.42 min



**b**

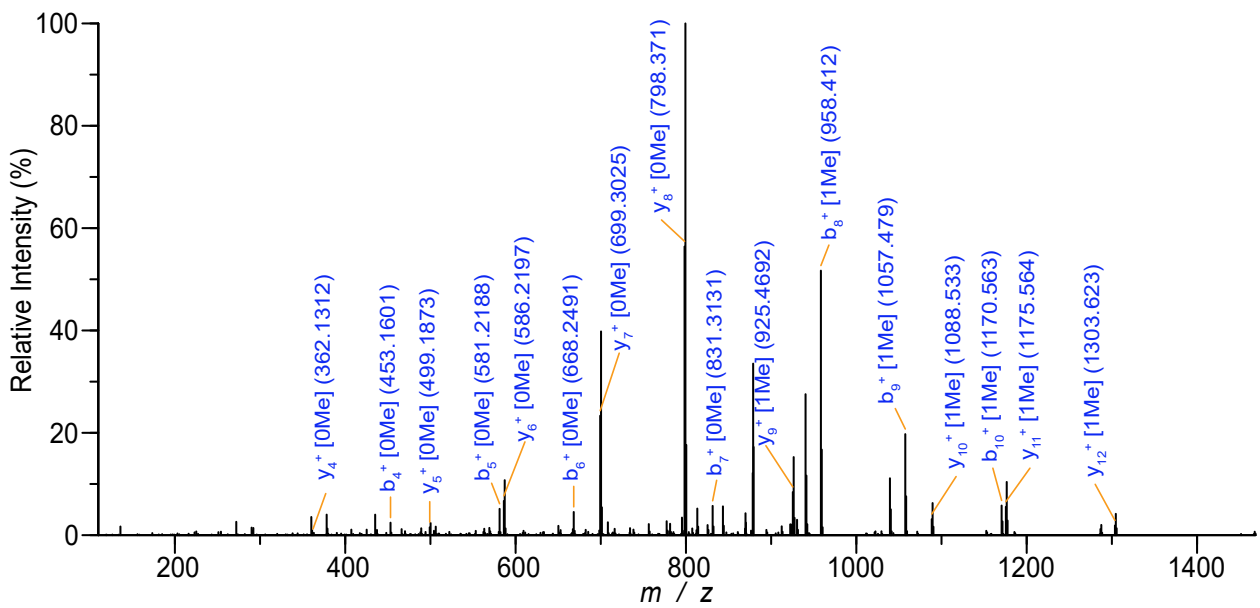
- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SonA (with SonM, at 24.5 min rxn)

Parent ion:  $[M+1Me+2H]^+$  (878.40)

RT: 11.43 min

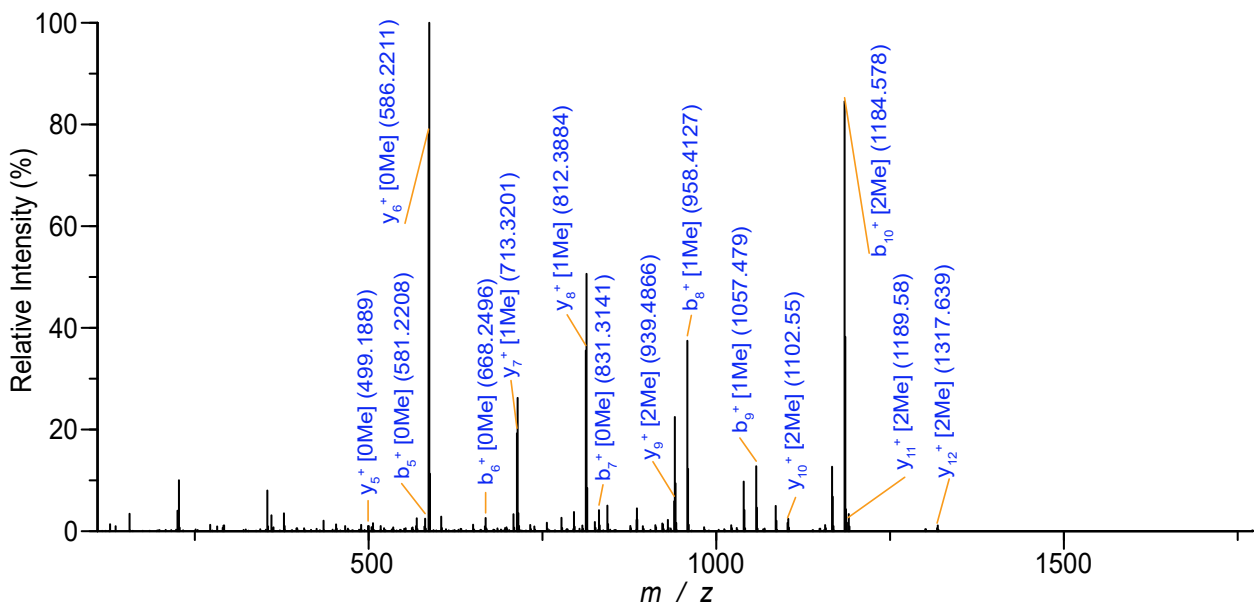


C

- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SonA (with SonM, at 24.5 min rxn)  
Parent ion:  $[M+2Me+2H]^{2+}$  (885.40)  
RT: 12.15 min

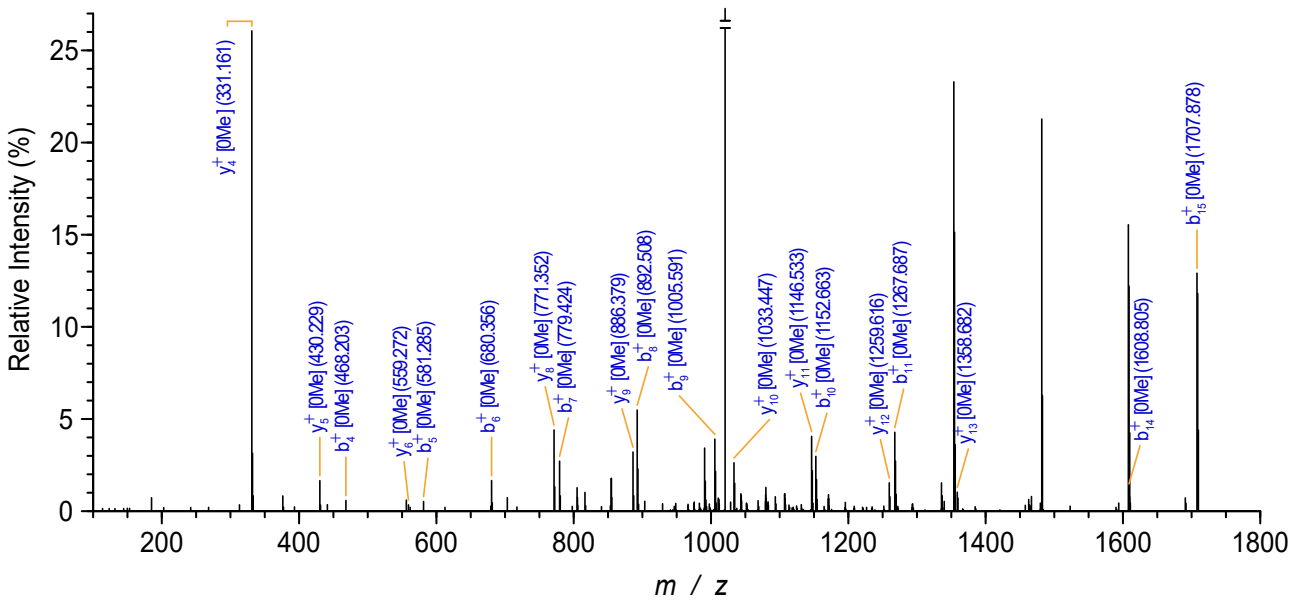


d

- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SspA<sub>NRRLS116</sub> (with SspM<sub>NRRLS116</sub> at 15 min rxn)  
 Parent ion: [M+OMe+2H]<sup>2+</sup> (1019.52)  
 HCD: 18, RT: 18.72 min



e

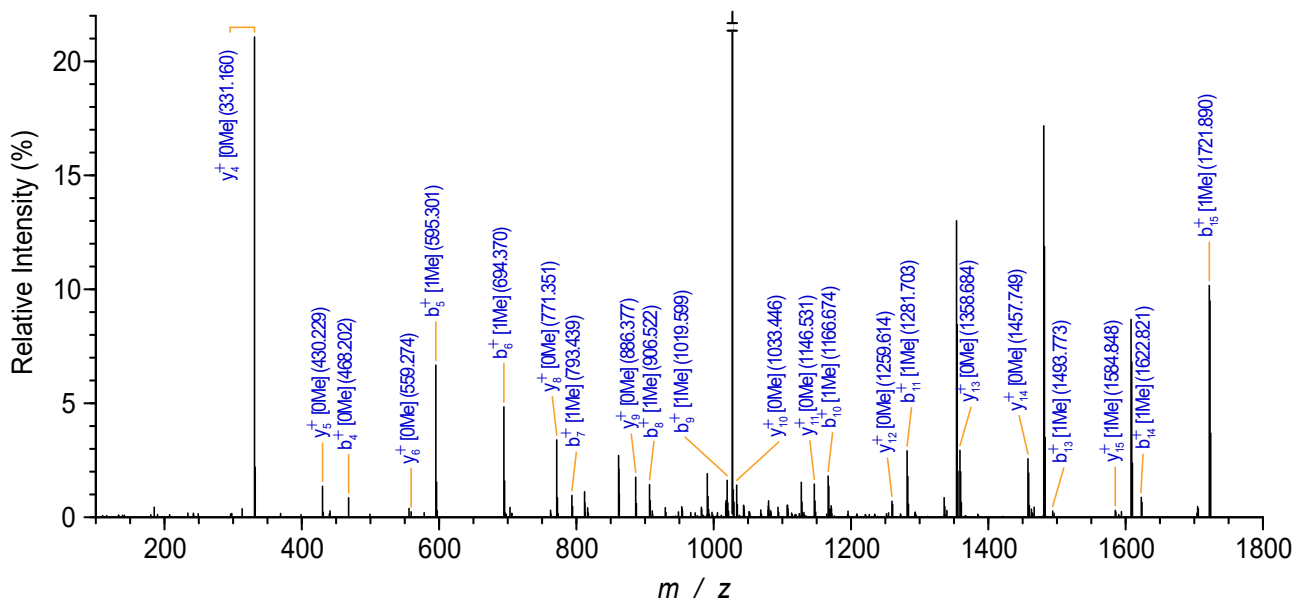
- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SspA<sub>NRRLS116</sub> (with SspM<sub>NRRLS116</sub> at 15 min rxn)

Parent ion: [M+1Me+2H]<sup>2+</sup> (1026.53)

HCD: 18, RT: 19.12 min

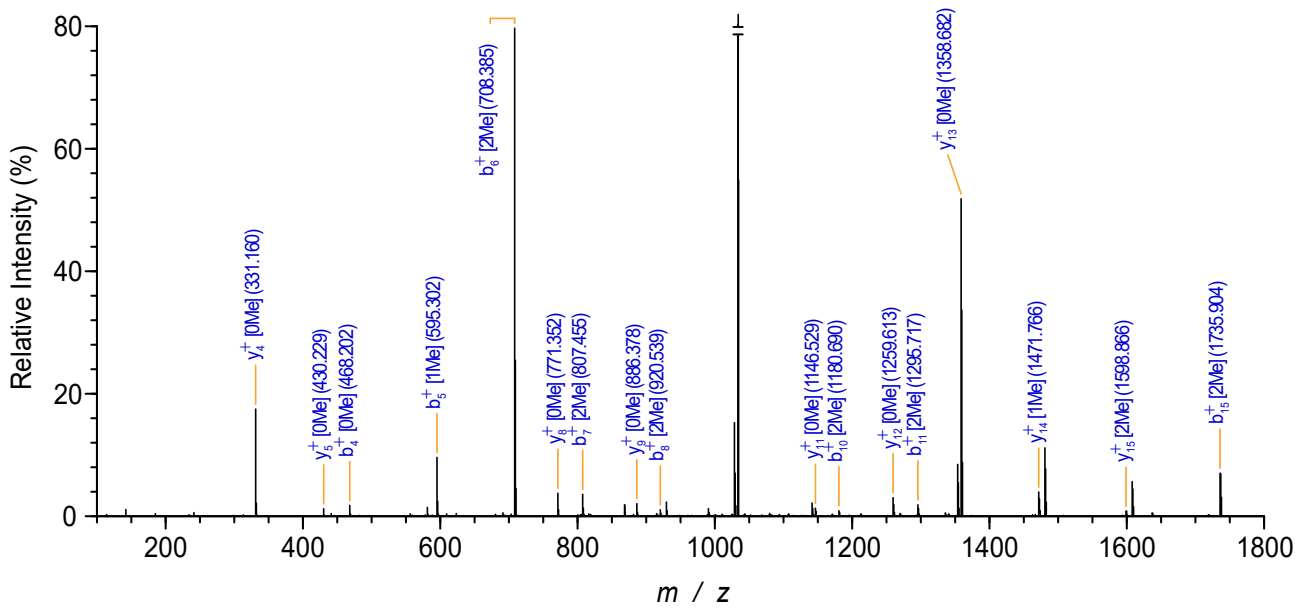


f

- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SspA<sub>NRRLS116</sub> (with SspM<sub>NRRLS116</sub> at 15 min rxn)  
 Parent ion: [M+OMe+2H]<sup>2+</sup> (1033.54)  
 HCD: 18, RT: 19.32 min



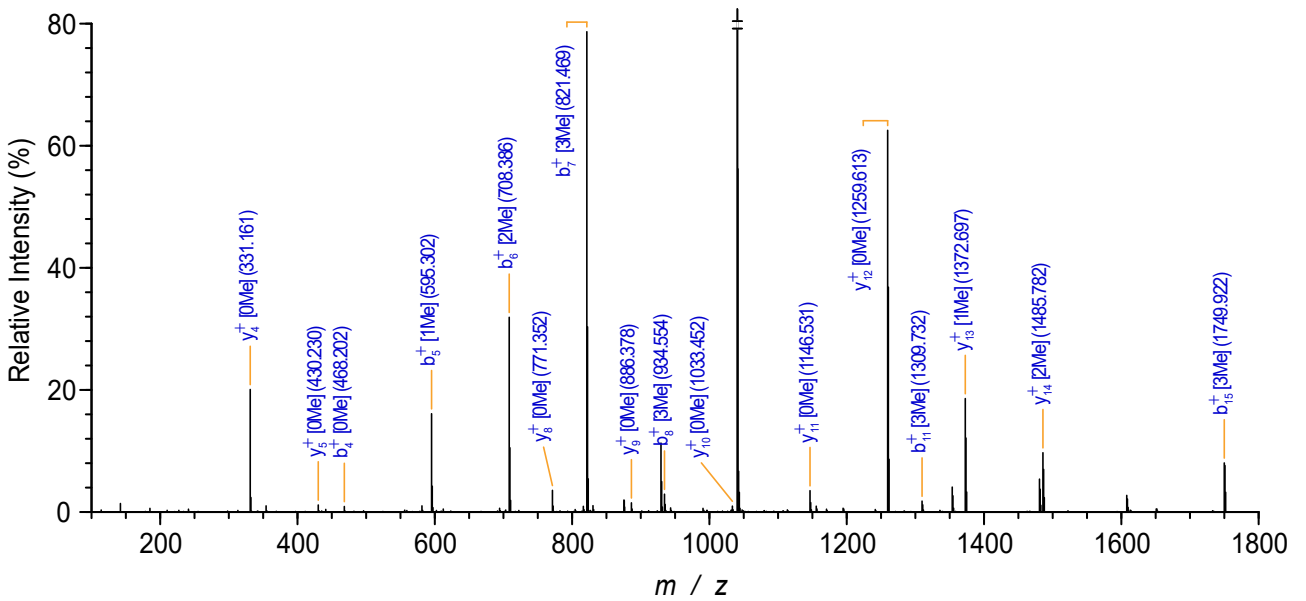


h

- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SspA<sub>NRRLS116</sub> (with SspM<sub>NRRLS116</sub> at 15 min rxn)  
 Parent ion: [M+OMe+3H]<sup>3+</sup> (1040.54)  
 HCD: 18, RT: 19.92 min





i

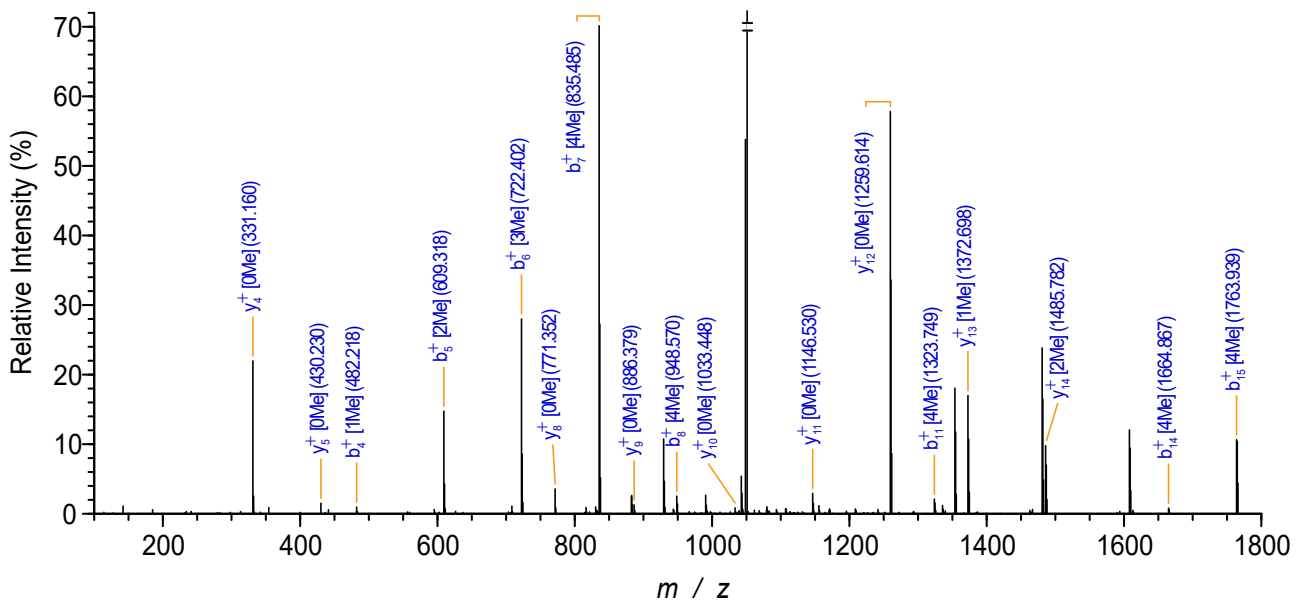
- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SspA<sub>NRRLS116</sub> (with SspM<sub>NRRLS116</sub> at 15 min rxn)

Parent ion: [M+4Me+2H]<sup>2+</sup> (1047.57)

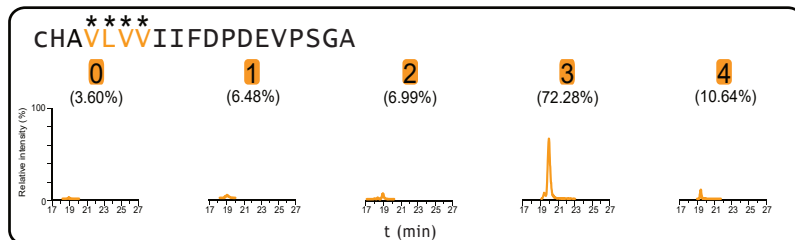
HCD: 18, RT: 19.92 min



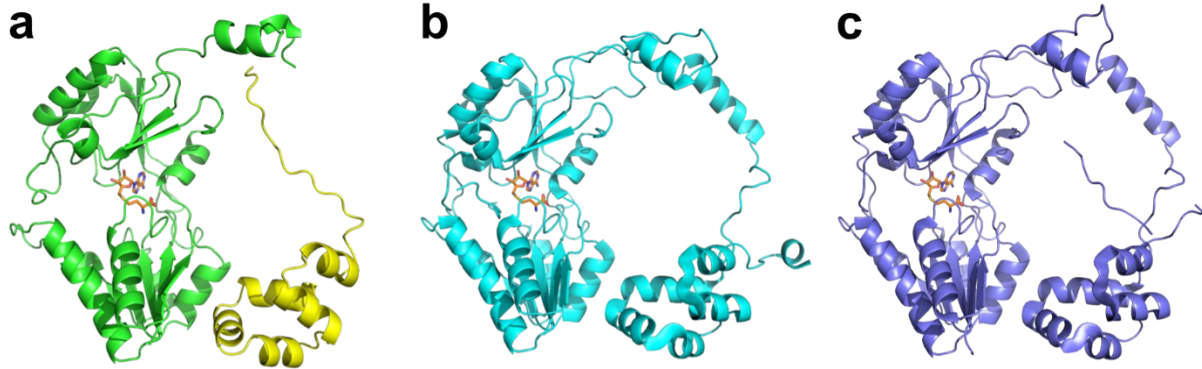
**j**

EIC from HPLC-MS: SspA<sub>NRRLS118</sub> (GluC digest)

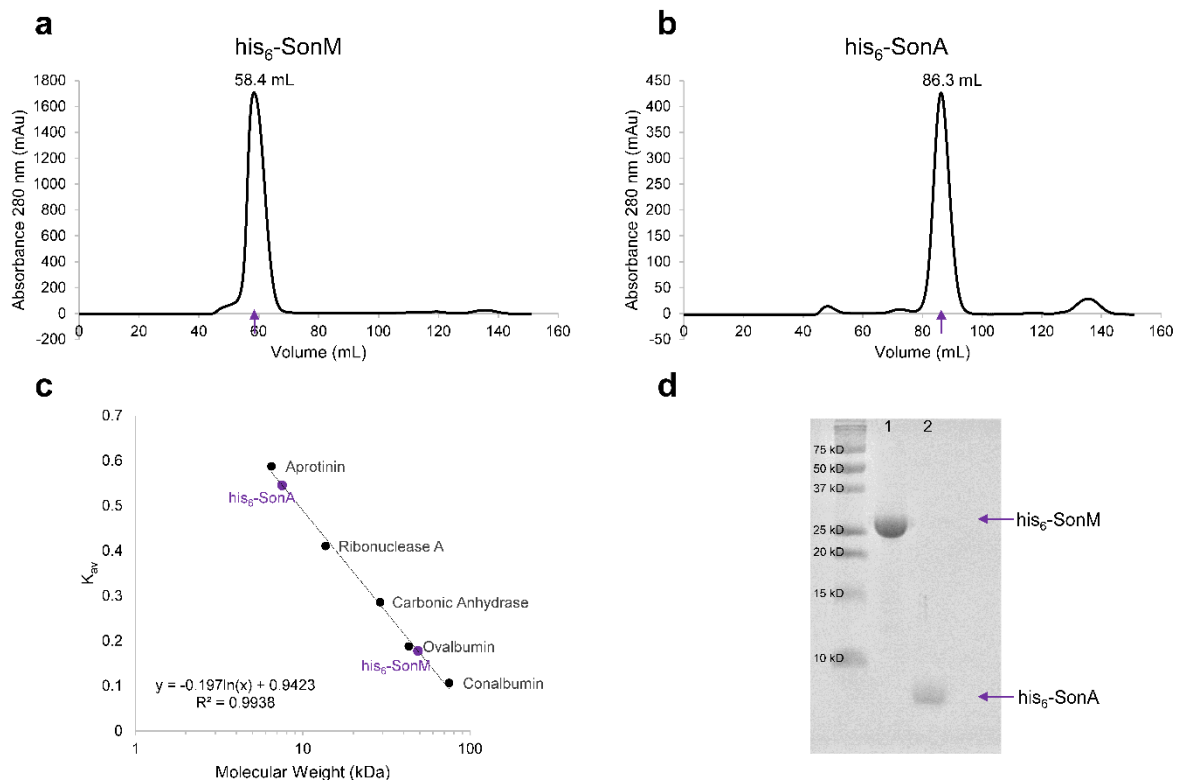
15 min



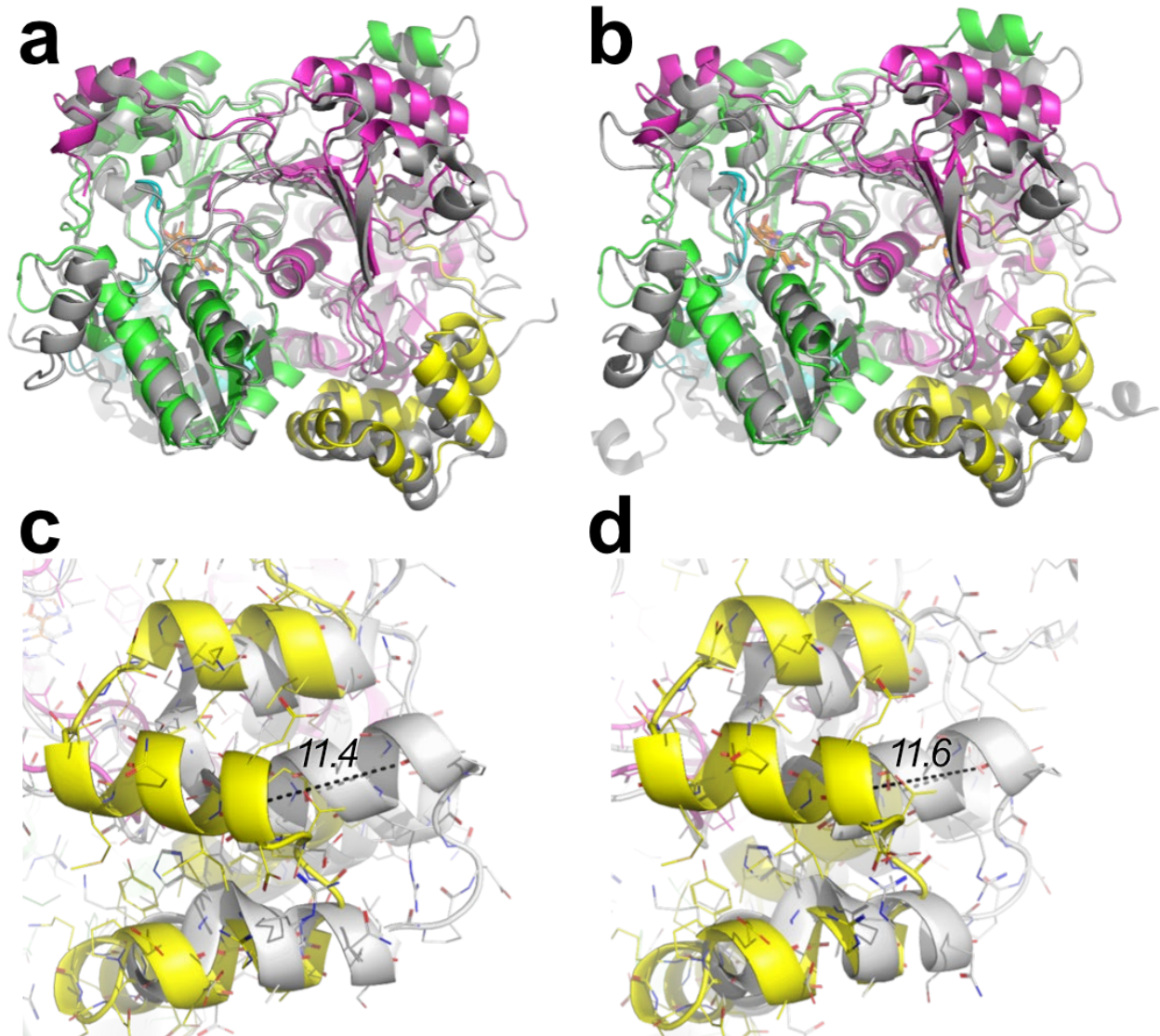
**Supplementary Figure 5. Global structural comparisons of borosin  $\alpha$ -N-methylating enzymes.** The (a) SonM—SonA-2Me—SAH (PDB: 7LTE) heteromeric configuration is different from the fused homologous fungal systems (b) OphMA<sup>2</sup> (PDB: 5N0X) and (c) dbOphMA<sup>3</sup> (PDB: 6MJG). A single heterodimer (SonM—SonA-2Me—SAH) or a single monomer of a homodimer (OphMA and dbOphMA) is represented as a cartoon, with SAH depicted as orange sticks.



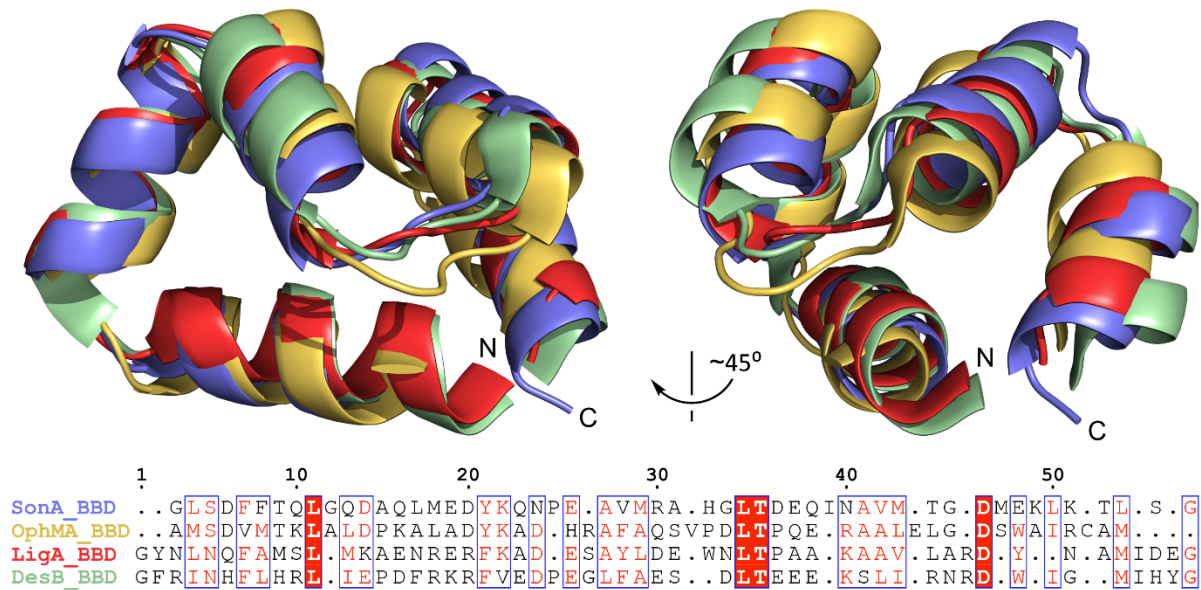
**Supplementary Figure 6. Oligomeric state of his<sub>6</sub>-SonM and his<sub>6</sub>-SonA proteins by size exclusion chromatography.** Size exclusion chromatogram of (a) his<sub>6</sub>-SonM and (b) his<sub>6</sub>-SonA after 24 hour expression in *E. coli* BL21(DE3) cells and nickel affinity purification. The volume at which the protein eluted is indicated with a purple arrow on the x-axis and labelled above the peak. (c) The calibration curve used to determine the oligomeric state of his<sub>6</sub>-SonM and his<sub>6</sub>-SonA in solution. The x-axis is the molecular weight in log scale, and the y-axis is the partition coefficient ( $K_{av} = (V_E - V_0)/(V_C - V_0)$ ). The molecular weight markers used were: aprotinin (6.5 kDa), ribonuclease A (13.7 kDa), carbonic anhydrase (29 kDa), ovalbumin (43 kDa), and conalbumin (75 kDa). All proteins and standards were run on a HiLoad 16/600 Superdex 75 pg column (Cytiva). The observed mass of the his<sub>6</sub>-SonM dimer was 48.5 kDa compared to the theoretical mass of 60.3 kDa. The extensive dimer interface of his<sub>6</sub>-SonM likely accounts for a smaller hydrodynamic radius and the slightly delayed elution time observed. The observed mass of the his<sub>6</sub>-SonA monomer is 7.5 kDa compared to the theoretical mass of 8.7 kDa. These results have been repeated at least twice from separate expressions with each protein. (d) SDS-PAGE of purified his<sub>6</sub>-SonM (lane 1) and his<sub>6</sub>-SonA (lane 2) proteins. A standard 15% (w/v) SDS-gel was loaded with 10 μg of each respective protein from the pooled elution fractions from size exclusion chromatography. Source data are provided as a Source Data file.



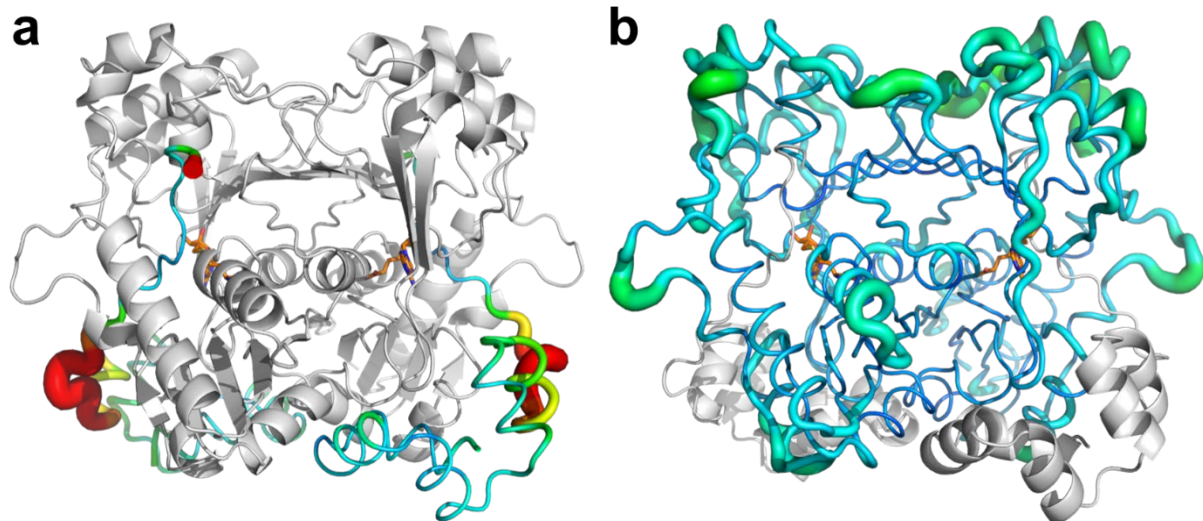
**Supplementary Figure 7. Structural overlay of SonM—SonA-2Me—SAH.** The SonM—SonA-2Me—SAH complex (SonA is shown in yellow and cyan cartoon, SonM in purple and green, PDB: 7LTE) shows structural similarity to homologous systems (a) OphMA<sup>2</sup> (PDB: 5N0X) and (b) dbOphMA<sup>3</sup> (PDB: 6MJG). Significant translation movement is visible for SonA (yellow) compared to counterparts in (c) OphMA and (d) dbOphMA. Key distances are depicted as black dashed lines and their lengths noted in italics in Ångstroms.



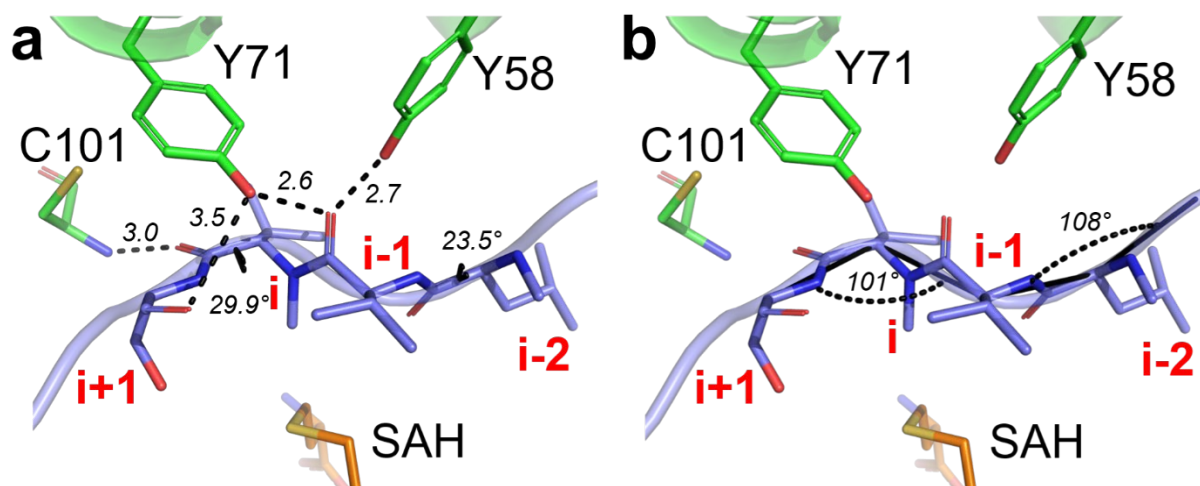
**Supplementary Figure 8. Structural overlay of the BBD with other protein domains.** The BBD from the SonM—SonA-2Me—SAH complex (PDB: 7LTE) shares structural homology to the BBD of the borosin methyltransferase OphMA<sup>2</sup> (PDB: 5N0X; root means squared distance (RMSD) of 2.3 Å for 306 atoms), to LigA of the protocatechuate 4,5-dioxygenase complex LigAB<sup>4</sup> (PDB: 1BOU; RMSD of 1.1 Å for 251 atoms), and to a tethered domain in the gallate dioxygenase DesB<sup>5</sup> (PDB: 3WRB; RMSD of 1.9 Å for 286 atoms). RMSDs (all atoms) were calculated using the 'super' function in PyMOL. The close structural homology is in contrast to the relatively low pairwise sequence identity of 24.3% and 45.3% sequence similarity amongst these domains.



**Supplementary Figure 9. Thermal motion B-factors for SonA and SonM.** Putty cartoon representations of the thermal motion B-factor variation for (a) SonA and (b) SonM in the structure SonM—SonA-2Me—SAH (PDB: 7LTE). B-factors are represented in a rainbow-color spectrum of dark blue (lowest mobility) to dark red (highest mobility).

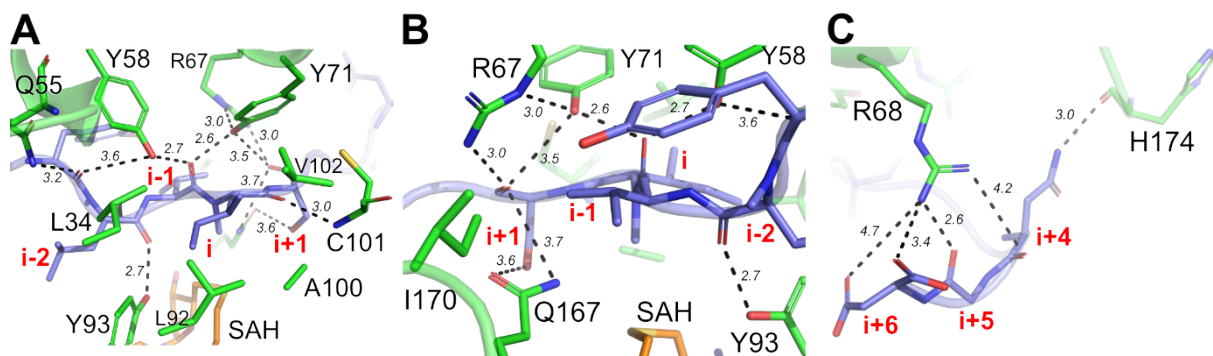


**Supplementary Figure 10. SonM—SonA-2Me—SAH active site coordination.** (a) SonM-Y71 and SonM-Y58 (green cartoon and sticks) coordinates with the 'i+1', 'i', and 'i-1' residues of the SonA core peptide (slate cartoon and sticks). Key distances are depicted as black dashed lines and their lengths noted in italics in Ångstroms. (b) Interestingly, the  $\Psi$  angle between residues 'i' and 'i+1' is 29.9° and consequently the main chain twists 101° (angle between 'i+1', 'i' and 'i-1'  $\alpha$ -carbon atoms). A similar observation was made in OphMA.<sup>2</sup>  $\Psi$  torsion angles values of  $\pm 30^\circ$  were reported to create electronic distortion in amide bonds that could increase the reactivity of the backbone NH group,<sup>6</sup> and could therefore help catalysis. This specific conformation may result from the active site pre-organization and is stabilized by a network of interactions between the side chains of SonM-Y71 and SonM-Y58 and SonA's main chain, and particularly with the carbonyl group of 'i-1', but also with a hydrogen bond between the carbonyl group of residue 'i' and the backbone NH of SonM-C101. Of note, the backbone NH group of SonM-C101 is located at the N-termini of an  $\alpha$ -helix, and helix macrodipoles have been found to be involved in stabilizing interactions.<sup>7</sup> Because we obtained the structure with the fully methylated peptide (i.e. post-catalysis), the twisted conformation of the main chain could also result from the presence of the methyl group. In fact, the  $\Psi$  angle between residues 'i-1' and 'i-2' is 23.5°, and also lead to a substantial twist in the main chain (108° angle between 'i', 'i-1' and 'i-2'  $\alpha$ -carbon atoms).

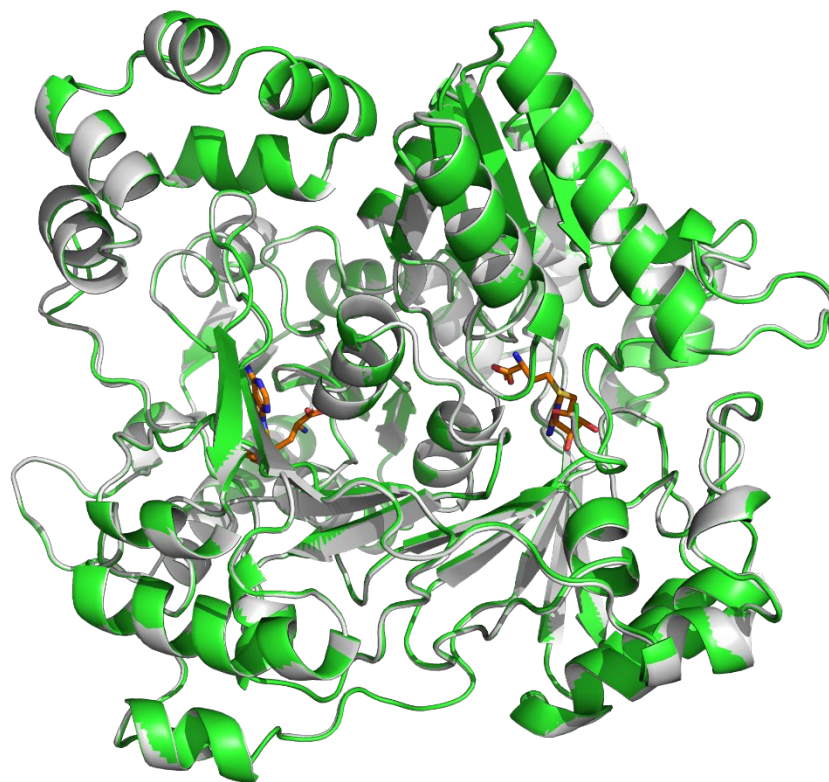




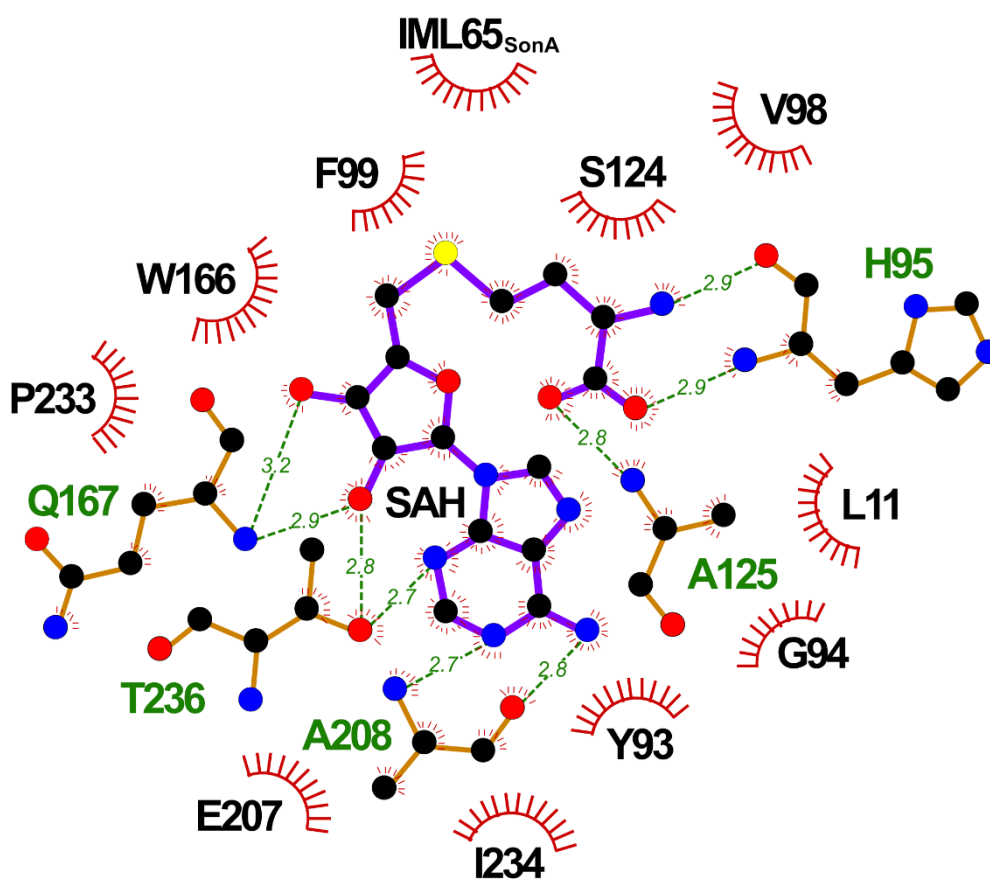
**Supplementary Figure 11. SonM interactions with the SonA core peptide in SonM—SonA-2Me—SAH.** (a) The side chain of residue 'i' of SonA (slate sticks and cartoon) sits in a well-defined pocket formed with SonM-A100, SonM-V102, SonM-L92, and SonM-L34 (green sticks and cartoon). Key distances are depicted as black dashed lines and their lengths noted in italics in Ångstroms. (b) Other core peptide side chains are accommodated by less defined binding pockets: the side chain of residue 'i-1' sits in a large cavity formed by polar and apolar residues (SonM-Q167, SonM-I170, SonM-R67, and i-3's side chain SonA-Y62) and the side chain of 'i-2' (SonA-MLE63) sits in a large hydrophobic pocket, while 'i+1' (SonA-S66) sits in a hydrophilic pocket and is hydrogen bonded to SonM-Q167 and SonM-F99. (c) Residues 'i+3' to 'i+7' are exposed to the solvent, and so are residues from 'i-3' to the N-termini of the BBD. Residue 'i+4' (SonA-N69) interacts with the carbonyl group of SonM-H174 (3.0 Å). SonM-R68, in addition to its interaction with the carbonyl groups of 'i+5' and the C-termini ('i+6'), it also interacts with the side chain of 'i+6' (SonA-D71) and therefore may contribute to the stabilization of the bound core peptide.



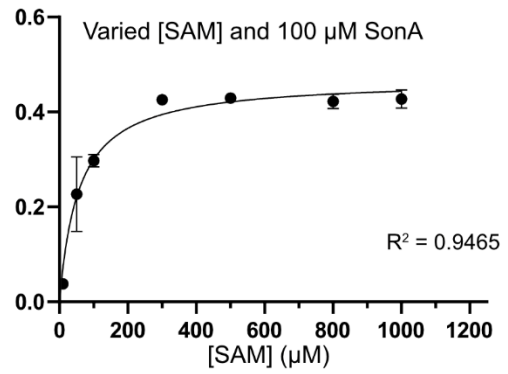
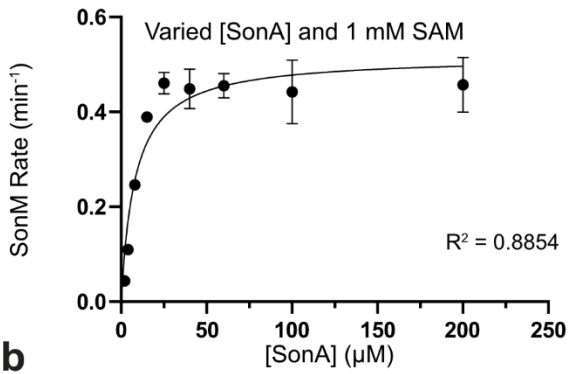
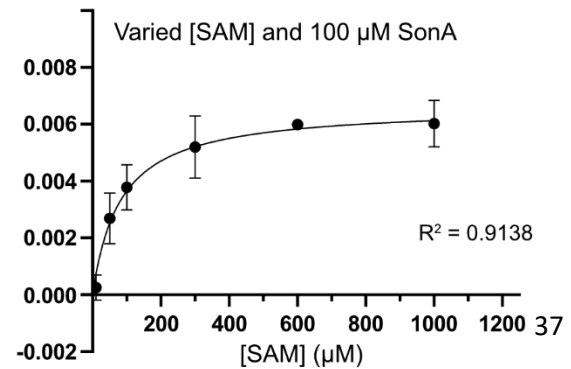
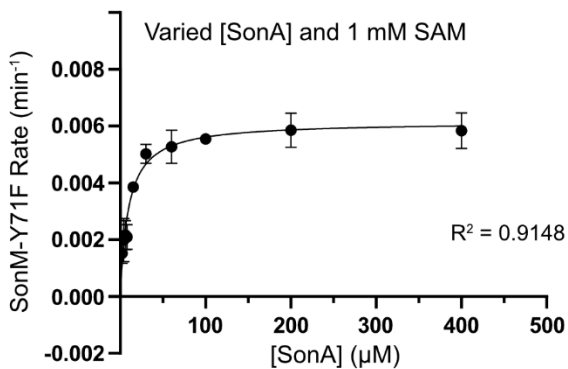
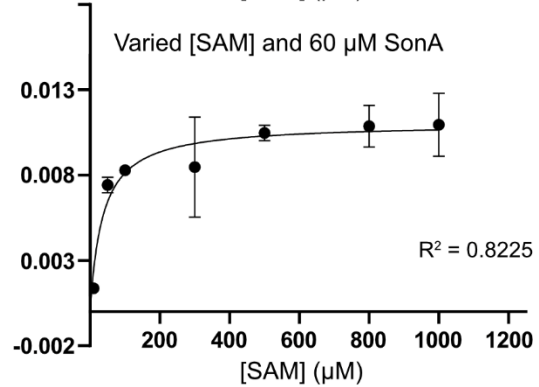
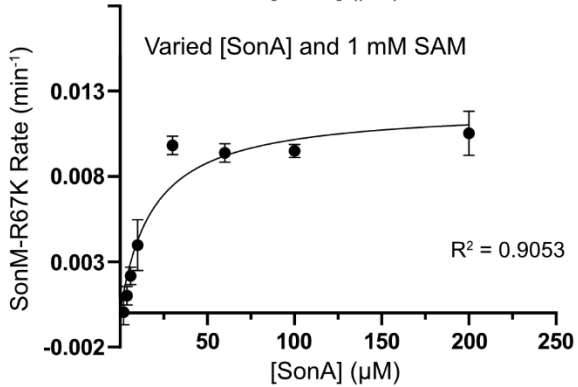
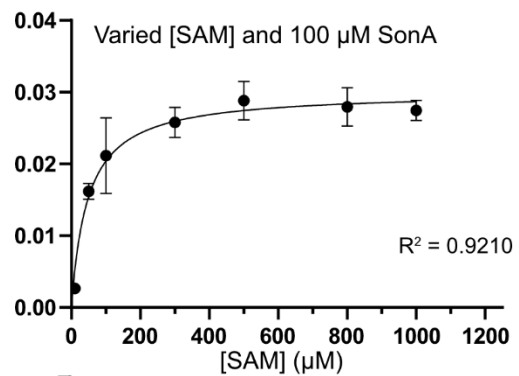
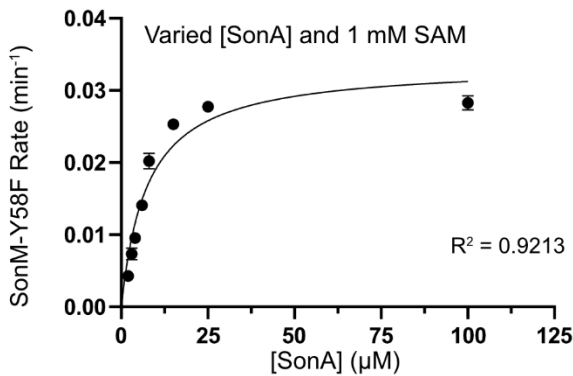
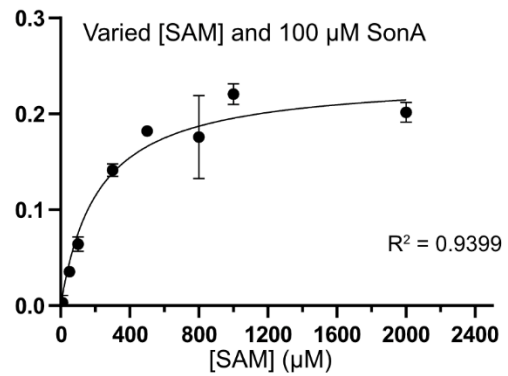
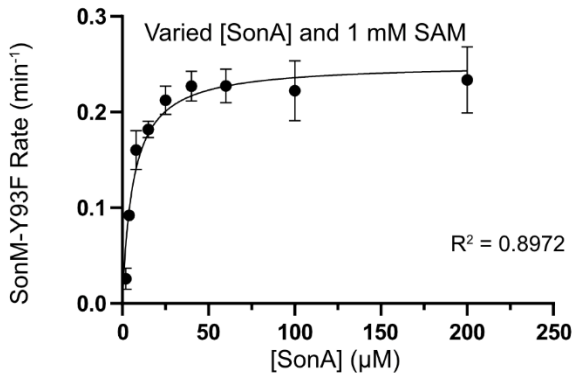
**Supplementary Figure 12. Superposition of SonM—SonA-2Me complexes.** The structure of *apo* SonM—SonA-2Me (grey cartoon, PDB: 7LTC) is highly similar to SonM—SonA-2Me—SAH (green cartoon, PDB: 7LTE), with an RMSD of 0.19 Å for 4638 atoms. The RMSD was calculated using the 'super' function in PyMOL. SAH is shown as orange sticks.



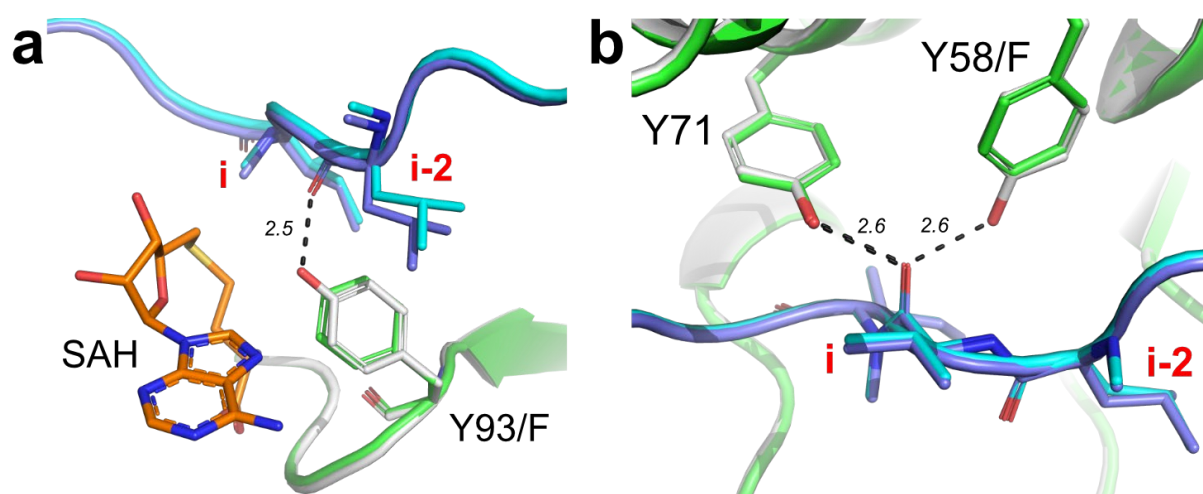
**Supplementary Figure 13. SAH makes extensive contacts in the SonM—SonA-2Me—SAH complex.** A LIGPLOT<sup>8</sup> of the extensive network of interactions made by SAH in SonM—SonA-2Me—SAH (PDB: 7LTE). SAH and key residues are displayed as ball and sticks, while other contacts are displayed as ‘eyelashes’. Key distances are depicted as green dashed lines and their lengths noted in italics in Ångstroms.



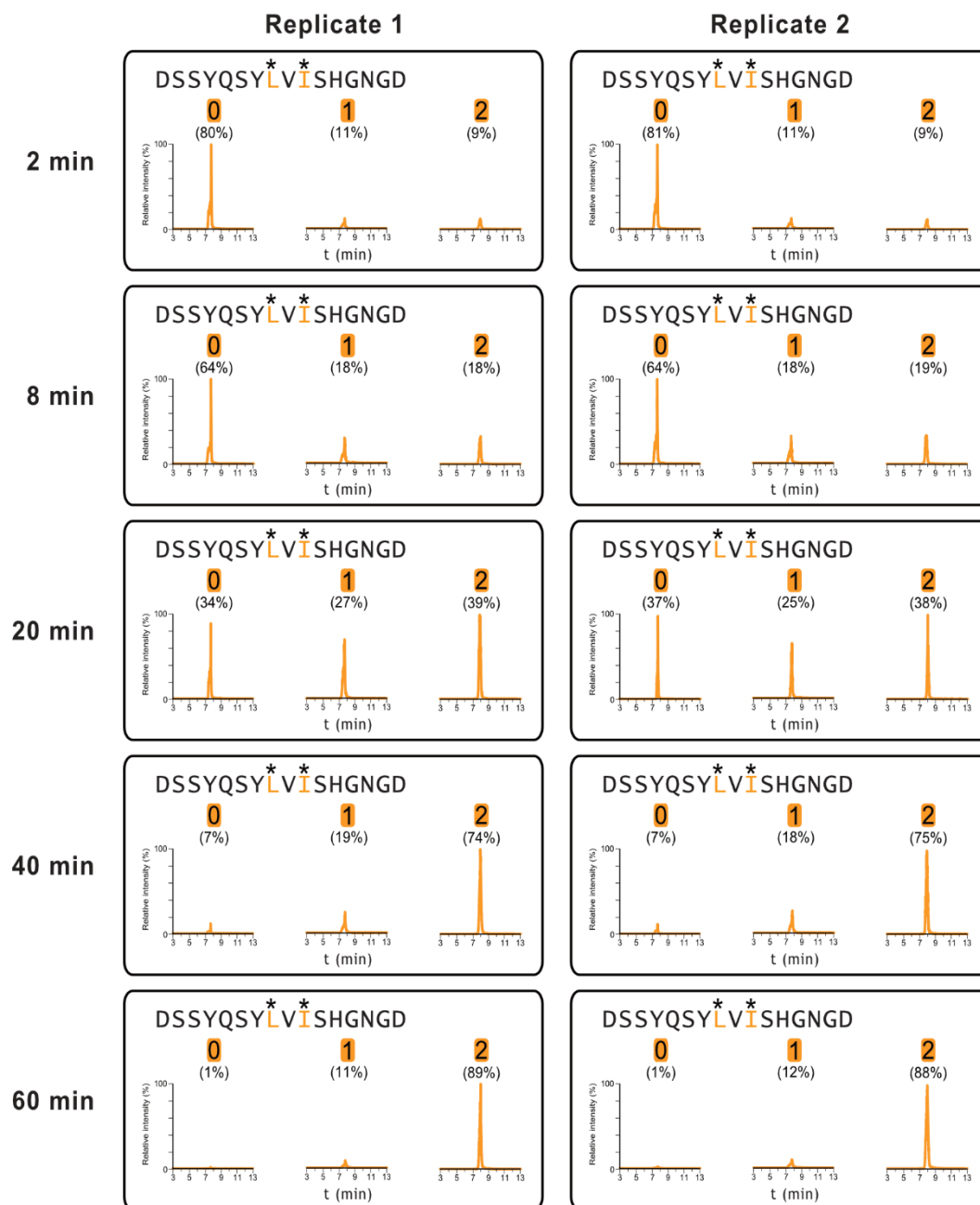
**Supplementary Figure 14. SonM – fitted Michaelis-Menten kinetic curves.** Michaelis-Menten substrate velocity curves of (a) wt SonM and (b) SonM mutants with varied [SonA] and saturating [SAM] (left) or varied [SAM] and saturating [SonA] (right). Each substrate concentration was assayed in triplicate (n=3); the enzyme assayed for each set of experiments is listed in the y-axes. The plotted point is the mean velocity measurement at that substrate concentration with the error bars representing the standard deviation between replicates. The overlaid curves were fit using nonlinear regression models in GraphPad Prism 8 and used to determine kinetic constants. The R<sup>2</sup> value for the fitted curve is included at the bottom right of each graph. No kinetics data is shown for SonM-R67A or SonM-Y58F/Y71F as these mutants had no measurable activity under the conditions of the kinetics assay used in this work. Source data are provided as a Source Data file.

**a****b**

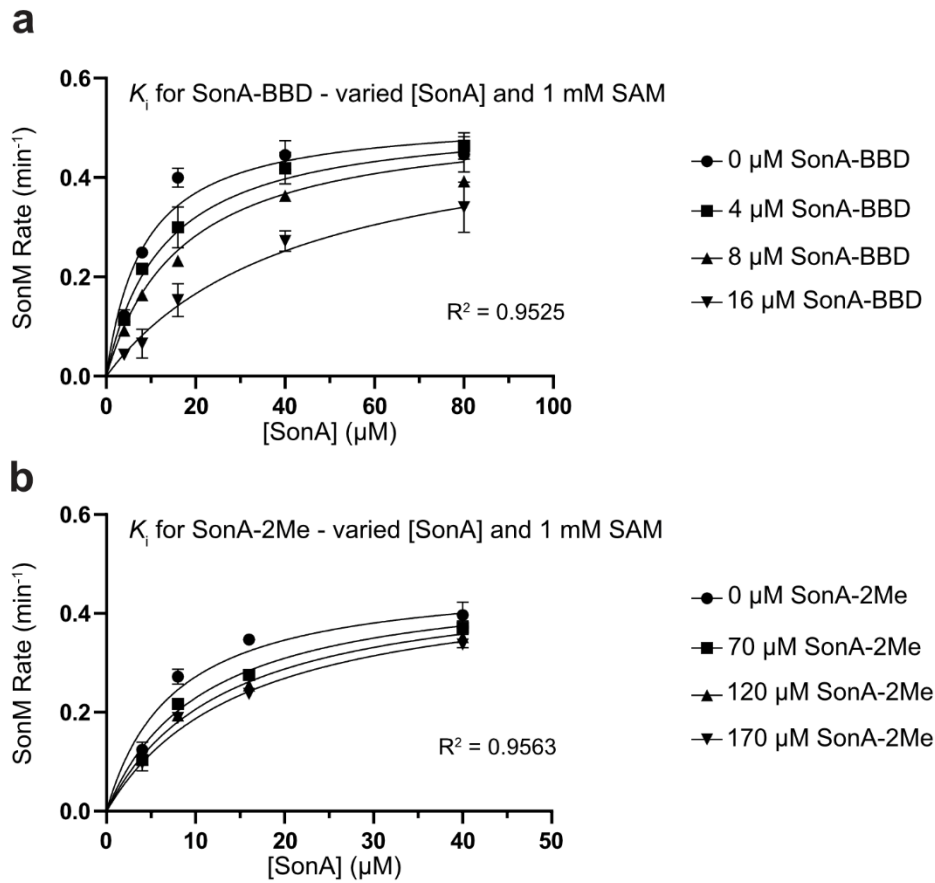
**Supplementary Figure 15. Active site tyrosine mutant structures.** (a) Superposition of the SonM-Y93F—SonA-2Me mutant structure complex (green, dark blue, PDB: 7LTH) with wt SonM—SonA-2Me—SAH (grey, cyan, PDB: 7LTE). The RMSD (all atoms) between the two structures is 0.16 Å for 4539 atoms, when using the ‘super’ function in PyMOL. (b) Superposition of the SonM-Y58F—SonA-2Me mutant structure complex (green, dark blue, PDB: 7LTF) with wt SonM—SonA-2Me—SAH (grey, cyan, PDB: 7LTE). The RMSD (all atoms) between the two structures is 0.08 Å for 4824 atoms, when using the ‘super’ function in PyMOL. Key residues are shown as sticks and key distances are depicted as black dashed lines, with their lengths noted in italics in Ångstroms. Both residues SonM-Y71 and SonM-Y58 are involved in an extensive network of interaction with the core peptide, including hydrogen bonding to the carbonyl group of ‘i-1’ SonM-IML65. This configuration is similar to OphMA, where the corresponding tyrosines OphMA-Y66 and OphMA-Y76 were proposed to stabilize  $sp^3$  hybridization and the developing negative charge on the carbonyl’s oxygen atom (oxyanion hole).<sup>2</sup> We note that in the SonM—SonA-2Me—SAH structure, the interaction angles between the carbonyl group and the hydroxyl groups of SonM-Y58 and SonM-Y71 are 122° and 108°, respectively. The interaction angle with SonM-Y71 is therefore close to the canonical angle value of 109.5° for  $sp^3$  hybridization to the carbonyl group.



**Supplementary Figure 16. SonM in vitro reactions analyzed by LC-MS/MS and compared to kinetic model simulations.** Relative abundances for each species of SonA (SonA-0Me, SonA-1Me, SonA-2Me) are depicted as extracted ion chromatograms from LC-MS data after incubation at 30°C with wt SonM, saturating [SAM], and the other enzymes and kit reagents used in the kinetics assay (see Methods section). All reactions were set up in duplicate under the same conditions and were quenched at time points as indicated on the left of each set of plots. The amino acid sequence of the AspN digested fragment is shown at the top with the methylated residues in orange with an asterisk (\*). The methylation state is indicated over each graph (0-2) in an orange box with the relative abundance (%) of the methylated species directly below. Relative abundance (intensity %) was determined by integrating under each peak from the extracted ion chromatogram. Each peak is plotted over its retention time (x-axis). This data is displayed in each panel of Fig. 4.

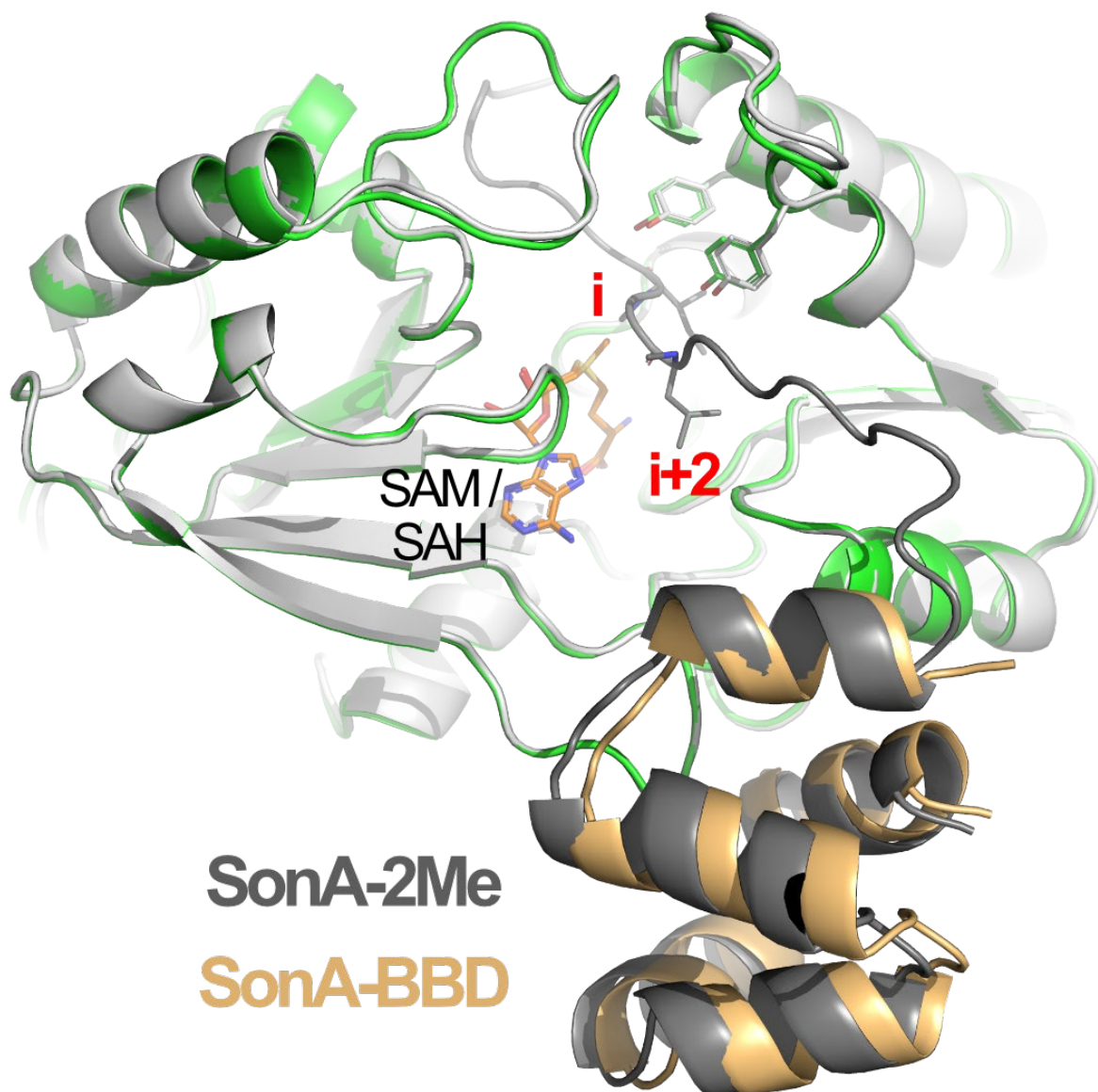


**Supplementary Figure 17. SonM – fitted Michaelis-Menten competitive inhibition kinetic curves.** Competitive inhibition curves for wt SonM with varied [SonA] saturating [SAM] and increasing (a) [SonA-BBD] or (b) [SonA-2Me], respectively. Each substrate concentration was assayed in triplicate (n=3). The plotted point is the mean velocity measurement at that substrate concentration with the error bars representing the standard deviation between replicates. The overlaid curves were fit using nonlinear regression models in GraphPad Prism 8 and used to determine kinetic constants. The  $R^2$  value for the fitted curve is included at the bottom right of each graph. Source data are provided as a Source Data file.

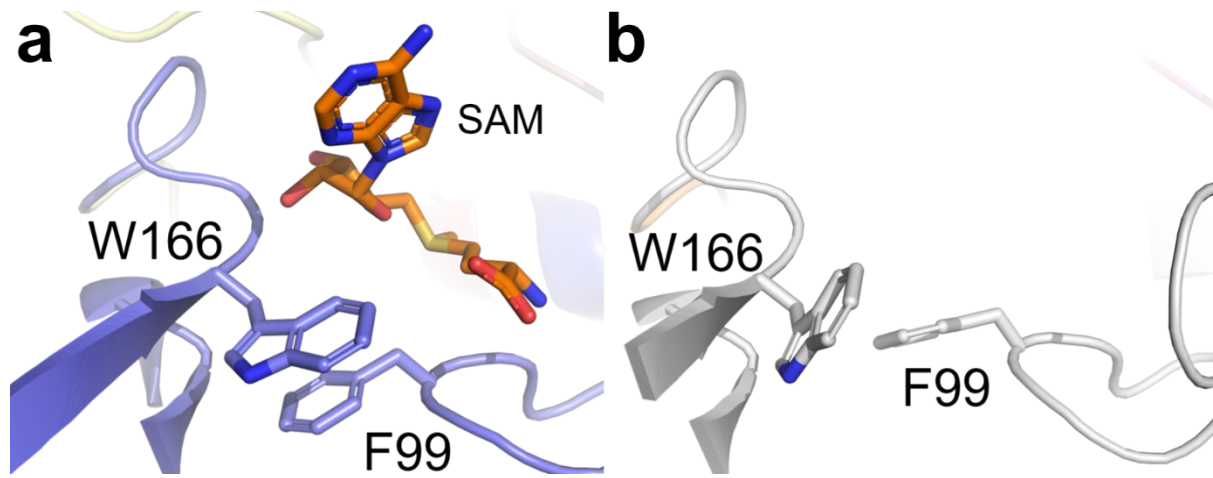




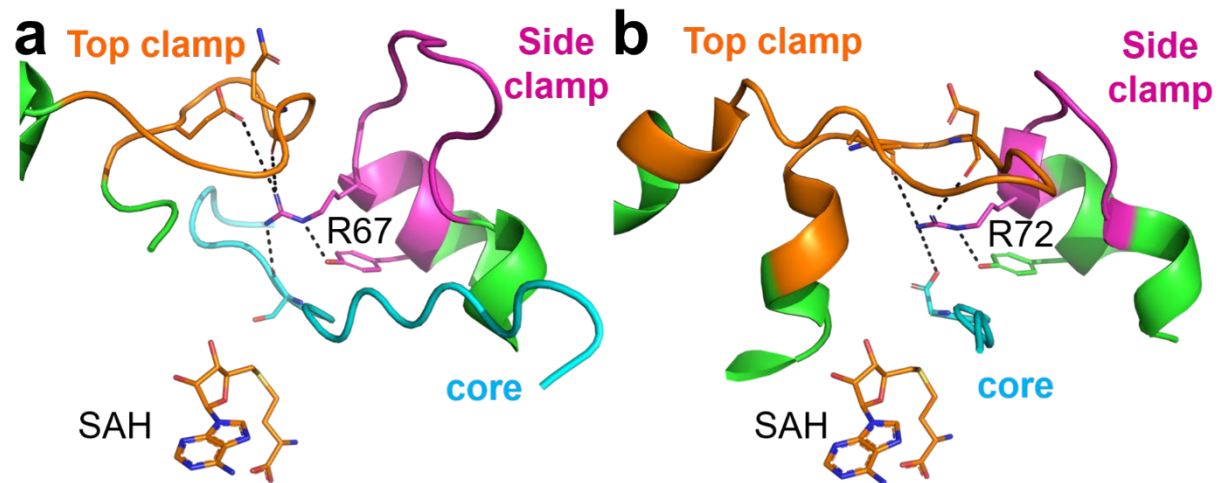
**Supplementary Figure 18. Superposition of SonM—SonA-2Me and SonM—SonA-BBD—(±)SAM complexes.** The SAM-bound heterodimer of SonM—SonA-BBD—(±)SAM (green and beige cartoon, PDB: 7LTR) is similar to SonM—SonA-2Me—SAH (grey and dark grey cartoon, PDB: 7LTE). SAM/SAH is shown as orange sticks.



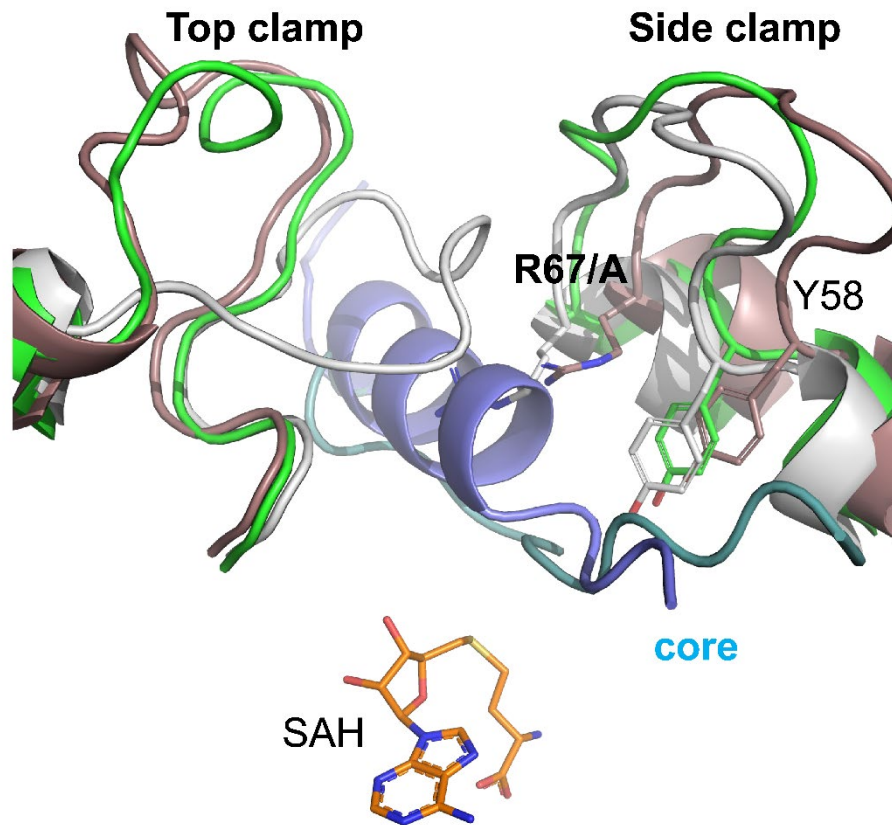
**Supplementary Figure 19. Bottom lock configurations of the SonM—SonA-BBD—( $\pm$ )SAM complex. (a) Closed and (b) open *bottom lock* configurations in SonM—SonA-BBD—( $\pm$ )SAM (PDB: 7LTR).**



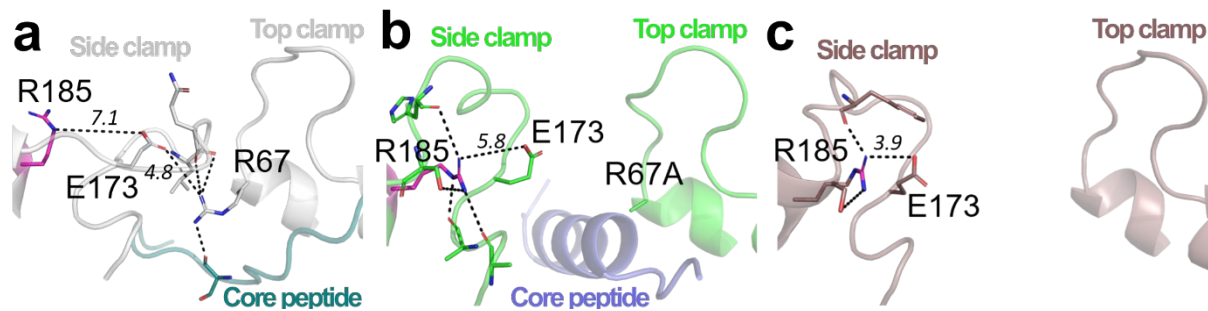
**Supplementary Figure 20. Top lock configurations between borosin  $\alpha$ -N-methyltransferases.** (a) View of the R67 active site interaction network of SonM—SonA-2Me—SAH (PDB: 7LTE) among residues in the top clamp, side clamp, and SonA core peptide. (b) View of the equivalent network of the homologous residue R72 in OphMA (PDB: 6MJG).



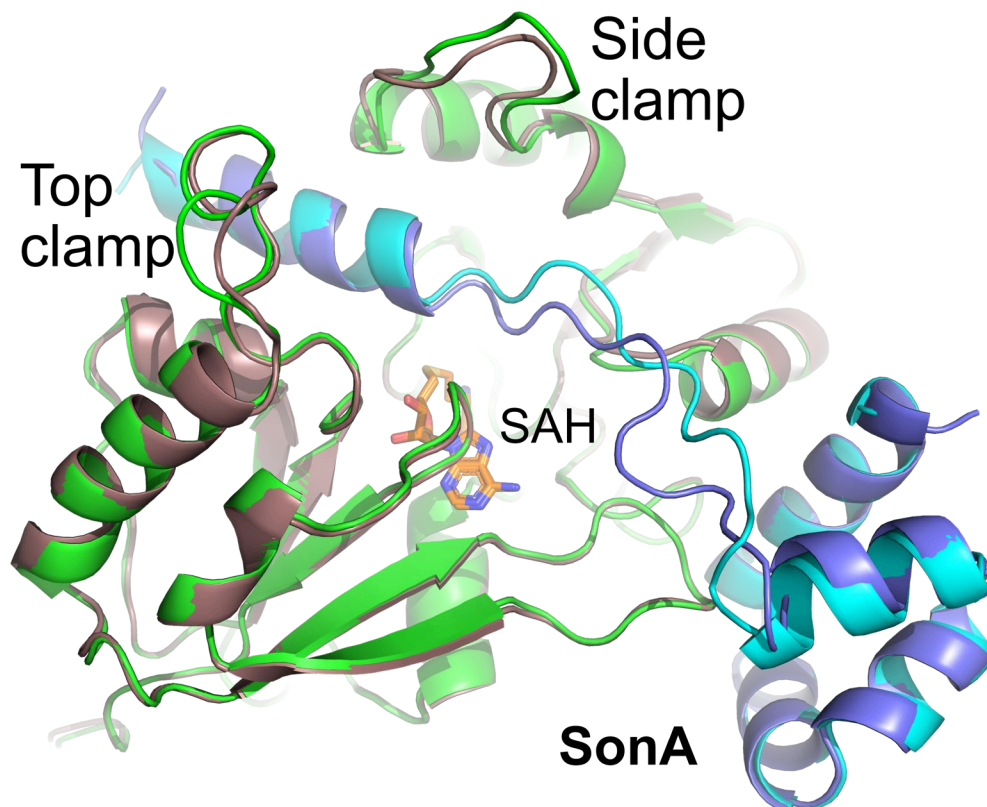
**Supplementary Figure 21. Active site configuration differences among split borosin complexes.** Superposition between the SonM—SonA-2Me—SAH (grey and teal cartoon, PDB: 7LTE), SonM-R67A—SonA-0Me—SAH (green and slate cartoon, PDB: 7LTS), and the heterodimer not bound to cofactor in SonM—SonA-BBD—( $\pm$ )SAM (maroon cartoon, PDB: 7LTR). In addition to the significant change in core peptide conformation, the top and side clamps of SonM-R67A—SonA-0Me—SAH are in intermediate positions as compared to the two other structures.



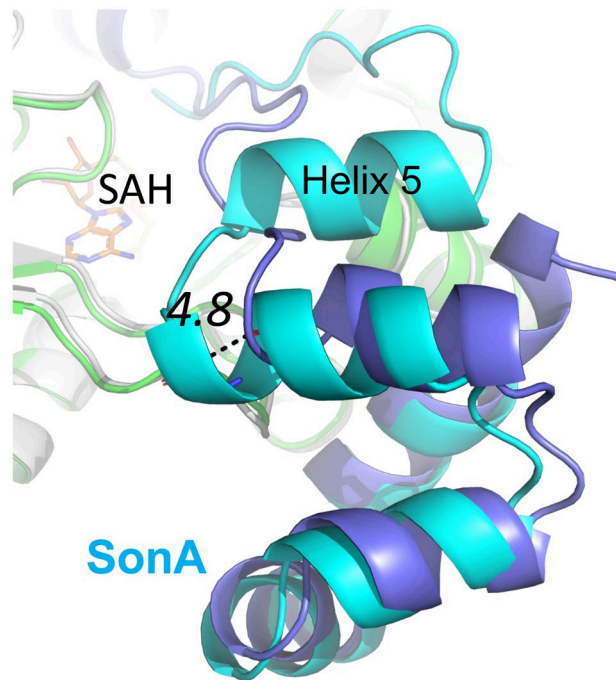
**Supplementary Figure 22. Interaction networks in different configurations of SonM and mutant complexes.** (a) Interaction network of R185 in the SonM—SonA-2Me—SAH (PDB: 7LTE) compared to the synonymous network in (b) SonM-R67A—SonA-0Me—SAH (PDB: 7LTS) and (c) Interaction network of R185 in the heterodimer not bound to cofactor in SonM—SonA-BBD—( $\pm$ )SAM (PDB: 7LTR). Similarly to SonM—SonA-BBD, SonM-R185 rotates  $\sim 5$  Å and interacts with SonM-E173 (5.8 Å), possibly compensating for the loss of the SonM-R67—SonM-E173 interaction. SonM-R185 also interacts with SonM-S182 (3.8 Å) and the carbonyl groups of SonM-A169 (3.5 Å), SonM-A171 (3.2 Å), and SonM-H180 (3.9 Å), contributing to stabilize the open top clamp conformation ( $\sim 14$  Å). Key distances are depicted as black dashed lines and their lengths noted in *italics* in Ångstroms.



**Supplementary Figure 23. Structural differences in the heterodimers of SonM-R67A—SonA-0Me—SAH.** Superposition of the two SonM-R67A—SonA-0Me—SAH heterodimers (green/cyan and maroon/slate cartoons, PDB: 7LTS). The RMSD (all atoms) between the two structures is 0.32 Å for 2000 atoms, when using the 'super' function in PyMOL. Small differences can be observed in the configuration of SonA as well the top and side clamps in SonM-R67A.



**Supplementary Figure 24. Structural differences in the BBD of SonM-R67A—SonA-0Me—SAH and *apo* SonM—SonA-2Me—SAH.** Superposition of SonM-R67A—SonA-0Me—SAH (grey and cyan cartoon, PDB: 7LTS) with *apo* SonM—SonA-2Me—SAH (green and slate cartoon, PDB: 7LTE). Helix 5 of the BBD is unwound in the SonM-R67A—SonA-0Me—SAH structure. Key distances are depicted as black dashed lines and their lengths noted in *italics> in Ångstroms.*



**Supplementary Figure 25. Mass spectrometric analysis of SonM mutant *in vitro* reactions.** HPLC-MS/MS spectra of the highest methylated species from AspN-digested SonA peptides after incubation at 30°C with the listed SonM mutant (a) SonM-Y93F, (b) SonM-R67K, (c) SonM-R67A, (d) SonM-Y58F, (e) SonM-Y71F, and (f) SonM-Y58F/Y71F, saturating [SAM], and the other enzymes and kit reagents used in the kinetics assay (see materials and methods). The amino acid sequence above each spectra depicts the *N*-methylated residues that could be confirmed by MS/MS fragmentation (solid orange circles) or are inferred *N*-methylated since the position is not completely defined by MS/MS (unfilled orange circles). Observed MS/MS fragmentation masses are listed above (b-ions) and below (y-ions) the amino acid sequence. The gray lines within the sequence mark the sites of fragmentation. Masses of methylation-containing ions are denoted in brackets, where 'Me' stands for methylation. The ppm difference from the observed masses to the theoretical expected masses are labeled in parentheses. A 10.0-ppm mass cutoff for annotated HPLC-MS/MS peaks was used. The protein, time of *in vitro* reaction, parent ion information and HPLC retention time (RT) are listed in the upper right corner of the LC-MS/MS spectra. Off-target methylations were not detected in any sample.



**a**

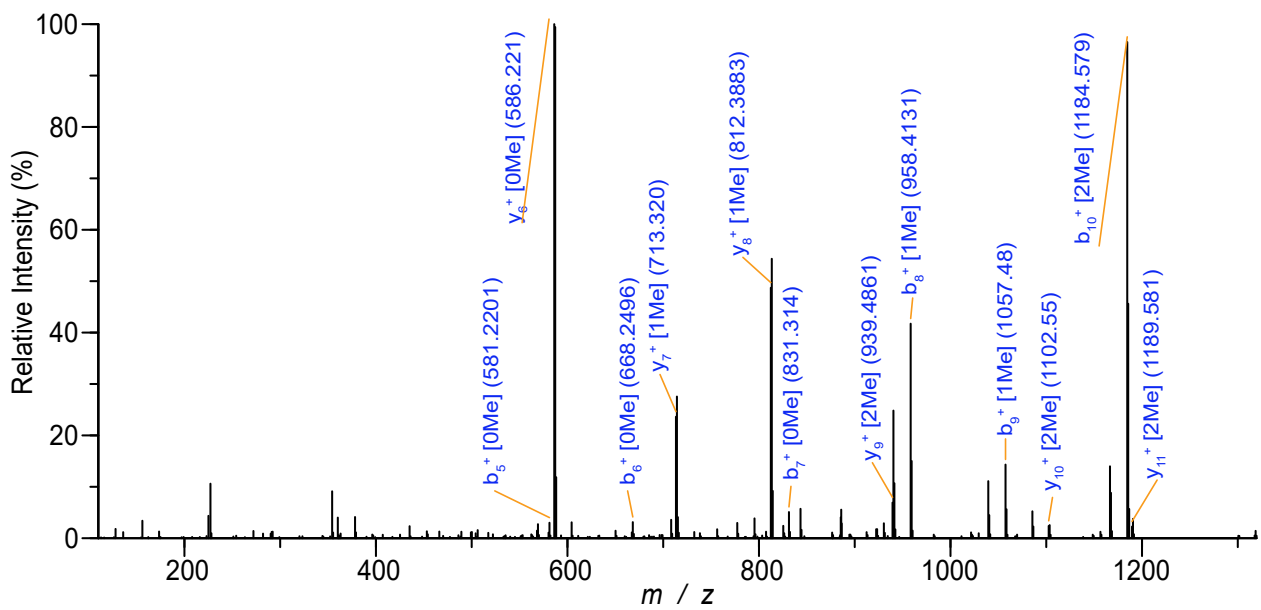
- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SonA (with SonM-Y93F, at 47.73 min rxn)

Parent ion:  $[M+2Me+2H]^{2+}$  (885.40)

RT: 12.16 min



**b**

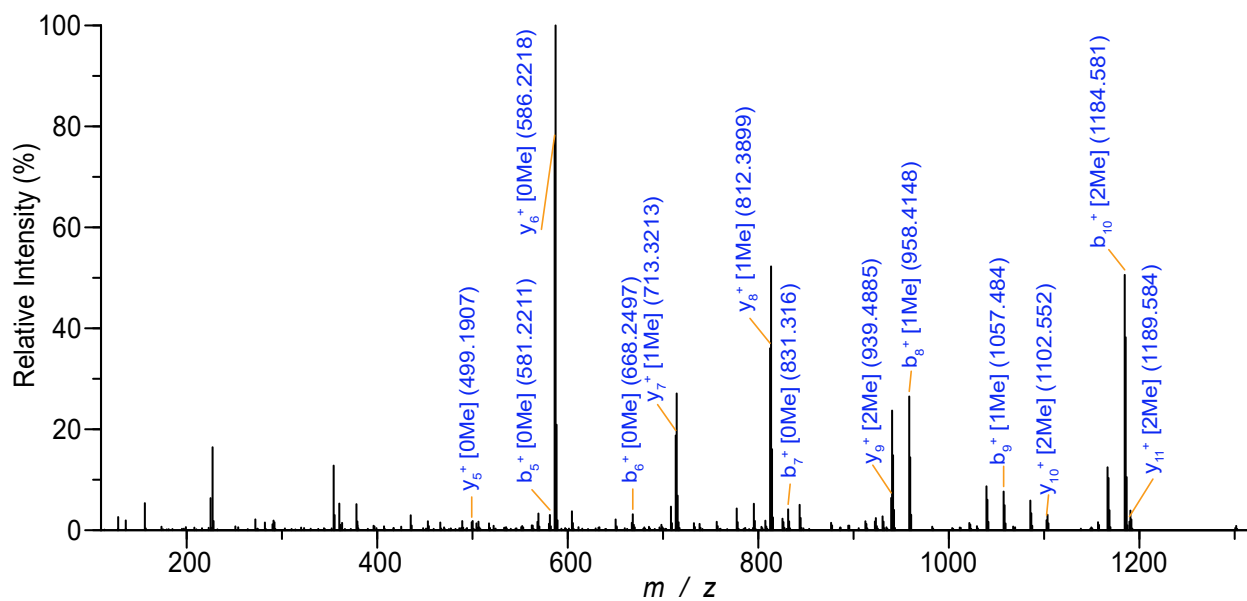
- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SonA (with SonM-R67K, at 135.5 min rxn)

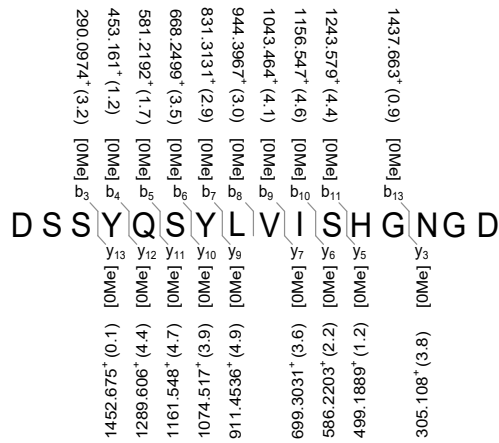
Parent ion:  $[M+2Me+2H]^{2+}$  (885.40)

RT: 12.11 min



C

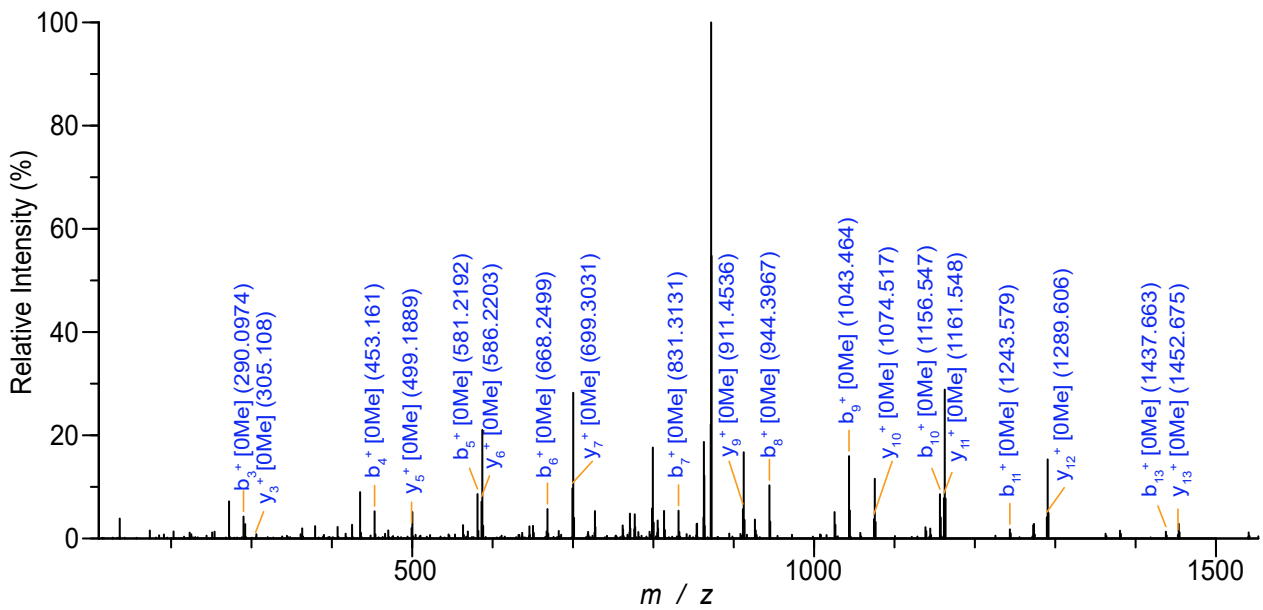
- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SonA (with SonM-R67A, at 128.9 min rxn)

Parent ion: [M+0Me+2H]<sup>2+</sup> (871.39)

RT: 11.54 min

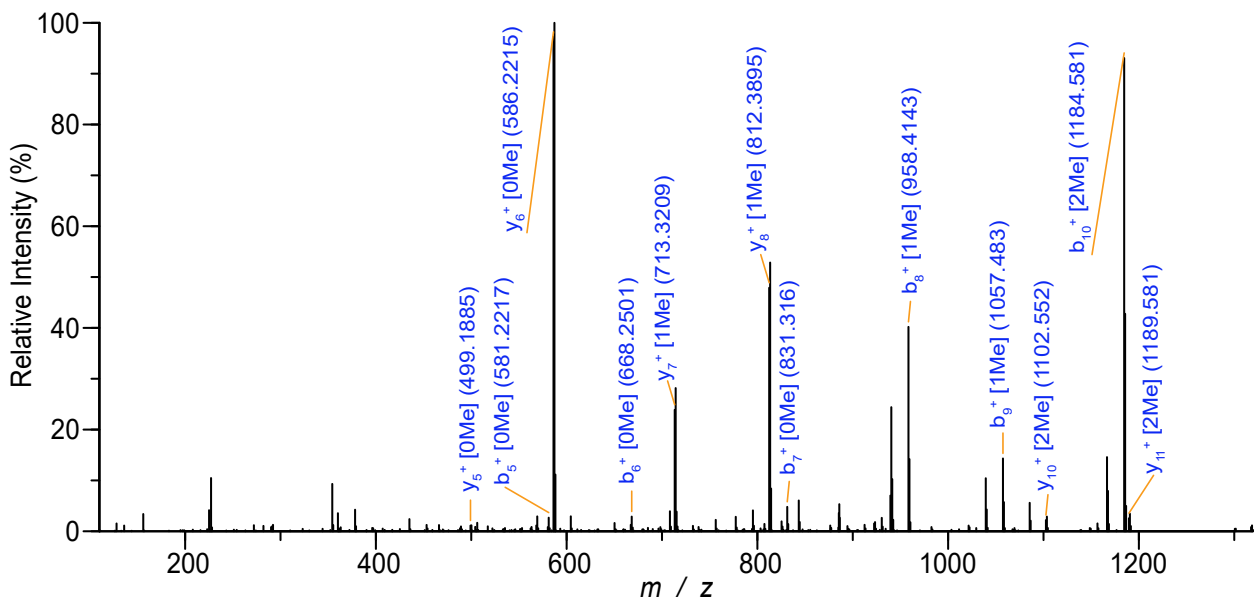


d

- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SonA (with SonM-Y58F, at 108.5 min rxn)  
Parent ion: [M+2Me+2H]<sup>2+</sup> (885.40)  
RT: 12.09 min



e

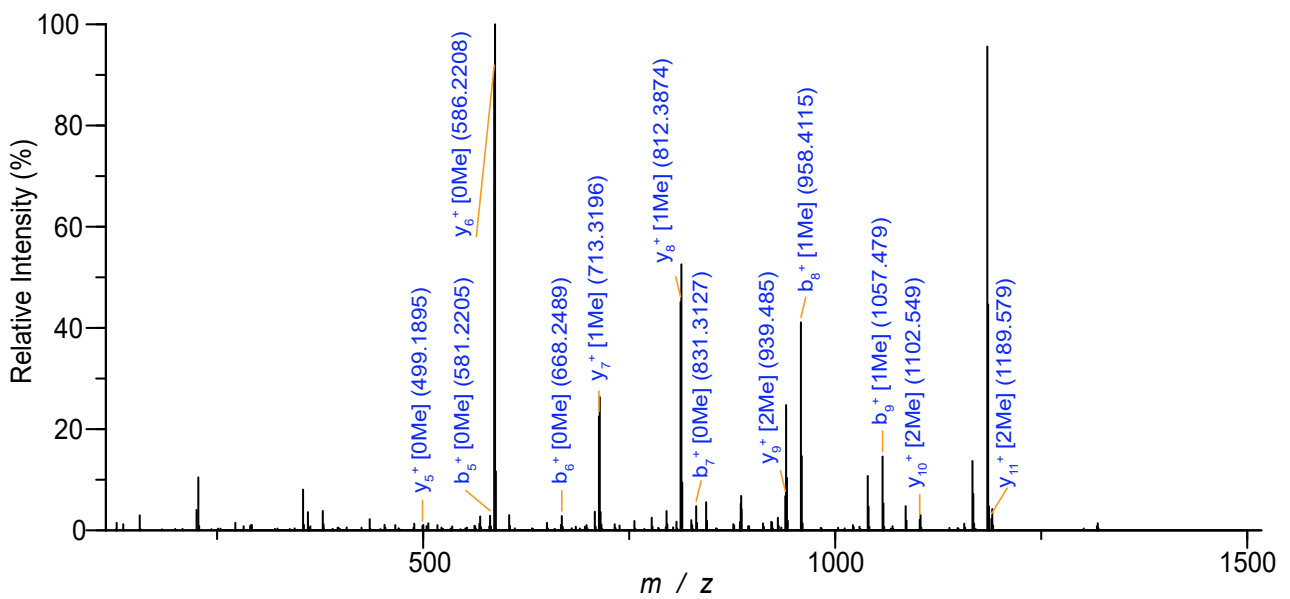
- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SonA (with SonM-Y71F, at 135.5 min rxn)

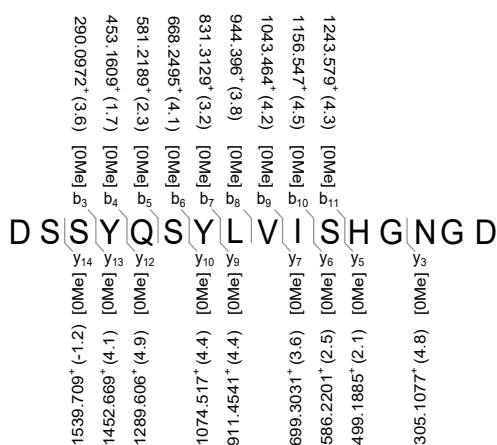
Parent ion: [M+2Me+2H]<sup>2+</sup> (885.40)

RT: 12.13 min



f

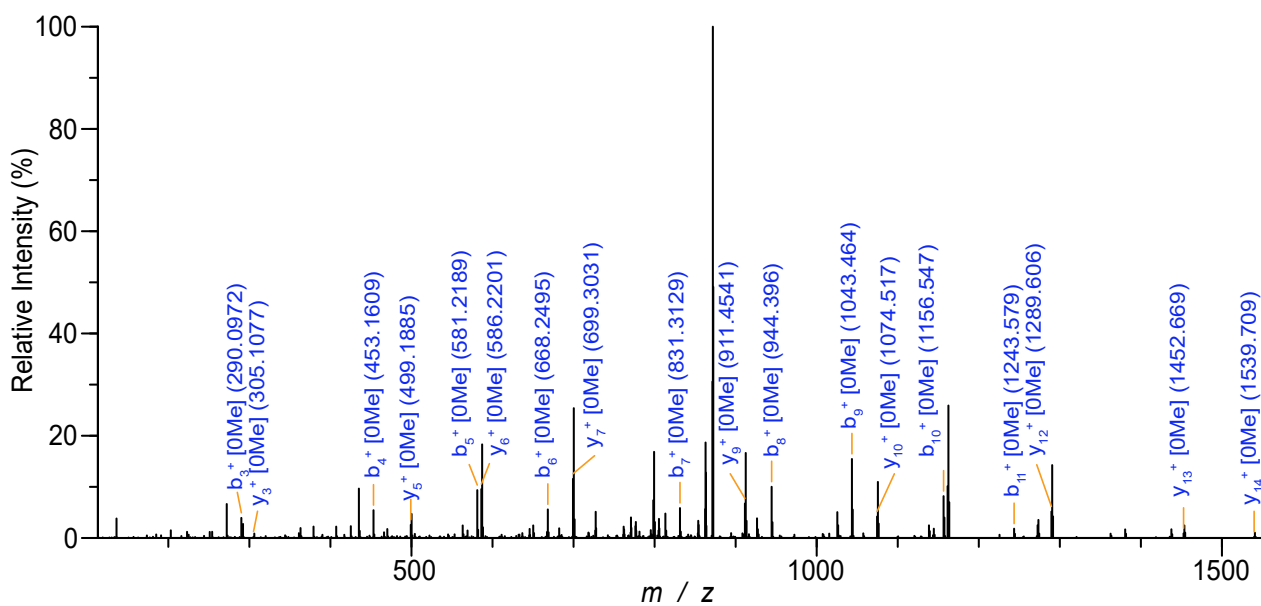
- Methylation localized by LC-MS/MS
- Methylation inferred by LC-MS/MS



SonA (with SonM-Y58F/Y71F, at 128.9 min rxn)

Parent ion:  $[M+0Me+2H]^{2+}$  (871.39)

RT: 11.47 min



## References

1. Montalbán-López, M. *et al.* New developments in RiPP discovery, enzymology and engineering. *Nat. Prod. Rep.* **37**, 919–961 (2020).
2. Song, H. *et al.* A molecular mechanism for the enzymatic methylation of nitrogen atoms within peptide bonds. *Sci. Adv.* **4**, eaat2720 (2018).
3. Ongpipattanakul, C. & Nair, S. K. Molecular basis for autocatalytic backbone *N*-methylation in RiPP natural product biosynthesis. *ACS Chem. Biol.* **13**, 2989–2999 (2018).
4. Sugimoto, K. *et al.* Crystal structure of an aromatic ring opening dioxygenase LigAB, a protocatechuate 4,5-dioxygenase, under aerobic conditions. *Structure* **7**, 953–965 (1999)
5. Sugimoto, K. *et al.* Molecular mechanism of strict substrate specificity of an extradiol dioxygenase, DesB, derived from *Sphingobium* sp. SYK-6. *PLoS One* **9**, e92249 (2014).
6. Strieter, E. R. & Andrew, T. L. Restricting the  $\psi$  torsion angle has stereoelectronic consequences on a scissile bond: An electronic structure analysis. *Biochemistry* **54**, 5748–5756 (2015).
7. Aqvist, J., Luecke, H., Quijcho, F. A. & Warshel, A. Dipoles localized at helix termini of proteins stabilize charges. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 2026–30 (1991).
8. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127–34 (1995).