Electronic Supplementary Material (ESI) for RSC Chemical Biology.

Supporting Information

A thorough analysis and categorization of bacterial interrupted adenylation domains, including previously unidentified families

Taylor A. Lundy,[a,§] Shogo Mori,[a,§] Sylvie Garneau-Tsodikova[a,*]

University of Kentucky, Department of Pharmaceutical Sciences, College of Pharmacy,
Lexington, KY 40536-0596, USA.

[§] These two authors contributed equally to this work.

* Correspondence to: sylviegtsodikova@uky.edu

Table of Contents **Pages**

Supplementary Tables and Figures **Pages**

1. Construction of data sets for taxonomic and phylogenetic tree analyses of interrupted A domains. The taxonomic and phylogenetic analyses were performed by using Genome Workbench (National Center for Biotechnology Information; NCBI). To construct the data sets, we performed protein BLAST searches for families 1-5 of interrupted A domains, with the following previously characterized or reported interrupted A domains as query sequences: CAJ34375.1 (1902-2806; TioS($(A_8M_IA_9)_4$),[1] ABV56588.1 (1964-2815; KtzH($(A_8M_{II}A_9)_4$)[2]/QDA77059.1 (4164-4980; ThxA2($(A_8M_{III}A_9)_6$),[3] CAJ34370.1 (1-768; TioN($A_2M_{IV}A_3$)),[4] AGC65517.1 (1-818; TtbB($(A_2M_VA_3)_5$),[5, 6] and AKQ09584.1 (508-1739; ColG($A_8M_{II}M_IA_9$))[7]/WP_014748210.1 (491-1780; DidJ($(A_8M_{III}M_IA_9)_{11}$).[8] For newly discovered families 6 and 7 of interrupted A domains, RKZ63928.1 (514-1309) and WP_083695640.1 (1564-2560) were used as query sequences, respectively. Data from the BLAST results that had a "Query Cover" value (the percentage of the query sequence that overlaps with the reference sequence) of ≥68.2% for family 1, ≥68.8% for family 2a, ≥73.6% for family 2b, ≥90.2% for family 3, and ≥81.5% for family 4, were extracted for further analyses. These thresholds were determined by analyzing several proteins individually around these coverage values to determine if they were corresponding interrupted A domains or if they were random hits. For other uncommon families (families 5-7), the top 20-30 hits were individually examined by analyzing conserved domains and/or identifying the a1-a10 motifs of A domains to determine if they were members of the corresponding families. From these data sets, identical (duplicated protein ID) proteins, proteins from unknown (environmental) sources, as well as proteins from other families (that were checked individually) were removed, which finally resulted data sets of n = 536, 149, 79, 64, 188, 11, 14, and 3 interrupted A domains for families 1, 2a, 2b, and 3-7, respectively. These data sets were used to build taxonomic trees (Figs. 3, S1-S7, and Tables S1-S7). Amino acid sequences from each data set were aligned by MUSCLE,[9, 10] whose results were used to build phylogenetic trees by the neighbor joining method (Figs. 4, S6-S12, and Tables S1-S7).[11] The generated tree data were visualized and processed by the web service Interactive Tree of Life (iTOL).[12] For building a phylogenetic tree of just the M domains, protein BLAST searches were performed with the data sets created above for constructing taxonomic and phylogenetic trees of each interrupted A domain family. The following truncated proteins were used as query sequences: CAJ34375.1 (2365-2752; TioS(M_I)$_4$), ABV56588.1 (2431-2719; KtzH(M_{II})$_4$), QDA77059.1 (4591-4931; ThxA2(M_{III})$_6$), CAJ34370.1 (136-385; TioN(M_{IV})), AGC65517.1 (158-436; TtbB(M_V)$_5$), and RKZ63928.1 (891-1185) for types I-VI M domains, respectively. These local BLAST searches resulted in n = 499, 148, 79, 63, 188, and 14 protein hits for types I-VI M domains, respectively. For family 5 (n = 11) and family 7 (n = 3) interrupted A domains (which embed two M domains in a single A domain), each M domain sequence was manually extracted. As a result, n = 22 and n = 6 of M domains were generated from families 5 and 7, respectively. All these M domains were aligned by MUSCLE. The result of the multiple sequence alignment was used to build a phylogenetic tree by the neighbor joining method followed by processing on the iTOL website (Fig. 5).

The size of each type of M domain (number of amino acids per M domain in Table 1) was calculated by the following method. First, the amino acid sequences from a8 to a9 of A domain conserved sequence motifs from representatives of families 1 (CAJ34375.1 (2313-2780; TioS($A_8M_IA_9$)$_4$)), 2a (ABV56588.1 (2379- 2749; KtzH($A_8M_{II}A_9$)$_4$)), and 2b (QDA77059.1 (4539-4955; ThxA2($A_8M_{III}A_9$)$_6$)) were extracted. Similarly, the amino acid sequences from a2 to a3 from representatives of families 3 (CAJ34370.1 (89-427; TioN($A_2M_{IV}A_3$))) and 4 (AGC65517.1 (100-468; TtbB($A_2M_VA_3$)$_5$)) were extracted. The reason for including A domain conserved motifs before and after the M domains was to ensure that the BLAST search recognized the correct regions accurately. Second, each extracted sequence was used as the query sequence to perform a protein BLAST search against the data sets of corresponding families created for generating the taxonomic and phylogenetic trees. Third, the number of amino acids of each hit of the BLAST search was counted, whose average and standard deviation were calculated. Since this average number of amino acids contained A domain portions before and after the corresponding M domain, the number of A domain amino acids of each representative was subtracted from the average number. This number ± standard deviation for each type of M domain is indicated as number of amino acids per M domain in Table 1. For type VI M domains that have only n = 14, the individual M domains were extracted to calculate the average number of amino acids and the standard deviation.

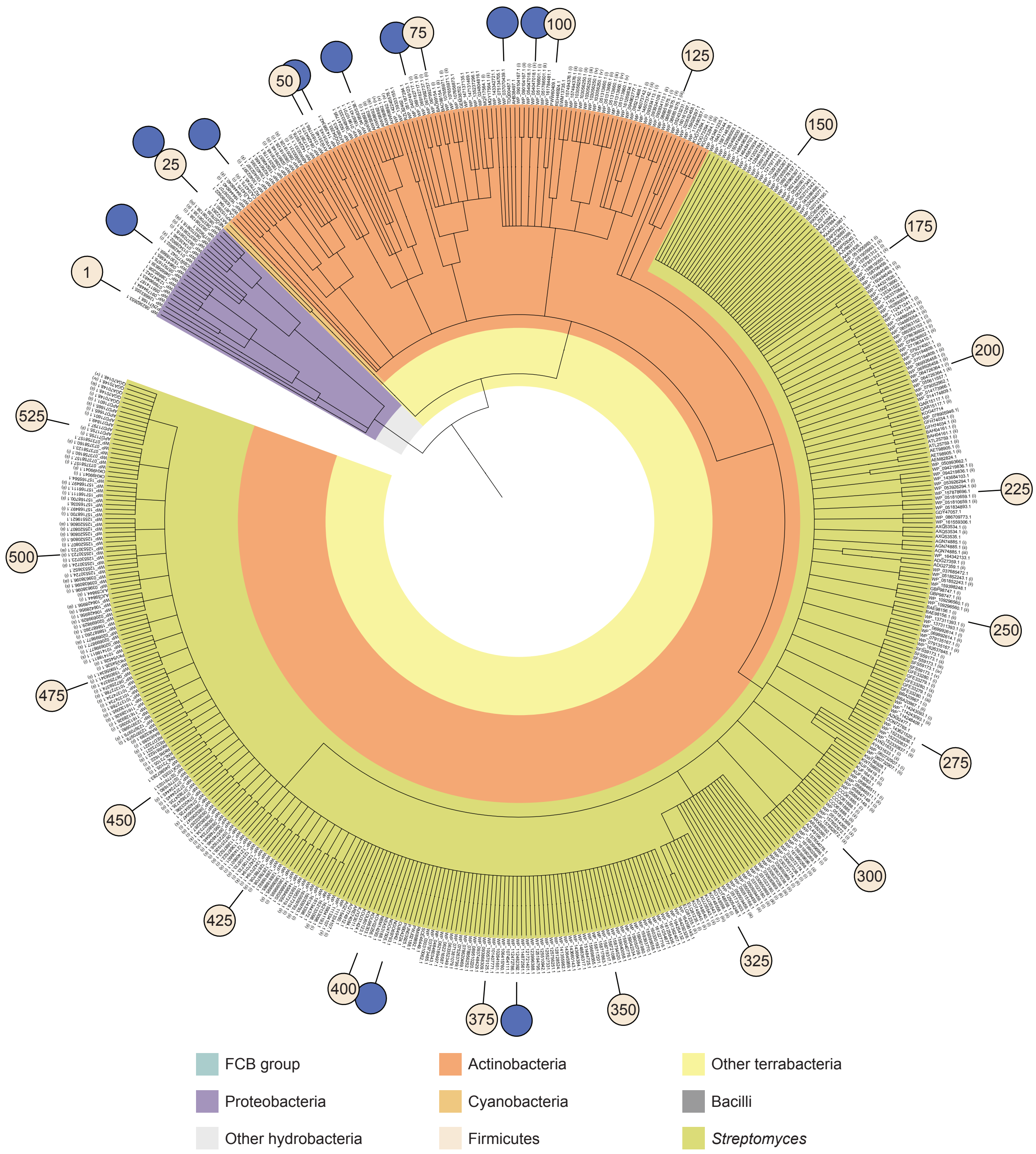


Fig. S1. Taxonomic tree of family 1 ($A_8M_bA_9$) interrupted A domains (n = 536). The interrupted A domains used for multiple sequence alignment (Fig. S13) are indicated by blue balloons. The interrupted A domains are numbered clockwise from the top left corner of the tree. Detailed information including these numbers, NCBI IDs, ranges of sequences used for the analysis, and organisms is shown in Table S1.

Fig. S2. Taxonomic tree of family 2a ($A_8M_{s(O,\ \mathrm{Ser/Thr})}A_9$) interrupted A domains (n = 149). The interrupted A domains used for multiple sequence alignment (Fig. S14) are indicated by blue balloons. The interrupted A domains are numbered clockwise from the top left corner of the tree. Detailed information including these numbers, NCBI IDs, ranges of sequences used for the analysis, and organisms is shown in Table S2.

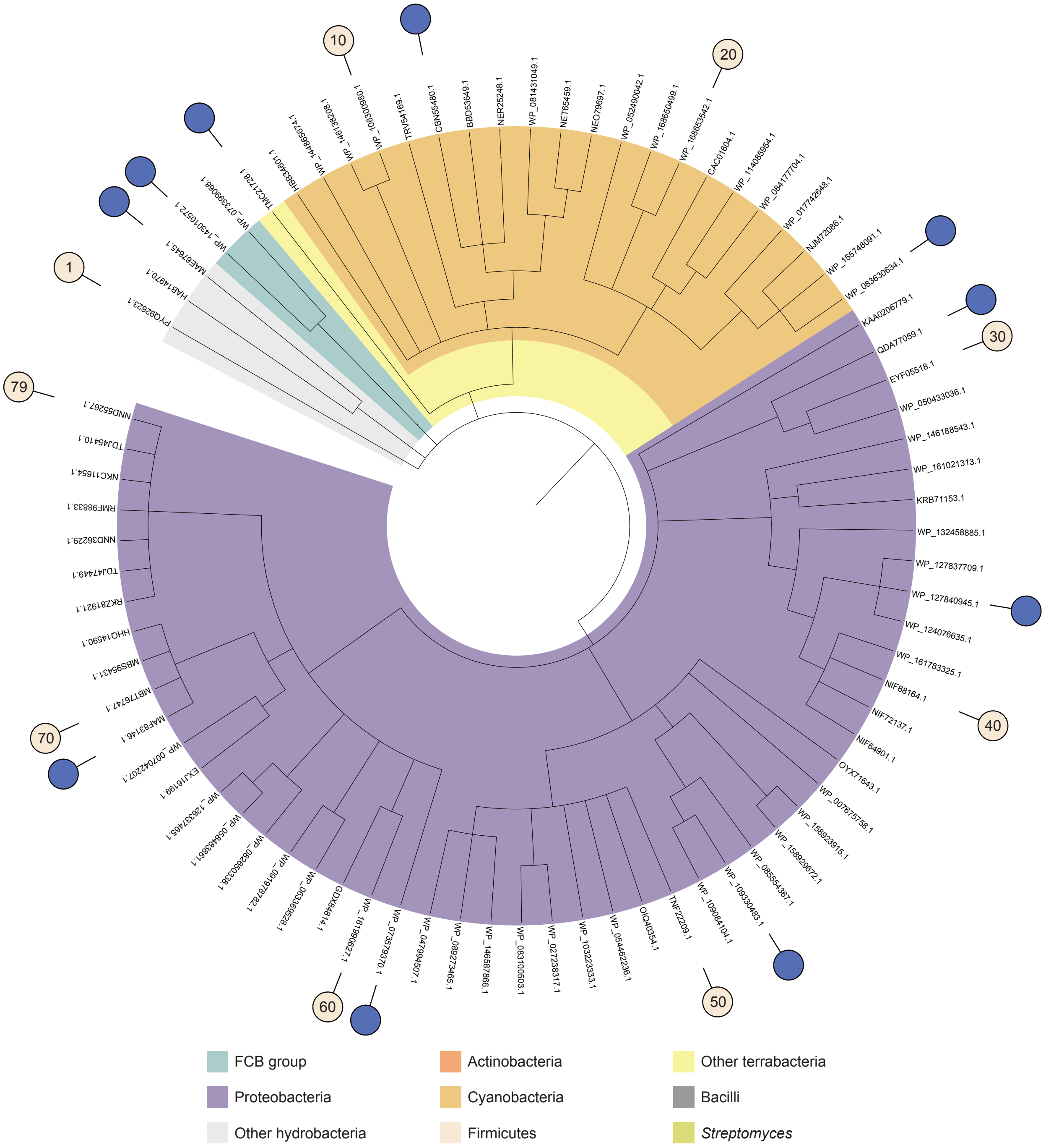


Fig. S3. Taxonomic tree of family 2b ($A_8M_{s(O,\ Tyr)}A_9$) interrupted A domains (n = 79). The interrupted A domains used for multiple sequence alignment (Fig. S15) are indicated by blue balloons. The interrupted A domains are numbered clockwise from the top left corner of the tree. Detailed information including these numbers, NCBI IDs, ranges of sequences used for the analysis, and organisms is shown in Table S3.

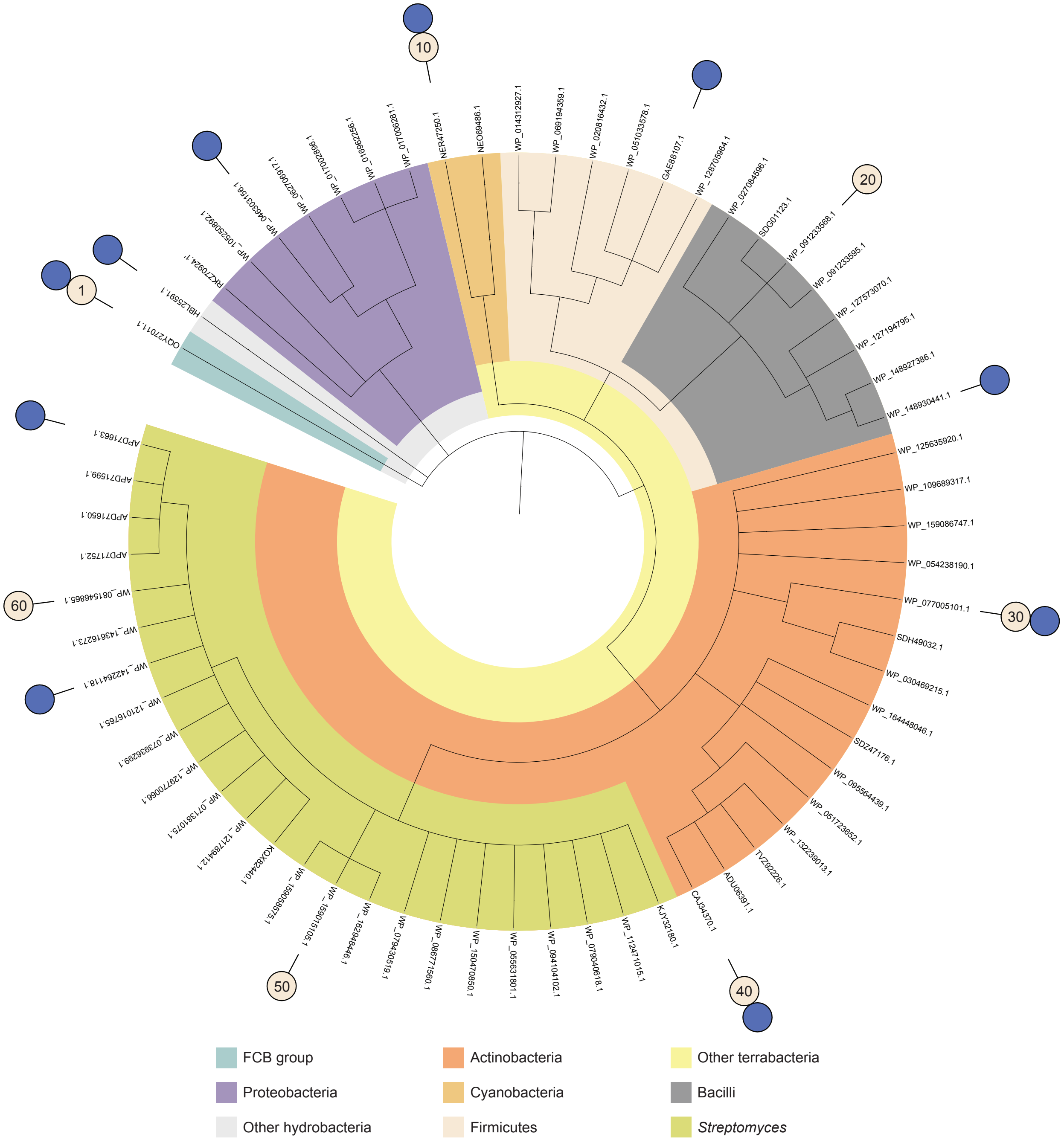


Fig. S4. Taxonomic tree of family 3 ($A_2M_{s(S)}A_3$) interrupted A domains (n = 64). The interrupted A domains used for multiple sequence alignment (Fig. S16) are indicated by blue balloons. The interrupted A domains are numbered clockwise from the top left corner of the tree. Detailed information including these numbers, NCBI IDs, ranges of sequences used for the analysis, and organisms is shown in Table S4.

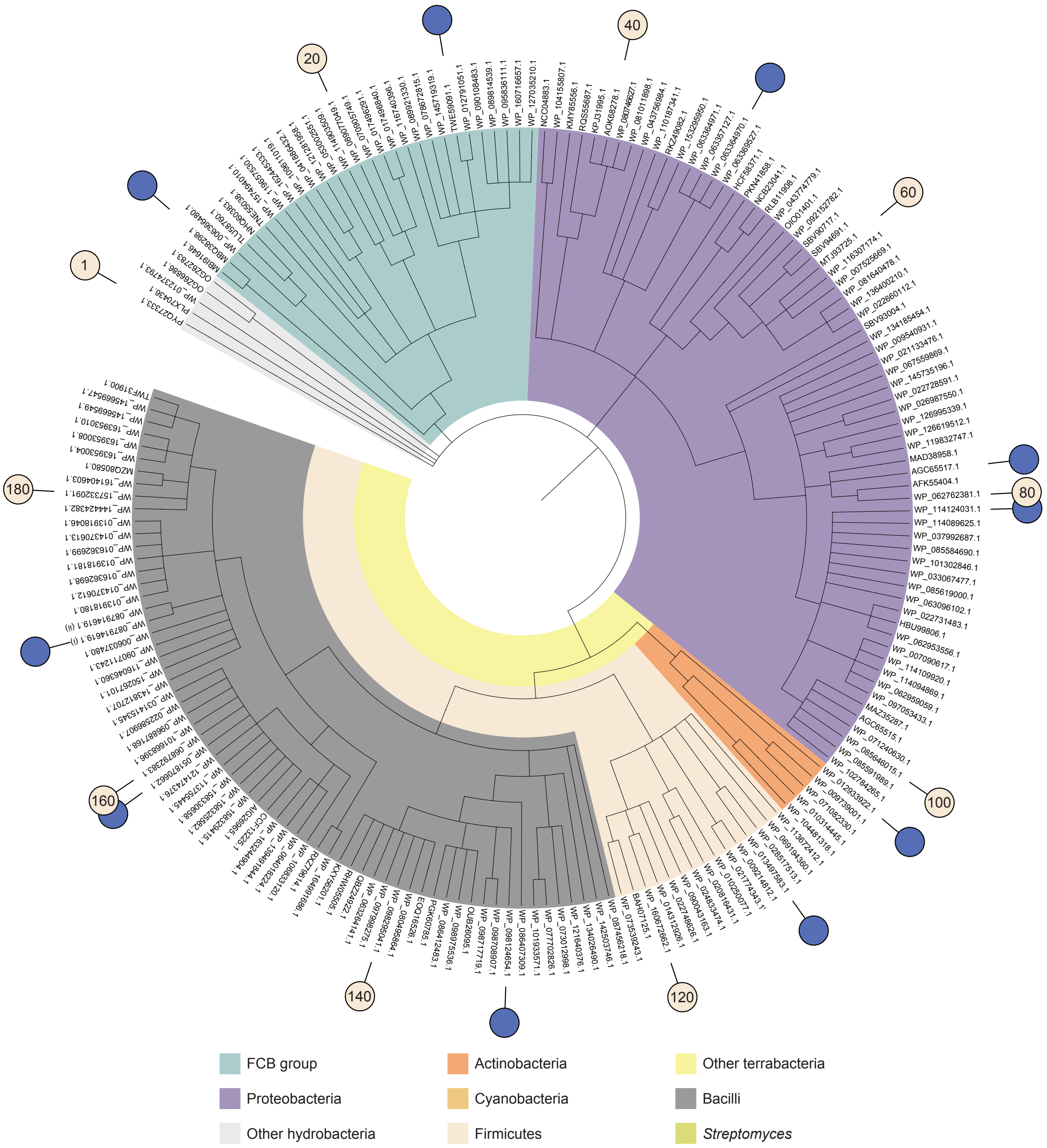


Fig. S5. Taxonomic tree of family 4 ($A_2M_bA_3$) interrupted A domains (n = 188). The interrupted A domains used for multiple sequence alignment (Fig. S17) are indicated by blue balloons. The interrupted A domains are numbered clockwise from the top left corner of the tree. Detailed information including these numbers, NCBI IDs, ranges of sequences used for the analysis, and organisms is shown in Table S5.

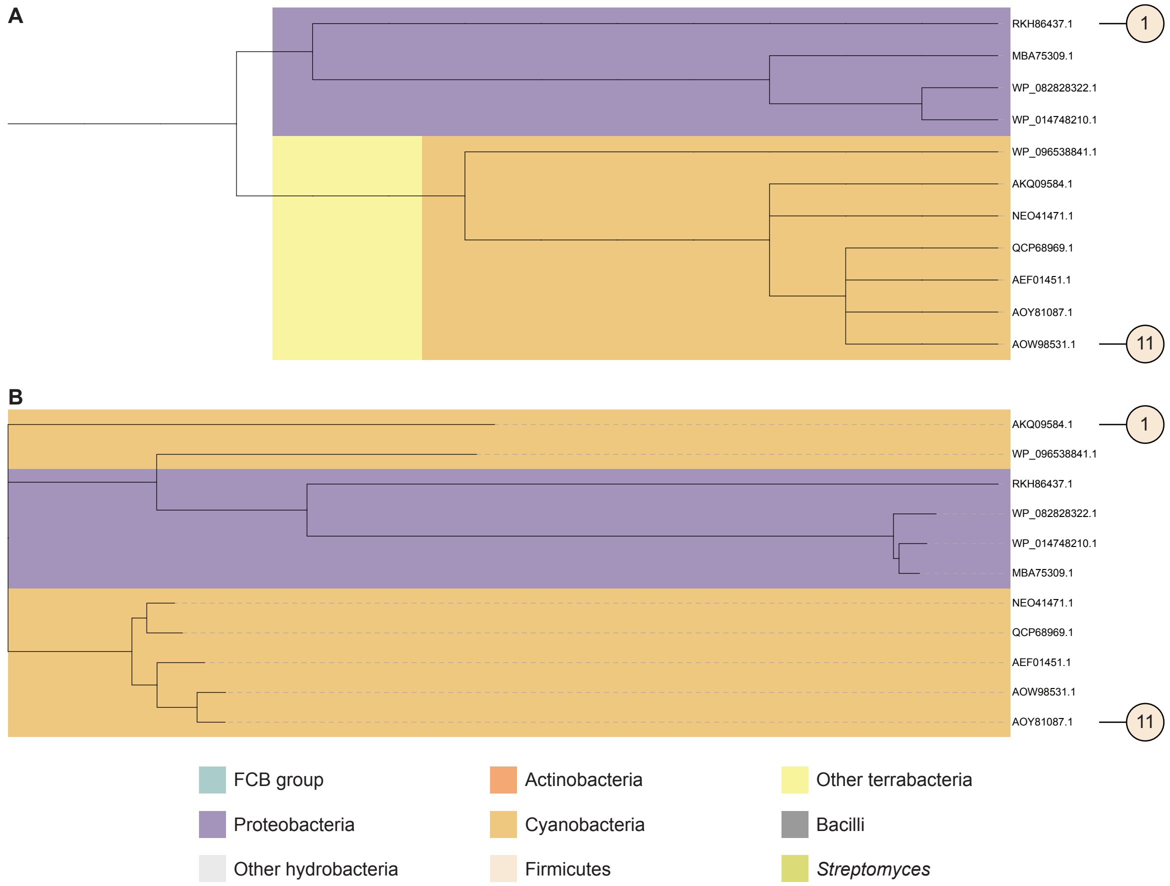


Fig. S6. A. Taxonomic and **B.** phylogenetic trees of families 5a/b ($A_8M_{s(O)}M_bA_9$) interrupted A domains (n = 11). The interrupted A domains are numbered from top to bottom. Detailed information including these numbers, NCBI IDs, ranges of sequences used for the analysis, and organisms is shown in Table S6.

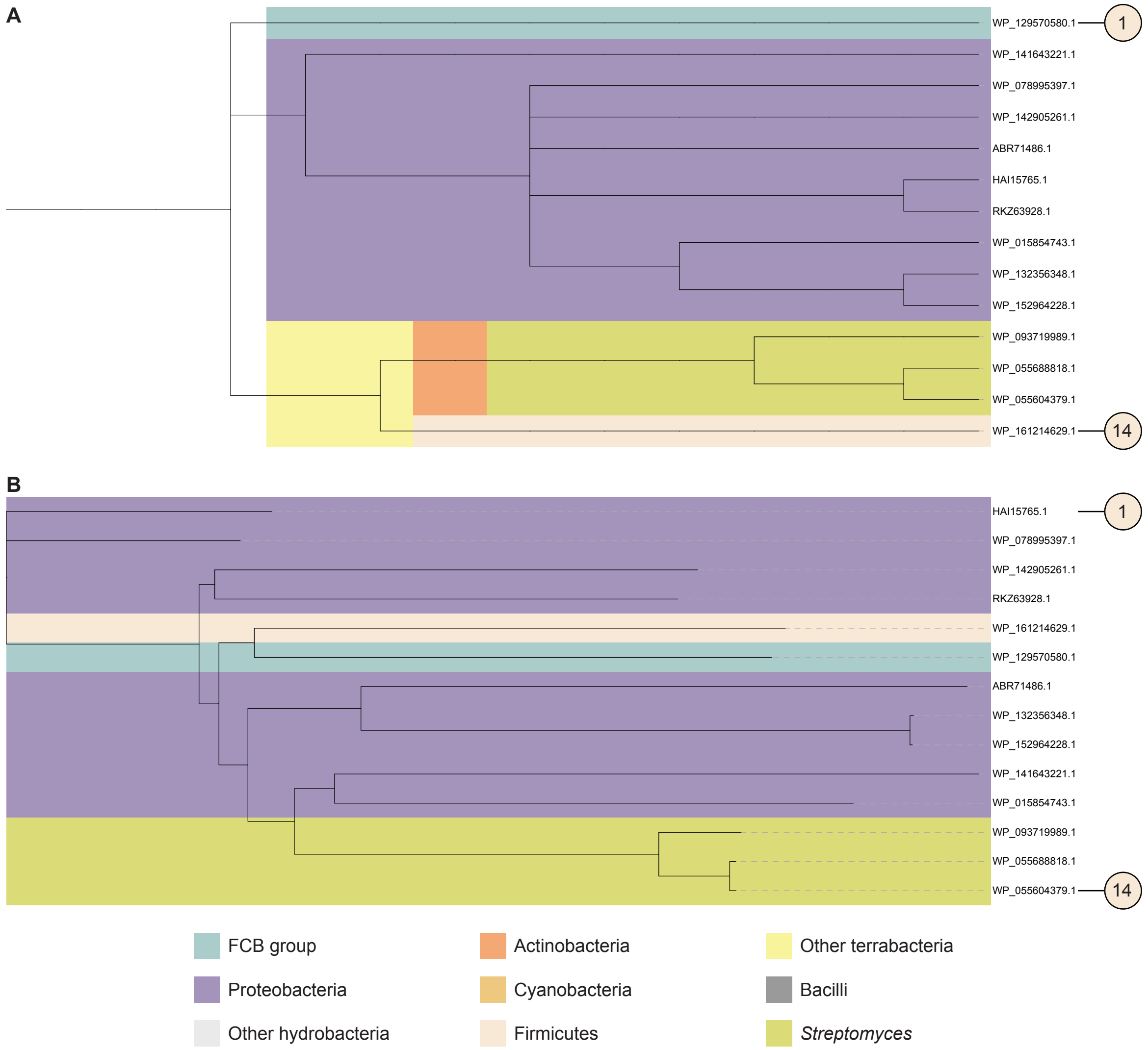


Fig. S7. A. Taxonomic and **B.** phylogenetic trees of family 6 ($A_6M_{s(O,\ \mathrm{arom})}A_7$) interrupted A domains (n = 14). The interrupted A domains are numbered from top to bottom. Detailed information including these numbers, NCBI IDs, ranges of sequences used for the analysis, and organisms is shown in Table S7.

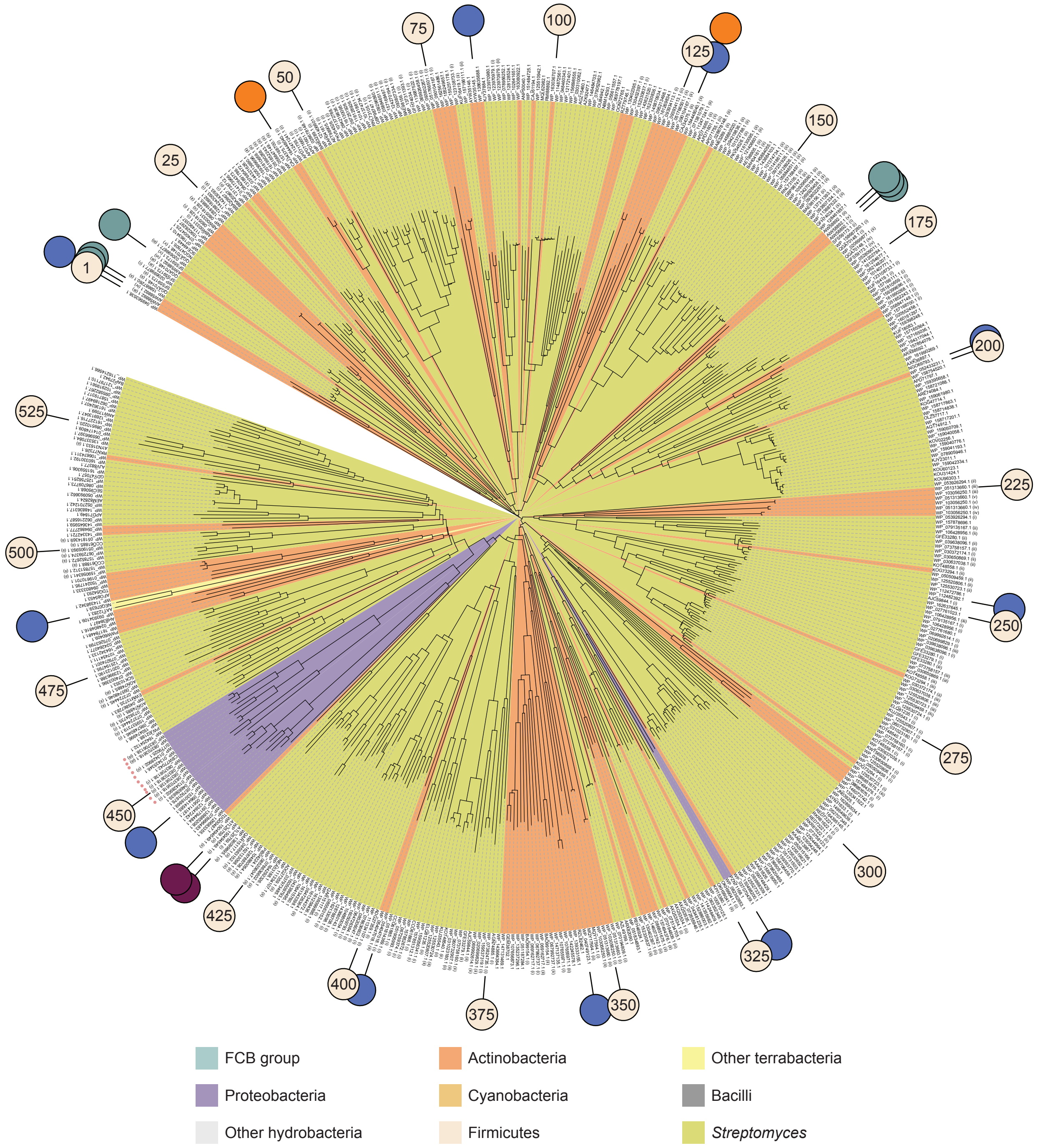


Fig. S8. Phylogenetic tree of family 1 ($A_8M_bA_9$) interrupted A domains (n = 536). The interrupted A domains used for multiple sequence alignment (Fig. S13) are indicated by blue balloons. Examples of proteins that contain multiple interrupted A domains are indicated by orange (two heterologous ones on CAJ34375.1; TioS), burgundy (two homologous ones on ATL25759.1), and greyish green (six combination ones on AIW58892.1) balloons. The interrupted A domains are numbered clockwise from the top left corner of the tree. Detailed information including these numbers, NCBI IDs, ranges of sequence used for the analysis, and organisms is shown in Table S1.

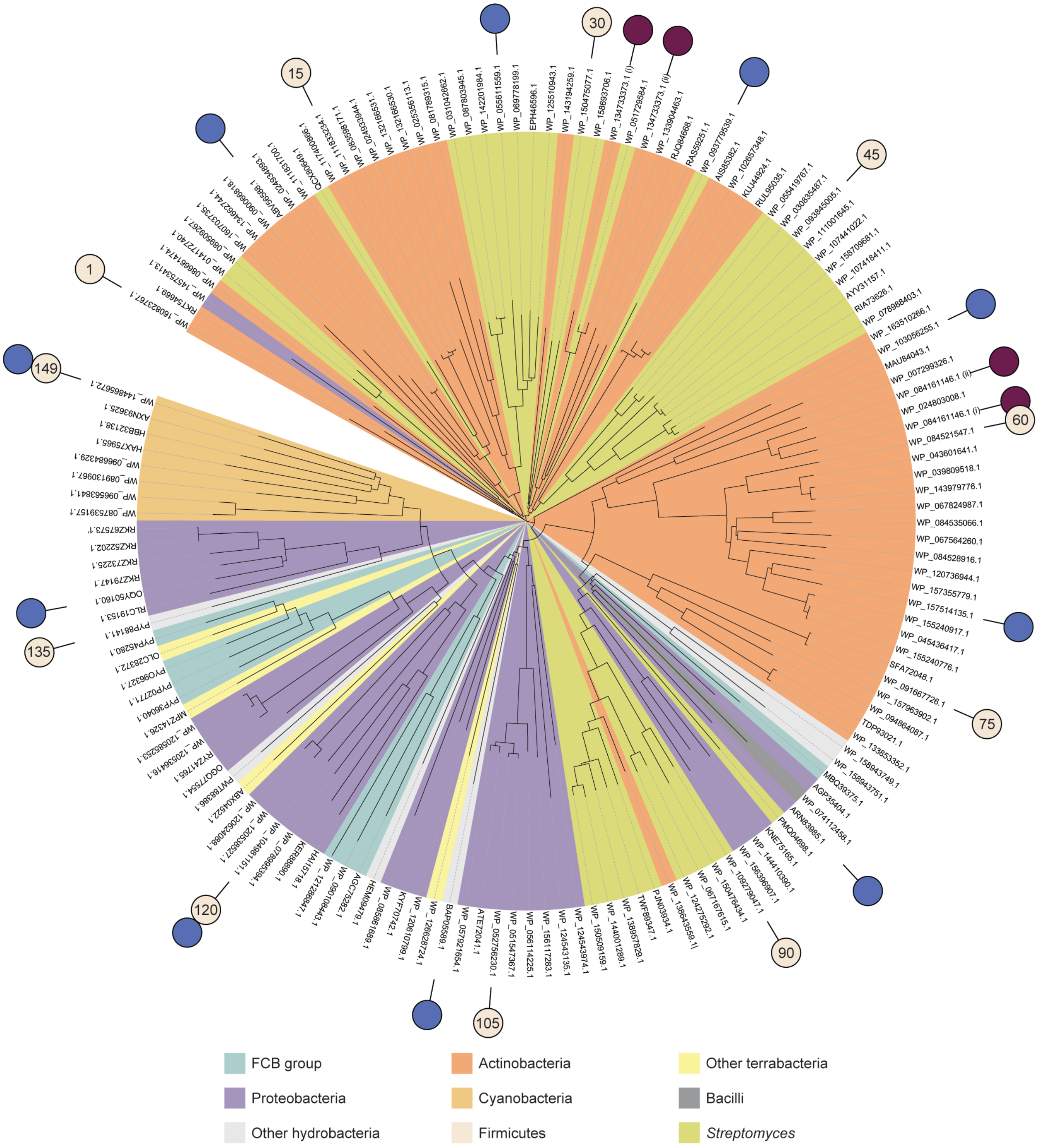


Fig. S9. Phylogenetic tree of family 2a ($A_8M_{s(O,\ \mathrm{Ser/Thr})}A_9$) interrupted A domains (n = 149). The interrupted A domains used for multiple sequence alignment (Fig. S14) are indicated by blue balloons. Proteins that contain two homologous interrupted A domains are indicated by burgundy balloons. The interrupted A domains are numbered clockwise from the top left corner of the tree. Detailed information including these numbers, NCBI IDs, ranges of sequences used for the analysis, and organisms is shown in Table S2.

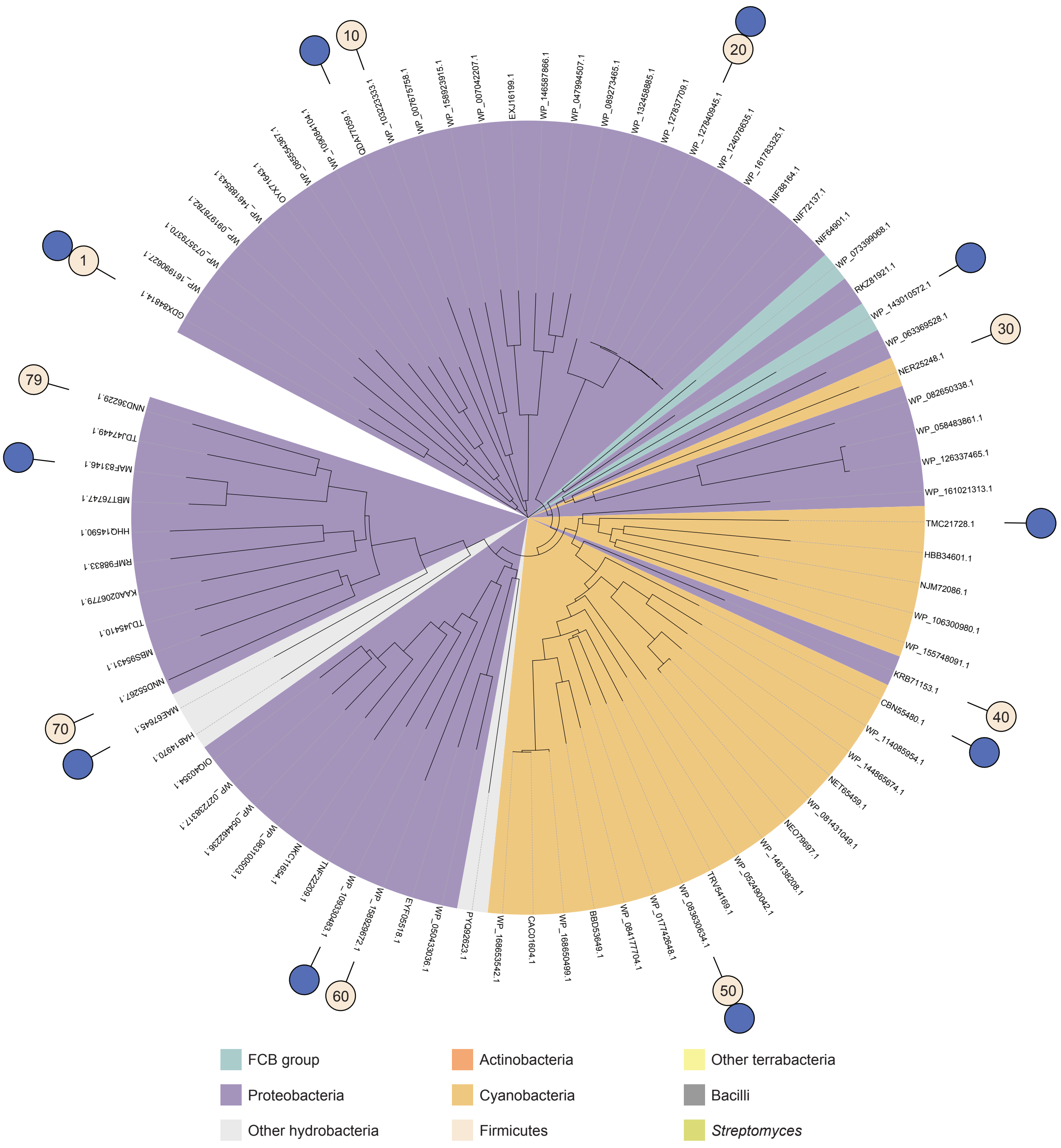


Fig. S10. Phylogenetic tree of family 2b ($A_8M_{s(O,\ \mathrm{Tyr})}A_9$) interrupted A domains (n = 79). The interrupted A domains used for multiple sequence alignment (Fig. S15) are indicated by blue balloons. The interrupted A domains are numbered clockwise from the top left corner of the tree. Detailed information including these numbers, NCBI IDs, ranges of sequences used for the analysis, and organisms is shown in Table S3.

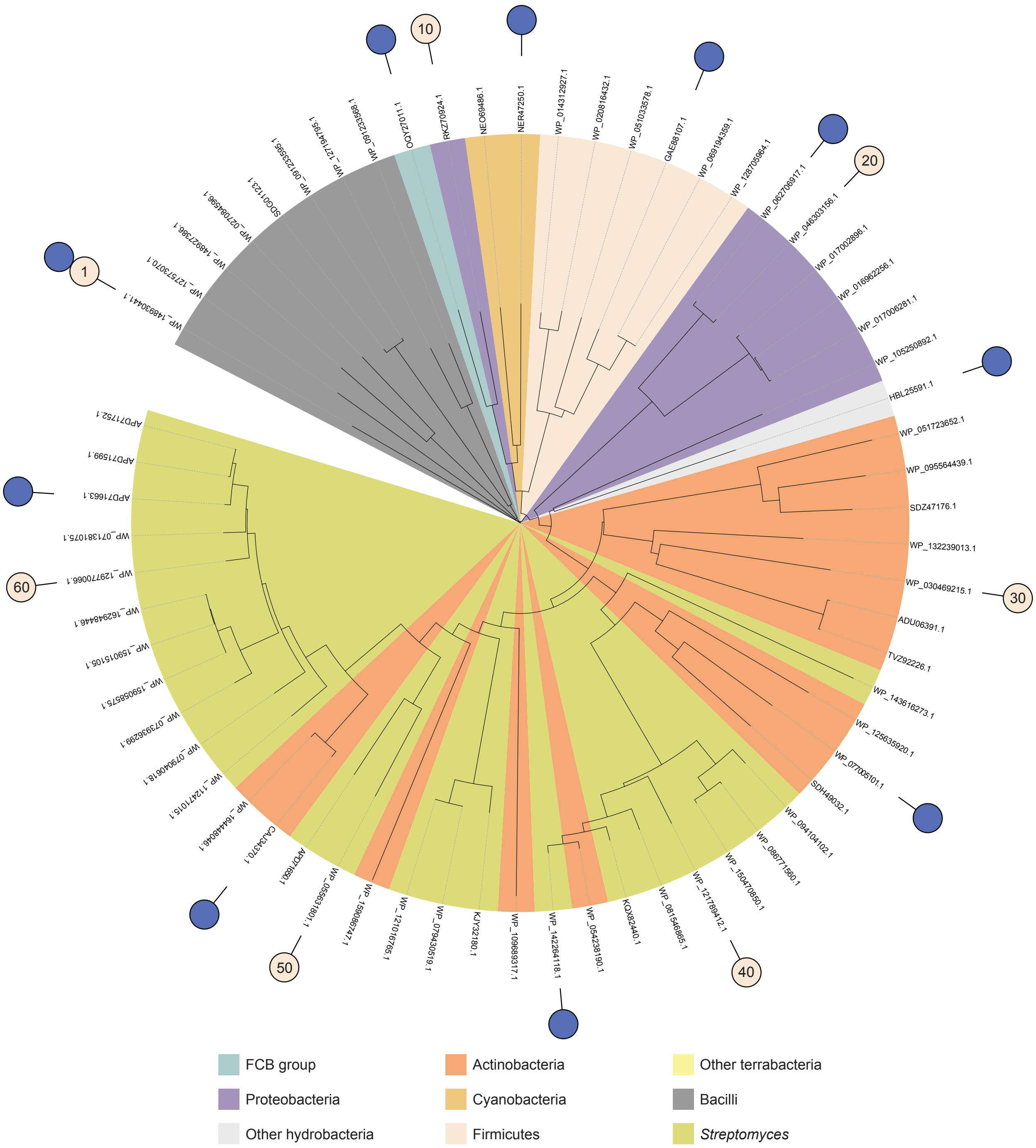


Fig. S11. Phylogenetic tree of family 3 ($A_2M_{S(S)}A_3$) interrupted A domains (n = 64). The interrupted A domains used for multiple sequence alignment (Fig. S16) are indicated by blue balloons. The interrupted A domains are numbered clockwise from the top left corner of the tree. Detailed information including these numbers, NCBI IDs, ranges of sequences used for the analysis, and organisms is shown in Table S4.

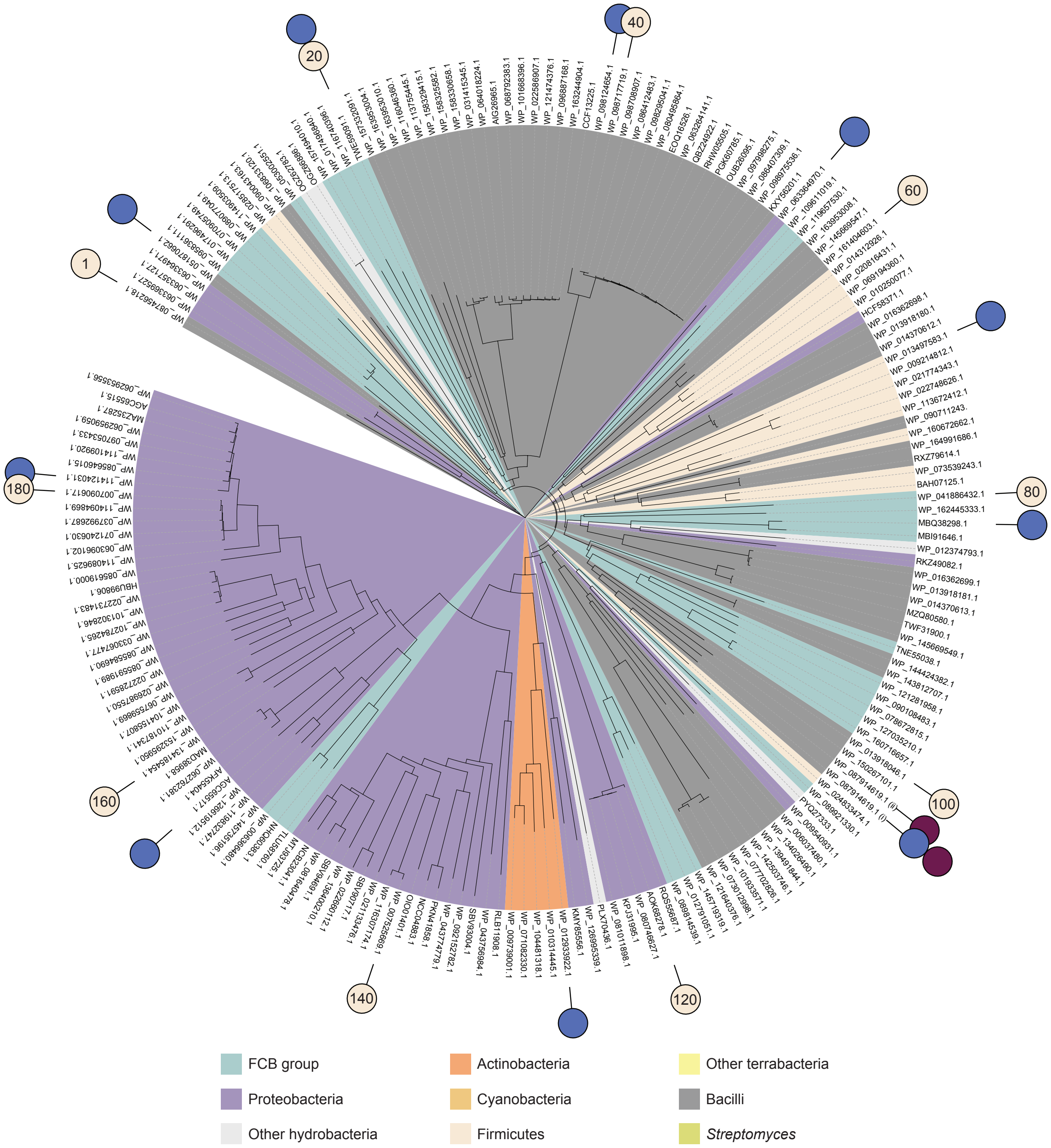


Fig. S12. Phylogenetic tree of family 4 ($A_2M_bA_3$) interrupted A domains (n = 188). The interrupted A domains used for multiple sequence alignment (Fig. S17) are indicated by blue balloons. Proteins that contain two homologous interrupted A domains are indicated by burgundy balloons. The interrupted A domains are numbered clockwise from the top left corner of the tree. Detailed information including these numbers, NCBI IDs, ranges of sequences used for the analysis, and organisms is shown in Table S5.

Table S1. Detailed information of family 1 ($A_8M_bA_9$) interrupted A domains used in taxonomic and phylogenetic trees.

# in taxonomic tree	**# in phylogenetic tree**	**NCBI ID**	**Range of sequence**	**Organism**
1	323	WP_082362653.1	5789-6723	*Chondromyces crocatus*
2	322	PZN71168.1	5821-6723	*Candidatus Methyloumidiphilus alinensis*
3	438	WP_126683355.1	2539-3463	*Dyella* sp. 4M-K27
4	441	WP_091794484.1	6097-7032	*Lysobacter* sp. yr284
5	443	WP_056114187.1	6132-7071	*Lysobacter* sp. Root690
6	439	WP_078996483.1	6128-7037	*Lysobacter enzymogenes*
7	442	WP_123647242.1	6128-7037	*Lysobacter enzymogenes*
8	440	WP_138885208.1	844-1753	*Lysobacter enzymogenes*
9	447	WP_046656639.1	6119-7059	*Lysobacter capsici*
11	446	WP_057921676.1	6121-7061	*Lysobacter capsici*
12	444	WP_096414585.1	6116-7056	*Lysobacter capsici*
12	445	WP_153018869.1	2887-3827	*Lysobacter capsici*
13	453	CBW77042.1	456-1357	*Paraburkholderia rhizoxinica* HKI 454
14	454	WP_013435346.1	468-1369	*Paraburkholderia rhizoxinica*
15	448	WP_013428902.1	1018-1921	*Paraburkholderia rhizoxinica*
16	455	WP_013428902.1	3528-4461	*Paraburkholderia rhizoxinica*
17	451	WP_083703618.1	455-1357	*Burkholderia* sp. b14
18	449	WP_083703618.1	5136-6034	*Burkholderia* sp. b14
19	457	WP_083703618.1	7640-8574	*Burkholderia* sp. b14
20	456	SIT79025.1	4300-5215	*Burkholderia* sp. b13
21	452	WP_083706138.1	455-1357	*Burkholderia* sp. b13
22	450	WP_083706138.1	5136-6036	*Burkholderia* sp. b13
23	458	WP_083706138.1	7640-8574	*Burkholderia* sp. b13
24	488	NEQ07639.1	1-919	*Moorea* sp. SIO4E2
25	487	AAT12283.1	7-923	*Lyngbya majuscula*
26	326	KPI15380.1	4088-4997	Actinobacteria bacterium OK074
27	96	CAJ61194.1	458-1326	*Frankia alni* ACN14a
28	93	WP_163080922.1	1993-2861	*Actinospica acidiphila*
29	100	WP_163036707.1	431-1326	*Actinospica acidiphila*
30	44	WP_164448040.1	442-1360	*Verrucosispora* sp. CWR15
31	129	WP_164448040.1	1892-2796	*Verrucosispora* sp. CWR15
32	43	CAJ34375.1	452-1370	*Micromonospora* sp. ML1
33	128	CAJ34375.1	1902-2806	*Micromonospora* sp. ML1
34	115	RZT79745.1	5-859	*Micromonospora violae*
35	117	WP_130402397.1	1472-2362	*Micromonospora violae*
36	342	WP_130402397.1	3948-4805	*Micromonospora violae*
37	116	WP_130402395.1	437-1323	*Micromonospora violae*
38	77	WP_162561486.1	1211-2077	*Salinispora pacifica*
39	78	WP_155249657.1	87-953	*Salinispora tropica*
40	79	WP_155251178.1	87-953	*Salinispora tropica*
41	49	WP_028678148.1	479-1356	*Salinispora arenicola*
42	135	WP_028678148.1	1890-2787	*Salinispora arenicola*
43	21	WP_080518128.1	1012-1904	*Salinispora arenicola*
44	23	WP_080518128.1	5527-6383	*Salinispora arenicola*
45	20	ABW00331.1	1008-1900	*Salinispora arenicola* CNS-205
46	22	ABW00331.1	5523-6379	*Salinispora arenicola* CNS-205
48	30	WP_131738712.1	1572-2452	*Actinomadura roseirufa*
49	32	WP_084264373.1	1532-2457	*Actinomadura macra*
50	33	WP_131891786.1	1548-2453	*Actinomadura* sp. H3C3
51	83	AYW35146.1	463-1311	*Actinomadura* sp. XM-4-3
52	84	AYW35145.1	462-1311	*Actinomadura* sp. XM-4-3
53	489	WP_114398342.1	370-1305	*Marinitenerispora sediminis*
54	491	TDQ54293.1	29-917	*Actinorugispora endophytica*
55	490	AFO85453.1	418-1316	*Marinactinospora thermotolerans*
56	493	WP_150241795.1	27-917	*Nocardiopsis quinghaiensis*
57	492	WP_094803333.1	21-917	*Nocardiopsis dassonvillei*

58	494	WP_019610701.1	21-917	*Nocardiopsis* sp. CNS-639
59	519	WP_106674311.1	528-1501	*Streptosporangium nondiastaticum*
60	1	WP_049563538.1	460-1347	*Nonomuraea* sp. SBT364
61	167	AIW58892.1	15-868	*Nonomuraea* sp. MJM5123
62	165	AIW58892.1	2404-3252	*Nonomuraea* sp. MJM5123
63	2	AIW58892.1	4831-5651	*Nonomuraea* sp. MJM5123
64	3	AIW58892.1	7237-8057	*Nonomuraea* sp. MJM5123
65	166	AIW58892.1	9592-1044	*Nonomuraea* sp. MJM5123
66	10	AIW58892.1	11006-1184	*Nonomuraea* sp. MJM5123
67	358	WP_142245578.1	436-1357	*Mycobacterium* sp. IS-836
68	85	WP_138635688.1	462-1311	*Actinomadura geliboluensis*
68	357	WP_139333155.1	436-1357	*Mycobacterium* sp. SP-6446
69	356	RZL82462.1	1514-2467	*Rhodococcus* sp.
70	368	WP_051187284.1	4488-5428	*Nocardia tenerifensis*
71	355	WP_040749123.1	7680-8576	*Nocardia transvalensis*
72	366	WP_069162717.1	2542-3485	*Nocardia altamirensis*
73	364	WP_069162717.1	4010-4963	*Nocardia altamirensis*
74	365	WP_067992737.1	2541-3456	*Nocardia pseudobrasiliensis*
75	362	WP_067992737.1	3983-4913	*Nocardia pseudobrasiliensis*
76	367	BAO99154.1	2490-3430	*Nocardia brasiliensis*
77	363	BAO99154.1	3957-4913	*Nocardia brasiliensis*
78	360	WP_107655971.1	463-1371	*Nocardia suismassiliense*
79	359	WP_107655971.1	1932-2836	*Nocardia suismassiliense*
80	370	WP_107656873.1	5663-6601	*Nocardia suismassiliense*
81	371	GEM38702.1	5650-6586	*Nocardia ninae* NBRC 108245
82	361	WP_147137135.1	463-1371	*Nocardia ninae*
83	372	WP_147131469.1	5660-6596	*Nocardia ninae*
84	369	WP_135237296.1	5670-6606	*Nocardia* sp. CS682
85	484	WP_024804816.1	429-1349	*Nocardia* sp. BMG51109
86	353	MQY17564.1	3520-4383	*Nocardia* sp. RB20
87	354	MQY17564.1	7518-8460	*Nocardia* sp. RB20
88	503	WP_143342721.1	2435-3279	*Crossiella equi*
89	337	WP_075134705.1	431-1340	*Actinophytocola xinjiangensis*
90	321	WP_037070439.1	170-1055	*Pseudonocardia acaciae*
91	437	TDQ00487.1	427-1347	*Labedaea rhizosphaerae*
92	485	AHB38497.1	459-1369	*Goodfellowiella coeruleoviolacea*
93	76	WP_090104167.1	433-1331	*Lentzea jiangxiensis*
94	127	WP_090104167.1	1871-2752	*Lentzea jiangxiensis*
95	397	WP_054047018.1	1501-2427	*Alloactinosynnema* sp. L-07
96	398	WP_054047018.1	2961-3880	*Alloactinosynnema* sp. L-07
97	75	WP_051766501.1	499-1323	*Saccharothrix syringae*
98	126	WP_051766501.1	1854-2738	*Saccharothrix syringae*
99	483	WP_161784481.1	2100-3007	*Actinokineospora spheciospongiae*
100	482	PWW60409.1	2100-3007	*Actinokineospora mzabensis*
101	344	REH46924.1	2016-2849	*Kutzneria buriramensis*
102	468	EWM13735.1	1033-1935	*Kutzneria* sp. 744
103	289	WP_157494376.1	8-834	*Kutzneria* sp. 744
104	339	WP_157494376.1	1360-2231	*Kutzneria* sp. 744
105	350	WP_103056250.1	432-1330	*Actinoalloteichus* sp. AHMU CJ021
106	352	WP_103056250.1	3982-4852	*Actinoalloteichus* sp. AHMU CJ021
107	226	WP_103056250.1	5415-6283	*Actinoalloteichus* sp. AHMU CJ021
108	230	WP_103056250.1	6860-7675	*Actinoalloteichus* sp. AHMU CJ021
109	228	WP_103056250.1	8210-9083	*Actinoalloteichus* sp. AHMU CJ021
110	349	WP_051313660.1	432-1330	*Actinoalloteichus cyanogriseus*
111	351	WP_051313660.1	3982-4852	*Actinoalloteichus cyanogriseus*
112	225	WP_051313660.1	5415-6283	*Actinoalloteichus cyanogriseus*
113	229	WP_051313660.1	6860-7675	*Actinoalloteichus cyanogriseus*
114	227	WP_051313660.1	8210-9083	*Actinoalloteichus cyanogriseus*
115	123	WP_161269834.1	2328-3222	*Amycolatopsis rubida*

116	504	WP_094862777.1	2591-3489	*Amycolatopsis antarctica*
117	122	WP_095213498.1	2333-3227	*Amycolatopsis* sp. M39
118	291	WP_146073742.1	204-1078	*Amycolatopsis* sp. CA-126428
119	341	WP_146073742.1	1596-2452	*Amycolatopsis* sp. CA-126428
120	124	WP_093934113.1	3229-4079	*Amycolatopsis thailandensis*
121	486	WP_093934109.1	421-1334	*Amycolatopsis thailandensis*
122	125	WP_093934113.1	4607-5484	*Amycolatopsis thailandensis*
123	290	WP_086851675.1	188-1080	*Amycolatopsis kentuckyensis*
124	340	WP_086851675.1	1598-2457	*Amycolatopsis kentuckyensis*
125	288	WP_086851675.1	2979-3871	*Amycolatopsis kentuckyensis*
126	200	WP_052433231.1	2021-2915	*Streptacidiphilus carbonis*
127	189	WP_160161297.1	1985-2915	*Embleya hyalina*
128	188	WP_020552458.1	1556-2487	*Streptomyces scabrisporus*
129	18	WP_159028297.1	779-1716	*Kitasatospora* sp. MY 5-36
130	19	WP_117490357.1	2234-3171	*Kitasatospora xanthocidica*
131	286	KOG73294.1	264-1146	*Kitasatospora aureofaciens*
132	242	KOG73294.1	1698-2598	*Kitasatospora aureofaciens*
133	265	KOG73294.1	3147-4020	*Kitasatospora aureofaciens*
134	332	WP_043481235.1	5576-6476	*Kitasatospora aureofaciens*
135	194	WP_164317044.1	1991-2914	*Streptomyces aureoverticillatus*
136	209	WP_158714838.1	2021-2911	*Streptomyces erythrochromogenes*
137	331	WP_152886003.1	2293-3198	*Streptomyces adustus*
138	320	WP_150498200.1	3471-4360	*Streptomyces chartreusis*
139	461	WP_150474996.1	2566-3504	*Streptomyces viridosporus*
140	296	WP_149549639.1	1536-2439	*Streptomyces marokkonensis*
141	108	WP_146484703.1	462-1321	*Streptomyces* sp. SSL-25
142	121	WP_145870934.1	1841-2726	*Streptomyces capillispiralis*
143	175	WP_141202784.1	467-1329	*Streptomyces griseorubiginosus*
144	164	WP_128461857.1	718-1619	*Streptomyces albidoflavus*
145	527	WP_126913047.1	4-891	*Streptomyces luteoverticillatus*
146	304	WP_125044964.1	14-899	*Streptomyces chrestomyceticus*
147	534	WP_121797110.1	46-953	*Streptomyces griseocarneus*
148	533	WP_102919561.1	28-938	*Streptomyces eurocidicus*
149	532	WP_095582267.1	28-930	*Streptomyces albireticuli*
150	333	WP_072655946.1	5684-6584	*Streptomyces viridifaciens*
151	113	WP_069778197.1	433-1327	*Streptomyces fodineus*
152	523	WP_065966397.1	459-1408	*Streptomyces sparsogenes*
153	177	WP_062148207.1	431-1332	*Streptomyces curacoi*
154	174	WP_052809011.1	463-1329	*Streptomyces natalensis*
155	297	WP_052707349.1	4589-5526	*Streptomyces rubellomurinus*
156	509	WP_052701242.1	4643-5568	*Streptomyces iranensis*
157	417	WP_043386495.1	1-875	*Streptomyces luteus*
158	480	WP_024264077.1	899-1841	*Streptomyces filamentosus*
159	475	WP_020123190.1	954-1868	*Streptomyces canus*
160	512	SEC95068.1	1-749	*Streptomyces melanosporofaciens*
161	520	RKN77326.1	427-1384	*Streptomyces klenkii*
162	210	OLZ57717.1	1990-2912	*Streptomyces amritsarensis*
163	98	MQL62982.1	431-1326	*Streptomyces vinaceus*
164	86	KJY37884.1	472-1327	*Streptomyces katrae*
165	306	KAA6213987.1	3-892	*Streptomyces albofaciens* JCM 4342
166	535	BAP27942.1	32-937	*Streptomyces blastmyceticus*
167	198	AXK36897.1	25-770	*Streptomyces armeniacus*
168	528	ANW17099.1	20-940	*Streptomyces clavuligerus*
169	94	ANP50040.1	421-1337	*Streptomyces griseochromogenes*
170	517	AJV88377.1	2349-3282	*Streptomyces drozdowiczii*
171	316	AAK81826.1	3464-4363	*Streptomyces lavendulae*
172	394	WP_051905993.1	1559-2423	*Streptomyces*
173	500	WP_051905993.1	2989-3882	*Streptomyces*
174	390	WP_157851312.1	1547-2445	*Streptomyces*

175	496	WP_157851312.1	3012-3909	*Streptomyces*
176	315	WP_158819050.1	1383-2284	*Streptomyces*
177	312	WP_158706459.1	14-904	*Streptomyces*
178	434	WP_150494549.1	214-1135	*Streptomyces kanamyceticus*
179	436	WP_150494549.1	4773-5519	*Streptomyces kanamyceticus*
180	35	WP_144321626.1	2016-2899	*Streptomyces spectabilis*
181	36	WP_150513690.1	2016-2899	*Streptomyces spectabilis*
182	292	WP_135341522.1	1312-2220	*Streptomyces palmae*
183	522	WP_135337084.1	458-1397	*Streptomyces palmae*
184	536	WP_116214566.1	32-940	*Streptomyces olivoreticuli*
185	294	WP_162890034.1	1471-2405	*Streptomyces olivoreticuli*
186	45	WP_112471241.1	441-1370	*Streptomyces triticisoli*
187	131	WP_112471241.1	1902-2803	*Streptomyces triticisoli*
188	405	WP_104880054.1	1492-2434	*Streptomyces dengpaensis*
189	427	WP_104880054.1	2996-3889	*Streptomyces dengpaensis*
190	409	WP_085563152.1	1491-2434	*Streptomyces griseofuscus*
191	430	WP_085563152.1	3002-3892	*Streptomyces griseofuscus*
192	401	WP_078636602.1	1483-2416	*Streptomyces antibioticus*
193	424	WP_078636602.1	2981-3880	*Streptomyces antibioticus*
194	16	WP_071963410.1	37-814	*Streptomyces cinnamoneus*
195	477	WP_079274001.1	897-1822	*Streptomyces cinnamoneus*
196	348	WP_070194809.1	459-1386	*Streptomyces oceani*
197	338	WP_070194809.1	4044-4976	*Streptomyces oceani*
198	73	WP_069926458.1	478-1366	*Streptomyces agglomeratus*
199	157	WP_069926458.1	1900-2798	*Streptomyces agglomeratus*
200	403	WP_064726364.1	1486-2425	*Streptomyces parvulus*
201	414	WP_064726364.1	2959-3885	*Streptomyces parvulus*
202	112	WP_055611557.1	462-1323	*Streptomyces phaeochromogenes*
203	109	WP_079052962.1	5755-6650	*Streptomyces phaeochromogenes*
204	31	WP_014173966.1	1555-2442	*Streptomyces bingchenggensis*
205	524	WP_014174809.1	485-1437	*Streptomyces bingchenggensis*
206	410	QAR15117.1	1491-2434	*Streptomyces costaricanus*
207	431	QAR15117.1	3002-3889	*Streptomyces costaricanus*
208	207	KOG47714.1	2033-2923	*Streptomyces virginiae*
209	218	WP_078905946.1	2025-2914	*Streptomyces virginiae*
210	71	GFH74034.1	476-1364	*Streptomyces diastaticus* subsp. *diastaticus*
211	163	GFH74034.1	1896-2797	*Streptomyces diastaticus* subsp. *diastaticus*
212	52	BAH04161.1	497-1383	*Streptomyces triostinicus*
213	144	BAH04161.1	1915-2816	*Streptomyces triostinicus*
214	433	ATL25759.1	25-964	*Streptomyces formicae*
215	435	ATL25759.1	4616-5557	*Streptomyces formicae*
216	53	AET98905.1	497-1376	*Streptomyces griseovariabilis* subsp. *bandungensis*
217	143	AET98905.1	1911-2809	*Streptomyces griseovariabilis* subsp. *bandungensis*
218	510	AEM82824.1	4615-5539	*Streptomyces violaceusniger* Tu 4113
219	511	WP_050993662.1	4644-5568	*Streptomyces violaceusniger*
220	56	WP_094219836.1	492-1365	*Streptomyces diastatochromogenes*
221	138	WP_094219836.1	1898-2797	*Streptomyces diastatochromogenes*
222	146	WP_143684103.1	906-1804	*Streptomyces diastatochromogenes*
223	231	WP_053926294.1	2236-3092	*Streptomyces chattanoogensis*
224	224	WP_053926294.1	3655-4551	*Streptomyces chattanoogensis*
225	232	WP_157878696.1	157-1012	*Streptomyces chattanoogensis*
226	182	WP_051810659.1	5-907	*Streptomyces ruber*
227	38	WP_051810659.1	1482-2368	*Streptomyces ruber*
228	345	WP_051834893.1	674-1572	*Streptomyces ruber*
229	515	GDY47057.1	4430-5354	*Streptomyces antimycoticus*
230	513	WP_086709773.1	4644-5568	*Streptomyces antimycoticus*
231	516	WP_161559306.1	4644-5568	*Streptomyces antimycoticus*

232	336	AXQ53534.1	479-1389	*Streptomyces koyangensis*
233	335	AXQ53534.1	1955-2855	*Streptomyces koyangensis*
234	334	AXQ53535.1	3637-4555	*Streptomyces koyangensis*
235	374	AGN74885.1	436-1359	*Streptomyces griseoviridis*
236	471	AGN74885.1	1909-2826	*Streptomyces griseoviridis*
237	466	AGN74885.1	4472-5369	*Streptomyces griseoviridis*
238	273	KUJ35543.1	432-1280	*Streptomyces rimosus* subsp. *rimosus*
238	479	WP_164342133.1	937-1847	*Streptomyces halstedii*
239	399	ADG27359.1	1522-2431	*Streptomyces anulatus*
239	385	KUJ35543.1	4974-5869	*Streptomyces rimosus* subsp. *rimosus*
240	420	ADG27359.1	2996-3892	*Streptomyces anulatus*
240	277	KOT48540.1	432-1280	*Streptomyces rimosus* subsp. *rimosus*
241	384	KOT48540.1	4942-5867	*Streptomyces rimosus* subsp. *rimosus*
241	34	WP_037685472.1	2021-2902	*Streptomyces aureocirculatus*
242	272	ELQ81235.1	479-1327	*Streptomyces rimosus* subsp. *rimosus* ATCC 10970
242	185	WP_051852243.1	44-936	*Streptomyces aureocirculatus*
243	388	ELQ81235.1	5021-5916	*Streptomyces rimosus* subsp. *rimosus* ATCC 10970
243	40	WP_051852243.1	1495-2408	*Streptomyces aureocirculatus*
244	190	WP_159398248.1	2267-3160	*Streptomyces aureocirculatus*
245	63	GBP98747.1	429-1345	*Streptomyces spongiicola*
246	154	GBP98747.1	1878-2779	*Streptomyces spongiicola*
247	64	WP_109296560.1	778-1416	*Streptomyces spongiicola*
248	155	WP_109296560.1	1949-2850	*Streptomyces spongiicola*
249	69	BAE98156.1	458-1366	*Streptomyces lasaliensis*
250	159	BAE98156.1	1898-2799	*Streptomyces lasaliensis*
251	70	WP_137311393.1	458-1366	*Streptomyces lasaliensis*
252	160	WP_137311393.1	1898-2799	*Streptomyces lasaliensis*
253	255	WP_069992614.1	437-1319	*Streptomyces qinglanensis*
254	378	WP_069992614.1	5062-5955	*Streptomyces qinglanensis*
255	252	WP_079135167.1	549-1439	*Streptomyces qinglanensis*
256	233	WP_079135167.1	2371-3288	*Streptomyces qinglanensis*
257	249	WP_162637645.1	685-1572	*Streptomyces qinglanensis*
258	168	SFS59173.1	15-863	*Streptomyces harbinensis*
259	6	SFS59173.1	3448-4268	*Streptomyces harbinensis*
260	8	SFS59173.1	5836-6644	*Streptomyces harbinensis*
261	173	SFS59173.1	8197-9004	*Streptomyces harbinensis*
262	11	SFS59173.1	9568-1040	*Streptomyces harbinensis*
263	259	GFE33280.1	560-1420	*Streptomyces libani* subsp. *rufus*
264	235	GFE33280.1	1985-2875	*Streptomyces libani* subsp. *rufus*
265	261	GFE33280.1	3407-4295	*Streptomyces libani* subsp. *rufus*
266	260	GFE33279.1	436-1333	*Streptomyces libani* subsp. *rufus*
267	380	GFE33279.1	5036-5937	*Streptomyces libani* subsp. *rufus*
268	28	BBA20967.1	1032-1958	*Streptomyces atratus*
269	29	BBA20967.1	4604-5530	*Streptomyces atratus*
270	26	WP_114243093.1	1068-1994	*Streptomyces atratus*
271	27	WP_114243093.1	4664-5557	*Streptomyces atratus*
272	329	WP_114248409.1	5644-6542	*Streptomyces atratus*
273	107	AZK92477.1	433-1322	*Streptomyces tsukubensis*
274	114	OON72765.1	5-862	*Streptomyces tsukubensis*
275	119	WP_143621520.1	269-1146	*Streptomyces tsukubensis*
276	120	WP_152330836.1	447-1324	*Streptomyces tsukubensis*
277	118	WP_152330837.1	1494-2390	*Streptomyces tsukubensis*
278	343	WP_152330837.1	3956-4858	*Streptomyces tsukubensis*
279	295	AYN31633.1	634-1397	*Streptomyces albus*
280	521	AYN31633.1	3029-3914	*Streptomyces albus*
281	275	WP_060732907.1	462-1326	*Streptomyces albus*
282	382	WP_060732907.1	5011-5915	*Streptomyces albus*

283	282	KWT56928.1	550-1432	*Streptomyces albus* subsp. *albus*
284	313	KWT56605.1	3-889	*Streptomyces albus* subsp. *albus*
285	180	KUF16419.1	2015-2887	*Streptomyces silvensis*
286	81	KUF16419.1	4980-5817	*Streptomyces silvensis*
287	191	KUF18083.1	4733-5654	*Streptomyces silvensis*
288	65	WP_058846511.1	478-1385	*Streptomyces silvensis*
289	151	WP_058846511.1	1919-2819	*Streptomyces silvensis*
290	186	WP_058847149.1	44-971	*Streptomyces silvensis*
291	41	WP_058847149.1	1560-2444	*Streptomyces silvensis*
292	395	CCO61885.1	1506-2433	*Streptomyces iakyrus*
293	391	CCO61888.1	1528-2437	*Streptomyces iakyrus*
294	501	CCO61885.1	2980-3886	*Streptomyces iakyrus*
295	497	CCO61888.1	3004-3901	*Streptomyces iakyrus*
296	396	WP_051814369.1	1506-2421	*Streptomyces iakyrus*
297	502	WP_051814369.1	2974-3880	*Streptomyces iakyrus*
298	392	WP_157852672.1	1551-2439	*Streptomyces iakyrus*
299	468	WP_157852672.1	2980-3903	*Streptomyces iakyrus*
300	196	ARX88592.1	325-1179	*Streptomyces alboflavus*
301	346	ARX82090.1	1992-2890	*Streptomyces alboflavus*
302	347	AZL49199.1	1992-2890	*Streptomyces alboflavus*
303	195	WP_157854078.1	1599-2490	*Streptomyces alboflavus*
304	183	WP_159399696.1	1-852	*Streptomyces alboflavus*
305	184	WP_161990268.1	18-941	*Streptomyces alboflavus*
306	197	WP_161990269.1	1588-2479	*Streptomyces alboflavus*
307	37	WP_159399696.1	1395-2313	*Streptomyces alboflavus*
308	39	WP_161990268.1	1484-2402	*Streptomyces alboflavus*
309	238	WP_030372174.1	1921-2821	*Streptomyces rimosus*
310	281	WP_030637038.1	550-1432	*Streptomyces rimosus*
311	266	WP_030372174.1	3370-4243	*Streptomyces rimosus*
312	283	WP_030650869.1	550-1432	*Streptomyces rimosus*
313	267	WP_030637038.1	3431-4304	*Streptomyces rimosus*
314	240	WP_030637038.1	1984-2882	*Streptomyces rimosus*
315	239	WP_030650869.1	1984-2882	*Streptomyces rimosus*
316	276	WP_031023180.1	479-1327	*Streptomyces rimosus*
317	263	WP_030650869.1	3431-4304	*Streptomyces rimosus*
318	383	WP_031023180.1	4991-5916	*Streptomyces rimosus*
319	285	WP_050509459.1	492-1374	*Streptomyces rimosus*
320	243	WP_050509459.1	1926-2826	*Streptomyces rimosus*
321	307	WP_050515166.1	20-907	*Streptomyces rimosus*
322	270	WP_050509459.1	3375-4248	*Streptomyces rimosus*
323	310	WP_051731503.1	3-889	*Streptomyces rimosus*
324	305	WP_125054248.1	16-901	*Streptomyces rimosus*
325	280	KOT48558.1	264-1146	*Streptomyces rimosus* subsp. *rimosus*
326	241	KOT48558.1	1698-2596	*Streptomyces rimosus* subsp. *rimosus*
327	264	KOT48558.1	3145-4018	*Streptomyces rimosus* subsp. *rimosus*
334	298	WP_164395145.1	10-773	*Streptomyces* sp. PRKS01-65
335	423	WP_164292326.1	1533-2456	*Streptomyces* sp. SID13588
336	459	WP_164254132.1	2580-3512	*Streptomyces* sp. S4.7
337	529	WP_161362407.1	15-943	*Streptomyces* sp. SID3343
338	412	WP_161258436.1	2-761	*Streptomyces* sp. SID1034
339	176	WP_161254677.1	467-1328	*Streptomyces* sp. SID685
340	526	WP_161227716.1	30-813	*Streptomyces* sp. SID8352
341	518	WP_160330192.1	425-1420	*Streptomyces* sp. MBT76
342	203	WP_159395658.1	2029-2919	*Streptomyces* sp. 3211
343	311	WP_159393770.1	14-900	*Streptomyces* sp. NRRL F-5755
344	206	WP_159061980.1	2031-2921	*Streptomyces* sp. WM6368
345	213	WP_159050709.1	1987-2909	*Streptomyces* sp. IMTB 1903
346	220	WP_159042334.1	2022-2911	*Streptomyces* sp. H021
347	216	WP_159040776.1	2021-2911	*Streptomyces* sp. XY593

348	214	WP_159040058.1	2024-2914	*Streptomyces* sp. XY533
349	201	WP_158754520.1	2036-2926	*Streptomyces* sp. NRRL F-2580
350	204	WP_158721088.1	2030-2920	*Streptomyces* sp. NRRL S-241
351	531	WP_158719317.1	2-871	*Streptomyces* sp. NBRC 110035
352	208	WP_158717863.1	2019-2909	*Streptomyces* sp. NRRL F-2664
353	211	WP_158717201.1	1987-2909	*Streptomyces* sp. NRRL F-4474
354	104	WP_156698555.1	431-1326	*Streptomyces* sp. Z38
355	95	WP_151484725.1	430-1342	*Streptomyces* sp. TRM68295
356	508	WP_148836317.1	457-1346	*Streptomyces* sp. sk2.1
357	373	WP_143664384.1	550-1517	*Streptomyces* sp. NRRL B-24572
358	406	WP_143660143.1	43-984	*Streptomyces* sp. WZ.A104
359	505	WP_143645959.1	426-1347	*Streptomyces* sp. OF1
360	330	WP_141355982.1	5647-6545	*Streptomyces* sp. 6-11-2
361	91	WP_128128524.1	1993-2861	*Streptomyces* sp. BSE7F
362	514	WP_125756251.1	4644-5568	*Streptomyces* sp. WAC05858
363	319	WP_125527331.1	3452-4363	*Streptomyces* sp. WAC 05379
364	97	WP_125510942.1	461-1327	*Streptomyces* sp. WAC 01529
365	476	WP_125194795.1	945-1852	*Streptomyces* sp. RP5T
366	474	WP_123996388.1	909-1835	*Streptomyces* sp. Ag82_G5-5
367	103	WP_121721401.1	431-1326	*Streptomyces* sp. E2N171
368	101	WP_114872561.1	431-1326	*Streptomyces* sp. E1N211
369	247	WP_112482392.1	440-1340	*Streptomyces* sp. ST1015
370	246	WP_112472786.1	440-1340	*Streptomyces* sp. ST1020
371	478	WP_107454111.1	898-1825	*Streptomyces* sp. A244
372	416	WP_106415760.1	1520-2414	*Streptomyces* sp. DH-12
373	92	WP_102641651.1	1993-2861	*Streptomyces* sp. SMS_SU21
374	178	WP_101407771.1	461-1330	*Streptomyces* sp. OK885
375	327	WP_100570125.1	5672-6576	*Streptomyces* sp. CB01373
376	318	WP_093868309.1	3477-4364	*Streptomyces* sp. KS_5
377	317	WP_093748429.1	3479-4366	*Streptomyces* sp. PAN_FS17
378	525	WP_089510220.1	476-1434	*Streptomyces* sp. NBS 14/10
379	90	WP_078956252.1	1969-2865	*Streptomyces* sp. NRRL B-1140
380	314	WP_078620469.1	3466-4367	*Streptomyces* sp. NRRL B-11253
381	481	WP_075263799.1	899-1841	*Streptomyces* sp. Tue 6075
382	130	WP_071381079.1	328-1223	*Streptomyces* sp. MUSC 1
383	328	WP_063352349.1	5690-6604	*Streptomyces* sp. MJM8645
384	506	WP_062216587.1	463-1356	*Streptomyces* sp. NBRC 109706
385	530	WP_062189497.1	24-933	*Streptomyces* sp. AVP053U2
386	102	WP_048460343.1	431-1326	*Streptomyces* sp. HNS054
387	419	WP_037973485.1	6-721	*Streptomyces* sp. NRRL WC-3742
388	105	WP_030310062.1	462-1331	*Streptomyces* sp. NRRL F-6131
389	15	SCE46449.1	1903-2858	*Streptomyces* sp. DvalAA-43
390	99	RMI88892.1	421-1316	*Streptomyces* sp. ZS0098
391	460	PWJ02188.1	883-1821	*Streptomyces* sp. NWU49
392	293	PJN03925.1	1529-2468	*Streptomyces* sp. CB01201
393	426	PIB04224.1	117-1009	*Streptomyces* sp. HG99
394	110	OIJ84942.1	432-1327	*Streptomyces* sp. MUSC 14
395	199	NGO69793.1	1282-2188	*Streptomyces* sp. SB3404
396	14	NGO47505.1	398-1391	*Streptomyces* sp. YC419
397	422	NEA53188.1	1533-2456	*Streptomyces* sp. SID13666
398	111	NBM14293.1	454-1347	*Streptomyces* sp. GC420
399	215	KOV02256.1	2017-2907	*Streptomyces* sp. XY511
400	223	KOU96303.1	2015-2904	*Streptomyces* sp. XY58
401	221	KOU60123.1	2015-2906	*Streptomyces* sp. IGB124
402	222	KOU31424.1	2015-2904	*Streptomyces* sp. WM6373
403	219	KJY23011.1	2015-2904	*Streptomyces* sp. NRRL S-104
404	205	ARE74084.1	2043-2933	*Streptomyces* sp. Sge12
405	212	AQT74912.1	1990-2912	*Streptomyces* sp. fd1-xmd
406	106	AGZ15460.1	431-1324	*Streptomyces* sp. MK498-98F14

407	400	WP_111334107.1	1523-2432	unclassified *Streptomyces*
408	421	WP_111334107.1	3001-3889	unclassified *Streptomyces*
409	217	WP_159041193.1	2024-2914	unclassified *Streptomyces*
410	408	WP_162133098.1	1491-2434	*Streptomyces* sp. TR1341
411	429	WP_162133098.1	3002-3889	*Streptomyces* sp. TR1341
412	413	WP_161344372.1	1-756	*Streptomyces* sp. SID5606
413	415	WP_161343398.1	1-752	*Streptomyces* sp. SID5606
414	404	WP_160509793.1	1486-2428	*Streptomyces* sp. BA2
415	418	WP_160509793.1	2959-3883	*Streptomyces* sp. BA2
416	376	WP_159537319.1	434-1354	*Streptomyces* sp. Tu 3180
417	463	WP_159537319.1	1886-2835	*Streptomyces* sp. Tu 3180
418	145	WP_145884055.1	717-1618	*Streptomyces* sp. BK340
419	137	WP_145880020.1	946-1845	*Streptomyces* sp. BK340
420	411	WP_136099895.1	1491-2434	*Streptomyces* sp. S816
421	432	WP_136099895.1	3017-3889	*Streptomyces* sp. S816
422	24	WP_128818885.1	187-1090	*Streptomyces* sp. S063
423	25	WP_128818892.1	1102-1975	*Streptomyces* sp. S063
424	407	WP_128788708.1	1491-2434	*Streptomyces* sp. endophyte_N2
425	428	WP_128788708.1	3002-3889	*Streptomyces* sp. endophyte_N2
426	67	WP_124270194.1	480-1368	*Streptomyces* sp. ADI96-02
427	156	WP_124270194.1	1903-2802	*Streptomyces* sp. ADI96-02
428	179	WP_123108733.1	2013-2884	*Streptomyces* sp. NEAU-LD23
429	80	WP_123108733.1	4947-5814	*Streptomyces* sp. NEAU-LD23
430	51	WP_121406601.1	487-1378	*Streptomyces* sp. 57
431	142	WP_121406601.1	1910-2810	*Streptomyces* sp. 57
432	72	WP_113687522.1	478-1366	*Streptomyces* sp. PT12
433	162	WP_113687522.1	1898-2799	*Streptomyces* sp. PT12
434	58	WP_097251926.1	494-1367	*Streptomyces* sp. Ag109_G2-15
435	149	WP_097251926.1	1900-2798	*Streptomyces* sp. Ag109_G2-15
436	462	WP_095748046.1	602-1546	*Streptomyces* sp. SA15
437	470	WP_095748046.1	3163-4098	*Streptomyces* sp. SA15
438	68	WP_093807334.1	476-1365	*Streptomyces* sp. Ncost-T6T-1
439	161	WP_093807334.1	1897-2798	*Streptomyces* sp. Ncost-T6T-1
440	74	WP_093802267.1	478-1366	*Streptomyces* sp. Wb2n-11
441	158	WP_093802267.1	1898-2798	*Streptomyces* sp. Wb2n-11
442	402	WP_085206047.1	1483-2421	*Streptomyces* sp. Amel2xC10
443	425	WP_085206047.1	2983-3882	*Streptomyces* sp. Amel2xC10
444	17	WP_074004725.1	2240-3177	*Streptomyces* sp. CB02056
445	473	WP_074003396.1	896-1821	*Streptomyces* sp. CB02056
446	375	WP_073824735.1	436-1372	*Streptomyces* sp. CB02130
447	465	WP_073824735.1	1902-2853	*Streptomyces* sp. CB02130
448	464	WP_073724445.1	832-1783	*Streptomyces* sp. TSRI0281
449	469	WP_073724445.1	3387-4065	*Streptomyces* sp. TSRI0281
450	254	WP_027761680.1	1-812	*Streptomyces* sp. CNT318
451	250	WP_027761023.1	751-1639	*Streptomyces* sp. CNT318
452	472	SCK10353.1	104-1047	*Streptomyces* sp. WMMB 714
453	467	WP_045867283.1	137-1053	*Streptomyces* sp. WMMB 714
454	62	RNL73105.1	107-918	*Streptomyces* sp. I6
455	153	RNL73105.1	1451-2352	*Streptomyces* sp. I6
456	300	RKR61622.1	43-950	*Streptomyces* sp. 3212.4
457	302	RKR61622.1	1518-2421	*Streptomyces* sp. 3212.4
458	299	RED73221.1	43-950	*Streptomyces* sp. 67
459	301	RED73221.1	1518-2421	*Streptomyces* sp. 67
460	50	BAI63289.1	491-1386	*Streptomyces* sp. SNA15896
461	136	BAI63289.1	1917-2818	*Streptomyces* sp. SNA15896
462	88	WP_123970979.1	435-1323	*Streptomyces* sp. TLI_185
463	89	WP_123970979.1	2335-3228	*Streptomyces* sp. TLI_185
464	87	WP_123970980.1	475-1322	*Streptomyces* sp. TLI_185
465	57	WP_116130095.1	492-1365	*Streptomyces* sp. Ag82_O1-9

466	59	WP_116128926.1	494-1367	*Streptomyces* sp. Ag82_O1-9
467	150	WP_116128926.1	1900-2798	*Streptomyces* sp. Ag82_O1-9
468	141	WP_116130095.1	1898-2797	*Streptomyces* sp. Ag82_O1-9
469	55	WP_101372788.1	492-1365	*Streptomyces* sp. 59
470	61	WP_101374734.1	494-1367	*Streptomyces* sp. 59
471	147	WP_101374734.1	1900-2798	*Streptomyces* sp. 59
472	139	WP_101372788.1	1898-2797	*Streptomyces* sp. 59
473	393	WP_067259374.1	266-1153	*Streptomyces* sp. DSM 15324
474	499	WP_067259374.1	1719-2617	*Streptomyces* sp. DSM 15324
475	389	WP_159056341.1	274-1172	*Streptomyces* sp. DSM 15324
476	495	WP_159056341.1	1739-2638	*Streptomyces* sp. DSM 15324
477	54	PKV54526.1	143-1016	*Streptomyces* sp. 73
478	140	PKV54526.1	1549-2448	*Streptomyces* sp. 73
479	60	WP_101418611.1	494-1367	*Streptomyces* sp. 73
480	148	WP_101418611.1	1900-2798	*Streptomyces* sp. 73
481	7	WP_020699877.1	436-1244	*Streptomyces* sp. AA0539
482	171	WP_020699877.1	2810-3617	*Streptomyces* sp. AA0539
483	12	WP_020699877.1	4181-5018	*Streptomyces* sp. AA0539
484	169	WP_158687260.1	1-830	*Streptomyces* sp. AA0539
485	4	WP_158687260.1	3416-4235	*Streptomyces* sp. AA0539
486	256	WP_020699829.1	437-1320	*Streptomyces* sp. AA1529
487	377	WP_020699829.1	5030-5923	*Streptomyces* sp. AA1529
488	253	WP_106428956.1	543-1431	*Streptomyces* sp. AA1529
489	234	WP_106428956.1	1967-2884	*Streptomyces* sp. AA1529
490	251	WP_106428956.1	3417-4305	*Streptomyces* sp. AA1529
491	248	AJC59844.1	485-1343	*Streptomyces* sp. 769
492	379	AJC59844.1	3964-4894	*Streptomyces* sp. 769
493	258	WP_039638096.1	546-1440	*Streptomyces* sp. 769
494	236	WP_039638096.1	2006-2886	*Streptomyces* sp. 769
495	257	WP_039638096.1	3418-4304	*Streptomyces* sp. 769
496	271	WP_125530724.1	473-1321	*Streptomyces* sp. WAC 06725
497	309	WP_125533652.1	3-889	*Streptomyces* sp. WAC 06725
498	386	WP_125530724.1	5015-5910	*Streptomyces* sp. WAC 06725
499	287	WP_125530723.1	550-1432	*Streptomyces* sp. WAC 06725
500	245	WP_125530723.1	1984-2884	*Streptomyces* sp. WAC 06725
501	269	WP_125530723.1	3433-4306	*Streptomyces* sp. WAC 06725
502	274	WP_125520807.1	479-1327	*Streptomyces* sp. WAC 06783
503	284	WP_125520806.1	550-1432	*Streptomyces* sp. WAC 06783
504	244	WP_125520806.1	1984-2884	*Streptomyces* sp. WAC 06783
505	387	WP_125520807.1	5021-5916	*Streptomyces* sp. WAC 06783
506	268	WP_125520806.1	3433-4306	*Streptomyces* sp. WAC 06783
507	308	WP_125519621.1	3-889	*Streptomyces* sp. WAC 06783
508	187	WP_157168700.1	44-955	*Streptomyces* sp. p1417
509	66	WP_157168497.1	478-1385	*Streptomyces* sp. p1417
510	193	WP_157165036.1	2045-2936	*Streptomyces* sp. p1417
511	42	WP_157168700.1	1570-2454	*Streptomyces* sp. p1417
512	181	WP_157166111.1	2027-2899	*Streptomyces* sp. p1417
513	82	WP_157166111.1	4992-5829	*Streptomyces* sp. p1417
514	152	WP_157168497.1	1919-2819	*Streptomyces* sp. p1417
515	192	WP_157165564.1	2289-3169	*Streptomyces* sp. p1417
516	324	OKH99041.1	4113-5033	*Streptomyces* sp. CB02923
517	325	OKH99041.1	7115-8028	*Streptomyces* sp. CB02923
518	279	WP_073758157.1	562-1432	*Streptomyces* sp. CB02923
519	237	WP_073758157.1	1993-2885	*Streptomyces* sp. CB02923
520	278	WP_073758160.1	437-1329	*Streptomyces* sp. CB02923
521	303	WP_073758123.1	14-903	*Streptomyces* sp. CB02923
522	381	WP_073758160.1	5000-5912	*Streptomyces* sp. CB02923
523	262	WP_073758157.1	3434-4316	*Streptomyces* sp. CB02923
524	46	APD71755.1	474-1370	*Streptomyces* sp.

525	134	APD71755.1	1902-2803	*Streptomyces* sp.
526	202	APD71797.1	2034-2924	*Streptomyces* sp.
527	507	APD71849.1	457-1348	*Streptomyces* sp.
528	48	APD71601.1	474-1370	*Streptomyces* sp.
529	47	APD71665.1	474-1370	*Streptomyces* sp.
530	132	APD71665.1	1902-2803	*Streptomyces* sp.
531	133	APD71601.1	1902-2803	*Streptomyces* sp.
532	170	QGA70148.1	15-859	*Streptomyces* sp.
533	5	QGA70148.1	3445-4264	*Streptomyces* sp.
534	9	QGA70148.1	5832-6640	*Streptomyces* sp.
535	172	QGA70148.1	8206-9013	*Streptomyces* sp.
536	13	QGA70148.1	9577-1041	*Streptomyces* sp.

Table S2. Detailed information of family 2a ($A_8M_{s(\mathit{O},\ \mathrm{Ser/Thr})}A_9$) interrupted A domains used in taxonomic and phylogenetic trees.

# in taxonomic tree	**# in phylogenetic tree**	**NCBI ID**	**Range of sequence**	**Organism**
1	108	BAP05589.1	21-867	uncultured *Candidatus Entotheonella* sp.
2	124	PWT88386.1	464-1334	Acidobacteria bacterium
3	81	WP_158943751.1	1874-2718	*Granulicella* sp. S190
4	80	WP_158943749.1	1485-2338	*Granulicella* sp. S190
5	135	PYP88141.1	1067-1927	Blastocatellia bacterium AA13
6	113	HEM09479.1	436-1292	Blastocatellia bacterium
7	116	WP_121286847.1	505-1373	*Pedobacter alluvionis*
8	115	WP_090108443.1	500-1356	*Chitinophaga* sp. CF118
9	114	AGC75282.1	2016-2902	*Nonlabens dokdonensis* DSW-6
10	82	MBQ39375.1	987-1857	Gemmatimonadaceae bacterium
11	134	PYP45280.1	4-861	Gemmatimonadetes bacterium
12	131	PYP02771.1	12-845	Gemmatimonadetes bacterium
13	130	PYP36040.1	41-862	Gemmatimonadetes bacterium
14	132	PYO96327.1	8-840	Gemmatimonadetes bacterium
15	3	WP_145753413.1	3173-4030	*Nitrospirillum amazonense*
16	112	WP_085861689.1	506-1328	unclassified *Rhizobium*
17	85	ARN83985.1	1581-2442	*Methylocystis bryophila*
18	88	WP_144410390.1	475-1306	*Chromobacterium vaccinii*
19	89	WP_156396907.1	1359-2238	*Noviherbaspirillum* sp. Root189
20	87	KNE75165.1	1502-2328	*Candidatus Burkholderia crenata*
21	100	WP_124543974.1	1635-2521	*Geomonas soli*
22	101	WP_124543135.1	1576-2441	*Geomonas soli*
23	136	RLC19153.1	407-1268	Deltaproteobacteria bacterium
24	125	OGQ77554.1	1056-1925	Deltaproteobacteria bacterium RIFCSPLOWO2_12_FULL_60_19
25	83	AGP35404.1	1-799	*Sorangium cellulosum* So0157-2
26	120	WP_104981151.1	410-1276	*Sorangium cellulosum*
27	111	KYF70742.1	1-612	*Sorangium cellulosum*
28	126	RYZ41765.1	4302-5166	Myxococcaceae bacterium
29	122	WP_120624088.1	403-1264	*Corallococcus* sp. CA040B
30	121	WP_120538527.1	403-1264	*Corallococcus* sp. CA054A
31	110	WP_120610799.1	450-1316	*Corallococcus* sp. CA053C
32	128	WP_120585253.1	1729-2593	*Corallococcus* sp. CA031C
33	127	WP_120536416.1	1721-2585	*Corallococcus* sp. CA047B
34	137	OQY50160.1	17-869	*Beggiatoa* sp. 4572_84
35	117	HAI15718.1	418-1277	Gammaproteobacteria bacterium
36	140	RKZ52202.1	6-841	Gammaproteobacteria bacterium
37	141	RKZ67573.1	6-841	Gammaproteobacteria bacterium
38	139	RKZ73225.1	14-838	Gammaproteobacteria bacterium
39	138	RKZ79147.1	15-856	Gammaproteobacteria bacterium
40	86	PMQ04698.1	1563-2443	*Dyella* sp. AD56
41	102	WP_156117283.1	1588-2441	*Xanthomonas sacchari*

42	118	KER88890.1	351-1210	*Xanthomonas arboricola pv. celebensis*
43	119	WP_078995394.1	423-1282	*Lysobacter enzymogenes*
44	103	WP_056114225.1	1579-2445	*Lysobacter* sp. Root690
45	105	WP_052756230.1	1584-2444	*Lysobacter capsici*
46	104	WP_051547367.1	1584-2444	*Lysobacter capsici*
47	106	ATE72041.1	1584-2444	*Lysobacter capsici*
48	107	WP_057921654.1	1584-2444	*Lysobacter capsici*
49	84	WP_074112458.1	501-1373	*Paenibacillus* sp. P46E
50	133	OLC28372.1	5-839	Armatimonadetes bacterium 13_1_40CM_64_14
51	109	WP_126628724.1	1483-2349	*Dictyobacter alpinus*
52	129	MPZ14326.1	15-885	Chloroflexi bacterium
53	123	ABX04522.1	434-1289	*Herpetosiphon aurantiacus* DSM 785
54	149	WP_144865672.1	3788-4647	*Hyella patelloides*
55	148	AXN93625.1	444-1334	*Symplocastrum muelleri* NIVA-CYA 644
56	147	HBB32138.1	459-1345	Cyanobacteria bacterium UBA9273
57	146	HAX75965.1	1072-1938	Cyanobacteria bacterium UBA11372
58	145	WP_096684329.1	499-1379	unclassified *Calothrix*
59	143	WP_096563841.1	429-1307	*Scytonema* sp. NIES-4073
60	144	WP_089130967.1	512-1381	*Tolypothrix* sp. NIES-4075
61	142	WP_087539157.1	429-1307	*Nostocales cyanobacterium* HT-58-2
62	53	WP_163510266.1	20-820	*Fodinicola* sp. GKU 173
63	19	WP_132166530.1	461-1294	*Kribbella antibiotica*
64	18	WP_132166531.1	2956-3782	*Kribbella antibiotica*
65	29	WP_143194259.1	3-845	*Micromonospora* sp. CB01531
66	41	KUJ44924.1	1556-2403	*Verrucosispora maris*
67	40	WP_102657348.1	1556-2403	*Verrucosispora* sp. ts21
68	42	RUL95035.1	1551-2398	*Verrucosispora* sp. FIM060022
69	39	AIS85382.1	1538-2241	*Verrucosispora* sp. MS100047
70	76	WP_157963902.1	440-1250	*Actinocorallia populi*
71	1	WP_160823767.1	20-848	*Actinomadura* sp. J1-007
72	94	WP_138643559.1	1842-2710	*Actinomadura* sp. 14C53
73	14	WP_117400866.1	1971-2825	*Actinomadura* sp. LHW52907
74	15	WP_111833234.1	1486-2345	*Actinomadura madurae*
75	17	WP_024933944.1	1485-2344	*Actinomadura madurae*
76	16	WP_083598171.1	1486-2345	*Actinomadura madurae*
77	12	WP_111831700.1	436-1265	*Actinomadura madurae*
78	11	WP_024934893.1	436-1265	*Actinomadura madurae*
79	55	MAU84043.1	436-1314	*Gordonia* sp.
80	56	WP_007299326.1	464-1317	*Rhodococcus imtechensis*
81	70	WP_157514135.1	436-1316	*Nocardia concava*
82	65	WP_084535066.1	432-1293	*Nocardia yamanashiensis*
83	67	WP_084528916.1	432-1310	*Nocardia crassostreae*
84	60	WP_084521547.1	430-1312	*Nocardia uniformis*
85	64	WP_067824987.1	409-1271	*Nocardia inohanensis*
86	66	WP_067564260.1	435-1306	*Nocardia acidivorans*
87	72	WP_045436417.1	436-1319	*Nocardia seriolae*
88	71	WP_155240917.1	436-1319	*Nocardia seriolae*
89	73	WP_155240776.1	436-1319	*Nocardia seriolae*
90	63	WP_143979776.1	432-1303	*Nocardia otitidiscaviarum*
91	62	WP_039809518.1	432-1303	*Nocardia otitidiscaviarum*
92	61	WP_043601641.1	432-1304	*Nocardia otitidiscaviarum*
93	69	WP_157355779.1	531-1413	*Nocardia* sp. ET3-3
94	68	WP_120736944.1	436-1316	*Nocardia* sp. CFHS0054
95	59	WP_084161146.1(part1)	2212-3085	*Nocardia* sp. BMG51109
96	58	WP_024803008.1	4554-5433	*Nocardia* sp. BMG51109
97	57	WP_084161146.1(part2)	4074-4959	*Nocardia* sp. BMG51109
98	35	WP_133904463.1	1446-2274	*Actinophytocola oryzae*
99	54	WP_103056255.1	28-900	*Actinoalloteichus* sp. AHMU CJ021

100	2	RKT54669.1	1044-1879	*Saccharothrix australiensis*
101	37	RAS59251.1	421-1269	*Lechevalieria atacamensis*
102	9	WP_090066818.1	1973-2817	*Lentzea flaviverrucosa*
103	4	WP_086661474.1	11-854	*Lentzea kentuckyensis*
104	78	TDP93021.1	456-1282	*Labedaea rhizosphaerae*
105	79	WP_133853352.1	497-1323	*Labedaea rhizosphaerae*
106	10	ABV56588.1	1964-2815	*Kutzneria* sp. 744
107	21	WP_081789315.1	3012-3857	*Kutzneria albida*
108	20	WP_025356113.1	424-1279	*Kutzneria albida*
109	77	WP_094864087.1	501-1360	*Amycolatopsis antarctica*
110	36	RJQ84668.1	1493-2349	*Amycolatopsis panacis*
111	74	SFA72048.1	443-1316	*Amycolatopsis marina*
112	75	WP_091667726.1	501-1374	*Amycolatopsis marina*
113	7	WP_160703735.1	1624-2526	*Amycolatopsis* sp. SID8362
114	8	WP_134662744.1	445-1301	*Amycolatopsis* sp. CFH S0078
115	32	WP_134733373.1(part1)	1760-2579	*Amycolatopsis* sp. CFH S0261
116	34	WP_134733373.1(part2)	440-1257	*Amycolatopsis* sp. CFH S0261
117	91	WP_150476434.1	1851-2709	*Streptomyces alboniger*
118	30	WP_150475077.1	434-1320	*Streptomyces prasinus*
119	90	WP_109279047.1	1754-2627	*Streptomyces orinoci*
120	49	WP_107418411.1	18-854	*Streptomyces mutomycini*
121	26	WP_069778199.1	1981-2828	*Streptomyces fodineus*
122	25	WP_055611559.1	1980-2821	*Streptomyces phaeochromogenes*
123	43	WP_055419767.1	8-832	*Streptomyces pactum*
124	5	WP_014172740.1	1010-1903	*Streptomyces bingchenggensis*
125	96	TWF89347.1	1604-2486	*Streptomyces capillispiralis*
126	27	EPH46596.1	1985-2824	*Streptomyces aurantiacus* JA 4570
127	98	WP_144001289.1	1584-2466	*Streptomyces spectabilis*
128	99	WP_150509159.1	1587-2469	*Streptomyces spectabilis*
129	48	WP_158709681.1	73-909	*Streptomyces* sp. NRRL S-15
130	31	WP_158693706.1	1531-2420	*Streptomyces* sp. CNT360
131	24	WP_142201984.1	1982-2830	*Streptomyces* sp. SLBN-109
132	28	WP_125510943.1	1980-2823	*Streptomyces* sp. WAC 01529
133	93	WP_124275292.1	1592-2436	*Streptomyces* sp. ADI93-02
134	46	WP_111001645.1	3-839	*Streptomyces* sp. NTH33
135	47	WP_107441022.1	18-854	*Streptomyces* sp. NRRL S-4
136	38	WP_093779539.1	26-865	*Streptomyces* sp. yr375
137	6	WP_089509267.1	1005-1904	*Streptomyces* sp. NBS 14/10
138	23	WP_087803945.1	1982-2830	*Streptomyces* sp. CS113
139	52	WP_078988403.1	12-844	*Streptomyces* sp. WM6372
140	92	WP_067167615.1	1610-2436	*Streptomyces* sp. ERV7
141	33	WP_051729584.1	552-1256	*Streptomyces* sp. NRRL F-3213
142	22	WP_031042662.1	1982-2830	*Streptomyces* sp. NRRL F-5650
143	44	WP_030835487.1	8-831	*Streptomyces* sp. NRRL S-475
144	51	RIA73626.1	7-838	*Streptomyces* sp. 3211.1
145	95	PJN03934.1	1796-2673	*Streptomyces* sp. CB01201
146	50	AYV31157.1	7-838	*Streptomyces* sp. ADI95-16
147	45	WP_093845005.1	8-831	unclassified *Streptomyces*
148	13	QCX80649.1	1983-2857	*Streptomyces* sp. YIM 121038
149	97	WP_138957829.1	1586-2468	*Streptomyces* sp. YIM 121038

Table S3. Detailed information of family 2b ($A_8M_{s(O,\text{ Tyr})}A_9$) interrupted A domains used in taxonomic and phylogenetic trees.

# in taxonomic tree	**# in phylogenetic tree**	**NCBI ID**	**Range of sequence**	**Organism**
1	57	PYQ92623.1	985-1823	Acidobacteria bacterium
2	68	HAB14970.1	854-1691	Verrucomicrobiales bacterium
3	69	MAE67645.1	16-876	Phycisphaeraceae bacterium
4	28	WP_143010572.1	544-1365	*Pedobacter steynii*

5	26	WP_073399068.1	1518-2363	*Mucilaginibacter* sp. OK098
6	35	TMC21728.1	549-1407	Chloroflexi bacterium
7	36	HBB34601.1	15-867	Cyanobacteria bacterium UBA9273
8	43	WP_144865674.1	2462-3319	*Hyella patelloides*
9	47	WP_146138208.1	1547-2393	*Chamaesiphon polymorphus*
10	38	WP_106300980.1	607-1455	*Chamaesiphon polymorphus*
11	49	TRV54169.1	2620-3483	*Microcystis panniformis* Mp_GB_SS_20050300_S99
12	41	CBN55480.1	1-688	*Kamptonema* sp. PCC 6506
13	53	BBD53649.1	2622-3498	*Planktothrix agardhii* NIES-204
14	30	NER25248.1	257-1086	*Symploca* sp. SIO1C2
15	45	WP_081431049.1	529-1380	*Moorea bouillonii*
16	44	NET65459.1	535-1399	*Moorea* sp. SIO1G6
17	46	NEO79697.1	529-1380	*Moorea* sp. SIO4G3
18	48	WP_052490042.1	502-1352	*Tolypothrix campylonemoides*
19	54	WP_168650499.1	2623-3487	*Dolichospermum* sp. UHCC 0352
20	56	WP_168653542.1	2623-3487	unclassified *Dolichospermum*
21	55	CAC01604.1	2622-3486	*Anabaena* sp. 90
22	42	WP_114085954.1	530-1376	*Nostoc* sp. ATCC 53789
23	52	WP_084177704.1	4880-5764	*Nostoc calcicola*
24	51	WP_017742648.1	3798-4647	*Scytonema hofmannii*
25	37	NJM72086.1	1171-2035	*Scytonema* sp. RU_4_4
26	39	WP_155748091.1	1171-2016	*Scytonema* sp. UIC 10036
27	50	WP_083630634.1	554-1411	*Scytonema* sp. HK-05
28	73	KAA0206779.1	30-865	Proteobacteria bacterium
29	9	QDA77059.1	4164-4980	*Jahnella* sp. MSr9139
30	59	EYF05518.1	1622-2497	*Chondromyces apiculatus* DSM 436
31	58	WP_050433036.1	1569-2444	*Chondromyces crocatus*
32	5	WP_146188543.1	1101-1914	*Limnohabitans* sp. 2KL-51
33	34	WP_161021313.1	1571-2413	*Duganella* sp. FT50W
34	40	KRB71153.1	707-1554	*Noviherbaspirillum* sp. Root189
35	18	WP_132458885.1	508-1342	*Paraburkholderia* sp. BL8N3
36	19	WP_127837709.1	510-1345	*Burkholderia gladioli*
37	20	WP_127840945.1	510-1345	*Burkholderia gladioli*
38	21	WP_124076635.1	510-1345	*Burkholderia gladioli*
39	22	WP_161783325.1	507-1342	*Burkholderia* sp. A1
40	23	NIF88164.1	510-1345	*Burkholderia* sp. Cy-637
41	24	NIF72137.1	510-1345	*Burkholderia* sp. Ap-962
42	25	NIF64901.1	510-1345	*Burkholderia* sp. Cy-647
43	6	OYX71643.1	1106-1916	Rhizobiales bacterium 32-66-11
44	11	WP_007675758.1	270-1093	Alphaproteobacterium BAL199
45	12	WP_158923915.1	250-1088	*Acidisphaera* sp. S103
46	60	WP_158929672.1	549-1333	*Acidisphaera* sp. S103
47	7	WP_085554367.1	1097-1919	*Azospirillum lipoferum*
48	61	WP_109330483.1	464-1314	*Azospirillum* sp. CFH 70021
49	8	WP_109084104.1	1099-1911	*Azospirillum* sp. TSH100
50	62	TNF22209.1	454-1278	Rhodobacteraceae bacterium
51	67	OIQ40354.1	473-1304	*Roseobacter* sp. MedPE-SWde
52	65	WP_054462236.1	473-1300	*Phaeobacter* sp. 11ANDIMAR09
53	10	WP_103223333.1	262-1086	*Labrenzia marina*
54	66	WP_027238317.1	473-1304	*Pseudophaeobacter arcticus*
55	64	WP_083100503.1	330-1162	*Pseudophaeobacter leonis*
56	15	WP_146587866.1	1071-1890	*Puniceibacterium confluentis*
57	17	WP_089273465.1	1062-1881	*Puniceibacterium sediminis*
58	16	WP_047994507.1	1069-1882	*Puniceibacterium* sp. IMCC21224
59	3	WP_073579370.1	1091-1923	*Vibrio quintilis*
60	2	WP_161990627.1	1063-1889	*Candidatus Methylobacter oryzae*
61	1	GDX84814.1	1068-1887	Methylococcaceae bacterium
62	29	WP_063369528.1	534-1376	*Pseudoalteromonas luteoviolacea*

63	4	WP_091978782.1	1081-1925	*Pseudoalteromonas denitrificans*
64	31	WP_082650338.1	2115-2941	*Legionella parisiensis*
65	32	WP_058483861.1	2102-2933	*Legionella spiritensis*
66	33	WP_126337465.1	2102-2933	*Legionella spiritensis*
67	14	EXJ16199.1	1079-1884	*Imhoffiella purpurea*
68	13	WP_007042207.1	1081-1886	*Thiorhodococcus drewsii*
69	77	MAF83146.1	3-871	Chromatiales bacterium
70	76	MBT76747.1	3-871	Chromatiales bacterium
71	71	MBS95431.1	21-870	Chromatiales bacterium
72	75	HHQ14590.1	9-847	Chromatiales bacterium
73	27	RKZ81921.1	26-710	Gammaproteobacteria bacterium
74	78	TDJ47449.1	4-861	Gammaproteobacteria bacterium
75	79	NND36229.1	4-853	Gammaproteobacteria bacterium
76	74	RMF98833.1	9-829	Gammaproteobacteria bacterium
77	63	NKC11654.1	468-1302	Gammaproteobacteria bacterium
78	72	TDJ45410.1	25-885	Gammaproteobacteria bacterium
79	70	NND55267.1	15-839	Gammaproteobacteria bacterium

Table S4. Detailed information of family 3 ($A_2M_{s(S)}A_3$) interrupted A domains used in taxonomic and phylogenetic trees.

# in taxonomic tree	**# in phylogenetic tree**	**NCBI ID**	**Range of sequence**	**Organism**
1	9	OQY27011.1	32-829	*Candidatus Cloacimonetes* bacterium 4572_55
2	25	HBL25591.1	895-1667	Acidobacteria bacterium
3	10	RKZ70924.1	38-775	Gammaproteobacteria bacterium
4	24	WP_105250892.1	226-965	*Shewanella* sp. WE21
5	20	WP_046303156.1	2853-3646	*Grimontia* sp. AD028
6	19	WP_062706917.1	2853-3646	*Grimontia marina*
7	21	WP_017002896.1	2862-3655	*Enterovibrio norvegicus*
8	22	WP_016962256.1	2862-3655	*Enterovibrio norvegicus*
9	23	WP_017006281.1	1378-2171	*Enterovibrio norvegicus*
10	12	NER47250.1	1209-2016	*Symploca* sp. SIO1A3
11	11	NEO69486.1	1589-2374	*Moorea* sp. SIO3H5
12	13	WP_014312927.1	258-1044	*Clostridium* sp. BNL1100
13	17	WP_069194359.1	260-1044	unclassified *Clostridium*
14	14	WP_020816432.1	258-1044	*Ruminiclostridium papyrosolvens*
15	15	WP_051033578.1	259-1081	*Hungateiclostridium cellulolyticum*
16	16	GAE88107.1	260-1044	*Hungateiclostridium straminisolvens* JCM 21531
17	18	WP_128705964.1	260-1044	*Hungateiclostridium* sp. N2K1
18	4	WP_027084596.1	260-1040	*Cohnella panacarvi*
19	5	SDG01123.1	252-1045	*Fontibacillus panacisegetis*
20	8	WP_091233568.1	248-1045	*Fontibacillus panacisegetis*
21	6	WP_091233595.1	303-1096	*Fontibacillus panacisegetis*
22	2	WP_127573070.1	1504-2299	*Paenibacillus xylaniclasticus*
23	7	WP_127194795.1	250-1047	*Paenibacillus anaericanus*
24	3	WP_148927386.1	260-1053	*Paenibacillus methanolicus*
25	1	WP_148930441.1	1288-2069	*Paenibacillus methanolicus*
26	34	WP_125635920.1	25-788	*Nonomuraea* sp. WAC 01424
27	45	WP_109689317.1	9-764	*Branchiibius hedensis*
28	49	WP_159086747.1	1-744	*Actinomyces* sp. Marseille-P3109
29	43	WP_054238190.1	32-820	Actinobacteria bacterium OV320
30	35	WP_077005101.1	13-762	*Saccharothrix* sp. ALI-22-I
31	36	SDH49032.1	1-753	*Lechevalieria fradiae*
32	30	WP_030469215.1	5-750	*Lentzea aerocolonigenes*
33	53	WP_164448046.1	11-778	*Verrucosispora* sp. CWR15
34	28	SDZ47176.1	5-699	*Asanoa ishikariensis*
35	27	WP_095564439.1	12-781	*Plantactinospora* sp. KBS50
36	26	WP_051723652.1	8-710	*Micromonospora chokoriensis*

37	29	WP_132239013.1	10-764	*Micromonospora* sp. CNZ303
38	32	TVZ92226.1	19-782	*Micromonospora* sp. CNZ297
39	31	ADU06391.1	19-782	*Micromonospora* sp. L5
40	52	CAJ34370.1	1-768	*Micromonospora* sp. ML1
41	46	KJY32180.1	6-781	*Streptomyces katrae*
42	54	WP_112471015.1	12-778	*Streptomyces triticisoli*
43	55	WP_079040618.1	12-778	*Streptomyces prasinus*
44	37	WP_094104102.1	7-765	*Streptomyces phaeoluteigriseus*
45	50	WP_055631801.1	8-769	*Streptomyces griseoruber*
46	39	WP_150470850.1	7-765	*Streptomyces galilaeus*
47	38	WP_086771560.1	7-765	*Streptomyces bobili*
48	47	WP_079430519.1	17-793	*Streptomyces*
49	59	WP_162948446.1	12-761	*Streptomyces europaeiscabiei*
50	58	WP_159015105.1	12-761	*Streptomyces europaeiscabiei*
51	57	WP_159058575.1	12-761	*Streptomyces europaeiscabiei*
52	42	KQX82440.1	18-803	*Streptomyces* sp. Root1310
53	40	WP_121789412.1	7-766	*Streptomyces* sp. Z022
54	61	WP_071381075.1	12-778	*Streptomyces* sp. MUSC 1
55	60	WP_129770066.1	13-785	*Streptomyces* sp. L-9-10
56	56	WP_073936299.1	12-765	*Streptomyces* sp. CB02400
57	48	WP_121016765.1	17-793	*Streptomyces* sp. 3211.6
58	44	WP_142264118.1	32-820	*Streptomyces* sp. SLBN-115
59	33	WP_143616273.1	14-752	*Streptomyces* sp. IF17
60	41	WP_081546865.1	32-836	*Streptomyces* sp. B9173
61	64	APD71752.1	12-778	*Streptomyces* sp.
62	51	APD71650.1	8-770	*Streptomyces* sp.
63	63	APD71599.1	12-778	*Streptomyces* sp.
64	62	APD71663.1	12-778	*Streptomyces* sp.

Table S5. Detailed information of family 4 ($A_2M_bA_3$) interrupted A domains used in taxonomic and phylogenetic trees.

# in taxonomic tree	**# in phylogenetic tree**	**NCBI ID**	**Range of sequence**	**Organism**
1	106	PYQ27333.1	1-777	Acidobacteria bacterium
2	124	PLX70436.1	1-783	*Denitrovibrio* sp.
3	84	WP_012374793.1	27-805	*Opitutus terrae*
4	16	OGZ66886.1	3-778	*Candidatus Staskawiczbacteria* bacterium RIFCSPHIGHO2_02_FULL_33_16
5	15	OGZ62783.1	3-778	*Candidatus Staskawiczbacteria* bacterium RIFCSPHIGHO2_01_FULL_34_27
6	83	MBI91646.1	37-800	Gemmatimonadaceae bacterium
7	82	MBQ38298.1	37-798	Gemmatimonadaceae bacterium
8	152	WP_006366480.1	1-814	*Chlorobium ferrooxidans*
9	150	TLU58760.1	1-813	*Chlorobium* sp.
10	151	NHQ60383.1	1-814	*Chlorobium* sp. BLA1
11	92	TNE55038.1	30-809	Bacteroidetes bacterium
12	17	WP_157494010.1	2179-2935	*Fulvivirga imtechensis*
13	57	WP_119657530.1	1567-2347	*Hymenobacter rubripertinctus*
14	81	WP_162445333.1	257-1024	*Rhodocytophaga* sp. 172606-1
15	56	WP_109611019.1	538-1309	*Mucilaginibacter oryzae*
16	80	WP_041886432.1	34-795	*Pedobacter lusitanus*
17	95	WP_121281958.1	530-1307	*Pedobacter alluvionis*
18	14	WP_053002551.1	1301-2014	*Kordia jejudonensis*
19	10	WP_114903509.1	3048-3803	*Kordia* sp. SMS9
20	9	WP_089077049.1	1334-2110	*Flavobacterium tructae*
21	8	WP_070905749.1	1334-2110	*Flavobacterium spartansii*
22	7	WP_017496291.1	1334-2110	*Flavobacterium* sp. WG21
23	18	WP_017496840.1	1310-2069	*Flavobacterium* sp. WG21
24	19	WP_116740396.1	540-1300	*Filimonas* sp. YR581

25	105	WP_089921330.1	1468-2236	*Chitinophaga rupis*
26	97	WP_078672815.1	532-1305	*Chitinophaga eiseniae*
27	116	WP_145719319.1	1529-2308	*Chitinophaga japonensis*
28	20	TWE59091.1	527-1286	*Chitinophaga pinensis*
29	117	WP_012791051.1	1527-2308	*Chitinophaga pinensis*
30	96	WP_090108483.1	541-1314	*Chitinophaga* sp. CF118
31	118	WP_089814539.1	1528-2308	*Chitinophaga* sp. YR627
32	6	WP_095836111.1	243-1024	*Chitinophaga* sp. MD30
33	99	WP_160716657.1	528-1307	*Chitinophaga* sp. O9
34	98	WP_127035210.1	528-1307	*Chitinophaga* sp. Mgbs1
35	138	NCC04883.1	1-774	Proteobacteria bacterium
36	163	WP_104155807.1	1-831	Proteobacteria bacterium 228
37	126	KMY85556.1	1-800	*Candidatus Paraburkholderia calva*
38	119	RQS55687.1	1128-1910	*Burkholderia* sp. Bp8984
39	122	KPJ31995.1	1098-1887	*Burkholderia multivorans*
40	120	AOK68278.1	1121-1910	*Burkholderia multivorans*
41	121	WP_080746627.1	1166-1955	*Burkholderia multivorans*
42	123	WP_081011898.1	1166-1955	*Burkholderia multivorans*
43	133	WP_043756984.1	1-789	*Imhoffiella purpurea*
44	162	WP_110187341.1	1-831	*Pokkaliibacter plantistimulans*
45	85	RKZ49082.1	1-801	Gammaproteobacteria bacterium
46	161	WP_153295950.1	1-804	Gammaproteobacterium SS-5
47	4	WP_063364971.1	273-1054	*Pseudoalteromonas luteoviolacea*
48	3	WP_063357127.1	273-1054	*Pseudoalteromonas luteoviolacea*
49	55	WP_063364970.1	1322-2134	*Pseudoalteromonas luteoviolacea*
50	2	WP_063369527.1	287-1057	*Pseudoalteromonas luteoviolacea*
51	65	HCF58371.1	29-712	Myxococcales bacterium
52	137	PKN41858.1	1-774	Deltaproteobacteria bacterium HGW-Deltaproteobacteria-18
53	148	NCB23041.1	2-817	Deltaproteobacteria bacterium
54	132	RLB11908.1	4-700	Deltaproteobacteria bacterium
55	136	WP_043774779.1	1-769	*Desulfomicrobium escambiense*
56	139	OIO01401.1	1-774	*Desulfovibrionaceae* bacterium CG1_02_65_16
57	135	WP_092152782.1	1-805	*Desulfovibrio legallii*
58	143	SBV90717.1	61-821	uncultured *Desulfovibrio* sp.
59	146	SBV94691.1	42-817	uncultured *Desulfovibrio* sp.
60	149	MTJ93725.1	47-811	*Desulfovibrio* sp.
61	141	WP_116307174.1	2-794	*Desulfovibrio* sp. HK-II
62	140	WP_007525669.1	2-791	*Desulfovibrio* sp. A2
63	147	WP_081640478.1	43-811	*Desulfovibrio desulfuricans*
64	145	WP_136400210.1	1-786	*Desulfovibrio desulfuricans*
65	144	WP_022660112.1	62-824	*Desulfovibrio desulfuricans*
66	134	SBV93004.1	24-779	uncultured Alphaproteobacteria bacterium
67	160	WP_134185454.1	1-796	*Pleomorphomonas* sp. CF100
68	107	WP_009540931.1	40-806	*Caenispirillum salinarum*
69	142	WP_021133476.1	1-770	*Phaeospirillum fulvum*
70	164	WP_067559869.1	1-798	*Oceanibaculum pacificum*
71	153	WP_145735196.1	42-812	*Nitrospirillum amazonense*
72	166	WP_022728591.1	1-800	*Fodinicurvata sediminis*
73	165	WP_026987550.1	1-800	*Fodinicurvata fenggangensis*
74	125	WP_126995339.1	1-802	*Azospirillum doebereinerae*
75	155	WP_126619512.1	1-801	*Azospirillum* sp. L-25-5w-1
76	154	WP_119832747.1	1-802	*Azospirillum* sp. K2W22B-5
77	159	MAD38958.1	1-794	*Tistrella* sp.
78	156	AGC65517.1	1-818	*Tistrella bauzanensis*
79	157	AFK55404.1	1-793	*Tistrella mobilis* KA081020-065
80	158	WP_062762381.1	1-794	*Tistrella mobilis*
81	181	WP_114124031.1	1-808	*Thalassospira xianhensis*
82	175	WP_114089625.1	1-815	*Thalassospira profundimaris*

83	178	WP_037992687.1	1-807	*Thalassospira permensis*
84	168	WP_085584690.1	17-805	*Thalassospira mesophila*
85	171	WP_101302846.1	1-802	*Thalassospira lohafexi*
86	169	WP_033067477.1	1-802	*Thalassospira australica*
87	174	WP_085619000.1	15-812	*Thalassospira alkalitolerans*
88	176	WP_063096102.1	1-807	*Thalassospira*
89	172	WP_022731483.1	1-802	*Thalassospira lucentensis*
90	173	HBU99806.1	1-802	*Thalassospira lucentensis*
91	188	WP_062953556.1	1-808	*Thalassospira lucentensis*
92	180	WP_007090617.1	1-807	*Thalassospira xiamenensis*
93	183	WP_114109920.1	1-808	*Thalassospira xiamenensis*
94	179	WP_114094869.1	1-807	*Thalassospira xiamenensis*
95	185	WP_062959059.1	1-808	*Thalassospira xiamenensis*
96	184	WP_097053433.1	1-808	*Thalassospira xiamenensis*
97	186	MAZ35287.1	1-808	*Thalassospira* sp.
98	187	AGC65515.1	1-784	*Thalassospira* sp. CNJ-328
99	177	WP_071240630.1	1-807	*Thalassospira* sp. MIT1004
100	182	WP_085646015.1	1-808	*Thalassospira* sp. MCCC 1A03138
101	167	WP_085591989.1	1-802	*Thalassospira* sp. MCCC 1A01428
102	170	WP_102784265.1	1-802	*Thalassospira* sp. GB04J01
103	127	WP_012933922.1	34-804	*Conexibacter woesei*
104	131	WP_009739001.1	32-824	*Frankia* sp. QA3
105	130	WP_071082330.1	1-812	*Frankia* sp. Cc1.17
106	128	WP_010314445.1	27-810	*Saccharopolyspora spinosa*
107	129	WP_104481318.1	1-804	*Actinokineospora auranticolor*
108	73	WP_113672412.1	467-1225	*Vallitalea guaymasensis*
109	63	WP_069194360.1	264-1024	*Clostridiales*
110	11	WP_028517513.1	1243-2038	*Ruminococcus flavefaciens*
111	69	WP_013497583.1	260-1014	*Ruminococcus albus*
112	70	WP_009214812.1	247-1016	*Oribacterium* sp. oral taxon 078
113	71	WP_021774343.1	249-1016	*Oribacterium* sp. oral taxon 078
114	64	WP_010250077.1	264-1024	*Hungateiclostridium cellulolyticum*
115	62	WP_020816431.1	266-1025	*Ruminiclostridium papyrosolvens*
116	104	WP_024833474.1	269-1041	*Ruminiclostridium josui*
117	12	WP_090043163.1	1839-2601	*Clostridium frigidicarnis*
118	72	WP_022748626.1	466-1225	*Clostridium saccharobutylicum*
119	61	WP_014312926.1	266-1025	*Clostridium* sp. BNL1100
120	75	WP_160672662.1	19-725	*Clostridium* sp. C8-1-8
121	79	BAH07125.1	265-1031	*Clostridium kluyveri NBRC 12016*
122	78	WP_073539243.1	271-1028	*Clostridium kluyveri*
123	1	WP_087456218.1	664-1376	*Tumebacillus avium*
124	111	WP_142503746.1	1152-1956	*Melghirimyces algeriensis*
125	109	WP_134026490.1	1187-1968	*Lysinibacillus* sp. YR326
126	115	WP_121640376.1	1167-1926	*Virgibacillus* sp. Bac330
127	114	WP_073012998.1	1167-1926	*Virgibacillus chiguensis*
128	112	WP_077702826.1	1167-1926	*Virgibacillus dokdonensis*
129	113	WP_101933571.1	1167-1926	*Virgibacillus dokdonensis*
130	52	WP_086407309.1	277-1039	*Bacillus*
131	39	WP_098124654.1	277-1039	*Bacillus pseudomycoides*
132	41	WP_098708907.1	277-1039	*Bacillus wiedmannii*
133	40	WP_098717719.1	277-1039	*Bacillus wiedmannii*
134	50	OUB26095.1	244-1006	*Bacillus thuringiensis serovar yunnanensis*
135	53	WP_098975536.1	277-1039	*Bacillus thuringiensis*
136	42	WP_086412483.1	277-1039	*Bacillus thuringiensis*
137	49	PGK60785.1	275-1037	*Bacillus thuringiensis*
138	45	EOQ16526.1	275-1037	*Bacillus cereus* VD184
139	44	WP_080495864.1	277-1039	*Bacillus cereus*
140	43	WP_098295041.1	277-1039	*Bacillus cereus*
141	51	WP_097998275.1	277-1039	*Bacillus cereus*

142	46	WP_063264141.1	275-1037	*Bacillus cereus*
143	47	QBZ24922.1	244-1006	*Bacillus cereus*
144	48	RHW05505.1	275-1037	*Bacillus cereus*
145	54	KXY56201.1	275-1037	*Bacillus cereus*
146	76	WP_164991686.1	267-1020	*Paenibacillaceae* bacterium
147	77	RXZ79614.1	295-1048	*Paenibacillaceae* bacterium
148	13	WP_106833120.1	210-978	*Brevibacillus porteri*
149	30	WP_064018224.1	1517-2294	*Brevibacillus* sp. SKDU10
150	110	WP_139491844.1	1202-1985	*Brevibacillus* sp. CHY01
151	37	WP_163244904.1	1517-2294	*Brevibacillus* sp. 7WMA2
152	38	CCF13225.1	1522-2299	*Brevibacillus laterosporus* GI-9
153	31	AIG26965.1	1517-2294	*Brevibacillus laterosporus* LMG 15441
154	26	WP_158329415.1	1517-2294	*Brevibacillus laterosporus*
155	27	WP_158325582.1	1517-2294	*Brevibacillus laterosporus*
156	28	WP_158330658.1	1517-2294	*Brevibacillus laterosporus*
157	25	WP_113755445.1	1517-2294	*Brevibacillus laterosporus*
158	35	WP_121474376.1	1517-2294	*Brevibacillus laterosporus*
159	5	WP_051870662.1	928-1702	*Brevibacillus laterosporus*
160	32	WP_068792383.1	1519-2294	*Brevibacillus laterosporus*
161	33	WP_101668396.1	1517-2294	*Brevibacillus laterosporus*
162	36	WP_096887168.1	1517-2294	*Brevibacillus laterosporus*
163	34	WP_022586907.1	1517-2294	*Brevibacillus laterosporus*
164	29	WP_031415345.1	1519-2294	*Brevibacillus laterosporus*
165	94	WP_143812707.1	272-1043	*Paenibacillus taichungensis*
166	101	WP_150267101.1	272-1044	*Paenibacillus* sp. SYSU G01001
167	24	WP_116046360.1	473-1232	*Paenibacillus paeoniae*
168	74	WP_090711243.1	272-1030	*Paenibacillus typhae*
169	108	WP_006037480.1	1195-1969	*Paenibacillus curdlanolyticus*
170	103	WP_087914619.1	268-1041	*Paenibacillus donghaensis*
171	102	WP_087914619.1	3652-4432	*Paenibacillus donghaensis*
172	67	WP_013918180.1	1339-2118	*Paenibacillus mucilaginosus*
173	68	WP_014370612.1	1339-2118	*Paenibacillus mucilaginosus*
174	66	WP_016362698.1	1343-2118	*Paenibacillus mucilaginosus*
175	87	WP_013918181.1	1563-2337	*Paenibacillus mucilaginosus*
176	86	WP_016362699.1	1563-2337	*Paenibacillus mucilaginosus*
177	88	WP_014370613.1	1563-2337	*Paenibacillus mucilaginosus*
178	100	WP_013918046.1	209-978	*Paenibacillus mucilaginosus*
179	93	WP_144424382.1	272-1043	*Paenibacillus* sp. A59
180	21	WP_157332091.1	1506-2286	*Paenibacillus* sp. N10
181	60	WP_161404603.1	1300-2065	*Paenibacillus* sp. 5J-6
182	89	MZQ80580.1	1466-2252	*Paenibacillus* sp. 5J-6
183	22	WP_163953004.1	1525-2300	*Paenibacillus* sp. SYP-B3998
184	58	WP_163953008.1	2550-3310	*Paenibacillus* sp. SYP-B3998
185	23	WP_163953010.1	485-1251	*Paenibacillus* sp. SYP-B3998
186	91	WP_145669549.1	1526-2294	*Paenibacillus* sp. 597
187	59	WP_145669547.1	1300-2065	*Paenibacillus* sp. 597
188	90	TWF31900.1	1509-2277	*Paenibacillus* sp. 597

Table S6. Detailed information of families 5a/b ($\mathrm{A}_8\mathrm{M}_{\mathrm{s}(O)}\mathrm{M}_\mathrm{b}\mathrm{A}_9$) interrupted A domains used in taxonomic and phylogenetic trees.

# in taxonomic tree	**# in phylogenetic tree**	**NCBI ID**	**Range of sequence**	**Organism**
1	3	RKH86437.1	542-1749	*Corallococcus* sp. AB032C
2	6	MBA75309.1	491-1780	*Tistrella* sp.
3	4	WP_082828322.1	491-1773	*Tistrella mobilis*
4	5	WP_014748210.1	491-1780	*Tistrella mobilis*
5	2	WP_096538841.1	5851-7126	*Nostoc linckia*
6	1	AKQ09584.1	508-1739	*Moorea bouillonii* PNG
7	7	NEO41471.1	1638-2864	*Moorea* sp. SIOASIH

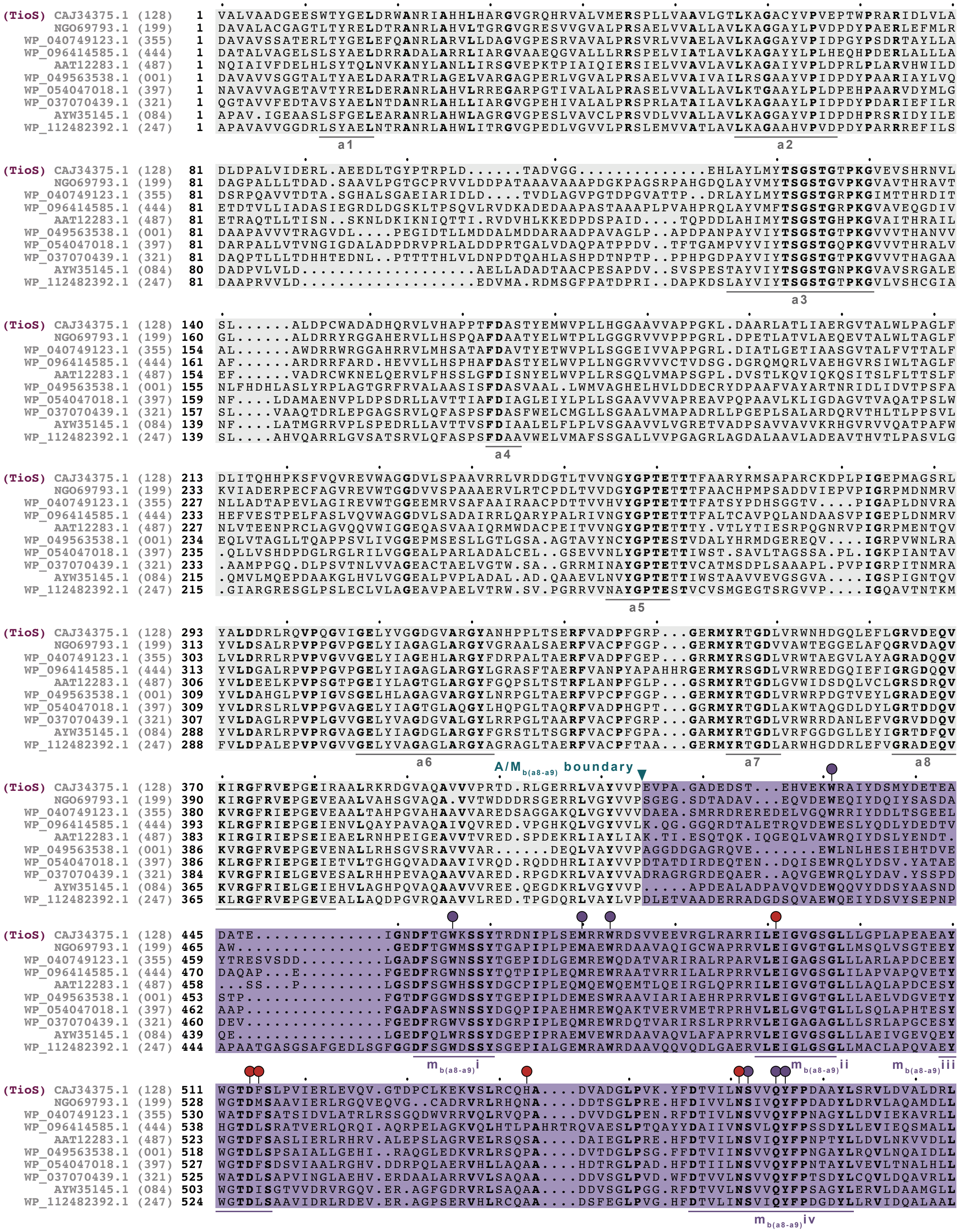
8	8	QCP68969.1	491-1783	*Moorea producens* ASI16Jul14-2
9	9	AEF01451.1	561-1789	*Moorea producens* 3L
10	11	AOY81087.1	554-1782	*Moorea producens* JHB
11	10	AOW98531.1	554-1782	*Moorea producens* PAL-8-15-08-1

Table S7. Detailed information of family 6 ($A_6M_{s(O,\text{ arom})}A_7$) interrupted A domains used in taxonomic and phylogenetic trees.

# in taxonomic tree	**# in phylogenetic tree**	**NCBI ID**	**Range of sequence**	**Organism**
1	6	WP_129570580.1	1978-2779	*Mucilaginibacter rubeus*
2	10	WP_141643221.1	485-1288	*Myxococcus* sp. AM401
3	2	WP_078995397.1	511-1305	*Lysobacter enzymogenes*
4	3	WP_142905261.1	3260-4060	Vibrionaceae bacterium R142
5	7	ABR71486.1	535-1316	*Marinomonas* sp. MWYL1
6	1	HAI15765.1	510-1304	Gammaproteobacteria bacterium
7	4	RKZ63928.1	514-1309	Gammaproteobacteria bacterium
8	11	WP_015854743.1	496-1300	*Dickeya paradisiaca*
9	8	WP_132356348.1	509-1311	*Photorhabdus khanii*
10	9	WP_152964228.1	509-1311	*Photorhabdus khanii*
11	12	WP_093719989.1	496-1279	*Streptomyces* sp. DI166
12	13	WP_055688818.1	505-1278	*Streptomyces prasinus*
13	14	WP_055604379.1	505-1278	*Streptomyces prasinus*
14	5	WP_161214629.1	1183-1973	*Clostridium butyricum*

2. Multiple sequence alignments and identification of boundaries of interrupted A domains.

Sequence alignments for the interrupted A domain families 1-7 (Figs. S13-S21 and Figs. 6-10) were generated using Multalin[13] with the ten representatives identified in section 1 for families 1-4 (Figs. S13-S17) and all available sequences for families 5-7 (Figs. S18-S21) since there are so few examples (n = 11, 14, and 3, respectively). For the interrupted A domains where representatives had been previously biochemically characterized (*i.e.*, families 1, 2a, 3, and 5a, (TioS($A_8M_IA_9$)$_4$, KtzH($A_8M_{II}A_9$)$_4$, TioN($A_2M_{IV}A_3$), and (ColG($A_8M_{II}M_IA_9$)), respectively), these representatives were marked as published and used as references and guides to mark the boundaries on the alignments presented in Figs. S13, S14, S16, and S18, respectively.[1, 2, 4, 7, 14] The representative for family 4 is TtbB($A_2M_VA_3$)$_5$,[5] which is analogous to TtcC($A_2M_VA_3$)$_5$ (AGC65515.1). TtcC($A_2M_VA_3$)$_5$ was first reported as an a2-a3 interrupted A domain responsible for incorporating *N*-methylated L-Tyr during thalassospiramide biosynthesis.[6] However, the sequence in the NCBI database was incomplete, so TtbB($A_2M_VA_3$)$_5$ was used in its place. While it was known that there was an M domain between a2-a3 in TtbB($A_2M_VA_3$)$_5$, it was unclear as to where the M domain started and where the A domain ended, therefore, we further refined the boundaries as described below. This alignment is presented in Fig. S17. For family 5c, its sole member is FrsG($A_8M_{II}A_{9\text{-}10}M_IA_9$)$_8$, which has a unique arrangement, the boundaries were pieced together by aligning the M domains from KtzH($A_8M_{II}A_9$)$_4$ and TioS($A_8M_IA_9$)$_4$ to their corresponding M domain counterparts (Fig. S19). For the refinement of family 4 (Fig. S17) and those with no reported annotation or characterization, families 2b (Fig. S10), 6 (Fig. S20), and 7 (Fig. S21), the boundaries were identified using the following general approach. First, all ten conserved A domain motifs were identified, which allowed the location of the interruption to be determined. Next, the region identified as the M domain by NCBI's "Identify Conserved Domains" function was marked along with the *S*-adenosyl-L-methionine (SAM) binding site, the hallmark of class I methyltransferases. From there, the M domain sequence was deleted, and the resulting "uninterrupted" A domain sequence was aligned using Multalin[13] to naturally occurring uninterrupted A domains PheA (PDB: 1AMU_A),[15] Ecm6(A_1) (accession: BAE98155.1),[16] and TioR(A_2) (accession: CAJ34374.1),[17] and checked for continuity. Then, more of the M domain was deleted as needed and the A domains were aligned again. This was repeated until (i) there were no large gaps in the uninterrupted sequences and (ii) the a1-a10 conserved motifs of the interrupted A domain aligned with those of the uninterrupted ones. This was determined to be A and not M domain sequence. Finally, the A domain sequence with the M removed was used in a BLAST search again to verify that there were no M domains found in the "Identify Conserved Domains" section.



```
(TioS) CAJ34375.1 (128)   1 VALVAADGEESWTYGELDRWANRIAHHLHARGVGRQHRVALVMERSPLLVAAVLGTLKAGACYVPVEPTWPRARIDLVLA
       NGO69793.1 (199)   1 DAVALACGAGTLTYRELDTRANRLAHVLTGRGVGRESVVGVALPRSAELVVALLAVLKAGGAYLPVDPDYPAERLEFMLR
    WP_040749123.1 (355)   1 DAVAVSSATERLTYGELEFQANRLARVLLDAGVGPESRVAVALPRSVRLVVAELAVLKAGGAYVPIDPGYPSDRTAYLLA
    WP_096414585.1 (444)   1 DATALVAGELSLSYAELDRRADALARRLIARGVAAEQGVALLLRRSPELVIATLAVLKAGAYYLPLHEQHPDERLALLLA
       AAT12283.1 (487)   1 NQIAIVFDELHLSYTQLNVKANYLANLLIRSGVEPKTPIAIQIERSIELVVAVLAVLKAGAIYVPLDPRLPLARVHWILD
    WP_049563538.1 (001)   1 DAVAVVSGGTALTYAELDARATRLAGELVARGAGPERLVGVALPRSAELVVAIVAILRSGAAYVPIDPDYPAARIAYLVQ
    WP_054047018.1 (397)   1 NAVAVVAGETAVTYRELDERANRLAHVLRREGARPGTIVALALPRASELVVALLAVLKTGAAYLPLDPEHPAARVDYMLG
    WP_037070439.1 (321)   1 QGTAVVFEDTAVSYAELNTDANRLAHLLIARGVGPEHIVALALPRSPRLATAILAVLKAGAAYLPIDPDYPDARIEFILR
       AYW35145.1 (084)   1 APAV.IGEAASLSFGELEARANRLAHWLAGRGVGPESLVAVCLPRSVDLVVALLAVLKAGGAYVPIDPDHPRSRIDYILE
    WP_112482392.1 (247)   1 APAVAVVGGDRLSYAELNTRANRLAHWLITRGVGPEDLVGVVLPRSLEMVVATLAVLKAGAAHVPVDPDYPARRREFILS

(TioS) CAJ34375.1 (128)  81 DLDPALVIDERL.AEEDLTGYPTRPLD......TADVGG..............EHLAYLMYTSGSTGTPKGVEVSHRNVL
       NGO69793.1 (199)  81 DAGPALLLTDAD.SAAVLPGTGCPRVVLDDPATAAAVAAAPDGKPAGSRPAHGDQLAYVMYTSGSTGVPKGVAVTHRAVT
    WP_040749123.1 (355)  81 DSRPQAVVTDTA.SGHALSGAEIARIDLD....TVDLAGVPGTDPGVATTP..DRLAYLMYTSGSTGRPKGIMTTHRDIT
    WP_096414585.1 (444)  81 ETDTVLLIADASIEGRDLDGSKLTPSQVLRVDKADEDAAPASTAAAPLPVAHPRQLAYVMFTSGSTGRPKGVAVEQGDIV
       AAT12283.1 (487)  81 ETRAQTLLTISN..SKNLDKIKNIQTTI.RVDVHLKKEDPDSPAID....TQPDDLAHIMYTSGSTGHPKGVAITHRAIL
    WP_049563538.1 (001)  81 DAAPAVVVTRAGVDL....PEGIDTLLMDDALMDDARAADPAVAGLP..APDPANPAYVIYTSGSTGTPKGVVVTHANVV
    WP_054047018.1 (397)  81 DARPALLVTVNGIGDALADPDRVPRLALDDPRTGALVDAQPATPPDV..TFTGAMPVYVIYTSGSTGQPKGVVVTHRALV
    WP_037070439.1 (321)  81 DAQPTLLLTDHHTEDNL..PTTTTHLVLDNPDTQAHLASHPDTNPTP..PPHPGDPAYVIYTSGSTGTPKGVVVTHAGAA
       AYW35145.1 (084)  80 DADPVLVLD..................AELLADADTAACPESAPDV..SVSPESTAYVIYTSGSTGNPKGVAVSRGALE
    WP_112482392.1 (247)  81 DAAPRVVLD..................EDVMA.RDMSGFPATDPRI..DAPKDSLAYVIYTSGSTGTPKGVLVSHCGIA

(TioS) CAJ34375.1 (128) 140 SL......ALDPCWADADHQRVLVHAPPTFDASTYEMWVPLLHGGAAVVAPPGKL.DAARLATLIAERGVTALWLPAGLF
       NGO69793.1 (199) 160 GL......ALDRRYRGGAHERVLLHSPQAFDAATYELWTPLLGGGRVVVPPPGRL.DPETLATVLAEQEVTALWLTAGLF
    WP_040749123.1 (355) 154 AL......AWDRRWRGGAHRRVLMHSATAFDAVTYETWVPLLSGGEIVVAPPGRL.DIATLGETIAASGVTALFVTTALF
    WP_096414585.1 (444) 161 AF......ARDRRFARD.HEVVLLHSPHAFDASTYELWVPLLNGGRVVVCTVDSG.DGRQMQRLVAEHGVRSIWLTAGLF
       AAT12283.1 (487) 154 EF......VADRCWKNELQERVLFHSSLGFDISNYELWVPLLRSGQLVMAPSGPL.DVSTLKQVIQKQSITSLFLTTSLF
    WP_049563538.1 (001) 155 NLFHDHLASLYRPLAGTGRFRVALAASISFDASVAAL.LWMVAGHELHVLDDECRYDPAAFVAYARTNRIDLIDVTPSFA
    WP_054047018.1 (397) 159 NF....LDAMAENVPLDPSDRLLAVTTIAFDIAGLEIYLPLLSGAAVVVAPREAVPQPAAVLKLIGDAGVTVAQATPSLW
    WP_037070439.1 (321) 157 SL....VAAQTDRLEPGAGSRVLQFASPSFDASFWELCMGLLSGAALVMAPADRLLPGEPLSALARDQRVTHLTLPPSVL
       AYW35145.1 (084) 139 NF....LATMGRRVPLSPEDRLLAVTTVSFDIAALELFLPLVSGAAVVLVGRETVADPSAVVAVVKRHGVRVVQATPAFW
    WP_112482392.1 (247) 139 SL....AHVQARRLGVSATSRVLQFASPSFDAAVWELVMAFSSGALLVVPGAGRLAGDALAAVLADEAVTHVTLPASVLG

(TioS) CAJ34375.1 (128) 213 DLITQHHPKSFVQVREVWAGGDVLSPAAVRRLVRDDGTLTVVNGYGPTETTTFAARYRMSAPARCKDPLPIGEPMAGSRL
       NGO69793.1 (199) 233 KVIADERPECFAGVREVWTGGDVVSPAAAERVLRTCPDLTVVDGYGPTETTTFAACHPMPSADDVIEPVPIGRPMDNMRV
    WP_040749123.1 (355) 227 NLLADTAPEVLAGIREVWTGGEEMRVSAFAAIRAACPDTTVVHVYGPTETTTFATSYPDHSGGTV....PIGAPLDNVRA
    WP_096414585.1 (444) 233 HEFVESTPELFASLVQVWAGGDVLSADAIRRLQARYPALRIVNGYGPTETTTFALTCAVPQLANDAASVPIGEPLDNMRV
       AAT12283.1 (487) 227 NLVTEENPRCLAGVQQVWIGGEQASVAAIQRMWDACPEITVVNGYGPTETTTY.VTLYTIESRPQGNRVPIGRPMENTQV
    WP_049563538.1 (001) 234 EQLVTAGLLTQAPPSVLIVGGEPMSESLLGTLGSA.AGTAVYNCYGPTESTVDALYHRMDGEREQV....IGRPVWNLRA
    WP_054047018.1 (397) 235 .QLLVSHDPDGLRGLRILVGGEALPARLADALCEL..GSEVVNLYGPTETTIWST.SAVLTAGSSA.PL.IGKPIANTAV
    WP_037070439.1 (321) 233 .AAMPPGQ.DLPSVTNLVVAGEACTAELVGTW.SA..GRRMINAYGPTETTVCATMSDPLSAAAPL.PVPIGRPITNMRA
       AYW35145.1 (084) 215 .QMVLMQEPDAAKGLHVLVGGEALPVPLADAL.AD.QAAEVLNVYGPTETTIWSTAAVVEVGSGVA....IGSPIGNTQV
    WP_112482392.1 (247) 215 .GIARGRESGLPSLECLVLAGEAVPAELVTRW.SV.PGRRVVNAYGPTESTVCVSMGEGTSRGVVP....IGQAVTNTKM

(TioS) CAJ34375.1 (128) 293 YALDDRLRQVPQGVIGELYVGGDGVARGYANHPPLTSERFVADPFGRP...GERMYRTGDLVRWNHDGQLEFLGRVDEQV
       NGO69793.1 (199) 313 YVLDSALRPVPPGVPGELYIAGAGLARGYVGRAALSAERFVACPFGGP...GERMYRTGDVVAWTEGGELAFQGRVDAQV
    WP_040749123.1 (355) 303 LVLDRRLRPVPVGVVGELYIAGEHLARGYFDRPALTAERFVADPFGEP...GRRMYRSGDLVRWTAEGVLAYAGRADQQV
    WP_096414585.1 (444) 313 YVLDGALRPVPPGVPGELYIAGAGLARGYLGRASFTAERFVANPYAPAHHRGERMYRSGDLVRWREDGQIEFIGRGDQQV
       AAT12283.1 (487) 306 YVLDEELKPVPSGTPGEIYLAGTGLARGYFGQPSLTSTRFLANPFGLP...GSRMYRTGDLGVWIDSDQLVCLGRSDRQV
    WP_049563538.1 (001) 309 YVLDAHGLPVPIGVSGELHLAGAGVARGYLNRPGLTAERFVPCPFGGP...GERMYRTGDLVRWRPDGTVEYLGRADEQV
    WP_054047018.1 (397) 309 YVLDRSLRLVPPGVAGELYIAGTGLAQGYLHQPGLTAQRFVADPHGPT...GGRMYRTGDLAKWTAQGDLDYLGRTDDQV
    WP_037070439.1 (321) 307 YVLDAGLRPVPLGVVGELYVAGDGVALGYLRRPGLTAARFVACPFGRP...GARMYRTGDLVRWRRDANLEFVGRVDDQV
       AYW35145.1 (084) 288 YVLDARLRPVPRGVAGELYIAGDGLARGYFGRSTLTSGRFVACPFGPA...GARMYRTGDLVRFGGDGLLEYIGRTDFQV
    WP_112482392.1 (247) 288 FVLDPALEPVPVGVVGELYVAGAGLARGYAGRAGLTAERFVACPFTAA...GERMYRTGDLARWHGDDRLEFVGRADEQV

(TioS) CAJ34375.1 (128) 370 KIRGFRVEPGEIRAALRKRDGVAQAVVVPRTD.RLGERRLVAYVVPEVPA.GADEDST...EHVEKWRAIYDSMYDETEA
       NGO69793.1 (199) 390 KIRGFRVEPGEVEAALVAHSGVAQA.VVTWDDDRSGERRLVGYVVPSGEG.SDTADAV...EQVDEWEQIYDQIYSASDA
    WP_040749123.1 (355) 380 KVRGFRIEPGEVEAALTAHPGVAHAAVVAREDSAGGAKQLVGYVVVDAEA.SMRRDREREDELVGQWRRIYDDLTSGEEL
    WP_096414585.1 (444) 393 KLRGFRIEPGEIENVLQAYPAVAQAIVQVRED.VPGHKQLVGYVVLK.QG.GGQRDTALETRQVDEWESLYQDLYDEDTV
       AAT12283.1 (487) 383 KIRGIRIEPSEIEAELRNHPEIGEAVVTVRED.SPDEKRLIAYLIAK.TI.ESQTQK.IQGEQLVAWRQIYDSLYENDT.
    WP_049563538.1 (001) 386 KVRGFRVEPGEVENALLRHSGVSRAVVVAR......DEQLVAYVVPAGGDDGAGRQVE.......EWLNLHESIEHTDVE
    WP_054047018.1 (397) 386 KLRGFRIEPGEIETVLTGHGQVADAAVIVRQD.RQDDHRLIAYVVPDTATDIRDEQTEN..DQISEWRQLYDSV.YATAE
    WP_037070439.1 (321) 384 KVRGFRIELGEVESALRHHPEVAQAAVVARED.RPGDKRLVAYVVADRAGRGRDEQAER..AQVGEWRQLYDAV.YSSPD
       AYW35145.1 (084) 365 KVRGFRIELGEIEHVLAGHPQVAQAAVVVREE.QEGDKRLVGYVVP.....APDEALADPAVQVDEWQQVYDDSYAASND
    WP_112482392.1 (247) 365 KLRGFRVEPGEVEALLAQDPGVRQAAVVLRED.TPGDQRLVAYLVPDLETVAADERRADGDSQVAEWQEIYDSVYSEPNP

(TioS) CAJ34375.1 (128) 445 DATE...............IGNDFTGWKSSYTRDNIPLSEMRRWRDSVVEEVRGLRARRILEIGVGSGLLGPLAPEAEAY
       NGO69793.1 (199) 465 AW.................GEDFTGWMSSYTGAPIPLPEMAEWRDAAVAQIGCWAPRRVLEIGVGTGLLMSQLVSGTEEY
    WP_040749123.1 (355) 459 YTRESVSDD.........LGADFSGWNSSYTGEPIDLGEMREWQDATVARIRALRPARVLEIGAGSGLLLARLAPDCEEY
    WP_096414585.1 (444) 470 DAQAP...E.........FGEDFRGWNSSYTQTPIPLEQMREWRAATVRRILALRPRRVLEIGVGSGLILAPVAPQVETY
       AAT12283.1 (487) 458 ...SS...P.........LGSDFSGWHSSYDGCPIPLEQMQEWQEMTLQEIRGLQPRRILEIGVGTGLLLAQLAPDCESY
    WP_049563538.1 (001) 453 STP...............FGTDFGGWDSSYTGEPIPLDEMESWRAAVIARIAEHRPRRVLEIGVGTGLLLAELVDGVETY
    WP_054047018.1 (397) 462 AAP...............FGADFASWNSSYDGQPIPAEHMREWQAKTVERVMETRPRHVLELGVGTGLLLSQLAPHTESY
    WP_037070439.1 (321) 460 DEV...............FGEDFRGWVSSYDGRPIPLEQMREWRDQTVARIRSLRPRRVLEIGAGAGLLLSRLAPGCESY
       AYW35145.1 (084) 439 QE................LGEDFQLWRSSYDGEPIPRAEMVEWRDAAVAQVLAFAPRRVLEIGVGSGLLLAEIVGEVQEY
    WP_112482392.1 (247) 444 APAATGASGSAFGEDLSGFGGDFSGWDSSYSGEPIALGEMRAWRDAAVQQVQDLGAERVLEIGLGSGLLMACLAPQVAEY

(TioS) CAJ34375.1 (128) 511 WGTDFSLPVIERLEVQV.GTDPCLKEKVSLRCQHA....DVADGLPVK.YFDTVILNSVVQYFPDAAYLSRVLDVALDRL
       NGO69793.1 (199) 528 WGTDMSAAVIERLRGQVEQVG..CADRVRLRHQNA....DDTSGLPRE.HFDIVVLNSIVQYFPDADYLDRVLAQAMDLL
    WP_040749123.1 (355) 530 WATDFSATSIDVLATRLRSSGQDWVRRVQLRVQPA....DVVDGLPEN.RFDTIVLNSVVQYFPNAGYLLDVIEKAVRLL
    WP_096414585.1 (444) 538 HGTDLSRATVERLQRQI.AQRPELAGKVQLHTLPAHRTRQVAESLPTQAYYDAIIVNSVLQYFPSSDYLLEVIEQSMALL
       AAT12283.1 (487) 523 WGTDFSASLIERLRHRV.ALEPSLAGRVELRSQSA....DAIEGLPRE.HFDTVILNSVIQYFPNPTYLLDVLNKVVDLL
    WP_049563538.1 (001) 518 WGTDLSPSAIALLGEHI..RAQGLEDKVRLRSQPA....DVTDGLPSG.FFDTVIINSVVQYFPNGAYLRQVLNQAIDLL
    WP_054047018.1 (397) 527 WGTDFSDSVIAALRGHV.DDRPQLAERVHLLAQAA....HDTRGLPAD.HFDTIILNSVVQYFPNTAYLVEVLTNALHLL
    WP_037070439.1 (321) 525 WATDLSAPVINGLAEHV.ERDAALARRVVLSAQAA....DDVSGLPVD.WFDTVVLNSVVQYFPTSDYLCQVLDQAVRLL
       AYW35145.1 (084) 503 WGTDISGTVVDRVRGQV.ER.AGFGDRVRLSAQAA....DDVSGLPGG.RFDTVVLNSVVQYFPSAGYLERVLDAAMGLL
    WP_112482392.1 (247) 524 WGTDLSAAVIDRLRDEV.ER.AGPSERVHLRCQAA....DSFEGLPVG.HFDTVLINSVIQYFPDGDYLLRVIDQALAAL
```

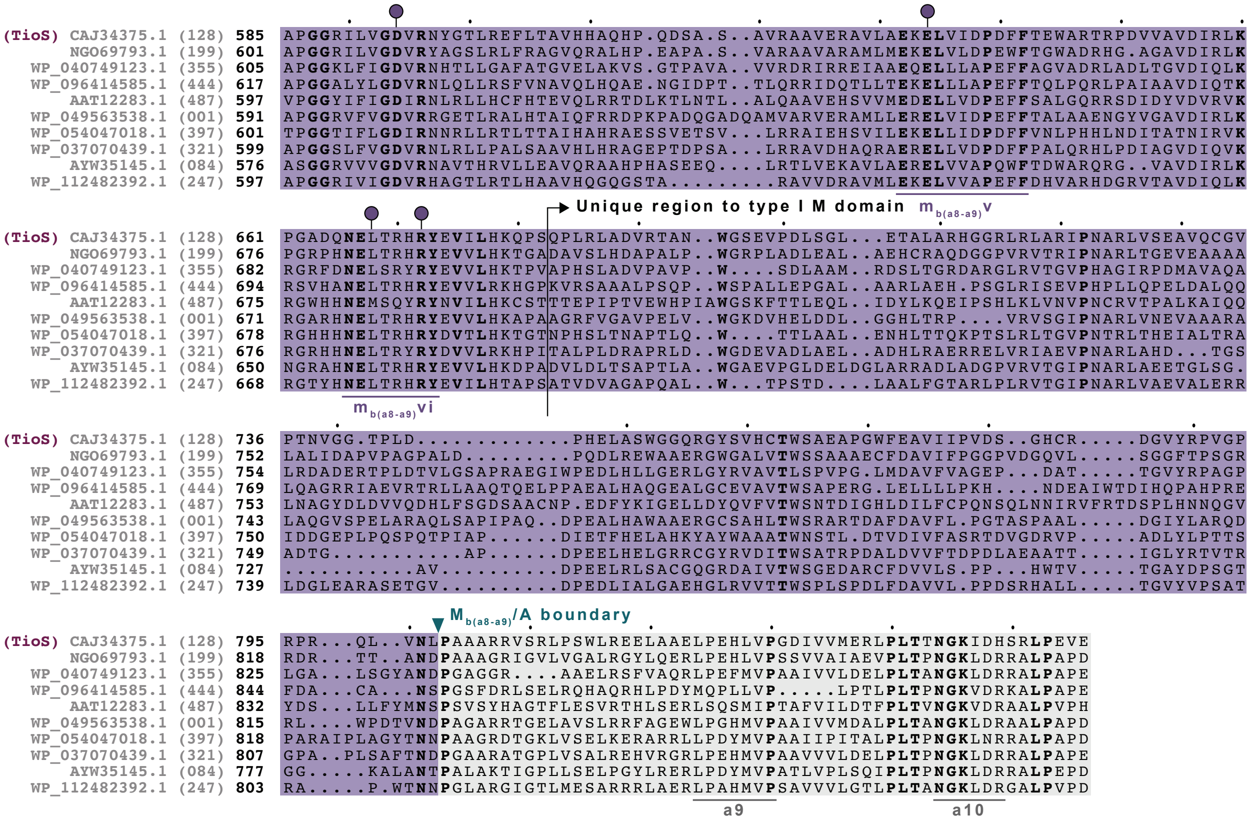
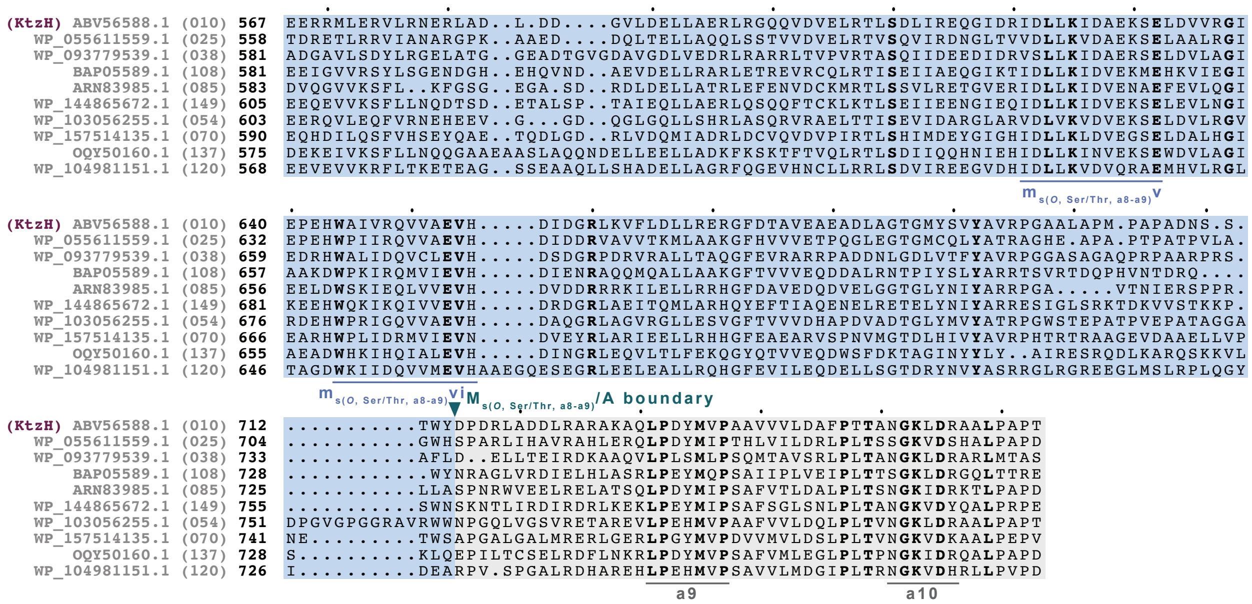


Fig. S13. Multiple sequence alignment of family 1 ($A_8M_bA_9$) interrupted A domains using the ten representative sequences indicated in Figs. S1 and S8. The numbering next to the accession number for each sequence matches the number in the taxonomic tree (Table S1). The A domain portion of the sequence is highlighted in light grey and the $M_{b(a8\text{-}a9)}$ portion in light purple. The divisions between the two are indicated by teal triangles. The conserved residues for binding of SAM and the amino acid bound Ppant arm, based on the structure of TioS($A_8M_IA_9$),[1] are indicated by red and dark purple balloons, respectively. The conserved motifs for the A domain are underlined by a dark grey bar and labeled a1-a10. The conserved M domain motifs (outlined in Table 1) are underlined by a dark purple bar and labeled $m_{b(a8\text{-}a9)}$i-vi. *Note*: Type I M domains have a unique region, which is indicated by the line/arrow in this figure. A shortened version of this annotated alignment is presented in Fig. 6A.

```
   (KtzH) ABV56588.1 (010)   1 AVLCEGEAVSYAELDRRAERLARRLAARGVGPETTVGVMLPRSVDLVVALLAVLKSGAAYLPVDADHPAERVAFVYADAR
      WP_055611559.1 (025)   1 ALIGEQETLTYGEFNRRANQLAHWLVEQGAGPERLVAVRIPRSVELMVAVYAVVKTGAAYLPVDTELPEDRVRQLMESAK
      WP_093779539.1 (038)   1 AVVCGTSEVTYADLALRVDRLADGLCRLGVRPGDVVSVAVPRSVELIAVLLAVLDVGAAYLPLDADQPAERVRYLLTDAR
          BAP05589.1 (108)   1 ALVYGKQQLTYEELNRRSSCLARYLQLQGIQPDQPVGLCVERSLEMVVGILGILRAGGAYLPLEPQYPGERLQYMLQDSR
          ARN83985.1 (085)   1 AVVFEDAQLAYGELNARANQLAHHLRSLGVGPDALVGICVERSLEMVVGLLAILKAGGAYLPLDPDYPRERIGFMLEDAK
      WP_144865672.1 (149)   1 AVVYDNEQLTYQELNARANQLAHYLQSLGVKPEVLVGICVERSLEMIVGILGILKAGGAYLPLDSTNPQERLALMLSDSQ
      WP_103056255.1 (054)   1 AVTSEADSISYSELNSRANQVARAVLARGIRPGDRVGLCLDRSVDLVVALVGVLKTGASYVPMDPRYPADRLRYTMDDAG
      WP_157514135.1 (070)   1 AITDGNQHLTYAQLEAEARQLAAMLAARGAGPESLVALALPKGIPLITALLAILETGAAYVPLDPTSPPERIAYILADAA
          OQY50160.1 (137)   1 AVVFENQQLTYRSLNARANQLAHYLKRQRVEPEILVGLAVERSLDMLVGLLGILKAGGAYVPLDPAYPKKRLAAILSDAQ
      WP_104981151.1 (120)   1 AVVYAGDVLSFDALNRRANRVAHALRACGVGPEAIVPLVYERSVEMMVGLLGILKAGAAYLPVDPHQPRARLALILGDAR
                                         a1                                              a2

   (KtzH) ABV56588.1 (010)  81 PLLTIAAEDAPV..PE....GEPRLAVDESIVDDGVRVTPPM.......PANPAYVLYTSGSTGRPKGVAVDHAAIVNRL
      WP_055611559.1 (025)  81 PLLVLDGA.LPD..VS....GYPEVNPERAL.........S.......PDNAAYVIHTSGSTGGPKGVQVAHRSIMNRL
      WP_093779539.1 (038)  81 PRCVVVAGAQPE..VDEAAGAHTPLHRLHDLETAGARGTEPTSNGHGPRPADTAYVIYTSGSTGRPKGVAVSHTGIVNHL
          BAP05589.1 (108)  81 VEWVLTQSRFTASLSPLLDDSVHLLSLDREWESITEGAQRQPALKRQVRLHHLAYILYTSGSTGRPKGVMLEHRALCNRI
          ARN83985.1 (085)  81 PRLVLTQARLRDRLP.ERVNALLLDAHDHTISTLSPDNPVPVAA.....PHNLAYVIYTSGSTGAPKGVAIAHQNVRRLF
      WP_144865672.1 (149)  81 VSLLVTQDKLKENLPPHKATVVCLDTDKKDIAQQRKDNPVSDVT.....PENLAYVIYTSGSTGTPKGVLVNHANVMRLF
      WP_103056255.1 (054)  81 IRLAIGTSETVPSLP.SAVSVLRIDEETAAIASFPDGDLSDEERGGRVSPDSVAYVIYTSGSTGRPKGVAVPHRGVVSLL
      WP_157514135.1 (070)  81 PTLVLVDGRTAELVAGSDVARLRID.....IADSNGHGSLDNPHPGR..PGTVAYVIYTSGTTGKPKGVAVTHENAVSLF
          OQY50160.1 (137)  81 VSVLLTQKALMAGLPQH.HAYTVCLDTDLGAISQERNENLINRA....TPNNLAYVIYTSGSTGLPKGVAIQHNSVINLS
      WP_104981151.1 (120)  81 ARHAVVQPGLEDAVRGALGAGAPELQLVSTRVSEGPEHNPSSGV....LPENLAYVIYTSGSTGTPKGTLVTHASAVNLA
                                                                                           a3

   (KtzH) ABV56588.1 (010) 148 LWMRDNYAV...TERDRVLHKTPLGFDVSVWELFLPLVTGATLVVARPDGHRDPAYLAELIRRTASTVAHFVPSMLAAFV
      WP_055611559.1 (025) 137 KWGLAHFDV...GAEDRVLLSTTASFDVSVPELFAPLQVGAAVVVARPDGRKDPAYLAELIQREGVTGADFVPSLLEAFV
      WP_093779539.1 (038) 159 AWMQRHFGL...RADDRVLHKTQLGFDVSVWELFWPLTTGAVLVVAEPGLHKDPAALARLMAEERITTVHFVPSMLDVLL
          BAP05589.1 (108) 161 LWMEDEYGL...SPEDRVLQKTPFSFDVSGWEFFWPLITGARLVLAAPEKHKDPYYLRDTICRSGITVLHFVPSMFHSFL
          ARN83985.1 (085) 155 AATQAEFGF...SSADVWTLFHSFAFDFSVWEIWGALLYGGQLVIVPHWVSRSPESLVELLEKHAVTVLNQTPSSFYQLD
      WP_144865672.1 (149) 156 AATQSWFNF...NASDVWTMFHNCAFDFSVWELWGALLHGGRLVVVPYLVTRSPEAFYKLLCQEQVTVLNQTPSAFRQLI
      WP_103056255.1 (054) 160 TGTQDRYGF...DAGDVWTLFHSVAFDFSVWEMWGALAFGGRLVVVPWEVSRSPGEFLELVVAQRVTVLNQTPSAFEQFV
      WP_157514135.1 (070) 154 QSTRSWCEF...GPGDVWTWCHSQAFDFSVWEIWGALLHGGRVVVVPREVAYSPAELWDLIVEQGVTILNQTPSAFYTLI
          OQY50160.1 (137) 156 NGLYQTIYAHYQNAPLRVSLNGSLAFDTSVKQII.QLLYGHTVDIIPESLRYDGKALLDYLISHQIDVFDCTPSQLELLI
      WP_104981151.1 (120) 157 HAERRMLGLE.AGPPLRVGLNAPLIFDASIKQLL.LMLFGHSLHVVPEEVRPDGRALVEFLRAARLDVLDLTPSQMKLLL
                                                         a4

   (KtzH) ABV56588.1 (010) 225 AEP....AAGDCTSLREVVCSGEVLPDSLRDRYYAVL.GA..RLNNLYGPTEAAIDVTA.A.PCPPQ...........GP
      WP_055611559.1 (025) 214 AEP....TARKCTSLRWIEVAGEAFPAELANRFVELLPGC..GAHNLYGPTEASVEVTGWQ.HVPGA...........DR
      WP_093779539.1 (038) 236 GET....SLSDCTSLRRVITSGETLSAGLCRAFQEAC.DA..PLHNLYGPTEAAIDVTSWECPRPLE...........GE
          BAP05589.1 (108) 238 GIA....EVERCNSIRKVFCSGEALTVQHTTRFFETFGDR..ELHNLYGPTEAAIDVSYYACRKE.............AS
          ARN83985.1 (085) 232 AFEGGR.DEPRLSSLRIVIFGGEALEPRRLAGWYARYGDDSPQLVNMYGITETTVHVTMRRMKREDACG.........AP
      WP_144865672.1 (149) 233 WAENIA.ETTQQLNLRLVIFGGEALEIQTLKPWFDRHGDRTPQLVNMYGITETTVHVTYRPLTKDDLNN.........T.
      WP_103056255.1 (054) 237 RAEAENPGLGRELCLRWVVFGGEALEVSRVRPW.LARHPDTPVLVNMYGITETTVHVTAHPLNPTGKGENGAEVDLVEGS
      WP_157514135.1 (070) 231 ESRRATVDDGSRL..RMVIFGGEALDPTRLQRWPILQRPAAPRLINMYGITETTVHVTEQALTPQHWRG........DAG
          OQY50160.1 (137) 235 ADGL...LSKSLSAPKIVLIGGEPINESTWQAL.AQNI....HFYNVYGPTECTVDATVCQVRT..............SA
      WP_104981151.1 (120) 235 FAEI...ERNIEVLPARVLLGGEALDADTWERLRALGD....RFHNVYGPTENTDISTMIPLSA..............SE
                                                                      a5

   (KtzH) ABV56588.1 (010) 285 V.TIGEPVWNTRVYVLD..QWLAPVPAGTAGELYLAGDQLARGYVNRAGLTADRFVANPFAPA....GGRMYRTGDVVRW
      WP_055611559.1 (025) 276 V.PIGAPIWNTQVYVLD..SALRPVPPGVTGELYLAGMGLARGYLGQSALTANRFVACPFGEP....GTRMYRTGDAVRW
      WP_093779539.1 (038) 298 VVPIGRPVDNTGIHVLD..EALRPVGPGVTGELCVTGVQLATGYVNSPGLTAERFVACPFGDP....GQRMYRTGDLAWW
          BAP05589.1 (108) 299 VVPIGKPIANIRLYILD..SRMNPSPIGSEGELCIAGIGLARGYLNRPDLTADRFFPDPFAST...PGAHLYRTGDLARW
          ARN83985.1 (085) 302 DNAVGRPLNDLLAYLLD..GRMQPVPIGAAGELYIGGAGLARGYLGRADLTAERFLPDPFGGA....GARLYRTGDLARH
      WP_144865672.1 (149) 302 GNAIGVPIPDLQVCILD..RYLQPLPIGVPGEMYIAGSGLARGYLHRPELTAERFIEKPNVGAIHEQDCRLYKTGDKARY
      WP_103056255.1 (054) 316 AGVIGHAVPGWRVYVLDRAGGLRPVPPGVVGEMYVAGAGVAQGYAGRAGLTASRFVADPFAGD....GRRMYRTGDLARV
      WP_157514135.1 (070) 301 ISPIGVPLANLRVYVLD..SWLRPAPVGVAGELYVAGNQVARGYLSRPALTASRFVADPFDPA....GGRLYRSGDVVRW
          OQY50160.1 (137) 293 QPVIGRPLPNVQTYILD..HQLQPVPNGEAGELHIGGVGLARGYLNHPELTAEKFIPNPFYND...PESRLYKTGDLTRY
      WP_104981151.1 (120) 294 RPTIGSPIANVRVHVLD..HHLSPVPVGVPGELYLSGKGLSRGYFGRSALTAERFLPSPFGPE...PGARLYRTGDLVRH
                                                                a6                                     a7

   (KtzH) ABV56588.1 (010) 358 TG................DGRLEYLGRADHQVKIRGQRIELGEIENVLAGHPGADRVLVTAGD..........SGTGDRQ
      WP_055611559.1 (025) 349 NQ................DGQVEYIGRTDFQVKIRGFRIELGEIETVLSRHPAVRQAAVIARE..........DVNGGRQ
      WP_093779539.1 (038) 372 NR................DGDLEYLGRSDDQVKINGVRVEVAEVEAALMRHPAVRRAAVVV.D..........GPPDARR
          BAP05589.1 (108) 374 LP................DGNIEYLGRIDNQIKLRGNRIEPGEIEVAFREHEAVSECAVVVQQ..........DERGERR
          ARN83985.1 (085) 376 RA................DGDIEFLGRIDHQVKIRGFRIELGEIEAVLRRHPSSGDAVVIAQE..........EGARGGR
      WP_144865672.1 (149) 380 LPALKDTASLKDSLRVAHNGSIEYLGRLDNQVKIRGFRIELGEIETAIAQHPGVRETVVIPRE..........DEPDNKR
      WP_103056255.1 (054) 392 DG................QGRLSYVGRADEQVKVRGFRIEPGEVAARLGEHPGVAQAVVVVRE..........DTPGDKR
      WP_157514135.1 (070) 375 TE................AGVLEYLGRADEQVKIRGYRIELGEIESLLAGQECVAGAVVTTHT....AGPATPDAPPSTQ
          OQY50160.1 (137) 368 LP................EGTLEFLGRTDSQVKLRGFRIELGEIETVLTQHPDIQQAVVTVREEALTPAISQTENATDKR
      WP_104981151.1 (120) 369 RE................DGLIEFIGRVDRQVKIRGFRVELGEIEAALRRHPGVREAVVVLRED..........QPGDPR
                                                              a8

                                  A/M s(O, Ser/Thr, a8-a9) boundary
   (KtzH) ABV56588.1 (010) 412 LVAYAVADPGTAAPAATLSRLRRQGRLDG..RELHELPNGMLVCARNRSNTAFLYDEIFVRDEYLRAGVSLPERPCVVDV
      WP_055611559.1 (025) 403 LAGYLVVDSEVAPAVARFHEMQSTGRLAG..VEFHELPNGMAVAGRNKLNMSFLYDEIFQRGEYARGGVTVGDGACVVDV
      WP_093779539.1 (038) 425 LTAQLTLDSRVAPGADRLVRLLELPEESR..PRVVTLAPRLPVCVHQESEARFMFREIFEDEVYVQQGVTVPDGGCVLDV
          BAP05589.1 (108) 428 LVAWFSPDARHARAILKHMEMEAQGLLAD..RKRHTLPNGMEIIHLNPGETDFLYQEIFEDNGYLKHGLAVLPGDCVFDV
          ARN83985.1 (085) 430 LVAYVTPNASQGAKLGRLLRMQKDGVIEN..LLQFEFPDGGLALHKNQSEIEFLHKEIFSDKCYAKHGIRIGPGDCVFDV
      WP_144865672.1 (149) 450 LVAYIVPDEKQAAPVLQLIRFQEQGRLTD..KSLYELPNGMAIAHLNKNETDFVYQEIFEQLSYLRHGITINDGDCIFDV
      WP_103056255.1 (054) 446 LVAYLTPDTTTAAPVARYTRLAATGELAATGAELHELPNGMLVLGRNRSNIAFLYQEIFANTEYVRGGIHLDPGAVVIDV
      WP_157514135.1 (070) 435 LVAYLVPDPRFAPVAAKCLELQRQGAIAA..AELHELPNGMPVVGRNRSNIEFLYEEIFETEEYCHGGVTIPGDGCIVDI
          OQY50160.1 (137) 432 LVSYIVPKPSRAPTIAGKPR.............YTLPNKLAIAHLNQNETDFLYKEIFEVKAYLKHGITIKDGDCIFDV
      WP_104981151.1 (120) 423 LAAYVVPSPKGSPAALGRPR.............HVLPNDMAVVQQNKNETDYLYEEIFRKKTYSLYGIEIPRGAVVFDV
                                                                                       m s(O, Ser/Thr, a8-a9) i   m s(O, Ser/Thr, a8-a9) ii

   (KtzH) ABV56588.1 (010) 490 GGHVGLFSLFVKTRRPDCRIYAFEPIPELAEMFRINAELHDIDAVVTNCGVGATAG.TARFTYYPDMSMLSGR..FADER
      WP_055611559.1 (025) 481 GGHVGMFGLWAGSVARGVRVYACEPMPESAEFYRINAYLHGLDVVVSDYGISDAAG.RAEFTYYPEMSLMSGR..FADEA
      WP_093779539.1 (038) 503 GANIGLFTLFVGLQAPDVRVYACEPVPEVFDVLRRNVNLHGVDAVLLPYALADEPRESADFTYYPGVSIISGR..YASTA
          BAP05589.1 (108) 506 GANIGLFSLCMARE..QVQVHAFEPMPEVYEVMRLNTRMYSDTIRTYPCGLSDAPG.QETFTYYPNVTILSGA..HADLE
          ARN83985.1 (085) 508 GANIGMFSLYAAQQADDVKLFAFEPIPEVFEALRANAALHGLNARLFDCALGDAER.LVTFRYYPHATVLSGM....GGA
      WP_144865672.1 (149) 528 GANIGLFTLFAAKICSDVRVYAFEPIPPVFELLRLNTDLYGLNVKLFDVGLSAEST.EAKFTYYPHVSVISGR..FADAS
      WP_103056255.1 (054) 526 GGHVGLFSLFVTHQVPDARVYAFEPIPELARLYRLNTALHGADTVLSECGVASEPG.VAEFTYYPEMSIMSGR..FADQD
      WP_157514135.1 (070) 513 GAHVGMFSLYAKRKSPGVKLFAVEPMPELRRMYAANAALHDLDATLVPYAIGAEAT.TDTFTYYPEMSVLSGR..GTGSR
          OQY50160.1 (137) 498 GTNIGLFSLFAHINGNNVKLYGFEPNPFVFEILELNTKLYEVDAKLFNCGLSRQTT.TANFTFYPKFSFLSGL..YADTA
      WP_104981151.1 (120) 489 GANIGMYSLYVTLHAEDARIFAFEPLAPIYATLAANAALYGRDIKLFHHGLSDREE.TVAFTYYPRYSMMSGAEAFARPE
                                                 m s(O, Ser/Thr, a8-a9) iii                         m s(O, Ser/Thr, a8-a9) iv
```



```
(KtzH) ABV56588.1 (010) 567 EERRMLERVLRNERLAD..L.DD....GVLDELLAERLRGQQVDVELRTLSDLIREQGIDRIDLLKIDAEKSELDVVRGI
   WP_055611559.1 (025) 558 TDRETLRRVIANARGPK..AAED....DQLTELLAQQLSSTVVDVELRTVSQVIRDNGLTVVDLLKVDAEKSELAALRGI
   WP_093779539.1 (038) 581 ADGAVLSDYLRGELATG..GEADTGVGDAVGDLVEDRLRARRLTVPVRTASQIIDEEDIDRVSLLKIDAERSELDVLAGI
       BAP05589.1 (108) 581 EEIGVVRSYLSGENDGH..EHQVND..AEVDELLRARLETREVRCQLRTISEIIAEQGIKTIDLLKIDVEKMEHKVIEGI
       ARN83985.1 (085) 583 DVQGVVKSFL..KFGSG..EGA.SD..RDLDELLATRLEFENVDCKMRTLSSVLRETGVERIDLLKIDVENAEFEVLQGI
   WP_144865672.1 (149) 605 EEQEVVKSFLLNQDTSD..ETALSP..TAIEQLLAERLQSQQFTCKLKTLSEIIEENGIEQIDLLKIDVEKSELEVLNGI
   WP_103056255.1 (054) 603 EERQVLEQFVRNEHEEV..G...GD..QGLGQLLSHRLASQRVRAELTTISEVIDARGLARVDLVKVDVEKSELDVLRGV
   WP_157514135.1 (070) 590 EQHDILQSFVHSEYQAE..TQDLGD..RLVDQMIADRLDCVQVDVPIRTLSHIMDEYGIGHIDLLKLDVEGSELDALHGI
       OQY50160.1 (137) 575 DEKEIVKSFLLNQQGAAEAASLAQQNDELLEELLADKFKSKTFTVQLRTLSDIIQQHNIEHIDLLKINVEKSEWDVLAGI
   WP_104981151.1 (120) 568 EEVEVVKRFLTKETEAG..SSEAAQLLSHADELLAGRFQGEVHNCLLRRLSDVIREEGVDHIDLLKVDVQRAEMHVLRGL

(KtzH) ABV56588.1 (010) 640 EPEHWAIVRQVVAEVH.....DIDGRLKVFLDLLRERGFDTAVEAEADLAGTGMYSVYAVRPGAALAPM.PAPADNS.S.
   WP_055611559.1 (025) 632 EPEHWPIIRQVVAEVH.....DIDDRVAVVTKMLAAKGFHVVVETPQGLEGTGMCQLYATRAGHE.APA.PTPATPVLA.
   WP_093779539.1 (038) 659 EDRHWALIDQVCLEVH.....DSDGRPDRVRALLTAQGFEVRARRPADDNLGDLVTFYAVRPGGASAGAQPRPAARPRS.
       BAP05589.1 (108) 657 AAKDWPKIRQMVIEVH.....DIENRAQQMQALLAAKGFTVVVEQDDALRNTPIYSLYARRTSVRTDQPHVNTDRQ....
       ARN83985.1 (085) 656 EELDWSKIEQLVVEVH.....DVDDRRRKILELLRRHGFDAVEDQDVELGGTGLYNIYARRPGA.....VTNIERSPPR.
   WP_144865672.1 (149) 681 KEEHWQKIKQIVVEVH.....DRDGRLAEITQMLARHQYEFTIAQENELRETELYNIYARRESIGLSRKTDKVVSTKKP.
   WP_103056255.1 (054) 676 RDEHWPRIGQVVAEVH.....DAQGRLAGVRGLLESVGFTVVVDHAPDVADTGLYMVYATRPGWSTEPATPVEPATAGGA
   WP_157514135.1 (070) 666 EARHWPLIDRMVIEVN.....DVEYRLARIEELLRHHGFEAEARVSPNVMGTDLHIVYAVRPHTRTRAAGEVDAAELLVP
       OQY50160.1 (137) 655 AEADWHKIHQIALEVH.....DINGRLEQVLTLFEKQGYQTVVEQDWSFDKTAGINYYLY..AIRESRQDLKARQSKKVL
   WP_104981151.1 (120) 646 TAGDWKIIDQVVMEVHAAEGQESEGRLEELEALLRQHGFEVILEQDELLSGTDRYNVYASRRGLRGREEGLMSLRPLQGY

(KtzH) ABV56588.1 (010) 712 ...........TWYDPDRLADDLRARAKAQLPDYMVPAAVVVLDAFPTTANGKLDRAALPAPT
   WP_055611559.1 (025) 704 ...........GWHSPARLIHAVRAHLERQLPDYMIPTHLVILDRLPLTSSGKVDSHALPAPD
   WP_093779539.1 (038) 733 ...........AFLD..ELLTEIRDKAAQVLPLSMLPSQMTAVSRLPLTANGKLDRARLMTAS
       BAP05589.1 (108) 728 ............WYNRAGLVRDIELHLASRLPEYMQPSAIIPLVEIPLTTSGKLDRGQLTTRE
       ARN83985.1 (085) 725 ...........LLASPNRWVEELRELATSQLPDYMIPSAFVTLDALPLTSNGKIDRKTLPAPD
   WP_144865672.1 (149) 755 ...........SWNSKNTLIRDIRDRLKEKLPEYMIPSAFSGLSNLPLTANGKVDYQALPRPE
   WP_103056255.1 (054) 751 DPGVGPGGRAVRWWNPGQLVGSVRETAREVLPEHMVPAAFVVLDQLPLTVNGKLDRAALPAPT
   WP_157514135.1 (070) 741 NE.........TWSAPGALGALMRERLGERLPGYMVPDVVMVLDSLPLTVNGKVDKAALPEPV
       OQY50160.1 (137) 728 S..........KLQEPILTCSELRDFLNKRLPDYMVPSAFVMLEGLPLTPNGKIDRQALPAPD
   WP_104981151.1 (120) 726 I..........DEARPV.SPGALRDHAREHLPEHMVPSAVVLMDGIPLTRNGKVDHRLLPVPD
```

Fig. S14. Multiple sequence alignment of family 2a ($A_8M_{s(O,\ Ser/Thr)}A_9$) interrupted A domains using the ten representative sequences indicated in Figs. S2 and S9. The numbering next to the accession number for each sequence matches the number in the taxonomic tree (Table S2). The A domain portion of the sequence is highlighted in light grey and the $M_{s(O,\ Ser/Thr,\ a8\text{-}a9)}$ portion in light blue. The divisions between the two are indicated by teal triangles. The conserved motifs for the A domain are underlined by a dark grey bar and labeled a1-a10. The conserved M domain motifs (outlined in Table 1) are underlined by a dark blue bar and labeled $m_{s(O,\ Ser/Thr,\ a8\text{-}a9)}$i-vi. A shortened version of this annotated alignment is presented in Fig. 7A.

```
                             .        .         .                     .         .         .         .
(ThxA2) QDA77059.1 (029)   1 FEAQAARAPGAVAAVSEGRP............LTYRELDARASEVASRLRGEHGLQPGERVGVVMGRNAAWVAVLLGVLK
        GDX84814.1 (061)   1 FEQQVLRTPNQVALMSPTVS............LTYEQLNQQANGLAEKLQSLNL....ERAVVAVYSERIDVVLMLACLK
    WP_143010572.1 (004)   1 FEAQVKKTPEAIALTFENTS............LTYLELSCISGQFAAYLSRIYHPVPNDLIGLKLPRTEWMVIGMLGILK
    WP_127840945.1 (037)   1 FEALARSQPTAPAAVFGTQV............LDYDALNRRANRIAHALIER.GVRADQRVAICVERGLDMVAGLLGILK
        NET65459.1 (016)   1 FEEQVEKTPDAIAVVFEEQK............LTYSQLNSKANQLAHYL.QKLGVVPETLVGICVERSVEMIVGLLAILK
    WP_083630634.1 (027)   1 FEQQVAKTPDNIAVSFESYREFLTANSNGLKTLSYQQLNEKANQLAHYLIQNHNIKPDTLIGICVERSSEMIIGLLGILK
    WP_109330483.1 (048)   1 FEERVRHSPDAVAALFAGTG........GDRSLTYAGLNAAANRLARRL.KALGVGPESRVAIAMEPSLARLAAVLAVLK
        TMC21728.1 (006)   1 FEQQVEQTPDAIAIVFEETV............LTYRELNRQANQLAHRLLGE.GVSPDVLVGVCMEHSLELVIALLAILK
        MAE67645.1 (003)   1 FEQRAAEAPDAIAIVFNDTS............LSYRELNERANRLAHHLQHL.GVGPDVLVGSCLPRSCELIVCILGILK
        MAF83146.1 (069)   1 FENHAASRPYVVALTIGERH............LTYGDLNTLANQFAHALIGA.GVVAEVRVGLHLPPGAELIAAMLGIMK
                                                             ------                                        --
                                                               a1                                          a2

                              .         .          .         .                                        .
(ThxA2) QDA77059.1 (029)  69 AGGVYVPVDPSH.PRERIELVLADAHCGCVVTGTDG...............................VTITRRRERPAEP
        GDX84814.1 (061)  65 IGAVYLPIYRDL.PVQRIQFMLQDSDCKLILHRTETP...................LPEALSTYANLAWSNEVENEAANP
    WP_143010572.1 (004)  69 TGAAYLPIDPDY.PQERLDYIIADSACKVLVDEEEL.............................KRFQQAIDQPDGAEF
    WP_127840945.1 (037)  68 AGAAYVPLDPAY.PAARLEHMVRDSAPVALLSQQALAPLAGT......LPV......GEAPVLLLDDGGFEGLPE..HDP
        NET65459.1 (016)  68 AGGAYVPLDPAY.PQERLEYMFGDSQVSVLLTQQQLLA...Q......LPQ......HQAKVVFLDRDWEKIATETPEKV
    WP_083630634.1 (027)  81 AGGAYVPLDPTY.PQERLAYMLSDSGVSLLLTTEKLAT...G......LPE......HQVHVVCLDTGWSVIAEQSLENP
    WP_109330483.1 (048)  72 AGGAYVPLDPAF.PRDLLHELVDDCAAAVVLTDAALAPTLAGCAAAHLLTL......DEGLSAAAGESAEDLPENLPEEP
        TMC21728.1 (006)  68 AGGAYVSLDPELLPQERLTFLLEDAQVSLILTQTHVRERLPRVAVKMISVVATSGGQDPCAASTVPQQGSPSEHWHADNL
        MAE67645.1 (003)  68 AGGTYVPLDPAY.PHARLASMLEDARPPVLVTTSDLRDQLPG...............TGATVVSLDTDWSAISGQPPDTP
        MAF83146.1 (069)  68 AGGAYVPMNPDL.PTNRLAQILELAEPAVIITQPTAADLPG................HAAQLIYADSDPVKASGS..DDP
                             ----------

                                .         .         .         .         .         .         .         .
(ThxA2) QDA77059.1 (029) 117 LEDGTSRSLAYILYTSGSTGKPKGVMIEHRGLVNMALDQIRVFGITPDDRILQFSSIAFDASLAELFSALLSGACLVLPG
        GDX84814.1 (061) 125 KREISPDDVAYILYTSGSTGHPKGVKVKHAGFVNMALAQISGFTVTAEDTVLQFASMAFDASLSEVFMALLSGAKLVTIE
    WP_143010572.1 (004) 119 LVEAKPSDLAYLIYTSGSTGYPKGVMVSHGNLSNFLGHVKRNYAPDAPLSLPFVASNSFDISVFQLFCSLLSGGRSVLLT
    WP_127840945.1 (037) 133 ELPIDAASLAYVIYTSGSTGTPKGVMVEHKGLSNLAQAQIRVFDVQPDSRVLQFASFSFDACISEIAMALCRGATLYLAS
        NET65459.1 (016) 132 KSEVSPQNLAYIIYTSGSTGKPKGVLVEHKGLCNLVTVQIQAFQVNSNSRVLQFASLSFDASIWEIVMALGSGASLYMGS
    WP_083630634.1 (027) 145 ITETTAANLAYVIYTSGSTGKPKGVMVEHKGLCNLTQALIQTFGLHSDSRVLQFASFSFDASIAEILMALGSGATLYLGT
    WP_109330483.1 (048) 145 DGTAGPEHAAYLIYTSGSTGRRKGVLVTHRSLRLLVEAQVDAFRIRAGGRVLQFASFSFDASVSEIFTALLAGATLCLAA
        TMC21728.1 (006) 148 NCSVQPENLAYLIYTSGSTGRPKGVMVTHRSISNNMHWRQSVFQFKPGDSIVQRTPLSFDVSVGEIFFPLLTTGARVIMAH
        MAE67645.1 (003) 132 DCRATGDDLMYVIFTSGSTGPPKGVMVTHRNVVHLYRGLPPSLCFDATDIWTMFHSCAFGFSVWEIWGALIHGGRLIVLP
        MAF83146.1 (069) 130 HMPVASDDLCYVMFTSGSTGAPKGVMVTHGNLAPLFDDIGARLDIDAGDVWTMFHPLSFGYSVWEIWGALCHGGRLVIVP
                                     -----------------                                 -----
                                            a3                                          a4

                                .         .           .         .          .           .         .
(ThxA2) QDA77059.1 (029) 197 EERIRDPGLLLPFMAEQRI..TVVSLPPSFLGALGRAPLPT.LRALIT..MGEAANSGDALHYASRLRYFNA.......Y
        GDX84814.1 (061) 205 QAIINDTTEFVAYLARYGV..THITLPPVYLQALQRQPLPT.LKTLIT..AGEAANIDDALFYAKQCNYFNA.......Y
    WP_143010572.1 (004) 199 KDQVQDPAQFVEVLKGVNAIDTVPGVYQGLVSYLTDQGLSA.DFAHISHLFIGGDSIPDALLHQLSKIFSHA..KLTVTY
    WP_127840945.1 (037) 213 REALLPGEPLLATLSSHAI..THVTLPPVAAGILAPHADSE.ALRALRTLAVAGEACPAALV.RRW..AKGR..RMINAY
        NET65459.1 (016) 212 RDTLMPGAGLSQWLQSKKI..THITIPPSALAVMP....KE.ELPDLKTIVVAGEACPPELI.SQW..SVGR..QFVNAY
    WP_083630634.1 (027) 225 KDSLLPGQPLIEQLRNHCI..THITLPPSALAVMP....GE.KLPALQTIIVAGEACSAELI.KQW..SIGR..NFFNAY
    WP_109330483.1 (048) 225 REALLPSAEMLRLLARWRI..DTVTLPPSVLARLP....DG.ELPDLRTLVVAGEPCPADLA.ARW..APGR..RFLNAY
        TMC21728.1 (006) 228 PEGHKDSTYLLKLAAEQQA..TLLNLVPSLLQVCVEEPGLE.QCLSLRHMLCGAEAFPAELQ.ERFFACNHA..ELCNMY
        MAE67645.1 (003) 212 TESCLSPSALLDELCRRRVTVLSQTPSAFRQLLLPEVPTETADRLCLRMIVFSGEPVVSGDLRRWFDRARGRAVQLIDTY
        MAF83146.1 (069) 210 PELRTDPARLFALVCSQHVTMMSQTPSAFRQNLLSDIFADGIAATALRCVVLSGEAVDTEALQRWFRQHGDAEPRIVNTY
                                                                                                      ---
                                                                                                       a5

                                  .         .           .         .              .         .         .
(ThxA2) QDA77059.1 (029) 265 GPTEASVCATIHEVMPGDEYAG..GIPIGRPIANTEILIL.....DEHLKPVPIGVPGEICVAGVGLARGYLDRPELTAA
        GDX84814.1 (061) 273 GPTETSVCASFYRVNPQGDYSQ..RIPVGSAISNLEFYLL.....DDALNQVSIGVTGEIYIAGIGVAEGYINAPELTAA
    WP_143010572.1 (004) 276 GPTEGTIFCTHLIYNPGEINRGIKGSIIGKPIDGAKIYLT.....DQQGLLCPVGVIGEICIGGKGLSQGYLNRPQLTSA
    WP_127840945.1 (037) 285 GPTEATVCASMQVCEPGETDG...AVPIGAPMDNVRIHLL.....DELGQPVPIGVAGEIHIGGIGLARGYLNLPEMTRE
        NET65459.1 (016) 280 GPSESTVCATMAECSP..ECS...VPAIGRPISNTQIYIL.....DRNLQPVPIGVAGEIYIGGVSLARGYLNQPQLTQE
    WP_083630634.1 (027) 293 GPTEASVCATIAKCTL..EDE...KISIGKAIANTQVYIL.....DENLQPVPIGVPGELHIGSVGLARGYLNRPELTQQ
    WP_109330483.1 (048) 293 GPTETTVCATIAAVEP..DGA...RPPIGRPSGDSRVYLL.....DGDLNPVPAGVPGHLHVGGPGVSRGYAGRPGQTAA
        TMC21728.1 (006) 302 GPTEATIEVTRWICERENTES...TVPIGRPIANTQIYLL.....DIHLQPVPIGVKGEIYIGGLALARGYVNRADLTAE
        MAE67645.1 (003) 292 AITEAGGQIACRLIEPSDIDDSG.TLDIGDILSGTTVHIL.....DEQLRPVPTGSTGELHVAGPGLARGYLNRPALTRD
        MAF83146.1 (069) 290 AITETAGQLTFAACETADRSAGV.AVGVGRPLAHAELVIISEECADQEWLPVPTGQAGELFVGGPSVARGYIGVEEQTSQ
                             -----                                                    ----------------
                              a5                                                            a6

                               .           .         .         .         .         .         .         .
(ThxA2) QDA77059.1 (029) 338 AFVDRPG..RPGERIYRTGDVGRFRSDGAIEYAGRRDAQVKVRGHRVEPGEIEAALRAHARVADACVVPWQRGEGALDLA
        GDX84814.1 (061) 346 SFVSIPQ..FSNSALYKTGDLGCWNDAGQVEYFGRKDQQVKVSGYRIETAEVDFALRQFPAINQAVVTTQRDANGQNALC
    WP_143010572.1 (004) 351 QFVDCLF..EPGQQMYKTGDLGRWLVDGEIEFMGRKDNQVKIRGYRIELGEIEHVLRNHEEVNDVVVLPKKDQDGNSFLA
    WP_127840945.1 (037) 357 RFVEDPFSGAAQARLYRTGDLGRWLPDGSIEYLGRNDTQVKLRGFRIEPAEVEARLRLHPQVKDAIVMAREDAPGDPRLA
        NET65459.1 (016) 350 KFISHPLRNLTEERLYKTGDLARYLPDGNIEFIGRIDNQVKIRGFRIEPGEIESVLRRHPQVQQAVVIDREDTPGNKRLV
    WP_083630634.1 (027) 363 KFIPNPFSNEAYSRLYKTGDLGRYLNDGSIEFLGRIDHQVKIRGYRIELGEVENVLSQQSGVHKAVVVTREDQPDNKRLV
    WP_109330483.1 (048) 363 AFIPDPFAAEPGARLYRTGDLARFREDGSIEYLGRLDEQIKIRGIRIEPGEVEGVLRADPAVYEAAVLPVGEGDATS.LV
        TMC21728.1 (006) 374 RFVPNPWSILPGERLYRTGDLARYRSDGAIEYLGRIDQQVKLRGHRIELGEIEAVLLQHPAVDEAAVLVQGERPEDKRLT
        MAE67645.1 (003) 366 RFPDDPFSRTDGGRLYRTGDRARILPDGRIALIGRVDDQVKIRGYRVEPGEIEAALSDHDAVHESAVVCRRDASGDNRLF
        MAF83146.1 (069) 369 RFITISIDGCQATRWYRTGDRARLTSEGELEFLGRTDEQIKLRGYRIEPGEIAAVLREHPAIDDAAVALRSDYGGEPRLV
                                            -------           ---------------------
                                              a7                       a8

                                  ▼A/Ms(O, Tyr, a8-a9) boundary
                                 .                           .         .         .         .         .
(ThxA2) QDA77059.1 (029) 416 AYVIERSR..................VEPWPSVAEFFVYDDVLYGSMAGHEERNRAYRAAFERSLAGKTVVEIGPGPHAI
        GDX84814.1 (061) 424 AYYTRAKK..................VELFPSIAEFYVYDDVVYGSMATDYARNDRYRAAFGKVLNGKTVVEIGPGPEVI
    WP_143010572.1 (004) 429 AFIEKKKR..................VFLTPSTSEYFVYDDLLYYALTSNDKRNFHYKEVFKQLLKDKVVLDLGAGGDAI
    WP_127840945.1 (037) 437 AYVTATDR..................VELWPSVSEFYVYDALLYHAMANHTTRNRCYANAFARHLPGKTVVEIGPGPHAV
        NET65459.1 (016) 430 AYFVPEPE...LEQNIQKMKG.....IELWPSIAEYYVYDELLYYAMTNDHRRNQSYQVAINQLVKDRVVVEIGTGKDAI
    WP_083630634.1 (027) 443 AYVVPEQK...NLESSQVAENLSTDKVELWPSVAEYYVYDELLYYAMTNDHRRNDSYKVAINQLVLDKIVVEIGTGKDAI
    WP_109330483.1 (048) 442 AFVVPDHG...RPAAKP.........QEWWPSIAEFFVYDELAYHAMTSDERRNIAYRAAMAAMVPGRVVLEVGTGPEAL
        TMC21728.1 (006) 454 AYLVLQPE...QEK............VELWPSVGEYPVYDEHIYSEMTHNDARNALYKAALAQTVAGKIVVDVGTGRNAI
        MAE67645.1 (003) 446 GYVVPIPEYADAPKKQSTANADDHDRIELWPSLGEHQLYDEVLYDLMAADPVRLDQYRKAIDQHVRDRVVVDIGTGQDAL
        MAF83146.1 (069) 449 GYVVPHAGLLPEPAT..........RPEFWPSVGPYQLYDEFLYDLMSSEAERLKSYREAFDRSVRGKVVLDIGTGEHAL
                                                                                                   ----------
                                                                                                   ms(O, Tyr, a8-a9)i

                               .         .         .         .         .         .         .         .
(ThxA2) QDA77059.1 (029) 478 LSRMCVEAGARRVYAVEYLKESYEKARETVRSLGLEDRITVVHGDVREVELPEPADWCVSEIVGGIGGSEGAALLVNSAR
        GDX84814.1 (061) 486 LSRLCLEAGAEKIYAIELLEETYLKAKQTLKRLGLEDKIQLIHGNAMTTELPEKVDYCISEIVGGIGGSEGAAKIINSAR
    WP_143010572.1 (004) 491 LSQFCIEAGAKKVYSVEILLETFNKAKKVIEDLGLQDKIILIHGDISNIQLPEEVDYCVSEIVGSIGGSEGASKLINETR
    WP_127840945.1 (037) 499 LSRMAIDAGARKVYAIELLERSYLAAREKVRSLGLEDRIVVLHGDATRIELPEPVDYCISEIIGNIAGSEGSAVIINNAR
        NET65459.1 (016) 502 LARFCAQAGANKIYAIERDRETSQLAKACVDKLGLSAQIEIIHGDAMTVNLPELADVCVSEIVGAIGGSEGAAVIINNAR
    WP_083630634.1 (027) 520 LSRFCAQGGAKKIYAIERDKETCRKAIACIEKLGLADKITVIHGDATLVDLPELADVCVSEIVGAIGGSEGAAVIINNAR
    WP_109330483.1 (048) 510 LSRFCVEAGARKVYAIEKLEDSFNRATARVRELGLEDRITVIHGDATEVTLPEPAEVCVSEIVGAIGGSEGSAVIMNGVR
        TMC21728.1 (006) 519 LARLCVEAGAKRVYAIEMMEESYQQAQAWVNRLGMQEQIKVLHGDATQVQIPELAEVCVSEIFGAIGGAEGSGVILNNIG
        MAE67645.1 (003) 526 LARFCVEAGARKVYAIEIIDEAFRRARSLVESLGLADRIDVIHGDATRLQLPEPAEVCVHGLFGNIGSADGIVPIMNDAR
        MAF83146.1 (069) 519 LARMCIDAGAHKVYAVEILEDAYRKARTLINDQGLADRIIVMQGDMASVDLPERVDVCTQGIIGNIGSADGIVPIWNSAR
                                                                        -----------------------
                                                                              ms(O, Tyr, a8-a9)ii
```

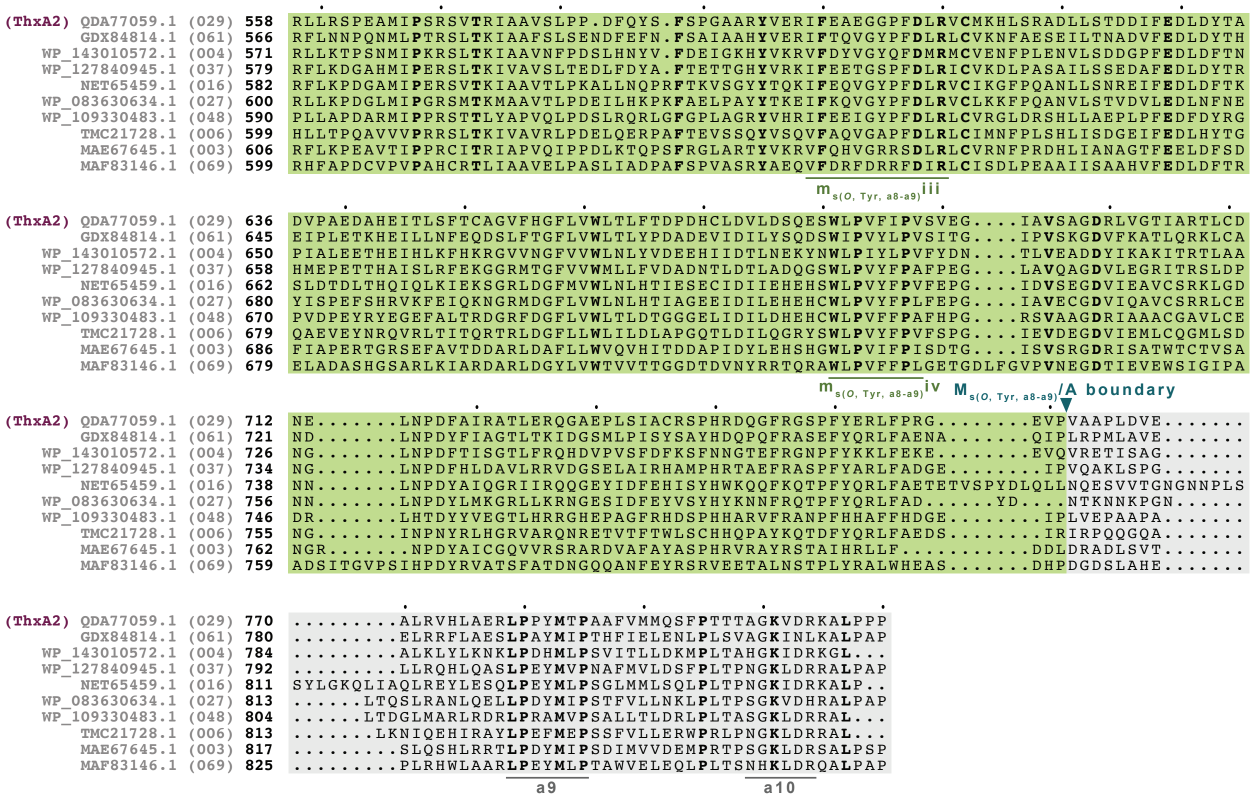
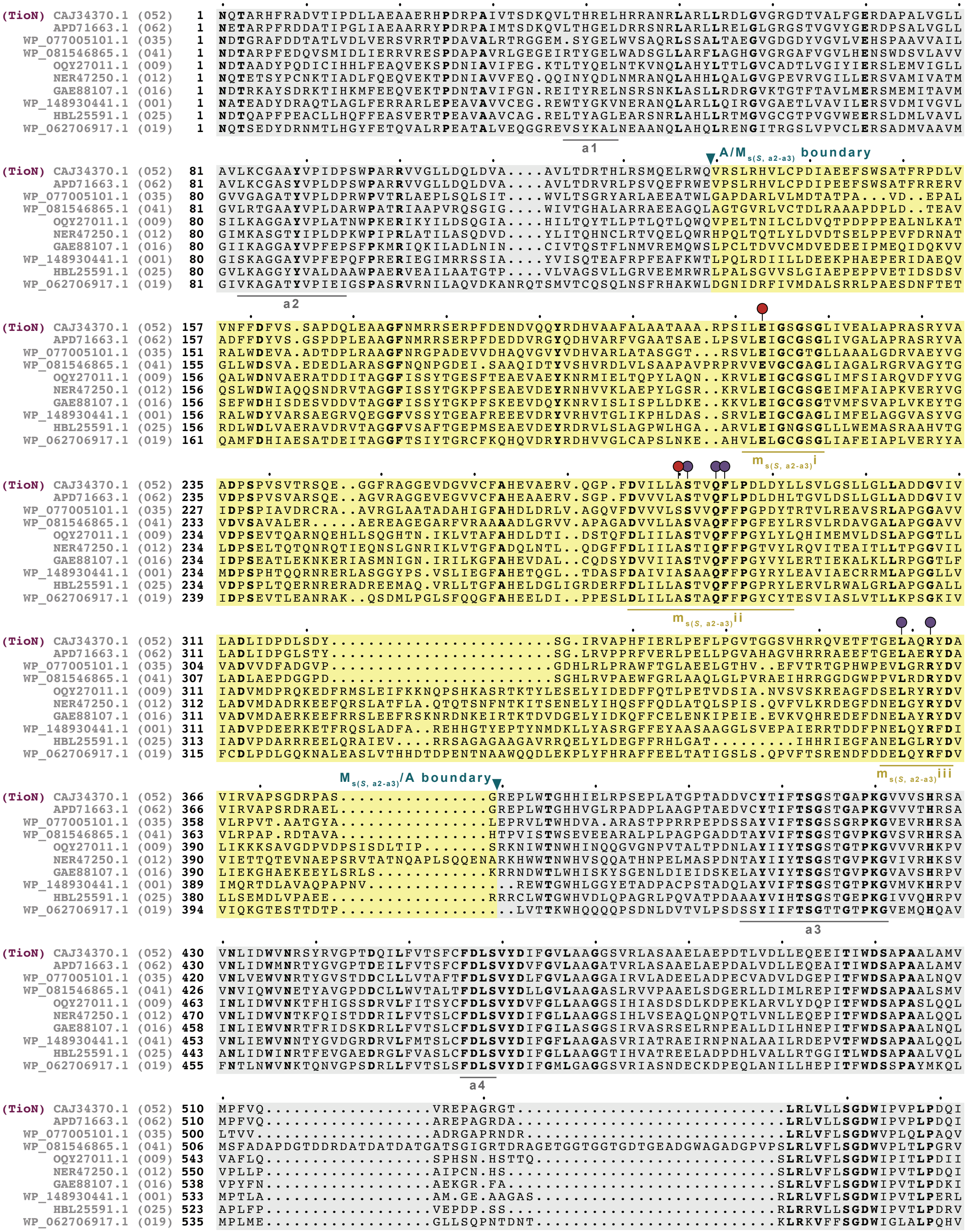
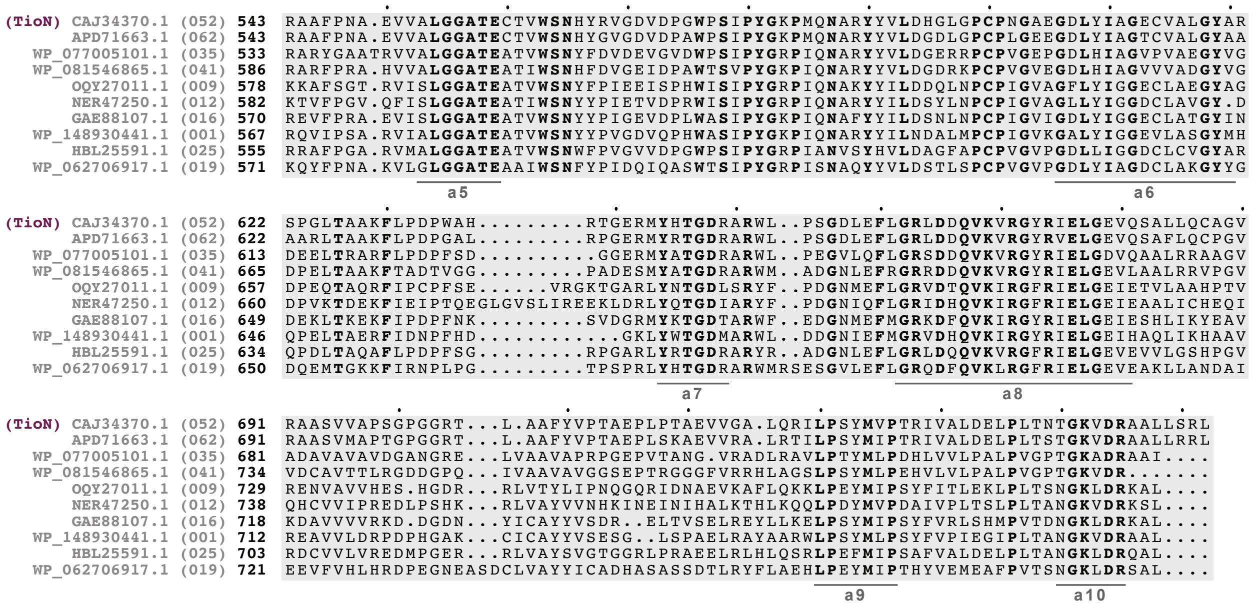


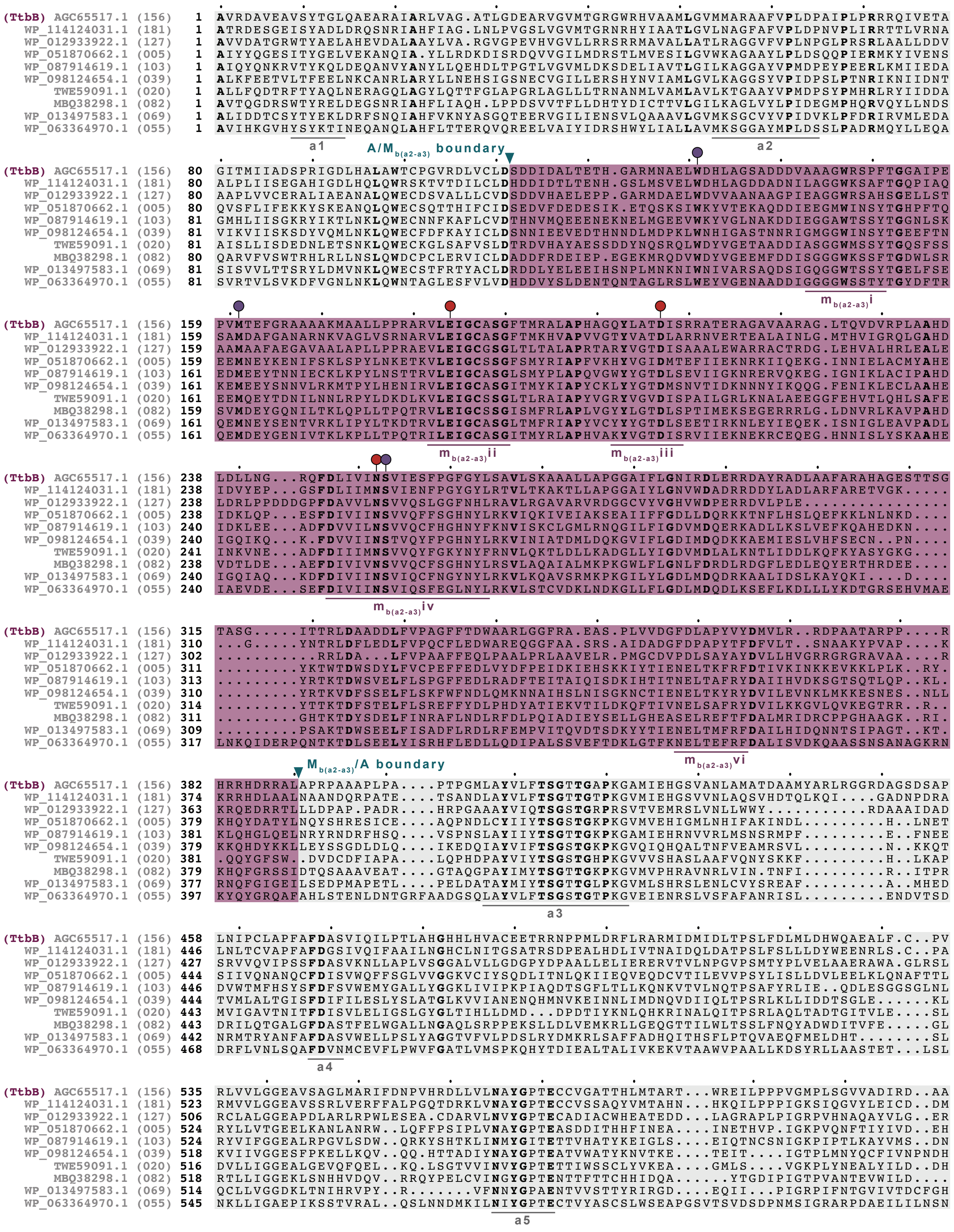
Fig. S15. Multiple sequence alignment of family 2b ($A_8M_{s(O,\ Tyr)}A_9$) interrupted A domains using the ten representative sequences indicated in Figs. S3 and S10. The numbering next to the accession number for each sequence matches the number in the taxonomic tree (Table S3). The A domain portion of the sequence is highlighted in light grey and the $M_{s(O,\ Tyr,\ a8\text{-}a9)}$ portion in light green. The divisions between the two are indicated by teal triangles. The conserved motifs for the A domain are underlined by a dark grey bar and labeled a1-a10. The conserved M domain motifs (outlined in Table 1) are underlined by a dark green bar and labeled $m_{s(O,\ Tyr,\ a8\text{-}a9)}$i-iv. A shortened version of this annotated alignment is presented in Fig. 7B.


(TioN) CAJ34370.1 (052)
APD71663.1 (062)
WP_077005101.1 (035)
WP_081546865.1 (041)
OQY27011.1 (009)
NER47250.1 (012)
GAE88107.1 (016)
WP_148930441.1 (001)
HBL25591.1 (025)
WP_062706917.1 (019)
a1
A/M$_{s(S, a2-a3)}$ boundary
a2
m$_{s(S, a2-a3)}$i
m$_{s(S, a2-a3)}$ii
m$_{s(S, a2-a3)}$iii
M$_{s(S, a2-a3)}$/A boundary
a3
a4

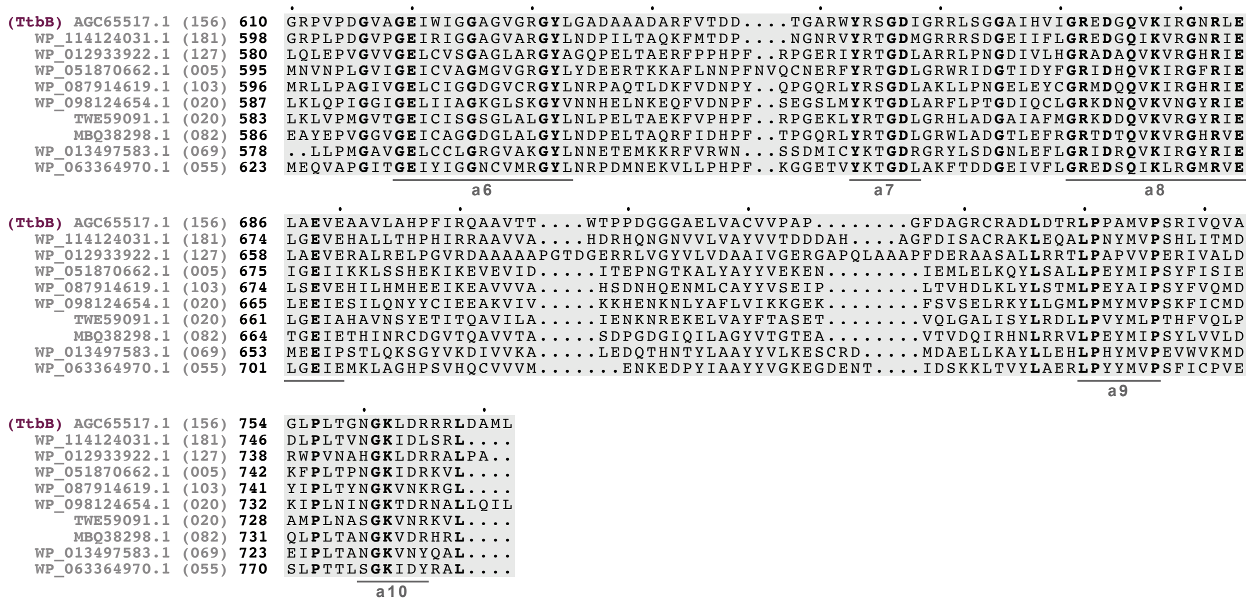


```
(TioN) CAJ34370.1 (052) 543 RAAFPNA.EVVALGGATECTVWSNHYRVGDVDPGWPSIPYGKPMQNARYYVLDHGLGPCPNGAEGDLYIAGECVALGYAR
       APD71663.1 (062) 543 RAAFPNA.EVVALGGATECTVWSNHYGVGDVDPAWPSIPYGKPMQNARYYVLDGDLGPCPLGEEGDLYIAGTCVALGYAA
     WP_077005101.1 (035) 533 RARYGAATRVVALGGATEATVWSNYFDVDEVGVDWPSIPYGRPIQNARYYVLDGERRPCPVGEPGDLHIAGVPVAEGYVG
     WP_081546865.1 (041) 586 RARFPRA.HVVALGGATEATIWSNHFDVGEIDPAWTSVPYGKPIQNARYYVLDGDRKPCPVGVEGDLHIAGVVVADGYVG
       OQY27011.1 (009) 578 KKAFSGT.RVISLGGATEATVWSNYFPIEEISPHWISIPYGRPIQNAKYYILDDQLNPCPIGVAGFLYIGGECLAEGYAG
       NER47250.1 (012) 582 KTVFPGV.QFISLGGATEATIWSNYYPIETVDPRWISIPYGRPIQNARYYILDSYLNPCPIGVAGLLYIGGDCLAVGY.D
       GAE88107.1 (016) 570 REVFPRA.EVISLGGATEATVWSNYYPIGEVDPLWASIPYGKPIQNAFYYILDSNLNPCPIGVIGDLYIGGECLATGYIN
     WP_148930441.1 (001) 567 RQVIPSA.RVIALGGATEATVWSNYYPVGDVQPHWASIPYGKPIQNARYYILNDALMPCPIGVKGALYIGGEVLASGYMH
       HBL25591.1 (025) 555 RRAFPGA.RVMALGGATEATVWSNWFPVGVVDPGWPSIPYGRPIANVSYHVLDAGFAPCPVGVPGDLLIGGDCLCVGYAR
     WP_062706917.1 (019) 571 KQYFPNA.KVLGLGGATEAAIWSNFYPIDQIQASWTSIPYGRPISNAQYYVLDSTLSPCPVGVPGDLYIAGDCLAKGYYG

(TioN) CAJ34370.1 (052) 622 SPGLTAAKFLPDPWAH.........RTGERMYHTGDRARWL..PSGDLEFLGRLDDQVKVRGYRIELGEVQSALLQCAGV
       APD71663.1 (062) 622 AARLTAAKFLPDPGAL.........RPGERMYRTGDRARWL..PSGDLEFLGRLDDQVKVRGYRVELGEVQSAFLQCPGV
     WP_077005101.1 (035) 613 DEELTRARFLPDPFSD..........GGERMYATGDRARWL..PEGVLQFLGRSDDQVKVRGYRIELGDVQAALRRAAGV
     WP_081546865.1 (041) 665 DPELTAAKFTADTVGG.........PADESMYATGDRARWM..ADGNLEFRGRRDDQVKVRGYRIELGEVLAALRRVPGV
       OQY27011.1 (009) 657 DPEQTAQRFIPCPFSE......VRGKTGARLYNTGDLSRYF..PDGNMEFLGRVDTQVKIRGFRIELGEIETVLAAHPTV
       NER47250.1 (012) 660 DPVKTDEKFIEIPTQEGLGVSLIREEKLDRLYQTGDIARYF..PDGNIQFLGRIDHQVKIRGFRIELGEIEAALICHEQI
       GAE88107.1 (016) 649 DEKLTKEKFIPDPFNK.........SVDGRMYKTGDTARWF..EDGNMEFMGRKDFQVKIRGYRIELGEIESHLIKYEAV
     WP_148930441.1 (001) 646 QPELTAERFIDNPFHD............GKLYWTGDMARWL..DDGNIEFMGRVDHQVKIRGYRIELGEIHAQLIKHAAV
       HBL25591.1 (025) 634 QPDLTAQAFLPDPFSG.........RPGARLYRTGDRARYR..ADGNLEFLGRLDQQVKVRGFRIELGEVEVVLGSHPGV
     WP_062706917.1 (019) 650 DQEMTGKKFIRNPLPG.........TPSPRLYHTGDRARWMRSESGVLEFLGRQDFQVKLRGFRIELGEVEAKLLANDAI

(TioN) CAJ34370.1 (052) 691 RAASVVAPSGPGGRT...L.AAFYVPTAEPLPTAEVVGA.LQRILPSYMVPTRIVALDELPLTNTGKVDRAALLSRL
       APD71663.1 (062) 691 RAASVMAPTGPGGRT...L.AAFYVPTAEPLSKAEVVRA.LRTILPSYMIPTRIVALDELPLTSTGKVDRAALLRRL
     WP_077005101.1 (035) 681 ADAVAVAVDGANGRE...LVAAVAPRPGEPVTANG.VRADLRAVLPTYMLPDHLVVLPALPVGPTGKADRAAI....
     WP_081546865.1 (041) 734 VDCAVTTLRGDDGPQ...IVAAVAVGGSEPTRGGGFVRRHLAGSLPSYMLPEHVLVLPALPVGPTGKVDR.......
       OQY27011.1 (009) 729 RENVAVVHES.HGDR...RLVTYLIPNQGQRIDNAEVKAFLQKKLPEYMIPSYFITLEKLPLTSNGKLDRKAL....
       NER47250.1 (012) 738 QHCVVIPREDLPSHK...RLVAYVVNHKINEINIHALKTHLKQQLPDYMVPDAIVPLTSLPLTANGKVDRKSL....
       GAE88107.1 (016) 718 KDAVVVVRKD.DGDN...YICAYYVSDR..ELTVSELREYLLKELPSYMIPSYFVRLSHMPVTDNGKLDRKAL....
     WP_148930441.1 (001) 712 REAVVLDRPDPHGAK...CICAYYVSESG..LSPAELRAYAARWLPSYMLPSYFVPIEGIPLTANGKVDRAAL....
       HBL25591.1 (025) 703 RDCVVLVREDMPGER...RLVAYSVGTGGRLPRAEELRLHLQSRLPEFMIPSAFVALDELPLTANGKLDRQAL....
     WP_062706917.1 (019) 721 EEVFVHLHRDPEGNEASDCLVAYYICADHASASSDTLRYFLAEHLPEYMIPTHYVEMEAFPVTSNGKLDRSAL....
```

Fig. S16. Multiple sequence alignment of family 3 ($A_2M_{s(S)}A_3$) interrupted A domains using the ten representative sequences indicated in Figs. S4 and S11. The numbering next to the accession number for each sequence matches the number in the taxonomic tree (Table S4). The A domain portion of the sequence is highlighted in light grey and the $M_{s(S,\ a2\text{-}a3)}$ portion in light yellow. The divisions between the two are indicated by teal triangles. Suspected residues for binding of SAM and the amino acid bound Ppant arm, based on the structure of TioS($A_8M_IA_9$),[1] are indicated by red and dark purple balloons, respectively. The conserved motifs for the A domain are underlined by a dark grey bar and labeled a1-a10. The conserved M domain motifs (outlined in Table 1) are underlined by a dark yellow bar and labeled $m_{s(S,\ a2\text{-}a3)}$i-iii. A shortened version of this annotated alignment is presented in Fig. 8A.

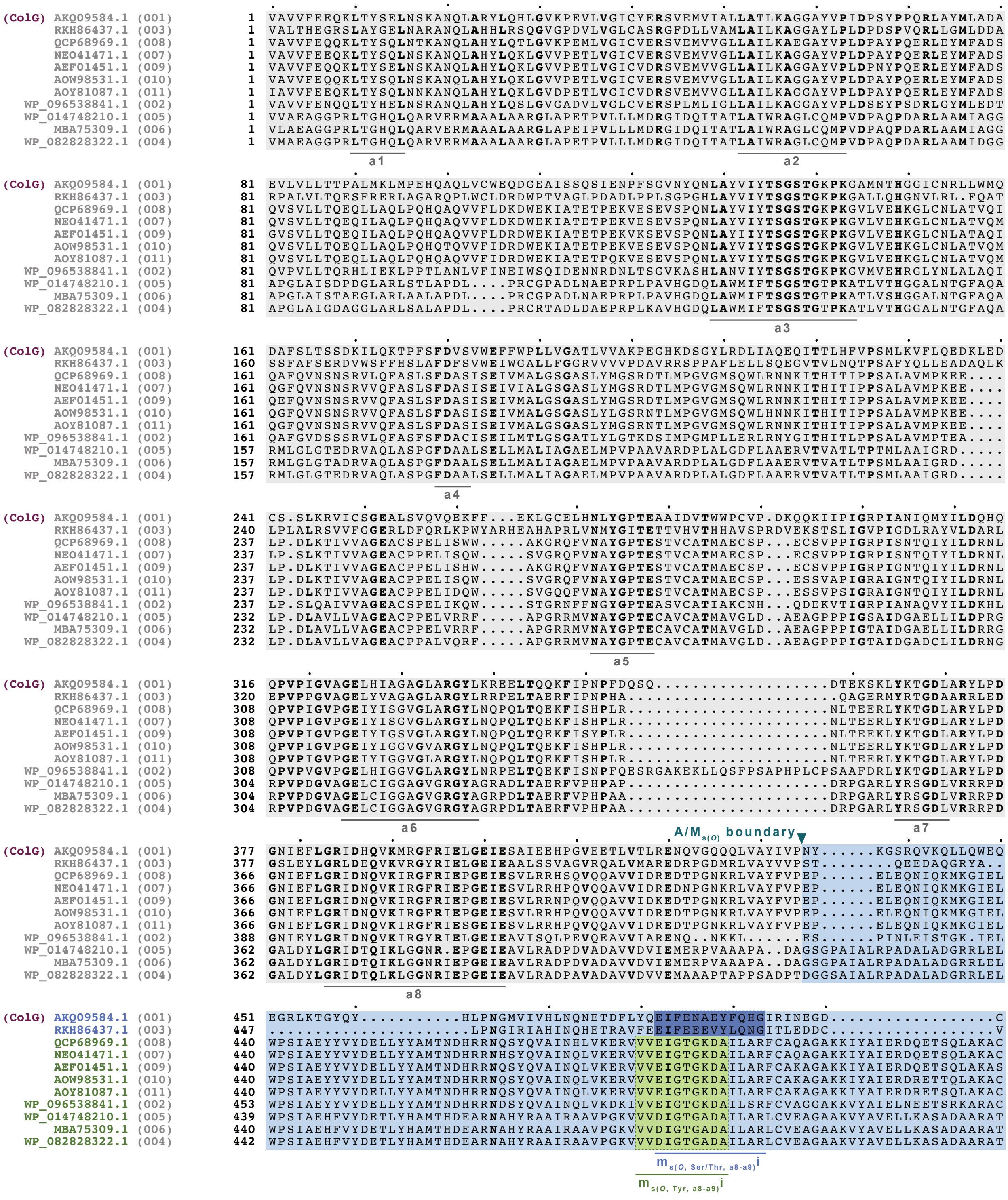


```
(TtbB) AGC65517.1 (156)   1 AVRDAVEAVSYTGLQAEARAIARLVAG.ATLGDEARVGVMTGRGWRHVAAMLGVMMARAAFVPLDPAIPLPRRRQIVETA
   WP_114124031.1 (181)   1 ATRDESGEISYADLDRQSNRIAHFIAG.LNLPVGSLVGVMTGRNRHYIAATLGVLNAGFAFVPLDPNVPLIRRTTLVRNA
   WP_012933922.1 (127)   1 AVVDATGRWTYAELAHEVDALAAYLVA.RGVGPEVHVGVLLRRSRRMAVALLATLRAGGVFVPLNPGLPRSRLAALLDDV
   WP_051870662.1 (005)   1 AIYYQGESITYGELVEKANQIA.YLLRDKDISRDQVVGILMDRSTLMVESILGVWKAGAAYLPLDSQQPIERMKYIVENS
   WP_087914619.1 (103)   1 AIQYQNKRVTYKQLDEQANVYANYLLQEHDLTPGTLVGVMLDKSDELIAVTLGILKAGGAYVPMDPEYPEERLKMIIEDA
   WP_098124654.1 (039)   1 ALKFEETVLTFEELNKCANRLARYLLNEHSIGSNECVGILLERSHYNVIAMLGVLKAGGSYVPIDPSLPTNRIKNIIDNT
       TWE59091.1 (020)   1 ALLFQDTRFTYAQLNERAGQLAGYLQTTFGLAPGRLAGLLLTRNANMLVAMLAVLKTGAAYVPMDPSYPMHRLRYIIDDA
       MBQ38298.1 (082)   1 AVTQGDRSWTYRELDEGSNRIAHFLIAQH.LPPDSVVTFLLDHTYDICTTVLGILKAGLVYLPIDEGMPHQRVQYLLNDS
   WP_013497583.1 (069)   1 ALIDDTCSYTYEKLDRFSNQIAHFVKNYASGQTEERVGILIENSVLQAAAVLGVMKSGCVYVPIDVKLPFDRIRVMLEDA
   WP_063364970.1 (055)   1 AVIHKGVHYSYKTINEQANQLAHFLTTERQVQREELVAIYIDRSHWYLIALLAVMKSGGAYMPLDSSLPADRMQYLLEQA

(TtbB) AGC65517.1 (156)  80 GITMIIADSPRIGDLHALAWTCPGVRDLVCLDSDDIDALTETH.GARMNAELWDHLAGSADDVAAAGWRSPFTGGAIPE
   WP_114124031.1 (181)  80 ALPLIISEGAHIGDLHQLQWRSKTVTDILCLDHDDIDTLTENH.GSLMSVELWDHLAGDDADNILAGGWKSAFTGQPIAQ
   WP_012933922.1 (127)  80 AAPLVVCERALIAEANALQWECDSVALLLCVDSDDVHAELEPP.GARMDAELWDVVAANAAGPIEAGGWRSAHSGELLST
   WP_051870662.1 (005)  80 QVSFLIFEKKYSKEANQLQWECSQTTHIFCIDSEDVFDEDESIK.ETQSKSIWKYVTEKAQDDIEEGMWINSYTGHPFTQ
   WP_087914619.1 (103)  81 GMHLIISGKRYIKTLNKLQWECNNFKAFLCVDTHNVMQEEENEKNELMGEEVWKYVGLNAKDDIEGGAWTSSYTGGNLSK
   WP_098124654.1 (039)  81 VIKVIISKSDYVQMLNKLQWECFDFKAYICLDSNNIEEVEDTHNNDLMDPKLWNHIGASTNNRIGMGGWINSYTGEEFTN
       TWE59091.1 (020)  81 AISLLISDEDNLETSNKLQWECKGLSAFVSLDTRDVHAYAESSDDYNQSRQLWDYVGETAADDIASGGWMSSYTGQSFSS
       MBQ38298.1 (082)  80 QARVFVSWTRHLRLLNSLQWDCPCLERVICLDADDFRDEIEP.EGEKMRQDVWDYVGEEMFDDISGGGWKSSFTGDWLSR
   WP_013497583.1 (069)  81 SISVVLTTSRYLDMVNKLQWECSTFRTYACLDRDDLYELEEIHSNPLMNKNIWNIVARSAQDSIGQGGWTSSYTGELFSE
   WP_063364970.1 (055)  81 SVRTVLSVKDFVGNLNKLQWNTAGLESFVLVDHDDVYSLDENTQNTAGLRSLWEYVGENAADDIGGGGWSSTYTGYDFTR

(TtbB) AGC65517.1 (156) 159 PVMTEFGRAAAAKMAALLPPRARVLEIGCASGFTMRALAPHAGQYLATDISRRATERAGAVAARAG.LTQVDVRPLAAHD
   WP_114124031.1 (181) 159 SAMDAFGANARNKVAGLVSRNARVLEIGCASGFTMKHVAPVVGTYVATDLARRNVERTEALAINLG.MTHVIGRQLGAHD
   WP_012933922.1 (127) 159 AAMAAFAEGAVAALAPLLPPRAEVLEIGCGSGLTLTALAPRTARYVGTDISAAALEWARRACTRDG.LEHVALHRLEALE
   WP_051870662.1 (005) 159 EEMNEYKENIFSKLSPYLNKETKVLEIGCSSGFSMYRIAPFVKHYVGIDMTEFIIEKNRKIIQEKG.INNIELACMYAHE
   WP_087914619.1 (103) 161 EDMEEYTNNIECKLKPYLNSTTRVLEIGCASGLSMYPLAPQVGWYYGTDLSEVIIGKNRERVQKEG.IGNIKLACIPAHD
   WP_098124654.1 (039) 161 KEMEEYSNNVLRKMTPYLHENIRVLEIGCASGITMYKIAPYCKLYYGTDMSNVTIDKNNNYIKQQG.FENIKLECLAAHE
       TWE59091.1 (020) 161 EEMQEYTDNILNNLRPYLDKDLKVLEIGCSSGLTLRAIAPYVGRYVGVDISPAILGRLKNALAEEGGFEHVTLQHLSAFE
       MBQ38298.1 (082) 159 SVMDEYGQNILTKLQPLLTPQTRVLEIGCSSGISMFRLAPLVGYYLGTDLSPTIMEKSEGERRRLG.LDNVRLKAVPAHD
   WP_013497583.1 (069) 161 QEMNEYSENTVRKLIPYLTKDTRVLEIGCASGLTMFRIAPYVCYYYGTDLSEETIRKNLEYIEQEK.ISNIGLEAVPADL
   WP_063364970.1 (055) 161 QEMDEYGENIVTKLKPLLTPQTRILEIGCASGITMYRLAPHVAKYVGTDISRVIIEKNEKRCEQEG.HNNISLYSKAAHE

(TtbB) AGC65517.1 (156) 238 LDLLNG...RQFDLIVINSVIESFPGFGYLSAVLSKAAALLAPGGAIFLGNIRDLERRDAYRADLAAFARAHAGESTTSG
   WP_114124031.1 (181) 238 IDVYEP...GSFDLIIMNSVIENFPGYGYLRTVLTKAKTLLAPGGAIYLGNVWDADRRDDYLADLARFARETVGK.....
   WP_012933922.1 (127) 238 LDRLPPDDDGPFDAVVLNSVVQSLGGFNHLRAVLRGAVARVRDGGCVYVGHVWDPERRDVLPLE...............
   WP_051870662.1 (005) 238 IDKLQP...ESFDIVIINSVVQFFSGHNYLRKVIQKIVEIAKSEAIIFFGDLLDQRKKTNFLHSLQEFKKLNLNKD....
   WP_087914619.1 (103) 240 IDKLEE...ADFDVVILNSVVQCFHGHNYFKNVISKCLGMLRNQGILFIGDVMDQERKADLLKSLVEFKQAHEDKN....
   WP_098124654.1 (039) 240 IGQIKQ...K.FDVVIINSTVQYFPGHNYLRKVINIATDMLDQKGVIFLGDIMDQDKKAEMIESLVHFSECN..PN....
       TWE59091.1 (020) 241 INKVNE...ADFDIIIMNSVVQYFGKYNYFRNVLQKTLDLLKADGLLYIGDVMDLALKNTLIDDLKQFKYASYGKG....
       MBQ38298.1 (082) 238 VDTLDE...AEFDIVIVNSVVQCFSGHNYLRSVLAQAIALMKPKGWLFLGNLFDRDLRDGFLEDLEQYERTHRDEE....
   WP_013497583.1 (069) 240 IGQIAQ...KDFDIVIINSVIQCFNGYNYLRKVLKQAVSRMKPKGILYLGDLMDQDRKAALIDSLKAYQKI....D....
   WP_063364970.1 (055) 240 IAEVDE...SEFDIVIINSVIQSFEGLNYLRKVLSTCVDKLNDKGLLFLGDLMDQDLKSEFLKDLLLYKDTGRSEHVMAE

(TtbB) AGC65517.1 (156) 315 TASG.....ITTRLDAADDLFVPAGFFTDWAARLGGFRA.EAS.PLVVDGFDLAPYVYDMVLR..RDPAATARPP....R
   WP_114124031.1 (181) 310 ...G.....YNTRLDFLEDLFVPQCFLEDWAREQGGFAA.SRS.AIDADGFDPAPYTFDFVLT..SNAAKYPVAP....K
   WP_012933922.1 (127) 302 ...........RRLDA...LFVPAAFFEQLPAALPRLAAVELR.PMGCDVPDLSAYAYDVLLHVGRRGRGRAVAA....R
   WP_051870662.1 (005) 311 .........YKTWTDWSDYLFVCPEFFEDLVYDFPEIDKIEHSKKIYTIENELTKFRFDTIVKINKKEVKKLPLK..RY.
   WP_087914619.1 (103) 313 .........YRTKTDWSVELFLSPGFFEDLRADFTEITAIQISDKIHTITNELTAFRYDAIIHVDKSGTSQTLQP..KL.
   WP_098124654.1 (039) 310 .........YRTKVDFSSELFLSKFWFNDLQMKNNAIHSLNISGKNCTIENELTKYRYDVILEVNKLMKKESNES..NLL
       TWE59091.1 (020) 314 .........YTTKTDFSTELFLSREFFYDLPHDYATIEKVTILDKQFTIVNELSAFRYDVILKKGVLQVKEGTRR..R..
       MBQ38298.1 (082) 311 .........GHTKTDYSDELFINRAFLNDLRFDLPQIADIEYSELLGHEASELREFTFDALMRIDRCPPGHAAGK..RI.
   WP_013497583.1 (069) 309 .........PSAKTDWSEELFISADFLRDLRFEMPVITQVDTSDKYYTIENELTMFRFDAILHIDQNNTSIPAGT..KT.
   WP_063364970.1 (055) 317 LNKQIDERPQNTKTDLSEELYISRHFLEDLLQDIPALSSVEFTDKLGTFKNELTEFRFDALISVDKQAASSNSANAGKRN

(TtbB) AGC65517.1 (156) 382 HRRHDRRALAPRPAAAPLPA....PTPGMLAYVLFTSGTTGAPKGAMIEHGSVANLAMATDAAMYARLRGGRDAGSDSAP
   WP_114124031.1 (181) 374 KRRHDLAALNAANDQRPATE....TSANDLAYVLFTSGTTGTPKGVMIEHGSVVNLAQSVHDTQLKQI....GADNPDRA
   WP_012933922.1 (127) 363 KRQEDRRTLLLDPAP.PADR....HRPGAAAYVIQTSGSTGRPRSVTVEMRSLVNLLWWY...........RDAAAIDAD
   WP_051870662.1 (005) 379 KHQYDATYLNQYSHRESICE....AQPNDLCYIIYTSGSTGKPKGVMVEHIGMLNHIFAKINDL.........H..LNET
   WP_087914619.1 (103) 381 KLQHGLQELNRYRNDRFHSQ....VSPNSLAYIIYTSGTTGKPKGAMIEHRNVVRLMSNSRMPF.........E..FNEE
   WP_098124654.1 (039) 379 KKQHDYKKLLEYSSGDLDLQ....IKEDQIAYVIFTSGSTGKPKGVQIQHQALTNFVEAMRSVL.........N..KKQT
       TWE59091.1 (020) 381 .QQYGFSW..DVDCDFIAPA....LQPHDPAYVIYTSGSTGHPKGVVVSHASLAAFVQNYSKKF.........H..LKAP
       MBQ38298.1 (082) 379 KHQFGRSSIDTQSAAAVEAT....GTAQGPAYIMYTSGTTGKPKGVMVPHRAVNRLVIN.TNFI.........R..ITPR
   WP_013497583.1 (069) 377 RNQFGIGEILSEDPMAPETL....PELDATAYMIYTSGTTGLPKGVMLSHRSLENLCVYSREAF.........A..MHED
   WP_063364970.1 (055) 397 KYQYGRQAFAHLSTENLDNTGRFAADGSQLAYVLFTSGSTGTPKGVEIENRSLVSFAFANRISL.........ENDVTSD

(TtbB) AGC65517.1 (156) 458 LNIPCLAPFAFDASVIQILPTLAHGHHLHVACEETRRNPPMLDRFLRARMIDMIDLTPSLFDLMLDHWQAEALF.C..PV
   WP_114124031.1 (181) 446 LNLTCVAPFAFDGSIVQIFAAILNGHCLNITGSATRSDPEALHDLIVTNAIDQLDATPSLFSLLLDYWEENRLS.C..PA
   WP_012933922.1 (127) 427 SRVVQVIPSSFDASVKNLLAPLVSGGALVLLGDGPYDPAALLELIERERVTVLNPGVPSMTYPLVELAAERAWA.GLRSL
   WP_051870662.1 (005) 444 SIIVQNANQCFDISVWQFFSGLVVGGKVCIYSQDLITNLQKIIEQVEQDCVTILEVVPSYLISLLDVLEELKLQNAFTTL
   WP_087914619.1 (103) 446 DVWTMFHSYSFDFSVWEMYGALLYGGKLIVIPKPIAQDTSGFLTLLKQNKVTVLNQTPSAFYRLIE..QDLESGGSGLNL
   WP_098124654.1 (039) 444 TVMLALTGISFDIFILESLYSLATGLKVVIANENQHMNVKEINNLIMDNQVDIIQLTPSRLKLLIDDTSGLE......KL
       TWE59091.1 (020) 443 MVIGAVTNITFDISVLELIGSLGYGLTIHLLDMD...DPDTIYKNLQHKRINALQITPSRLAQLTADTGITVLE....SL
       MBQ38298.1 (082) 443 DRILQTGALGFDASTFELWGALLNGAQLSRPPEKSLLDLVEMKRLLGEQGTTILWLTSSLFNQYADWDITVFE.....GL
   WP_013497583.1 (069) 442 NRMTRYANFAFDASVWELLPSLYAGGTVFVLPDSLRYDMKRLSAFFADHQITHSFLPTQVAEQFMELDHT........SL
   WP_063364970.1 (055) 468 DRFLVNLSQAFDVNMCEVFLPWVFGATLVMSPKQHYTDIEALTALIVKEKVTAAWVPAALLKDSYRLLAASTET...LSL

(TtbB) AGC65517.1 (156) 535 RLVVLGGEAVSAGLMARIFDNPVHRDLLVLNAYGPTECCVGATTHLMTART...WREILPPPVGMPLSGVVADIRD..AA
   WP_114124031.1 (181) 523 RMVVLGGEAVSSRLVERFFALPGQTDRKLVNAYGPTECCVSSAQYVMTAHN...HKQILPPPIGKSIQGVYLEICD..DM
   WP_012933922.1 (127) 506 RCLALGGEAPDLARLRPWLESEA.CDARVLNVYGPTECADIACWHEATEDD...LAGRAPLPIGRPVHNAQAYVLG..ER
   WP_051870662.1 (005) 524 RYLLVTGEELKANLANRW..LQFFPSIPLVNAYGPTEASDDITHHFINEA....INETHVP.IGKPVQNFTIYIVD..EH
   WP_087914619.1 (103) 524 RYVIFGGEALRPGVLSDW..QRKYSHTKLINMYGITETTVHATYKEIGLS....EIQTNCSNIGKPIPTLKAYVMS..DN
   WP_098124654.1 (039) 518 KVIIVGGESFPKELLKQV..QQ.HTTADIYNAYGPTEATVWATYKNVTKE....TEIT....IGTPLMNYQCFIVNPNDH
       TWE59091.1 (020) 516 DVLLIGGEALGEVQFQEL..KQ.LSGTVVINVYGPTETTIWSSCLYVKEA....GMLS....VGKPLYNEALYILD..DH
       MBQ38298.1 (082) 518 RTLLIGGEKLSNHHVDQV..RRQYPELCVINGYGPTENTTFTTCHHIDQA......YTGDIPIGTPVANTEVWILD....
   WP_013497583.1 (069) 514 QCLLVGGDKLTNIHRVPY..R.......VFNNYGPAENTVVSTTYRIRGD....EQII...PIGRPIFNTGVIVTDCFGH
   WP_063364970.1 (055) 545 NKLLIGAEPIKSSTVRAL..QSLNNDMKILNIYGPTECTVYASCSLWSEAPGSVTSVDSDPNMSIGRARPDAEILILNSN
```

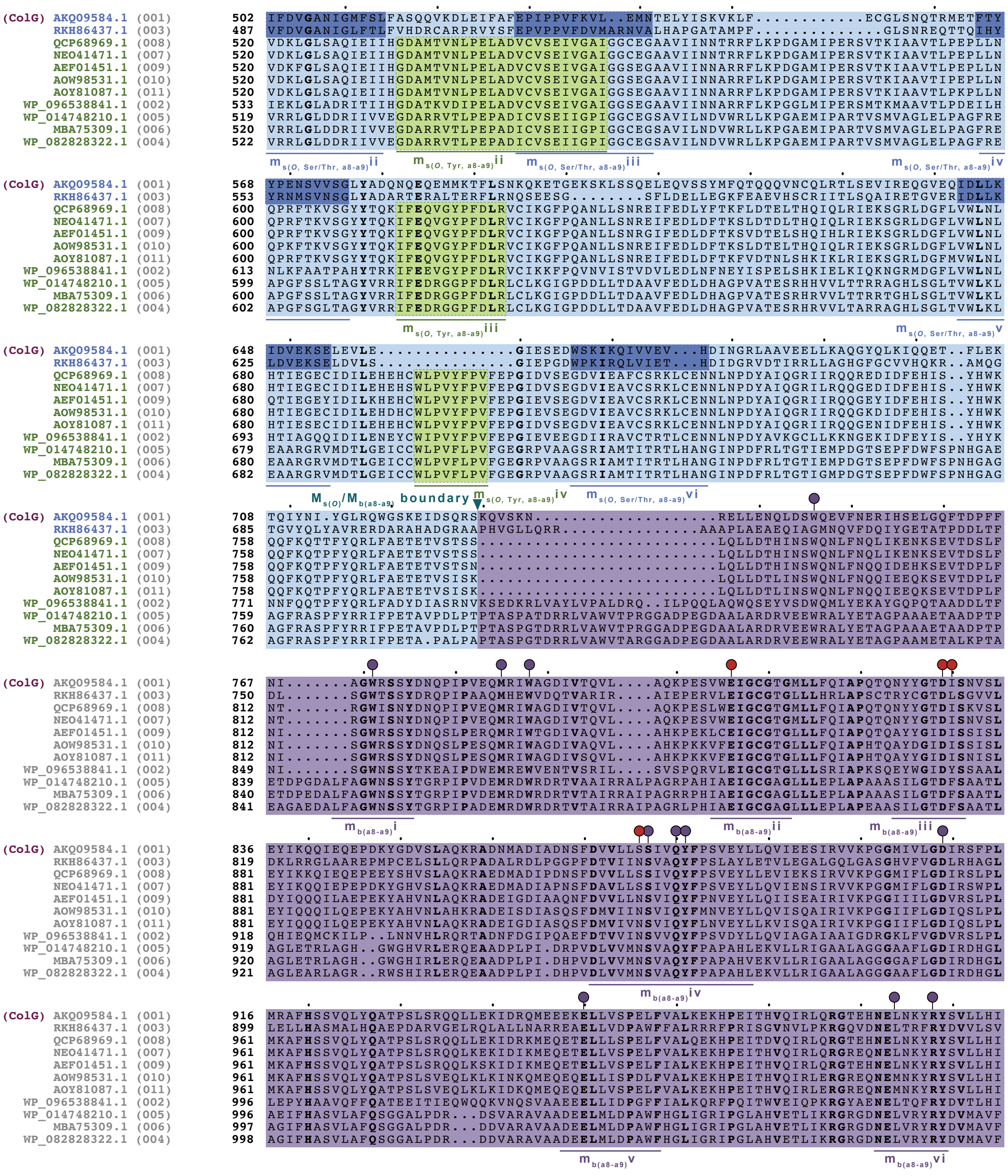


```
(TtbB) AGC65517.1 (156) 610  GRPVPDGVAGEIWIGGAGVGRGYLGADAAADARFVTDD....TGARWYRSGDIGRRLSGGAIHVIGREDGQVKIRGNRLE
   WP_114124031.1 (181) 598  GRPLPDGVPGEIRIGGAGVARGYLNDPILTAQKFMTDP....NGNRVYRTGDMGRRRSDGEIIFLGREDGQIKVRGNRIE
   WP_012933922.1 (127) 580  LQLEPVGVVGELCVSGAGLARGYAGQPELTAERFPPHPF..RPGERIYRTGDLARRLPNGDIVLHGRADAQVKVRGHRIE
   WP_051870662.1 (005) 595  MNVNPLGVIGEICVAGMGVGRGYLYDEERTKKAFLNNPFNVQCNERFYRTGDLGRWRIDGTIDYFGRIDHQVKIRGFRIE
   WP_087914619.1 (103) 596  MRLLPAGIVGELCIGGDGVCRGYLNRPAQTLDKFVDNPY..QPGQRLYRSGDLAKLLPNGELEYCGRMDQQVKIRGHRIE
   WP_098124654.1 (020) 587  LKLQPIGGIGELIIAGKGLSKGYVNNHELNKEQFVDNPF..SEGSLMYKTGDLARFLPTGDIQCLGRKDNQVKVNGYRIE
       TWE59091.1 (020) 583  LKLVPMGVTGEICISGSGLALGYLNLPELTAEKFVPHPF..RPGEKLYRTGDLGRHLADGAIAFMGRKDDQVKVRGYRIE
       MBQ38298.1 (082) 586  EAYEPVGGVGEICAGGDGLALGYLNDPELTAQRFIDHPF..TPGQRLYRTGDLGRWLADGTLEFRGRTDTQVKVRGHRVE
   WP_013497583.1 (069) 578  ..LLPMGAVGELCCLGRGVAKGYLNNETEMKKRFVRWN...SSDMICYKTGDRGRYLSDGNLEFLGRIDRQVKIRGYRIE
   WP_063364970.1 (055) 623  MEQVAPGITGEIYIGGNCVMRGYLNRPDMNEKVLLPHPF..KGGETVYKTGDLAKFTDDGEIVFLGREDSQIKLRGMRVE

(TtbB) AGC65517.1 (156) 686  LAEVEAAVLAHPFIRQAAVTT....WTPPDGGGAELVACVVPAP........GFDAGRCRADLDTRLPPAMVPSRIVQVA
   WP_114124031.1 (181) 674  LGEVEHALLTHPHIRRAAVVA....HDRHQNGNVVLVAYVVTDDDAH....AGFDISACRAKLEQALPNYMVPSHLITMD
   WP_012933922.1 (127) 658  LAEVERALRELPGVRDAAAAAPGTDGERRLVGYVLVDAAIVGERGAPQLAAAPFDERAASALLRRTLPAPVVPERIVALD
   WP_051870662.1 (005) 675  IGEIIKKLSSHEKIKEVEVID.....ITEPNGTKALYAYYVEKEN........IEMLELKQYLSALLPEYMIPSYFISIE
   WP_087914619.1 (103) 674  LSEVEHILHMHEEIKEAVVVA.....HSDNHQENMLCAYYVSEIP........LTVHDLKLYLSTMLPEYAIPSYFVQMD
   WP_098124654.1 (020) 665  LEEIESILQNYYCIEEAKVIV.....KKHENKNLYAFLVIKKGEK........FSVSELRKYLLGMLPMYMVPSKFICMD
       TWE59091.1 (020) 661  LGEIAHAVNSYETITQAVILA.....IENKNREKELVAYFTASET........VQLGALISYLRDLLPVYMLPTHFVQLP
       MBQ38298.1 (082) 664  TGEIETHINRCDGVTQAVVTA.....SDPGDGIQILAGYVTGTEA........VTVDQIRHNLRRVLPEYMIPSYLVVLD
   WP_013497583.1 (069) 653  MEEIPSTLQKSGYVKDIVVKA.....LEDQTHNTYLAAYYVLKESCRD.....MDAELLKAYLLEHLPHYMVPEVWVKMD
   WP_063364970.1 (055) 701  LGEIEMKLAGHPSVHQCVVVM.......ENKEDPYIAAYYVGKEGDENT....IDSKKLTVYLAERLPYYMVPSFICPVE

(TtbB) AGC65517.1 (156) 754  GLPLTGNGKLDRRRLDAML
   WP_114124031.1 (181) 746  DLPLTVNGKIDLSRL....
   WP_012933922.1 (127) 738  RWPVNAHGKLDRRALPA..
   WP_051870662.1 (005) 742  KFPLTPNGKIDRKVL....
   WP_087914619.1 (103) 741  YIPLTYNGKVNKRGL....
   WP_098124654.1 (020) 732  KIPLNINGKTDRNALLQIL
       TWE59091.1 (020) 728  AMPLNASGKVNRKVL....
       MBQ38298.1 (082) 731  QLPLTANGKVDRHRL....
   WP_013497583.1 (069) 723  EIPLTANGKVNYQAL....
   WP_063364970.1 (055) 770  SLPTTLSGKIDYRAL....
```

Fig. S17. Multiple sequence alignment of family 4 ($A_2M_bA_3$) interrupted A domains using the ten representative sequences indicated in Figs. S5 and S12. The numbering next to the accession number for each sequence matches the number in the taxonomic tree (Table S5). The A domain portion of the sequence is highlighted in light grey and the $M_{b(a2\text{-}a3)}$ portion in light pink. The divisions between the two are indicated by teal triangles. Suspected conserved residues for binding of SAM and the amino acid bound Ppant arm, based on the structure of TioS($A_8M_IA_9$),[1] are indicated by red and dark purple balloons, respectively. The conserved motifs for the A domain are underlined by a dark grey bar and labeled a1-a10. The conserved M domain motifs (outlined in Table 1) are underlined by a dark pink bar and labeled $m_{b(a2\text{-}a3)}$i-vi. *Note*: There is no $m_{b(a2\text{-}a3)}$v, but the numbering was kept to match the same motifs found in $m_{b(a8\text{-}a9)}$. A shortened version of this annotated alignment is presented in Fig. 6B.



```
(ColG) AKQ09584.1 (001)     1 VAVVFEEQKLTYSELNSKANQLARYLQHLGVKPEVLVGICYERSVEMVIALLATLKAGGAYVPIDPSYPPQRLAYMLADA
       RKH86437.1 (003)     1 VALTHEGRSLAYGELNARANQLAHHLRSQGVGPDVLVGLCASRGFDLLVAMLAILKAGGAYLPLDPDSPVQRLLGMLDDA
       QCP68969.1 (008)     1 VAVVFEQQKLTYSQLNTKANQLAHYLQTLGVKPEMLVGICVERSVEMVVGLLAILKAEGAYVPLDPAYPQERLEYMFADS
       NEO41471.1 (007)     1 VAVVFEEQKLTYSQLNSKANQLAHYLQKLGVVPETLVGICVERSVEMVVGLLAILKAGGAYVPLDPAYPQERLEYMFADS
       AEF01451.1 (009)     1 VAVVFEQQKLTYSELNSKANQLAHYLQKLGVVPETLVGICVDRSVEMVVGLLAILKAGGAYVPLDPNYPQERLEYMFADS
       AOW98531.1 (010)     1 VAVVFEQQKLTYSQLNSKANQLAHYLQKLGVVPETLVGICVERSVEMVVGLLAILKAGGAYVPLDPAYPQERLEYMFADS
       AOY81087.1 (011)     1 IAVVFEEQKLTYSQLNNKANQLAHYLQKLGVDPETLVGICVDRSVEMVVGLLAILKAGGAYVPLDPAYPQERLEYMFADS
   WP_096538841.1 (002)     1 VAVVFENQQLTYHELNSRANQLAHYLQSLGVGADVLVGLCVERSPLMLIGLLAILKAGGAYVPLDSEYPSDRLGYMLEDT
   WP_014748210.1 (005)     1 VVAEAGGPRLTGHQLQARVERMAAALAARGLAPETPVLLLMDRGIDQITATLAIWRAGLCQMPVDPAQPDARLAAMIAGG
       MBA75309.1 (006)     1 VLAEAGGPRLTGHQLQARVERMAAALAARGLAPEIPVLLLMDRGIDQITATLAIWRAGLCQMPVDPAQPDARLAAMIAGG
   WP_082828322.1 (004)     1 VMAEAGGPRLTGHQLQARVERMAAALAARGLAPETPVLLLMDRGIDQITATLAIWRAGLCQMPVDPAQPDARLAAMIDGG

(ColG) AKQ09584.1 (001)    81 EVLVLLTTPALMKLMPEHQAQLVCWEQDGEAISSQSIENPFSGVNYQNLAYVIYTSGSTGKPKGAMNTHGGICNRLLWMQ
       RKH86437.1 (003)    81 RPALVLTQESFRERLAGARQPLWCLDRDWPTVAGLPDADLPPLSGPGHLAYVIYTSGSTGKPKGALLQHGNVLRL.FQAT
       QCP68969.1 (008)    81 QVSVLLTQEQLLAQLPQHQAQVVFLDKDWEKIATETPEKVESEVSPQNLAYVIYTSGSTGKPKGVLVEHKGLCNLATVQI
       NEO41471.1 (007)    81 QVSVLLTQEQILAQLPQHQAQVVFLDKDWEKIATETPEKVESEVSPQNLAYVIYTSGSTGKPKGVLVEHKGLCNLATVQM
       AEF01451.1 (009)    81 GVSVLLTQEQILAQLPQHQAQVVFLDRDWEKIATETPEKVKSEVSPQNLAYIIYTSGSTGKPKGVLVEHKGLCNLATAQI
       AOW98531.1 (010)    81 QVSVLLTQEQLLAQLPQHQTQVVFIDRDWEKIATETPEKVESEVSPQNLAYVIYTSGSTGKPKGVLVEHKGLCNLATVQM
       AOY81087.1 (011)    81 QVSVLLTQEQLLAQLPQHQAQVVFIDRDWEKIATETPQKVESEVSPQNLAYVIYTSGSTGKPKGVLVEHKGLCNLATVQM
   WP_096538841.1 (002)    81 QVPVLLTQRHLIEKLPPTLANLVFINEIWSQIDENNRDNLTSGVKASHLANVIYTSGSTGKPKGVMVEHRGLYNLALAQI
   WP_014748210.1 (005)    81 APGLAISDPDGLARLSTLAPDL....PRCGPADLNAEPRPLPAVHGDQLAWMIFTSGSTGTPKATLVTHGGALNTGFAQA
       MBA75309.1 (006)    81 APGLAISTAEGLARLAALAPDL....PRCGPADLNAEPRPLPAVHGDQLAWMIFTSGSTGTPKATLVSHGGALNTGFAQA
   WP_082828322.1 (004)    81 APGLAIGDAGGLARLSALAPDL....PRCRTADLDAEPRPLKAVHGDQLAWMIFTSGSTGTPKATLVTHGGALNTGFAQA

(ColG) AKQ09584.1 (001)   161 DAFSLTSSDKILQKTPFSFDVSVWEFFWPLLVGATLVVAKPEGHKDSGYLRDLIAQEQITTLHFVPSMLKVFLQEDKLED
       RKH86437.1 (003)   160 SSFAFSERDVWSFFHSLAFDFSVWEIWGALLFGGRVVVVPDAVRRSPPAFLELLSQEGVTVLNQTPSAFYQLLEADAQLK
       QCP68969.1 (008)   161 QAFQVNSNSRVLQFASLSFDASISEIVMALGSGASLYMGSRDTLMPGVGMSQWLRNNKITHITIPPSALAVMPKEE....
       NEO41471.1 (007)   161 QGFQVNSNSRVVQFASLSFDASISEIVIALGSGASLYMGSRDTLMPGVGMSQWLRNNKITHITIPPSALAVMPKEE....
       AEF01451.1 (009)   161 QEFQVNSNSRVVQFASLSFDASISEIVMALGSGASLYMGSRDTLMPGVGMSQWLRNNKITHITIPPSALAVMPKEE....
       AOW98531.1 (010)   161 QGFQVNSNSRVVQFASLSFDASISEIVMALGSGASLYLGSRNTLMPGVGMSQWLHNNKITHITIPPSALAVMPKEE....
       AOY81087.1 (011)   161 QGFQVNSNSRVVQFASLSFDASISEIVMALGSGASLYMGSRNTLMPGVGMSQWLRNNKITHITIPPSALAVMPKEE....
   WP_096538841.1 (002)   161 QAFGVDSSSRVLQFASFSFDACISEILMTLGSGATLYLGTKDSIMPGMPLLERLRNYGITHITLPPSALAVMPTEA....
   WP_014748210.1 (005)   157 RMLGLGTEDRVAQLASPGFDAALSELLMALIAGAELMPVPAAVARDPLALGDFLAAERVTVATLTPTMLAAIGRD.....
       MBA75309.1 (006)   157 RMLGLGTADRVAQLASPGFDAALSELLMALIAGAELMPVPAAVARDPLALGDFLAAERVTVATLTPSMLAAIGRD.....
   WP_082828322.1 (004)   157 RMLGLGTEDRVAQLASPGFDAALSELLMALIAGAELMPVPAAVARDPLALGDFLAAERVTVATLTPSMLAAIGRD.....

(ColG) AKQ09584.1 (001)   241 CS.SLKRVICSGEALSVQVQEKFF...EKLGCELHNLYGPTEAAIDVTWWPCVP.DKQQKIIPIGRPIANIQMYILDQHQ
       RKH86437.1 (003)   240 LPLALRSVVFGGERLDFQRLKPWYARHEAHAPRLVNMYGITETTVHVTHHAVSPRDVEKSTSLIGVPIGDLRAYVLDARL
       QCP68969.1 (008)   237 LP.DLKTIVVAGEACSPELISWW.....AKGRQFVNAYGPTESTVCATMAECSP...ECSVPPIGRPISNTQIYILDRNL
       NEO41471.1 (007)   237 LP.SLKTIVVAGEACPPELISQW.....SVGRQFVNAYGPTESTVCATMAECSP...ECSVPPIGRPISNTQIYILDRNL
       AEF01451.1 (009)   237 LP.DLKTIVVAGEACPPELISHW.....AKGRQFVNAYGPTESTVCATMAECSP...ECSVPPIGRPIGNTQIYILDRNL
       AOW98531.1 (010)   237 LP.DLKTIVVAGEACPPELISQW.....SVGRQFVNAYGPTESTVCATMAECSP...ESSVAPIGRAIGNTQIYILDRNL
       AOY81087.1 (011)   237 LP.DLKTIVVAGEACPPELIDQW.....SVGQQFVNAYGPTESTVCATMAECSP...ESSVPSIGRAIGNTQIYILDRNL
   WP_096538841.1 (002)   237 LP.SLQAIVVAGEACSPELIKQW.....STGRNFFNGYGPTEASVCATIAKCNH...QDEKVTIGRPIANAQVYILDKHL
   WP_014748210.1 (005)   232 LP.DLAVLLVAGEACPPELVRRF.....APGRRMVNAYGPTECAVCATMAVGLD..AEAGPPPIGSAIDGAELLILDPRG
       MBA75309.1 (006)   232 LP.DLAVLLVAGEACPPELVRRF.....APGRRMVNAYGPTECAVCATMAVGLD..AEAGPPPIGTAIDGAELLILDRRG
   WP_082828322.1 (004)   232 LP.DLAVLLVAGEACPPALVQRF.....APGRRMVNAYGPTECAVCATMAVGLD..AEAGPPPIGTAIDGADCLILDRNG

(ColG) AKQ09584.1 (001)   316 QPVPIGVAGELHIAGAGLARGYLKREELTQQKFIPNPFDQSQ..................DTEKSKLYKTGDLARYLPD
       RKH86437.1 (003)   320 EPVPPGVAGELYVAGAGLARGYLRRPELTAERFIPNPHA......................QAGERMYRTGDAARRLED
       QCP68969.1 (008)   308 QPVPIGVPGEIYISGVGLARGYLNQPQLTQEKFISHPLR....................NLTEERLYKTGDLARYLPD
       NEO41471.1 (007)   308 QPVPIGVPGEIYISGVGLARGYLNQPQLTQEKFISHPLR....................NLTEERLYKTGDLARYLPD
       AEF01451.1 (009)   308 QPVPIGVPGEIYIGSVGLARGYLNQPQLTQEKFISYPLR....................NLTEEKLYKTGDLARYLPD
       AOW98531.1 (010)   308 QPVPIGVPGEIYIGGVGVARGYLNQPQLTQEKFISHPLR....................NLTEERLYKTGDLARYLPD
       AOY81087.1 (011)   308 QPVPIGVPGEIYIGGVGLARGYLNQPQLTQEKFISHPLR....................NLTEERLYKTGDLARYLPD
   WP_096538841.1 (002)   308 QPVPVGVPGELHIGGVGLARGYLNRPELTQEKFISNPFQESRGAKEKLLQSFPSAPHPLCPSAAFDRLYKTGDLARYLPD
   WP_014748210.1 (005)   304 RPVPDGVAGELCIGGAGVGRGYAGRADLTAERFVPHPAP....................DRPGARLYRSGDLVRRRPD
       MBA75309.1 (006)   304 RPVPDGVAGELCIGGAGVGRGYAGRPDLTAERFVPHPAA....................DRPGARLYRSGDLVRRRPD
   WP_082828322.1 (004)   304 RPVPDGVAGELCIGGAGVGRGYAGRPDLTAERFVPHPAA....................DRPGARLYRSGDLVRRRPD

(ColG) AKQ09584.1 (001)   377 GNIEFLGRIDHQVKMRGFRIELGEIESAIEEHPGVEETLVTLRENQVGQQQLVAYIVPNY......KGSRQVKQLLQWEQ
       RKH86437.1 (003)   377 GSLEYLGRLDEQVKVRGYRIELGEIEAALSSVAGVRQALVMAREDRPGDMRLVAYVVPST........QEEDAQGRYA..
       QCP68969.1 (008)   366 GNIEFLGRIDNQVKIRGFRIEPGEIESVLRRHSQVQQAVVIDREDTPGNKRLVAYFVPEP......ELEQNIQKMKGIEL
       NEO41471.1 (007)   366 GNIEFLGRIDNQVKIRGFRIEPGEIESVLRRHPQVQQAVVIDREDNPGNKRLVAYFVPEP......ELEQNIQKMKGIEL
       AEF01451.1 (009)   366 GNIEFLGRIDNQVKIRGFRIEPGEIESVLRRNPQVQQAVVIDKEDTPGNKRLVAYFVPEP......ELEQNIQKMKGIEL
       AOW98531.1 (010)   366 GNIEFLGRIDNQVKIRGFRIEPGEIESVLRRHPQVQQAVVIDREDTPGNKRLVAYFVPEP......ELEQNIQKMKGIEL
       AOY81087.1 (011)   366 GNIEFLGRIDNQVKIRGFRIESGEIESVLRRHPQVQQAVVIDREDTPGNKRLVAYFVPEP......ELEQNIQKMKGIEL
   WP_096538841.1 (002)   388 GNIEYLGRIDNQVKIRGYRIELGEIEAVISQLPEVQEAVVIARENQ..NKKL......ES......PINLEISTGK.IEL
   WP_014748210.1 (005)   362 GALDYLGRIDTQIKLGGNR.EPGEIEAVLRADPDVADAVVDVIEMERPVAAAPA..DAGSGPAIALRPADALADGRRLEL
       MBA75309.1 (006)   362 GALDYLGRIDTQIKLGGNRIEPGEIEAVLRADPDVADAVVDVIEMERPVAAAPA..DAGSGPAIALRPADALADGRRLEL
   WP_082828322.1 (004)   362 GALDYLGRIDTQLKLGGNRIEPGEIEAVLRADPAVADAVVDVVEMAAAPTAPPSADPTDGGSAIALRPADALADGRRLEL

(ColG) AKQ09584.1 (001)   451 EGRLKTGYQY...........HLPNGMVIVHLNQNETDFLYQEIFENAEYFQHGIRINEGD.................C
       RKH86437.1 (003)   447 ......................LPNGIRIAHINQHETRAVFEEIFEEEVYLQNGITLEDDC.................V
       QCP68969.1 (008)   440 WPSIAEYYVYDELLYYAMTNDHRRNQSYQVAINHLVKERVVVEIGTGKDAILARFCAQAGAKKIYAIERDQETSQLAKAC
       NEO41471.1 (007)   440 WPSIAEYYVYDELLYYAMTNDHRRNHSYQVAINQLVKERVVVEIGTGKDAILARFCAQAGAKKIYAIERDQETSQLAKAC
       AEF01451.1 (009)   440 WPSIAEYYVYDELLYYAMTNDHRRNQSYQVAINQLVKERVVVEIGTGKDAILARFCAKAGAKKIYAIERDQETSQLAKAS
       AOW98531.1 (010)   440 WPSIAEYYVYDELLYYAMTNDHRRNQSYQVAINQLVKERVVVEIGTGKDAILARFCAKAGAKKIYAIERDRETTQLAKAC
       AOY81087.1 (011)   440 WPSIAEYYVYDELLYYAMTNDHRRNQSYQVAINQLVKERVVVEIGTGKDAILARFCAKAGAKKIYAIERDRETSQLAKAC
   WP_096538841.1 (002)   453 WPSIAEYYVYDELLYYAMTNDHRRNDSYQVAINQLVKDKIVVEIGTGKDAILSRFCVAGGAKKVYAIELNEETSRKARAC
   WP_014748210.1 (005)   439 WPSIAEHFVYDETLYHAMTHDEARNAHYRAAIRAAVPGKVVVDIGTGADAILARLCVEAGAAKVYAVELLKASADAARAT
       MBA75309.1 (006)   440 WPSIAEHFVYDETLYHAMTHDEARNAHYRAAIRAAVPGKVVVDIGTGADAILARLCVEAGAAKVYAVELLKASADAARAT
   WP_082828322.1 (004)   442 WPSIAEHFVYDETLYHAMTHDEARNAHYRAAIRAAVPGKVVVDIGTGADAILARLCVEAGAAKVYAVELLKASADAARAT
```



```
(ColG) AKQ09584.1 (001)   502 IFDVGANIGMFSLFASQQVKDLEIFAFEPIPPVFKVL..EMNTELYISKVKLF...........ECGLSNQTRMETFTY
       RKH86437.1 (003)   487 VFDVGANIGLFTLFVHDRCARPRVYSFEPVPPVFDVMARNVALHAPGATAMPF.............GLSNREQTTQIHY
       QCP68969.1 (008)   520 VDKLGLSAQIEIIHGDAMTVNLPELADVCVSEIVGAIGGCEGAAVIINNTRRFLKPDGAMIPERSVTKIAAVTLPEPLLN
       NEO41471.1 (007)   520 VDKLGLSAQIEIIHGDAMTVNLPELADVCVSEIVGAIGGCEGAAVIINNTRRFLKPDGAMIPERSVTKIAAVTLPEPLLN
       AEF01451.1 (009)   520 VDKFGLSAQIEIIHGDAMTVNLPELADVCVSEIVGAIGGSEGAAVIINNARRFLKPDGAMIPERSVTKIAAVTLPEPLLN
       AOW98531.1 (010)   520 VDKLGLSAQIEIIHGDAMTVNLPELADVCVSEIVGAIGGSEGAAVIINNARRFLKPDGAMIPERSVTKIAAVTIPEPLLN
       AOY81087.1 (011)   520 VDKLGLSAQIEIIHGDAMTVNLPELADVCVSEIVGAIGGSEGAAVIINNARRFLKPDGAMIPERSVTKIAAVTLPKPLLN
   WP_096538841.1 (002)   533 IEKLGLADRITIIHGDATKVDIPELADVCVSEIVGAIGGSEGAAVIINNARRFLKPGGLMIPERSMTKMAAVTLPDEILH
   WP_014748210.1 (005)   519 VRRLGLDDRIIVVEGDARRVTLPEPADICVSEIIGPIGGCEGSAVILNDVWRLLKPGAEMIPARTVSMVAGLELPAGFRE
       MBA75309.1 (006)   520 VRRLGLDDRIIVVEGDARRVTLPEPADICVSEIIGPIGGCEGSAVILNDVWRLLKPGAEMIPARTVSMVAGLELPAGFRE
   WP_082828322.1 (004)   522 VRRLGLDDRIIVVEGDARRVTLPEPADICVSEIIGPIGGCEGSAVILNDVWRLLKPGAEMIPARTVSMVAGLELPAGFRE

(ColG) AKQ09584.1 (001)   568 YPENSVVSGLYADQNQEQEMMKTFLSNKQKETGEKSKLSSQELEQVSSYMFQTQQQVNCQLRTLSEVIREQGVEQIDLLK
       RKH86437.1 (003)   553 YRNMSVNSGLYADARTERALTERFLRNQSEESG........SFLDELLEGKFEAEVHSCRIITLSQAIRETGVERIDLLK
       QCP68969.1 (008)   600 QPRFTKVSGYYTQKIFEQVGYPFDLRVCIKGFPQANLLSNREIFEDLYFTKFLDTELTHQIQLRIEKSGRLDGFLVWLNL
       NEO41471.1 (007)   600 QPRFTKVSGYYTQKIFEQVGYPFDLRVCIKGFPQANLLSNREIFEDLDFTKSLDTDLTHQIQLRIEKSGRLDGFLVWLNL
       AEF01451.1 (009)   600 QPRFTKVSGYYTQKIFEQVGYPFDLRVCIKGFPQANLLSNRGIFEDLDFTKSVDTELSHQIHLRIEKSGRLDGFLVWLNL
       AOW98531.1 (010)   600 QPKFTKVSGYYTQKIFEQVGYPFDLRVCIKGFPQANLLSNREIFEDLDFTKSLDTELTHQIQLRIEKSGRLDGFLVWLNL
       AOY81087.1 (011)   600 QPRFTKVSGYYTQKIFEQVGYPFDLRVCIKGFPQANLLSNREIFEDLDFTKFVDTNLSHKIKLRIEKSGRLDGFMVWLNL
   WP_096538841.1 (002)   613 NLKFAATPAHYTRKIFEEVGYPFDLRVCIKKFPQVNVISTVDVLEDLNFNEYISPELSHKIELKIQKNGRMDGFLVWLNL
   WP_014748210.1 (005)   599 APGFSSLTAGYVRRIFEDRGGPFDLRLCLKGIGPDDLLTDAAVFEDLDHAGPVATESRHHVVLTTRRAGHLSGLTVWLKL
       MBA75309.1 (006)   600 APGFSSLTAGYVRRIFEDRGGPFDLRLCLKGIGPDDLLTDAAVFEDLDHAGPVATESRHRVVLTTRRTGHLSGLTVWLKL
   WP_082828322.1 (004)   602 APGFSGLTAGYVRRIFEDRGGPFDLRLCLKGIGPDDLLTDAAVFEDLDHAGPVATESRHRVVLTARRAGHLSGLTVWLKL

(ColG) AKQ09584.1 (001)   648 IDVEKSELEVLE...............GIESEDWSKIKQIVVEV...HDINGRLAAVEELLKAQGYQLKIQQET..FLEK
       RKH86437.1 (003)   625 LDVEKSELDVLS...............GIEPGDWPKIRQLVIET...HDIDGRVDTIRRLLAGHGFGCVVHQKR..AMQG
       QCP68969.1 (008)   680 HTIEGECIDILEHEHCWLPVYFPVFEPGIDVSEGDVIEAFCSRKLCENNLNPDYAIQGRIIRQQREDIDFEHIS..YHWK
       NEO41471.1 (007)   680 HTIEGECIDILEHEHSWLPVYFPVFEPGIDVSEGDVIEAVCSRKLCENNLNPDYAIQGRILRQQGEDIDFEHIS..YHWK
       AEF01451.1 (009)   680 QTIEGEYIDILKHEHCWLPVYFPVFEPGIEVSEGDVIEAVCSRKLCENNLNPDYAIQGRIIRQQGEYIDFEHIS..YHWK
       AOW98531.1 (010)   680 HTIEGECIDILEHDHCWLPVYFPVFEPGIEVSEGDVIEAVCSRKLCENNLNPDYAIQGRIIRQQGKDIDFEHIS..YHWK
       AOY81087.1 (011)   680 HTIESECIDILEHEHCWLPVYFPVFEPGIDVSEGDVIEAVCSRKLCENNLNPDYAIQGRIIRQQGENIDFEHIS..YHWK
   WP_096538841.1 (002)   693 HTIAGQQIDILENEYCWIPVYFPVFEPGIEVEEGDIIRAVCTRTLCENNLNPDYAVKGCLLKKNGEKIDFEYIS..YHYK
   WP_014748210.1 (005)   679 EAARGRVMDTLGEICCWLPVFLPVFGEGRPVAAGSRIAMTITRTLHANGINPDFRLTGTIEMPDGTSEPFDWFSPNHGAE
       MBA75309.1 (006)   680 EAARGRVMDTLGEICCWLPVFLPVFGEGRPVAAGSRIAMTITRTLHANGINPDFRLTGTIEMPDGTSEPFDWFSPNHGAG
   WP_082828322.1 (004)   682 EAARGRVMDTLGEICCWLPVFLPVFGEGRPVAAGSRIAMTITRTLHANGINPDFRLTGTIGMPDGTSEPFDWFSPNHGAG

(ColG) AKQ09584.1 (001)   708 TQIYNI.YGLRQWGSKEIDSQRSKQVSKN...................RELLENQLDSWQEVFNERIHSELGQFTDPFF
       RKH86437.1 (003)   685 TGVYQLYAVRERDARAHADGRAAPHVGLLQRR...............AAAPLAEAEQIAGMNQVFDQIYGETATPDDLTF
       QCP68969.1 (008)   758 QQFKQTTFYQRLFAETETVSTSS.........................LQLLDTHINSWQNLFNQLIKENKSEVTDSLF
       NEO41471.1 (007)   758 QQFKQTPFYQRLFAETETVSTSS.........................LQLLDTHINSWQNLFNQLIKENKSEVTDSLF
       AEF01451.1 (009)   758 QQFKQTPFYQRLFAETETVSTSN.........................LQLLDTHINSWQNLFNQQIDEHKSEVTDPLF
       AOW98531.1 (010)   758 QQFKQTPFYQRLFAETETVSISK.........................LQLLDTLINSWQNLFNQQIEEQKSEVTDPLF
       AOY81087.1 (011)   758 QQFKQTPFYQRLFAETETVSISK.........................LQLLDTLINSWQNLFNQQIEEQKSEVTDPLF
   WP_096538841.1 (002)   771 NNFQQTPFYQRLFADYDIASRNVKSEDKRLVAYLVPALDRQ..ILPQQLAQWQSEYVSDWQMLYEKAYGQPQTAADDLTF
   WP_014748210.1 (005)   759 AGFRASPFYRRIFPETAVPDLPTPTASPATDRRLVAWVTPRGGADPEGDAALARDRVEEWRALYETAGPAAAETAADPTP
       MBA75309.1 (006)   760 AGFRASPFYRRIFPETAVPDLPTPTASPGTDRRLVAWVTPRGGADPEGDAALARDRVEEWRALYETAGPAAAETAADPTP
   WP_082828322.1 (004)   762 AGFRASPFYRRIFPETA.PALPAPTASPGTDRRLVAWVTARGGADPEGDAALARDRVEEWRALYETAGPAAMETALKPTA

(ColG) AKQ09584.1 (001)   767 NI.......AGWRSSYDNQPIPVEQMRIWAGDIVTQVL....AQKPESVWEIGCGTGMLLFQIAPQTQNYYGTDISNVSL
       RKH86437.1 (003)   750 DL.......SGWTSSYDRQPIPAAQMHEWVDQTVARIR....AIEPERVLEIGCGTGLLLHRLAPSCTRYCGTDLSGVVL
       QCP68969.1 (008)   812 NT.......RGWISNYDNQPIPVEQMRIWAGDIVTQVL....AQKPESLWEIGCGTGMLLFQIAPQTQNYYGTDISKVSL
       NEO41471.1 (007)   812 NT.......RGWISNYDNQPIPVEQMRIWAGDIVTQVL....AQKPESVWEIGCGTGMLLFQIAPQTQNYYGTDISNVSL
       AEF01451.1 (009)   812 NI.......SGWRSSYDNQSLPERQMRIWTGDIVAQVL....AHKPEKLCEIGCGTGLLLFQIAPQTQAYYGIDISSISL
       AOW98531.1 (010)   812 NI.......SGWRSSYDNQSLPESQMRIWAGDIVAQVL....AHKPKKVFEIGCGTGLLLFQIAPHTQAYDGIDISSISL
       AOY81087.1 (011)   812 NI.......SGWRSSYDNQSLPESQMRIWAGDIVSQVL....AHKPKKVFEIGCGTGLLLFQIAPHTQAYYGIDISNISL
   WP_096538841.1 (002)   849 NI.......SGWNSSYTKEAIPDWEMREWVENTVSRIL....SVSPQRVLEIGCGTGLLLSRIAPKSQEYWGIDYSSAAL
   WP_014748210.1 (005)   839 ETDPGDALFAGWNSSYTGRPIPVDEMRDWRDRTVAAIRRALPAGRPAHIAEIGCGAGLLLEPLAPAAASILGTDFSAATL
       MBA75309.1 (006)   840 ETDPEDALFAGWNSSYTGRPIPVDEMRDWRDRTVTAIRRAIPAGRRPHIAEIGCGAGLLLEPLAPAAASILGTDFSAATL
   WP_082828322.1 (004)   841 EAGAEDALFAGWNSSYTGRPIPADEMRDWRDRTVTAIRRAIPAGRLPHIAEIGCGAGLLLEPLAPEAASILGTDFSAATL

(ColG) AKQ09584.1 (001)   836 EYIKQQIEQEPDKYGDVSLAQKRADNMADIADNSFDVVLLSSIVQYFPSVEYLLQVIEESIRVVKPGGMIVLGDIRSFPL
       RKH86437.1 (003)   819 DKLRRGLAAREPMPCELSLLQRPADALRDLPGGPFDTVIINSVAQYFPSLAYLETVLEGALGQLGASGHVFVGDLRHAGL
       QCP68969.1 (008)   881 EYIKKQIEQEPEEYSHVSLAQKRAEDMADIPDNSFDVVLLSSIVQYFPSVEYLLEVIEKSIRVVKPGGMIFLGDIRSLPL
       NEO41471.1 (007)   881 EYIKQQIEPEPDKYGHVSLAQKRAEDMADIAPNSFDAVLLSSIVQYFPSVEYLLQVIENSIRVVKPGGMIFLGDIRSWPL
       AEF01451.1 (009)   881 DYIQQQILAEPEKYAHVNLAQKRADEIGDIAAQNFDVVLLNSVIQYFPNVEYLLQIISEAIRIVKPGGIIILGDVRSLPL
       AOW98531.1 (010)   881 EYIQQQILQEPEKYAHVNLAHKRADEISDIASQSFDMVIINSVIQYFMNVEYLLQVISQAIRIVKPGGIIFLGDIQSLPL
       AOY81087.1 (011)   881 EYIQQQILQEPEKYAHVNLAQKRADEIGDIASQSFDMVIVNSVIQYFLNVEYLLKVISQAIRIVKPGGIIFLGDIQSLPL
   WP_096538841.1 (002)   918 QHIEQMCKILP..LNNVHLRQRTADNFDGIPQAEFDTVVINSVVQYFPSVDYLLQVIAGAIAAIGDRGKLFVGDVRSLPL
   WP_014748210.1 (005)   919 AGLETRLAGH..GWGHVRLERQEAADPLPI.DRPVDLVVMNSVAQYFPAPAHLEKVLLRIGAALAGGGAAFLGDIRDHGL
       MBA75309.1 (006)   920 AGLETRLAGH..GWGHIRLERQEAADPLPI.DHPVDLVVMNSVAQYFPAPAHVEKVLLRIGAALAGGGGAFLGDIRDHGL
   WP_082828322.1 (004)   921 AGLEARLAGR..RWSHIRLERQEAADPLPI.DHPVDLVVMNSVAQYFPAPAHLEKVLLRIGAALAGGGAAFLGDIRDHGL

(ColG) AKQ09584.1 (001)   916 MRAFHSSVQLYQATPSLSRQQLLEKIDRQMEEEKELLVSPELFVALKEKHPEITHVQIRLQRGTEHNELNKYRYSVLLHI
       RKH86437.1 (003)   899 LELLHASMALHQAEPDARVGELRQLALLNARREAELLVDPAWFFALRRRFPRISGVNVLPKRGQVDNELTRFRYDVVLSV
       QCP68969.1 (008)   961 MKAFHSSVQLYQATPSLSRQQLLEKIDRKMEQETELLLSPELFVALQEKHPEITDVQIRLQRGTEHNELNKYRYSVLLHI
       NEO41471.1 (007)   961 MKAFHSSVQLYQATPSLSRQQLKEKIDIKMEQETELLVSPELFVALKEKHPEITHVQIRLQRGREQNELNKYRYSVLLHI
       AEF01451.1 (009)   961 MKAFHSSVQLYQATPSLSRQQLKSKIDIKMEQETELLVSPELFIALKEKHPEITHVQIRLQRGTEHNELNKYRYSVLLHI
       AOW98531.1 (010)   961 MKAFHSSVQLYQATPSLSVEQLKLKINKQMEQEQELLISPDLFVALKEKHPEITHVQIRLERGREQNEMNKYRYSVLLHI
       AOY81087.1 (011)   961 MKAFHSSVQLYQATPSLSVEQLKLKIDKQMEQEQELLVSPELFVALKEKHPEITHVQIRLERGREQNEMNKYRYSVLLHI
   WP_096538841.1 (002)   996 LEPYHAAVQFFQATEEITIEQWQQKVNQSVAAEEELLIDPGFFIALKQRFPQITWVEIQPKRGYAENELTQFRYDVTLHI
   WP_014748210.1 (005)   996 AEIFHASVLAFQSGGALPDR...DSVARAVAADEELMLDPAWFHGLIGRIPGLAHVETLIKRGRGDNELVRYRYDVMAVF
       MBA75309.1 (006)   997 AGIFHASVLAFQSGGALPDR...DDVARAVAADEELMLDPAWFHGLIGRIPGLAHVETLIKRGRGDNELVRYRYDVMAVF
   WP_082828322.1 (004)   998 AGIFHASVLAFQSGGALPDR...DDVARAVAADEELMLDPAWFHGLIGRIPGLAHVETLIKRGRGDNELVRYRYDVMAVF
```

```
(ColG) AKQ09584.1 (001)    996 EAQPGK.VITPTVESGASLSVQQIENYLR..............DKGPESICFSGLVNGRVANDVDLVELLSQPESKQNV
       RKH86437.1 (003)    979 GQAAAELVPDWQGWESEEAVRQRLAR.................EAPPSLAVRDIPQARLASMRAAQGLLHRMADEATA
       QCP68969.1 (008)   1041 EAQPGK.VITPTVESGASLSVQQIETYLR..............DQQPESICFSGLVNGRVASDVGLVELLSQPEAKQNV
       NEO41471.1 (007)   1041 EAQPGK.VITPTVESGAAVSVQQIETYLR..............DIGPESICFSGLVNGRVANDVELVEWLSQPESKQNV
       AEF01451.1 (009)   1041 EAQPGT.VITPTVESGADLSVQQIETYLR..............SKGPESICFSGLVNGRVANDVELVELLSQPESKQNV
       AOW98531.1 (010)   1041 EAQAES.IIAPTVESGADMSAQEIETYLQ..............QKQPESICFSALVNGRVANDIGLVEFLSQPESKQNV
       AOY81087.1 (011)   1041 EAQAAS.IIAPTVESGADMSAQEIETYLR..............DQEPESICFSGLVNSRVANDVGLVEFLSQPESKLNV
    WP_096538841.1 (002)   1076 GADVQTPVLSELVLSEVEVAEVTVVPWLNWQLDRLSFSQIQNQLQQQPEVLGIRGVPNQRVQQAIQIWQWWENPPVVETV
    WP_014748210.1 (005)   1073 GPAP...........AAIRPEATLPW.................PGAEALA.ARLAGLTGALEVTDIPDARIADDTARG
       MBA75309.1 (006)   1074 GPAP...........APIRPELTLPW.................PGPEALA.TRLATVAGPVEVTDIPDARIAEDTARA
    WP_082828322.1 (004)   1075 GPAP...........APIRPELTLPW.................PGPEALA.TRLATVAGPVEVTDIPDARIADDTARA

(ColG) AKQ09584.1 (001)   1060 QQLRQFLESKPVNGIDPERLHQLSSDNGYSLELCWSAEGSPELMDGVFVRSELAKEGIVLTPLTQKSVAASNWNNY....
       RKH86437.1 (003)   1040 AGLARELDAQPLPGVDPARLEALAAERGYTVRFA.LASGNEGGLHAVFLRAPRAIDWSALYPAAGADAASLRAL......
       QCP68969.1 (008)   1105 QQLRQKLESKAVNGIDPERLYELSASLGYSLELCWSTQGSGELMDGVFVRSELAKEGIVLTPLTQKSVVAGNWHHY....
       NEO41471.1 (007)   1105 QQLRQFLESKQTKSIDPERLYELSASLGYSLELCWSGQGSPELMDGVFVRSELAKEGIVLTPLTQKSVVASNWNNY....
       AEF01451.1 (009)   1105 QQLRQSLQENLVNGIDPERLHQLSSELGYSLELCWSAQGSPELMDGVFVRSSLAKEGIVLTPLTQKSVVAVNWDNY....
       AOW98531.1 (010)   1105 QQLRQTLESKPVNGIDPERLYELSSDNGYSLELCWSAQGRGELMDGVFVRRELAQEGIVLTPLTQKSVVASNWHHY....
       AOY81087.1 (011)   1105 QQLRQKLESKPVNGIDPERLYELSSDLGYSLELCWSAQGGPELMDAVFVRSSLAKEGIVLTPLTQKSVVANNWHHY....
    WP_096538841.1 (002)   1156 EQLRQLLAKQPTEGINPEEFYQLGQQLGYTVHLSWWCSSQDGAYDVVFCRKQEQRVAF....WDNISNTAKSWTDY....
    WP_014748210.1 (005)   1122 QRLGLVSATATGEPAHPDSLHEMGERLGRAVRLAPAASGIAGAMDALFDVPGLADDPLARGWMRSPVSRGRLANDP....
       MBA75309.1 (006)   1123 MRLGLVPTTAAGDAADPDSLHELGLRLGRAVRLAPAASGIAGAMDALFDVPGLADDPLARGWMRSPRARGRLANDP....
    WP_082828322.1 (004)   1124 MRLGLVPTTAAGDAADPDSLHELGLRLGRAVRLAPAASGIVGAMDALFDAPGLADDPLARGWMRAPLPRGRLANDPRLAN

(ColG) AKQ09584.1 (001)   1136 ....GNNPLSSQFRKELIPQLREYLQSRLPEYMVPSGLMVLSQLPLTPNGKVDRKALPAPD
       RKH86437.1 (003)   1113 ....AHEPYLEERQAGLQHTVLERLREQLPAYMVPAHVVVLDALPLTPNGKVNKKALPAPD
       QCP68969.1 (008)   1181 ....GNNPLSSQLRKELIPQLREYIQSRLPEYMVPSGLMVLSQLPLTPNGKVDRKALPVPD
       NEO41471.1 (007)   1181 ....GNNPLSSQFRKELIPQLREYLQSRLPEYMVPSGLMVLSQLPLTPNGKVDRKALPELD
       AEF01451.1 (009)   1181 ....GNNPLSSQLRKQLIPQLREYLQSRLPEYMVPSGLMVLSQLPLTPNGKVDRKALPVPD
       AOW98531.1 (010)   1181 ....GNNPLSSELGKQLVPQLREYLELRLPEYMMPSELMVLSQLPLTPNGKVDRKALPVPD
       AOY81087.1 (011)   1181 ....GTNPLSFEVGKQLVPQLREYLELRLPEYMMPSELMVLSQLPLTPNGKVDRKALPVPD
    WP_096538841.1 (002)   1228 ....TNNPLYGKLVQKLVPQVRSFIEQKLPNYMVPQAFVLLNALPLTPNGKVDRRALPAPD
    WP_014748210.1 (005)   1198 ........LARRRKSALADRLRRRAREVLPAAMVPQAVMVLDRLPLGASGKLDRARLPRPE
       MBA75309.1 (006)   1199 ........LARRRKAALADRLRRRAREVLPAAMVPQAVMVLDRLPLGASGKLDRARLPRPE
    WP_082828322.1 (004)   1204 DPRLANDPLARRRKAALADRLRRRAREVLPAAMVPQAVMVL....................
```

Fig. S18. Multiple sequence alignment of families 5a/b ($A_8M_{s(O)}M_bA_9$) interrupted A domains. The numbering next to the accession number for each sequence matches the number in the taxonomic tree (Table S6). The A domain portion of the sequence is highlighted in light grey, the $M_{s(O)}$ portion in light blue, and $M_{b(a8\text{-}a9)}$ in light purple. The divisions between the different domains are indicated by teal triangles. Suspected conserved residues for binding of SAM and the amino acid bound Ppant arm, based on the structure of TioS($A_8M_IA_9$),[1] are indicated by red and dark purple balloons, respectively. The conserved motifs for the A domain are underlined by a dark grey bar and labeled a1-a10. The conserved $M_{s(O,\ \text{Ser/Thr, a8-a9})}$ domain motifs (outlined in Table 1) are boxed in dark blue, are underlined by dark blue bar, and are labeled $m_{s(O,\ \text{Ser/Thr, a8-a9})}$i-vi. The conserved $M_{s(O,\ \text{Tyr, a8-a9})}$ domain motifs (outlined in Table 1) are boxed in light green, are underlined by a dark green bar, and are labeled $m_{s(O,\ \text{Tyr, a8-a9})}$i-iv. The conserved $M_{b(a8\text{-}a9)}$ domain motifs (outlined in Table 1) are underlined by a dark purple bar and labeled $m_{b(a8\text{-}a9)}$i-vi. A shortened version of this annotated alignment is presented in Fig. 9.

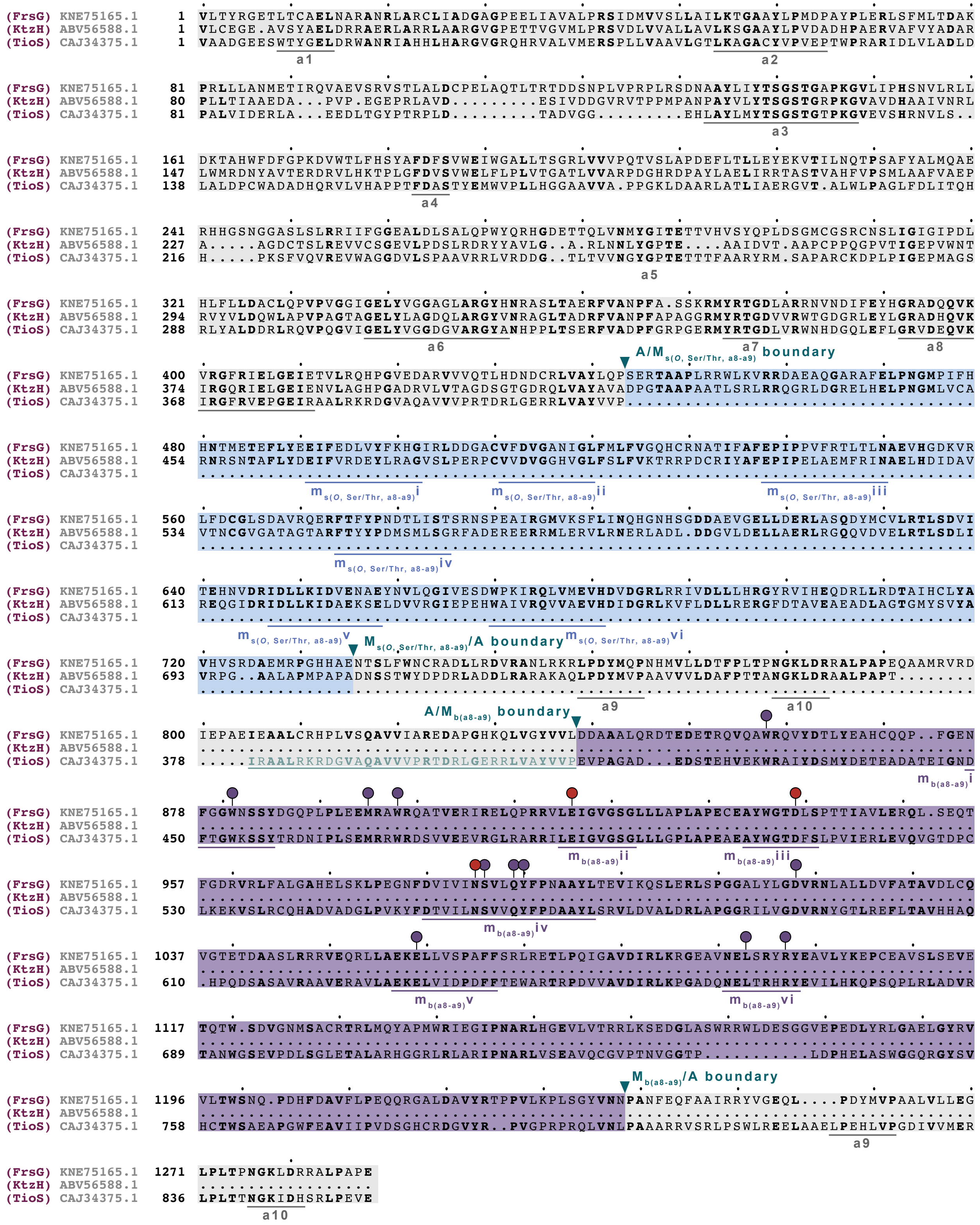
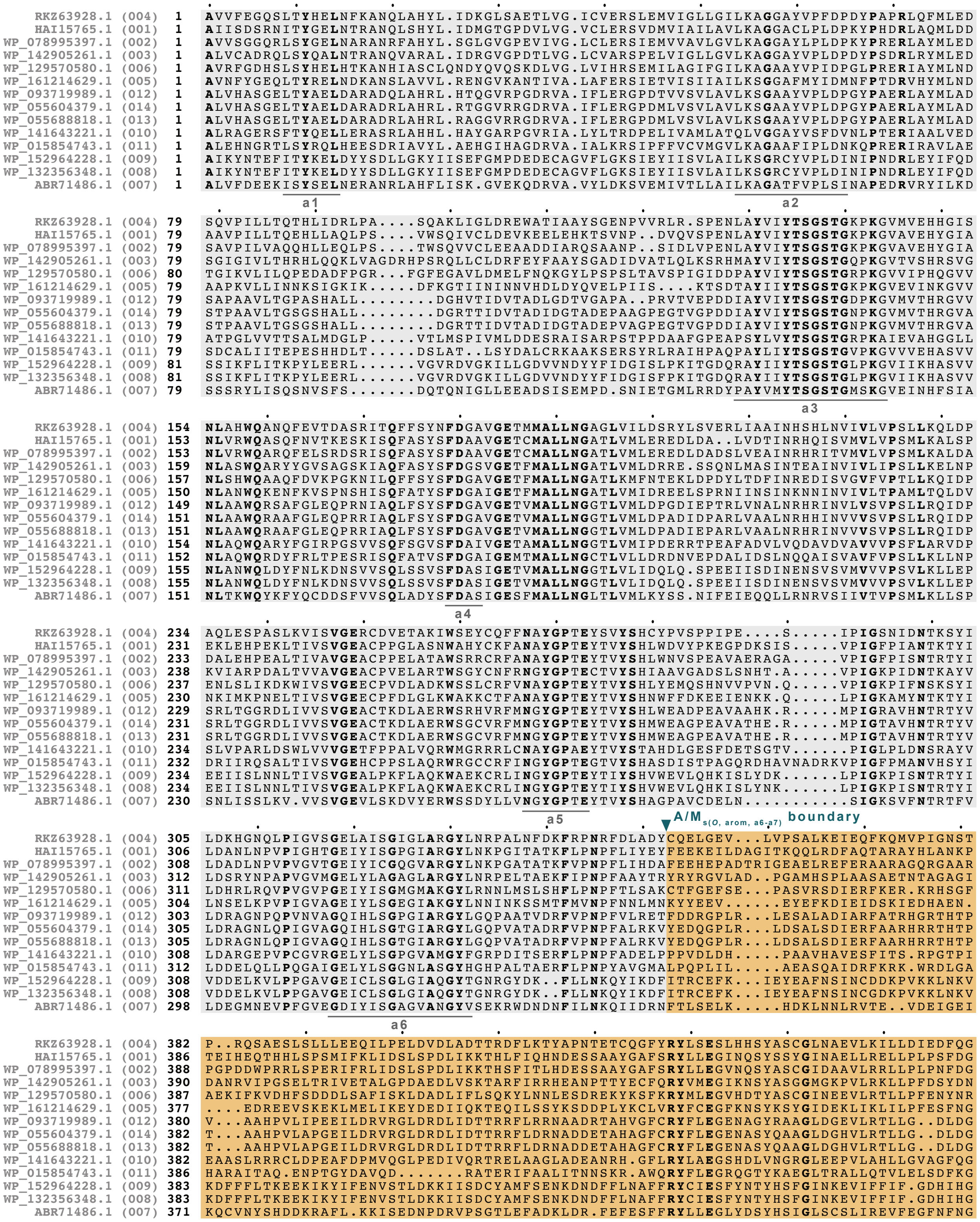
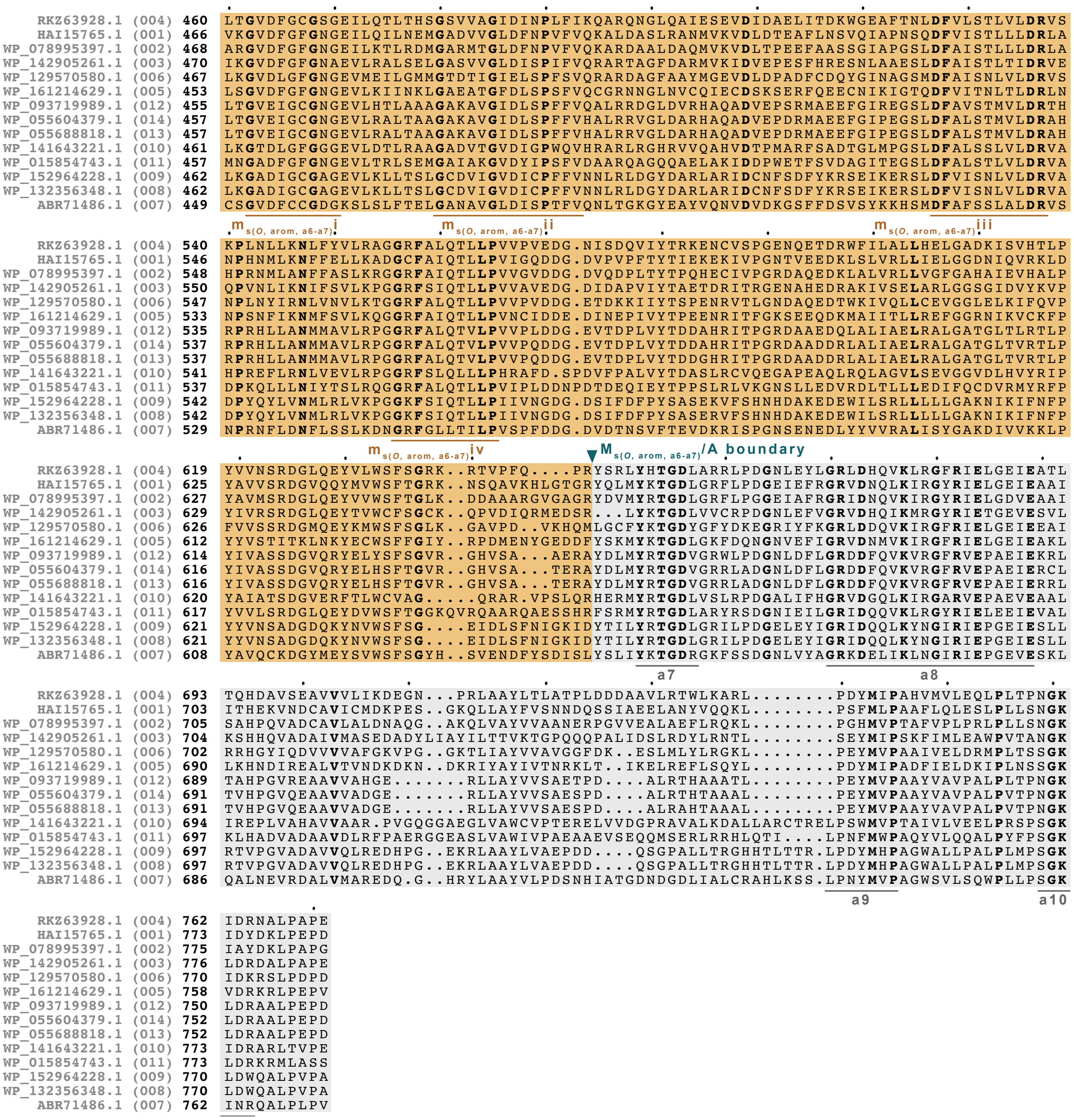


Fig. S19. Multiple sequence alignment of family 5c ($A_8M_{s(O,\ Ser/Thr)}A_{9\text{-}10}M_bA_9$) interrupted A domain. The A domain portion of the sequence is highlighted in light grey, the $M_{s(O,\ Ser/Thr)}$ portion in light blue, and $M_{b(a8\text{-}a9)}$ in light purple. The divisions between the different domains are indicated by teal triangles. Suspected conserved residues for binding of SAM and the amino acid bound Ppant arm, based on the structure of TioS($A_8M_IA_9$),[1] are indicated by red and dark purple balloons, respectively. The conserved motifs for the A domain are underlined by a dark grey bar and labeled a1-a10. The conserved $M_{s(O,\ Ser/Thr,\ a8\text{-}a9)}$ domain motifs (outlined in Table 1) are underlined by dark blue bar and labeled $m_{s(O,\ Ser/Thr,\ a8\text{-}a9)}$i-vi. The conserved $M_{b(a8\text{-}a9)}$ domain motifs (outlined in Table 1) are underlined by a dark purple bar and labeled $m_{b(a8\text{-}a9)}$i-vi. The region in pale teal is a repeat for alignment, but it exists only once in the TioS($A_8M_IA_9$) sequence. A shortened version of this annotated alignment is presented in Fig. 10A.



```
     RKZ63928.1 (004) 1   AVVFEGQSLTYHELNFKANQLAHYL.IDKGLSAETLVG.ICVERSLEMVIGLLGILKAGGAYVPFDPDYPAPRLQFMLED
     HAI15765.1 (001) 1   AIISDSRNITYGELNTRANQLSHYL.IDMGTGPDVLVG.VCLERSVDMVIAILAVLKAGGACLPLDPKYPHDRLAQMLDD
 WP_078995397.1 (002) 1   AVVSGGQRLSYGELNARANRFAHYL.SGLGVGPEVIVG.LCLERSIEMVVAVLGVLKAGGAYLPLDPKYPRERLAYMLED
 WP_142905261.1 (003) 1   ALVCADRQLSYQALNTRANQVARAL.IDRGVGPDTLVG.LCVARSPELVIGLLGVLKAGGAYVPLDPDYPSDRLRYMLED
 WP_129570580.1 (006) 1   AVRFGDHSLSYHELHTKANHIASCLQNDYQVQSKDLVG.LVIHRSEMILAGIFGILKAGAAYVPIDPGLPRERIAYMLND
 WP_161214629.1 (005) 1   AVNFYGEQLTYRELNDKANSLAVVL.RENGVKANTIVA.LAFERSIETVISIIAILKSGGAFMYIDMNFPTDRVHYMLND
 WP_093719989.1 (012) 1   ALVHASGELTYAELDARADQLAHRL.HTQGVRPGDRVA.IFLERGPDTVVSVLAVLKSGAAYVPLDPGYPAERLAYMLAD
 WP_055604379.1 (014) 1   ALVHASGELTYAELDARADRLAHRL.RTGGVRRGDRVA.IFLERGPDMLVSVLAVLKSGAAYVPLDPGYPAERLAYMLAD
 WP_055688818.1 (013) 1   ALVHASGELTYAELDARADRLAHRL.RAGGVRRGDRVA.IFLERGPDMLVSVLAVLKSGAAYVPLDPGYPAERLAYMLAD
 WP_141643221.1 (010) 1   ALRAGERSFTYQELLERASRLAHHL.HAYGARPGVRIA.LYLTRDPELIVAMLATQLVGGAYVSFDVNLPTERIAALVED
 WP_015854743.1 (011) 1   ALEHNGRTLSYRQLHEESDRIAVYL.AEHGIHAGDRVA.IALKRSIPFFVCVMGVLKAGAAFIPLDNKQPRERIRAVLAE
 WP_152964228.1 (009) 1   AIKYNTEFITYKELDYYSDLLGKYIISEFGMPDEDECAGVFLGKSIEYIISVLAILKSGRCYVPLDINIPNDRLEYIFQD
 WP_132356348.1 (008) 1   AIKYNTEFITYKELDYYSDLLGKYIISEFGMPDEDECAGVFLGKSIEYIISVLAILKSGRCYVPLDINIPNDRLEYIFQD
     ABR71486.1 (007) 1   ALVFDEEKISYSELNERANRLAHFLISK.GVEKQDRVA.VYLDKSVEMIVTLLAILKAGATFVPLSINAPEDRVRYILKD

     RKZ63928.1 (004) 79  SQVPILLTQTHLIDRLPA....SQAKLIGLDREWATIAAYSGENPVVRLR.SPENLAYVIYTSGSTGKPKGVMVEHHGIS
     HAI15765.1 (001) 79  AAVPILLTQEHLLAQLPS....VWSQIVCLDEVKEELEHKTSVNP..DVQVSPENLAYVIYTSGSTGKPKGVAVEHYGIA
 WP_078995397.1 (002) 79  SAVPILVAQQHLLEQLPS....TWSQVVCLEEAADDIARQSAANP..SIDLVPENLAYVIYTSGSTGKPKGVAVEHYGIA
 WP_142905261.1 (003) 79  SGIGIVLTHRHLQQKLVAGDRHPSRQLLCLDRFEYFAAYSGADIDVATLQLKSRHMAYVIYTSGSTGQPKGVTVSHRSVG
 WP_129570580.1 (006) 80  TGIKVLILQPEDADFPGR...FGFEGAVLDMELFNQKGYLPSPSLTAVSPIGIDDPAYVIYTSGSTGRPKGVVIPHQGVV
 WP_161214629.1 (005) 79  AAPKVLLINNKSIGKIK.....DFKGTIININNVHDLDYQVELPIIS....KTSDTAYIIYTSGSTGKPKGVEVINKGVV
 WP_093719989.1 (012) 79  SAPAAVLTGPASHALL........DGHVTIDVTADLGDTVGAPA..PRVTVEPDDIAYVIYTSGSTGDPKGVMVTHRGVV
 WP_055604379.1 (014) 79  STPAAVLTGSGSHALL........DGRTTIDVTADIDGTADEPAAGPEGTVGPDDIAYVIYTSGSTGNPKGVMVTHRGVA
 WP_055688818.1 (013) 79  STPAAVLTGSGSHALL........DGRTTIDVTADIDGTADEPVAGPEGTVGPDDIAYVIYTSGSTGNPKGVMVTHRGVA
 WP_141643221.1 (010) 79  ATPGLVVTTSALMDGLP.....VTLMSPIVMLDDESRAISARPSTPPDAAFGPEAPSYLVYTSGSTGRPKAIEVAHGGLL
 WP_015854743.1 (011) 79  SDCALIITEPESHHDLT.....DSLAT..LSYDALCRKAAKSERSYRLRAIHPAQPAYLIYTSGSTGVPKGVVVEHASVV
 WP_152964228.1 (009) 81  SSIKFLITKPYLEERL......VGVRDVGKILLGDVVNDYYFIDGISLPKITGDQRAYIIYTSGSTGLPKGVIIKHASVV
 WP_132356348.1 (008) 81  SSIKFLITKPYLEERL......VGVRDVGKILLGDVVNDYYFIDGISFPKITGDQRAYIIYTSGSTGLPKGVIIKHASVV
     ABR71486.1 (007) 79  SSSRYLISQSNVSFS.......DQTQNIGLLEEADSISEMPD.SNIETGMLRRDYPAYVMYTSGSTGMSKGVEINHFSIA

     RKZ63928.1 (004) 154 NLAHWQANQFEVTDASRITQFFSYNFDGAVGETMMALLNGAGLVILDSRYLSVERLIAAINHSHLNVIVLVPSLLKQLDP
     HAI15765.1 (001) 153 NLVRWQASQFNVTKESKISQFASYSFDAAVGETCMALLNGATLVMLEREDLDA..LVDTINRHQISVMVLVPSMLKALSP
 WP_078995397.1 (002) 153 NLVRWQARQFELSRDSRISQFASYSFDAAVGETCMALLNGATLVMLEREDLDADSLVEAINRHRITVMVLVPSMLKALDA
 WP_142905261.1 (003) 159 NLASWQARYYGVSAGSKIAQFASYSFDGSVGETFMALLNGATLVMLDRRE.SSQNLMASINTEAINVIVLIPSLLKELNP
 WP_129570580.1 (006) 157 NLSHWQAAQFDVKPGKNILQFFSYSFDGAVGETFMALLNGATLKMFNTEKLDPDYLTDFINREDISVGVFVPTLLKQIDP
 WP_161214629.1 (005) 150 NLANWQKENFKVSPNSHISQFATYSFDGAIGETFMALLNGATLVMIDREELSPRNIINSINKNNINVIVLTPAMLTQLDV
 WP_093719989.1 (012) 149 NLAAWQRSAFGLEQPRNIAQLFSYSFDGAVGETVMALLNGGTLVMLDPGDIEPTRLVNALNRHRINVLVSVPSLLRQIDP
 WP_055604379.1 (014) 151 NLAAWQRAAFGLEQPRRIAQLFSYSFDGAVGETVMALLNGGTLVMLDPADIDPARLVAALNRHHINVVVSVPSLLRQIDP
 WP_055688818.1 (013) 151 NLAAWQRAAFGLEQPRRIAQLFSYSFDGAVGETVMALLNGGTLVMLDPADIEPARLVAALNRHRINVVVSVPSLLRQIDP
 WP_141643221.1 (010) 154 NLAQWQARYFGIRPGSVVSQFSGVSFDAIVGETAMALLNGGTLVMIPDERRTPEAFADVLVQDAVDVAVVVPSFLARVDP
 WP_015854743.1 (011) 152 NLAQWQRDYFRLTPESRISQFATVSFDGAIGEMTMALLNGCTLVLLDRDNVEPDALIDSLNQQAISVAVFVPSLLKLLNP
 WP_152964228.1 (009) 155 NLANWQLDYFNLKDNSVVSQLSSVSFDASIGETVMALLNGGTLVLIDQLQ.SPEEIISDINENSVSVMVVVPSVLKLLEP
 WP_132356348.1 (008) 155 NLANWQLDYFNLKDNSVVSQLSSVSFDASIGETVMALLNGGTLVLIDQLQ.SPEEIISDINENSVSVMVVVPSVLKLLEP
     ABR71486.1 (007) 151 NLTKWQYKFYQCDDSFVVSQLADYSFDASIGESFMALLNGLTLVMLKYSS.NIFEIEQQLLRNRVSIIVTVPSMLKLLSP

     RKZ63928.1 (004) 234 AQLESPASLKVISVGERCDVETAKIWSEYCQFFNAYGPTEYSVYSHCYPVSPPIPE....S.....IPIGSNIDNTKSYI
     HAI15765.1 (001) 231 EKLEHPEKLTIVSVGEACPPGLASNWAHYCKFANAYGPTEYTVYSHLWDVYPKEGPDKSIS.....VPIGFPIANTKTYI
 WP_078995397.1 (002) 233 DALEHPEALTIVAVGEACPPELATAWSRRCRFANAYGPTEYTVYSHLWNVSPEAVAERAGA.....VPIGFPISNTRTYI
 WP_142905261.1 (003) 238 KVIARPDALTVVAVGEACPVELARTWSGYCNFRNGYGPTECTVYSHIAAVGADSLSNHT.A.....VPIGKPIDNTKAYV
 WP_129570580.1 (006) 237 ENLSLIKDKWIVSVGEECPVDLADKWSSLCRFVNAYGPTEYTVYSHLYEMQSHNVVPVN.Q.....VPIGKPIFNSKSYI
 WP_161214629.1 (005) 230 NKIMKPNELTIVSVGEECPFDLGLKWAKKCTFANAYGPTEYTVYSHNWFFDKEEIENKK.Q.....LPIGKAMYNTKTYI
 WP_093719989.1 (012) 229 SRLTGGRDLIVVSVGEACTKDLAERWSRHVRFMNGYGPTEYTVYSHLWEADPEAVAAHK.R.....MPIGRAVHNTRTYV
 WP_055604379.1 (014) 231 SRLTGGRDLIVVSVGEACTKDLAERWSGCVRFMNGYGPTEYTVYSHMWEAGPEAVATHE.R.....MPIGTAVHNTRTYV
 WP_055688818.1 (013) 231 SRLTGGRDLIVVSVGEACTKDLAERWSGCVRFMNGYGPTEYTVYSHMWEAGPEAVATHE.R.....MPIGTAVHNTRTYV
 WP_141643221.1 (010) 234 SLVPARLDSWLVVVGETFPPALVQRWMGRRRLCNAYGPAEYTVYSTAHDLGESFDETSGTV......PIGLPLDNSRAYV
 WP_015854743.1 (011) 232 DRIIRQSALTIVSVGEHCPPSLAQRWRGCCRFINGYGPTEGTVYSHASDISTPAGQRDHAVNADRKVPIGFPMANVHSYI
 WP_152964228.1 (009) 234 EEIISLNNLTIVSVGEALPKFLAQKWAEKCRLINGYGPTEYTIYSHVWEVLQHKISLYDK......LPIGKPISNTRTYI
 WP_132356348.1 (008) 234 EEIISLNNLTIVSVGEALPKFLAQKWAEKCRLINGYGPTEYTIYSHVWEVLQHKISLYNK......LPIGKPISNTRTYI
     ABR71486.1 (007) 230 SNLISSLKV.VVSVGEVLSKDVYERWSSDYLLVNGYGPTEYTVYSHAGPAVCDELN...........SIGKSVFNTRTFV

     RKZ63928.1 (004) 305 LDKHGNQLPIGVSGELAISGIGLARGYLNRPALNFDKFRPNRFDLADYCQELGEV...LVPSALKEIEQFKQMVPIGNST
     HAI15765.1 (001) 306 LDANLNPVPIGHTGEIYISGPGIARGYLNKPGITATKFLPNPFLIYEYFEEKEILDAGITKQQLRDFAQTARAYHLANKP
 WP_078995397.1 (002) 308 LDADLNPVPVGVTGEIYICGQGVARGYLNKPGATATKFVPNPFLIHDAFEEHEPADTRIGEAELREFERAARAGQRGAAR
 WP_142905261.1 (003) 312 LDSRYNPAPVGVMGELYLAGAGLARGYLNRPELTAEKFIPNPFAAYTRYRYRGVLAD..PGAMHSPLAASAETNTAGAGI
 WP_129570580.1 (006) 311 LDHRLRQVPVGVPGEIYISGMGMAKGYLRNNLMSLSHFLPNPFTLSAKCTFGEFSE...PASVRSDIERFKER.KRHSGF
 WP_161214629.1 (005) 304 LNSELKPVPIGVAGEIYLSGEGIAKGYLNNINKSSMTFMVNPFNNLMNKYYEEV.....EYEFKDIEIDSKIEDHAEN.
 WP_093719989.1 (012) 303 LDRAGNPQPVNVAGQIHLSGPGIARGYLGQPAATVDRFVPNPFVLRETFDDRGPLR...LESALADIARFATRHGRTHTP
 WP_055604379.1 (014) 305 LDRAGNLQPIGVAGQIHLSGTGIARGYLGQPVATADRFVPNPFALRKVYEDQGPLR...LDSALSDIERFAARHRRTHTP
 WP_055688818.1 (013) 305 LDRAGNLQPIGVAGQIHLSGTGIARGYLGQPVATADRFVPNPFALRKVYEDQGPLR...LDSALSDIERFAARHRRTHTP
 WP_141643221.1 (010) 308 LDARGEPVPCGVRGELYLSGPGVAMGYFGRPDITSERFLPNPFADELPPPVDLDH....PAAVHAVESFITS.RPGTPI
 WP_015854743.1 (011) 312 LDDELQLLPQGAIGELYLSGGNLASGYHGHPALTAERFLPNPYAVGMALPQPLIL....AEASQAIDRFKRK.WRDLGA
 WP_152964228.1 (009) 308 VDDELKVLPPGAVGEICLSGLGIAQGYTGNRGYDK..FLLNKQYIKDFITRCEFK...IEYEAFNSINCDDKPVKKLNKV
 WP_132356348.1 (008) 308 VDDELKVLPPGAVGEICLSGLGIAQGYTGNRGYDK..FLLNKQYIKDFITRCEFK...IEYEAFNSINCGDKPVKKLNKV
     ABR71486.1 (007) 298 LDEGMNEVPFGVEGDIYISGAGVANGYVSEKRWDNDNFILNKQIIDRNFTLSELK....HDKLNNLRVTE..VDEIGEI

     RKZ63928.1 (004) 382 P..RQSAESLSLLLEEQILPELDVDLADTTRDFLKTYAPNTETCQGFYRYLSESLHHSYASCGLNAEVLKILLDIEDFQG
     HAI15765.1 (001) 386 TEIHEQTHHLSPSMIFKLIDSLSPDLIKKTHLFIQHNDESSAAYGAFSRYLLEGINQSYSSYGINAELLRELLPLPSFDG
 WP_078995397.1 (002) 388 PGPDDWPRRLSPERIFRLIDSLSPDLIKKTHSFITLHDESSAAYGAFSRYLLEGVNQSYASCGIDAAVLRRLLPLPNFDG
 WP_142905261.1 (003) 390 DANRVIPGSELTRIVETALGPDAEDLVSKTARFIRRHEANPTTYECFQRYVMEGIKNSYASGGMGKPVLRKLLPFDSYDN
 WP_129570580.1 (006) 387 AEKIFKVDHFSDDDLSAFISKLDADLIFLSQKYLNNLESDREKYKSFKRYMLEGVHDTYASCGINEEVLRTLLPFENYNR
 WP_161214629.1 (005) 377 ....EDREEVSKEKLMELIKEYDEDIIQKTEQILSSYKSDDPLYKCLVRYFCEGFKNSYKSYGIDEKLIKLILPFESFNG
 WP_093719989.1 (012) 380 V...AAHPVLIPEEILDRVRGLDRDLIDTTRRFLRRNAADRTAHVGFCRYFLEGENAGYRAAGLDGEVLRTLLG..DLDG
 WP_055604379.1 (014) 382 T...AAHPVLAPGEILDRVRGLDRDLIDTTRRFLDRNADDETAHAGFCRYFLEGENASYQAAGLDGHVLRTLLG..DLDG
 WP_055688818.1 (013) 382 T...AAHPVLAPGEILDRVRGLDRDLIDTTRRFLDRNADDETAHAGFCRYFLEGENASYQAAGLDGHVLRTLLG..DLDG
 WP_141643221.1 (010) 382 EAASLRRRCLDPEAFDPMVQGLPEDIVQRTRELAAGLADEANRH.GFLRYLAEGSHDLVNGRGLEEPVLAHLLGVAGFQG
 WP_015854743.1 (011) 386 HARAITAQ.ENPTGYDAVQD.......RATERIFAALITNNSKR.AWQRYFLEGRQGTYKAEGLTRALLQTVLELSDFKG
 WP_152964228.1 (009) 383 KDFFFLTKEEKIKYIFENVSTLDKKIISDCYAMFSENKDNDFFLNAFFRYCIESFYNTYHSFGINKEVIFFIF.GDHIHG
 WP_132356348.1 (008) 383 KDFFFLTKEEKIKYIFENVSTLDKKIISDCYAMFSENKDNDFFLNAFFRYCIESFYNTYHSFGINKEVIFFIF.GDHIHG
     ABR71486.1 (007) 371 KQCVNYSHDDKRAFL.KKISEDNPDRVPSGTLEFADKLDR.FEFESFFRYLLEGLYDSYHSIGLSCSLIREVFEGFNFNG
```



```
RKZ63928.1 (004) 460 LTGVDFGCGSGEILQTLTHSGSVVAGIDINPLFIKQARQNGLQAIESEVDIDAELITDKWGEAFTNLDFVLSTLVLDRVS
HAI15765.1 (001) 466 VKGVDFGFGNGEILQILNEMGADVVGLDFNPVFVQKALDASLRANMVKVDLDTEAFLNSVQIAPNSQDFVISTLLLDRLA
WP_078995397.1 (002) 468 ARGVDFGFGNGEILKTLRDMGARMTGLDFNPVFVQKARDAALDAQMVKVDLTPEEFAASSGIAPGSLDFAISTLLLDRLA
WP_142905261.1 (003) 470 IKGVDFGFGNAEVLRALSELGASVVGLDISPIFVQRARTAGFDARMVKIDVEPESFHRESNLAAESLDFAISTLTIDRVE
WP_129570580.1 (006) 467 LKGVDLGFGNGEVMEILGMMGTDTIGIELSPFSVQRARDAGFAAYMGEVDLDPADFCDQYGINAGSMDFAISNLVLDRVS
WP_161214629.1 (005) 453 LSGVDFGFGNGEVLKIINKLGAEATGFDLSPSFVQCGRNNGLNVCQIECDSKSERFQEECNIKIGTQDFVITNLTLDRLN
WP_093719989.1 (012) 455 LTGVEIGCGNGEVLHTLAAAGAKAVGIDLSPFFVQALRRDGLDVRHAQADVEPSRMAEEFGIREGSLDFAVSTMVLDRTH
WP_055604379.1 (014) 457 LTGVEIGCGNGEVLRALTAAGAKAVGIDLSPFFVHALRRVGLDARHAQADVEPDRMAEEFGIPEGSLDFALSTMVLDRAH
WP_055688818.1 (013) 457 LTGVEIGCGNGEVLRALTAAGAKAVGIDLSPFFVHALRRVGLDARHAQADVEPDRMAEEFGIPEGSLDFALSTMVLDRAH
WP_141643221.1 (010) 461 LKGTDLGFGGGEVLDTLRAAGADVTGVDIGPWQVHRARLRGHRVVQAHVDTPMARFSADTGLMPGSLDFALSSLVLDRVA
WP_015854743.1 (011) 457 MNGADFGFGNGEVLTRLSEMGAIAKGVDYIPSFVDAARQAGQQAELAKIDDPWETFSVDAGITEGSLDFALSTLVLDRVA
WP_152964228.1 (009) 462 LKGADIGCGAGEVLKLLTSLGCDVIGVDICPFFVNNLRLDGYDARLARIDCNFSDFYKRSEIKERSLDFVISNLVLDRVA
WP_132356348.1 (008) 462 LKGADIGCGAGEVLKLLTSLGCDVIGVDICPFFVNNLRLDGYDARLARIDCNFSDFYKRSEIKERSLDFVISNLVLDRVA
ABR71486.1 (007) 449 CSGVDFCCGDGKSLSLFTELGANAVGLDISPTFVQNLTGKGYEAYVQNVDVDKGSFFDSVSYKKHSMDFAFSSLALDRVS

RKZ63928.1 (004) 540 KPLNLLKNLFYVLRAGGRFALQTLLPVVPVEDG.NISDQVIYTRKENCVSPGENQETDRWFILALLHELGADKISVHTLP
HAI15765.1 (001) 546 NPHNMLKNFFELLKADGCFAIQTLLPVIGQDDG.DVPVPFTYTIEKEKIVPGNTVEEDKLSLVRLLIELGGDNIQVRKLD
WP_078995397.1 (002) 548 HPRNMLANFFASLKRGGRFAIQTLLPVVGVDDG.DVQDPLTYTPQHECIVPGRDAQEDKLALVRLLVGFGAHAIEVHALP
WP_142905261.1 (003) 550 QPVNLIKNIFSVLKPGGRFSIQTLLPVVAVEDG.DIDAPVIYTAETDRITRGENAHEDRAKIVSELARLGGSGIDVYKVP
WP_129570580.1 (006) 547 NPLNYIRNLVNVLKTGGRFALQTLLPVVPVDDG.ETDKKIIYTSPENRVTLGNDAQEDTWKIVQLLCEVGGLELKIFQVP
WP_161214629.1 (005) 533 NPSNFIKNMFSVLKQGGRFAIQTLLPVNCIDDE.DINEPIVYTPEENRITFGKSEEQDKMAIITLLREFGGRNIKVCKFP
WP_093719989.1 (012) 535 RPRHLLANMMAVLRPGGRFALQTVLPVVPLDDG.EVTDPLVYTDDAHRITPGRDAAEDQLALIAELRALGATGLTLRTLP
WP_055604379.1 (014) 537 RPRHLLANMMAVLRPGGRFALQTVLPVVPQDDG.EVTDPLVYTDDAHRITPGRDAADDRLALIAELRALGATGLTVRTLP
WP_055688818.1 (013) 537 RPRHLLANMMAVLRPGGRFALQTVLPVVPQDDG.EVTDPLVYTDDGHRITPGRDAADDRLALIAELRALGATGLTVRTLP
WP_141643221.1 (010) 541 HPREFLRNLVEVLRPGGRFSLQLLLPHRAFD.SPDVFPALVYTDASLRCVQEGAPEAQLRQLAGVLSEVGGVDLHVYRIP
WP_015854743.1 (011) 537 DPKQLLLNIYTSLRQGGRFALQTLLPVIPLDDNPDTDEQIEYTPPSLRLVKGNSLLEDVRDLTLLLEDIFQCDVRMYRFP
WP_152964228.1 (009) 542 DPYQYLVNMLRLVKPGGKFSIQTLLPIIVNGDG.DSIFDFPYSASEKVFSHNHDAKEDEWILSRLLLLLGAKNIKIFNFP
WP_132356348.1 (008) 542 DPYQYLVNMLRLVKPGGKFSIQTLLPIIVNGDG.DSIFDFPYSASERVFSHNHDAKEDEWILSRLLLLLGAKNIKIFNFP
ABR71486.1 (007) 529 NPRNFLDNFLSSLKDNGRFGLLTILPVSPFDDG.DVDTNSVFTEVDKRISPGNSENEDLYYVVRALISYGAKDIVVKKLP

RKZ63928.1 (004) 619 YVVNSRDGLQEYVLWSFSGRK..RTVPFQ....PRYSRLYHTGDLARRLPDGNLEYLGRLDHQVKLRGFRIELGEIEATL
HAI15765.1 (001) 625 YAVVSRDGVQQYMVWSFTGRK..NSQAVKHLGTGRYQLMYKTGDLGRFLPDGEIEFRGRVDNQLKIRGYRIELGEIEAAI
WP_078995397.1 (002) 627 YAVMSRDGLQEYVVWSFTGLK..DDAAARGVGAGRYDVMYRTGDLGRFLPGGEIAFRGRIDNQVKIRGYRIELGDVEAAI
WP_142905261.1 (003) 629 YIVRSRDGLQEYTVWCFSGCK..QPVDIQRMEDSR...LYKTGDLVVCRPDGNLEFVGRVDHQIKMRGYRIETGEVESVL
WP_129570580.1 (006) 626 FVVSSRDGMQEYKMWSFSGLK..GAVPD..VKHQMLGCFYKTGDYGFYDKEGRIYFKGRLDDQVKIRGFRIELGEIEEAI
WP_161214629.1 (005) 612 YYVSTITKLNKYECWSFFGIY..RPDMENYGEDDFYSKMYKTGDLGKFDQNGNVEFIGRVDNMVKIRGFRIELGEIEAKL
WP_093719989.1 (012) 614 YIVASSDGVQRYELYSFSGVR..GHVSA...AERAYDLMYRTGDVGRWLPDGNLDFLGRDDFQVKVRGFRVEPAEIEKRL
WP_055604379.1 (014) 616 YIVASSDGVQRYELHSFTGVR..GHVSA...TERAYDLMYRTGDVGRRLADGNLDFLGRDDFQVKVRGFRVEPAEIERCL
WP_055688818.1 (013) 616 YIVASSDGVQRYELHSFTGVR..GHVSA...TERAYDLMYRTGDVGRRLADGNLDFLGRDDFQVKVRGFRVEPAEIERRL
WP_141643221.1 (010) 620 YAIATSDGVERFTLWCVAG.....QRAR.VPSLQRHERMYRTGDLVSLRPDGALIFHGRVDGQLKIRGARVEPAEVEAAL
WP_015854743.1 (011) 617 YVVLSRDGLQEYDVWSFTGGKQVRQAARQAESSHRFSRMYRTGDLARYRSDGNIEILGRIDQQVKLRGYRIELEEIEVAL
WP_152964228.1 (009) 621 YYVNSADGDQKYNVWSFSG....EIDLSFNIGKIDYTILYRTGDLGRILPDGELEYIGRIDQQLKYNGIRIEPGEIESLL
WP_132356348.1 (008) 621 YYVNSADGDQKYNVWSFSG....EIDLSFNIGKIDYTILYRTGDLGRILPDGELEYIGRIDQQLKYNGIRIEPGEIESLL
ABR71486.1 (007) 608 YAVQCKDGYMEYSVWSFSGYH..SVENDFYSDISLYSLIYKTGDRGKFSSDGNLVYAGRKDELIKLNGIRIEPGEVESKL

RKZ63928.1 (004) 693 TQHDAVSEAVVVLIKDEGN...PRLAAYLTLATPLDDDAAVLRTWLKARL........PDYMIPAHVMVLEQLPLTPNGK
HAI15765.1 (001) 703 ITHEKVNDCAVICMDKPES..GKQLLAYFVSNNDQSSIAEELANYVQQKL........PSFMLPAAFLQLESLPLLSNGK
WP_078995397.1 (002) 705 SAHPQVADCAVLALDNAQG..AKQLVAYVVAANERPGVVEALAEFLRQKL........PGHMVPTAFVPLPRLPLLSNGK
WP_142905261.1 (003) 704 KSHHQVADAIVMASEDADYLIAYILTTVKTGPQQQPALIDSLRDYLRNTL........SEYMIPSKFIMLEAWPVTANGK
WP_129570580.1 (006) 702 RRHGYIQDVVVVAFGKVPG..GKTLIAYVVAVGGFDK..ESLMLYLRGKL........PEYMVPAAIVELDRMPLTSSGK
WP_161214629.1 (005) 690 LKHNDIREALVTVNDKDKN..DKRIYAYIVTNRKLT..IKELREFLSQYL........PDYMIPADFIELDKIPLNSSGK
WP_093719989.1 (012) 689 TAHPGVREAAVVAHGE.......RLLAYVVSAETPD....ALRTHAAATL........PEYMVPAAYVAVPALPLTPNGK
WP_055604379.1 (014) 691 TVHPGVQEAAVVADGE.......RLLAYVVSAESPD....ALRTHTAAAL........PEYMVPAAYVAVPALPVTPNGK
WP_055688818.1 (013) 691 TVHPGVQEAAVVADGE.......RLLAYVVSAESPD....ALRAHTAAAL........PEYMVPAAYVAVPALPVTPNGK
WP_141643221.1 (010) 694 IREPLVAHAVVAAR.PVGQGGAEGLVAWCVPTERELVVDGPRAVALKDALLARCTRELPSWMVPTAIVLVEELPRSPSGK
WP_015854743.1 (011) 697 KLHADVADAAVDLRFPAERGGEASLVAWIVPAEAAEVSEQQMSERLRRHLQTI....LPNFMWPAQYVLQQALPYFPSGK
WP_152964228.1 (009) 697 RTVPGVADAVVQLREDHPG..EKRLAAYLVAEPDD....QSGPALLTRGHHTLTTR.LPDYMHPAGWALLPALPLMPSGK
WP_132356348.1 (008) 697 RTVPGVADAVVQLREDHPG..EKRLAAYLVAEPDD....QSGPALLTRGHHTLTTR.LPDYMHPAGWALLPALPLMPSGK
ABR71486.1 (007) 686 QALNEVRDALVMAREDQ.G..HRYLAAYVLPDSNHIATGDNDGDLIALCRAHLKSS.LPNYMVPAGWSVLSQWPLLPSGK

RKZ63928.1 (004) 762 IDRNALPAPE
HAI15765.1 (001) 773 IDYDKLPEPD
WP_078995397.1 (002) 775 IAYDKLPAPG
WP_142905261.1 (003) 776 LDRDALPAPE
WP_129570580.1 (006) 770 IDKRSLPDPD
WP_161214629.1 (005) 758 VDRKRLPEPV
WP_093719989.1 (012) 750 LDRAALPEPD
WP_055604379.1 (014) 752 LDRAALPEPD
WP_055688818.1 (013) 752 LDRAALPEPD
WP_141643221.1 (010) 773 IDRARLTVPE
WP_015854743.1 (011) 773 LDRKRMLASS
WP_152964228.1 (009) 770 LDWQALPVPA
WP_132356348.1 (008) 770 LDWQALPVPA
ABR71486.1 (007) 762 INRQALPLPV
```

Fig. S20. Multiple sequence alignment of family 6 ($A_6M_{s(\mathit{O},\ \mathrm{arom})}A_7$) interrupted A domains. The numbering next to the accession number for each sequence matches the number in the taxonomic tree (Table S7). The A domain portion of the sequence is highlighted in light grey and the $M_{s(\mathit{O},\ \mathrm{arom},\ \mathrm{a6\text{-}a7})}$ portion in light orange. The divisions between the different domains are indicated by teal triangles. The conserved motifs for the A domain are underlined by a dark grey bar and labeled a1-a10. The conserved $M_{s(\mathit{O},\ \mathrm{arom},\ \mathrm{a6\text{-}a7})}$ domain motifs (outlined in Table 1) are underlined by dark orange bar and labeled $m_{s(\mathit{O},\ \mathrm{arom},\ \mathrm{a6\text{-}a7})}$ i-iv. A shortened version of this annotated alignment is presented in Fig. 8B.

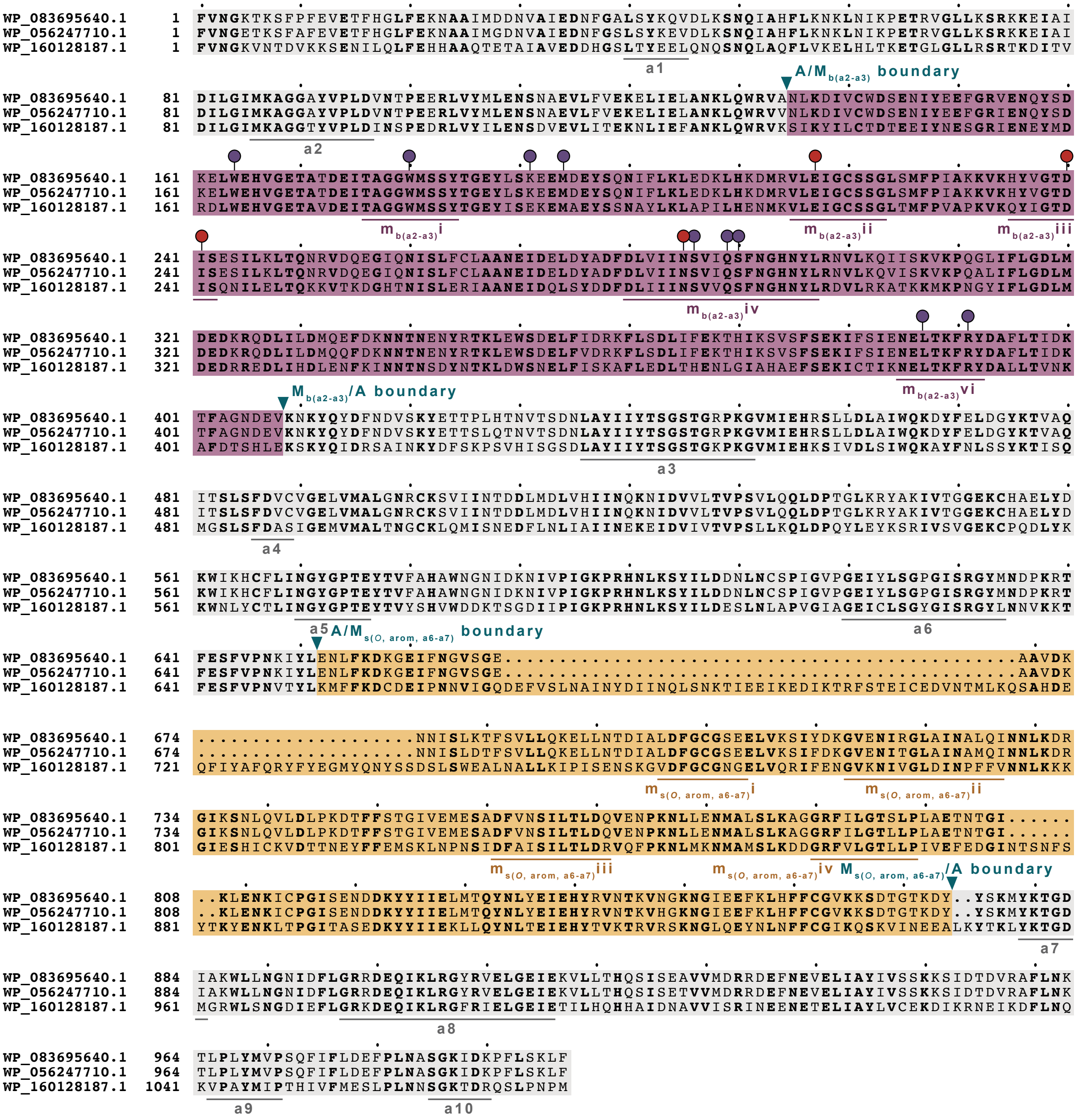


Fig. S21. Multiple sequence alignment of family 7 ($A_2M_bA_{3\text{-}6}M_{s(O,\text{ arom})}A_7$) interrupted A domains. The A domain portion of the sequence is highlighted in light grey, the $M_{b(a2\text{-}a3)}$ portion in light pink, and the $M_{s(O,\text{ arom, a6-a7})}$ portion in light orange. The divisions between the different domains are indicated by teal triangles. Suspected conserved residues for binding of SAM and the amino acid bound Ppant arm, based on the structure of TioS($A_8M_IA_9$),[1] are indicated by red and dark purple balloons, respectively. The conserved motifs for the A domain are underlined by a dark grey bar and labeled a1-a10. The conserved M domain motifs (outlined in Table 1) are underlined by a dark pink bar and labeled $m_{b(a2\text{-}a3)}$i-vi. *Note*: There is no $m_{b(a2\text{-}a3)}$v but the numbering was kept to match the same motifs found in $M_{b(a8\text{-}a9)}$. The conserved $M_{s(O,\text{ arom, a6-a7})}$ domain motifs (outlined in Table 1) are underlined by dark orange bar and labeled $m_{s(O,\text{ arom, a6-a7})}$ i-iv. A shortened version of this annotated alignment is presented in Fig. 10B.

3. Conserved M domain motifs and assignment of M domain types. The conserved motifs for all six types of M domains are listed in Table 1 in the main text. For types I and V, $\mathrm{M}_{\mathrm{b(a8\text{-}a9)}}$ and $\mathrm{M}_{\mathrm{b(a2\text{-}a3)}}$, respectively, the conserved motifs we found matched those in the literature for *N*-methylating M domains.[18, 19] For type II, $\mathrm{M}_{\mathrm{s}(O,\ \mathrm{Ser/Thr},\ \mathrm{a8\text{-}a9})}$, there was one that matched previously published motifs for *O*-methylating M domains,[19] but the rest were assigned manually based on the conserved residues observed in the alignment (Fig. S14). For type III, $\mathrm{M}_{\mathrm{s}(O,\ \mathrm{Tyr},\ \mathrm{a8\text{-}a9})}$, there was one motif that matched previously published motifs for *O*-methylating M domains,[19] but the others were assigned manually based on the conserved residues observed in the alignment (Fig. S15). For type IV, $\mathrm{M}_{\mathrm{s}(S,\ \mathrm{a2\text{-}a3})}$, there were no reported motifs for *S*-methylating M domains that matched, so they were assigned manually based on the conserved amino acids present in the alignment (Fig. S16). For type VI, $\mathrm{M}_{\mathrm{s}(O,\ \mathrm{arom},\ \mathrm{a6\text{-}a7})}$, there were no reported motifs for *O*-methylating M domains that matched this domain, so they were assigned manually based on the conserved amino acids present in the alignment (Fig. S20). To help with the assignment of types of M domains, two approaches were used. First, the phylogenetic tree of M domains (Fig. 5), which showed distinct grouping of each type of M domain. Second, for each type of M domain, a representative was selected and only the M domain portion was used in a BLAST search to identify the conserved domains listed in Table S8.

Table S8. Conserved regions of M domain types.

Type of M domain	**Representative used**	**Name of hit**	**Accession**	**Description**	**Superfamily**
I	$\mathrm{TioS(M_I)_4}$	Methyltransf_12	pfam08242	Methyltransferase domain	c|17173; AdoMet_MTases
		AdoMet_Mtases	cd02440	SAM-dependent methyltransferases class I	c|17173; AdoMet_MTases
		PRK06922	PRK06922	SAM-dependent methyltransferases class I	c|30734; PRK06922
		SmtA	COG0500	SAM-dependent methyltransferase (Secondary metabolites biosynthesis)	c|17173; AdoMet_MTases
		bacter_Hen1	TIGR04074	3' terminal RNA ribose 2'-*O*-methyltransferase Hen1	c|37382; bacter_Hen1
II	$\mathrm{KtzH(M_{II})_4}$	fkbM_fam	TIGR01444	Methyltransferase, FkbM family	c|17173; AdoMet_MTases
		Methyltransf_21	pfam05050	Methyltransferase, FkbM domain	c|17173; AdoMet_MTases
		TrmN6	COG4123	RNA(Val) A37 N6-methylase TrnN6	c|17173; AdoMet_MTases
III	$\mathrm{DidJ(M_{III})_{11}}$	CbiT	TIGR02469	Precorrin-6Y C5, 15-methyltransferase, CbiT subunit	c|17173; AdoMet_MTases
		ksgA	PRK14896	16S ribosomal RNA methyltransferase A	c|17173; AdoMet_MTases
		CobL	COG2242	Precorrin-6B methylase 2	c|28096; CobL
		Methyltransf_18	pfam12847	Methyltransferase domain	c|17173; AdoMet_MTases
		AdoMet_Mtases	cd02440	SAM-dependent methyltransferases class I	c|17173; AdoMet_MTases

		PKS_MT	smart00828	Methyltransferase in polyketide synthase (PKS) enzymes	c|29216; PKS_MT
IV	TioN(M_{IV})	Methyltransf_25	pfam13649	Methyltransferase domain	c|17173; AdoMet_MTases
		hemK_rel_arch	TIGR00537	Hemk-related putative methylase	c|17173; AdoMet_MTases
		AdoMet_Mtases	cd02440	SAM-dependent methyltransferases class I	c|17173; AdoMet_MTases
V	TtbB(M_V)$_5$	AdoMet_Mtases	cd02440	SAM-dependent methyltransferases class I	c|17173; AdoMet_MTases
		Methyltransf_12	pfam08242	Methyltransferase domain	c|17173; AdoMet_MTases
		RF_mod_PrmC	TIGR03534	Protein-(glutamine-N5) methyltransferase	c|37287; RF_mod_PrmC
		PRK14968	PRK14968	Putative methyltransferase	c|17173; AdoMet_MTases
		UbiG	COG2227	2-polyprenyl-3-methyl-5-hydroxy-6-metoxy-1,4-benzoquinol methylase	c|34436; UbiG
VI	RKZ63928.1(M_{VI})	Methyltransf_11	pfam08241	Methyltransferase domain	c|17173; AdoMet_MTases
		UbiG	COG2227	2-polyprenyl-3-methyl-5-hydroxy-6-metoxy-1,4-benzoquinol methylase	c|17173; AdoMet_MTases
		AdoMet_Mtases	cd02440	SAM-dependent methyltransferases class I	c|17173; AdoMet_MTases
		UbiG	TIGR01983	Ubiquinone biosynthesis *O*-methyltransferase	c|37019; UbiG
		PLN02396	PLN02396	Hexaprenyldihydroxybenzoate methyltransferase	c|31874; PLN02396

4. References

1. Mori, S.; Pang, A. H.; Lundy, T. A.; Garzan, A.; Tsodikov, O. V.; Garneau-Tsodikova, S. Structural basis for backbone *N*-methylation by an interrupted adenylation domain. *Nat. Chem. Biol.* **2018,** 14, 428-430.
2. Zolova, O. E.; Garneau-Tsodikova, S. KtzJ-dependent serine activation and *O*-methylation by KtzH for kutznerides biosynthesis. *J. Antibiot. (Tokyo)* **2014,** 67, 59-64.
3. Oueis, E.; Klefisch, T.; Zaburannyi, N.; Garcia, R.; Plaza, A.; Muller, R. Two biosynthetic pathways in *Jahnella thaxteri* for thaxteramides, distinct types of lipopeptides. *Org. Lett.* **2019,** 21, 5407-5412.
4. Al-Mestarihi, A. H.; Villamizar, G.; Fernandez, J.; Zolova, O. E.; Lombó, F.; Garneau-Tsodikova, S. Adenylation and *S*-methylation of cysteine by the bifunctional enzyme TioN in thiocoraline biosynthesis. *J. Am. Chem. Soc.* **2014,** 136, 17350-17354.
5. Ross, A. C.; Xu, Y.; Lu, L.; Kersten, R. D.; Shao, Z.; Al-Suwailem, A. M.; Dorrestein, P. C.; Qian, P. Y.; Moore, B. S. Biosynthetic multitasking facilitates thalassospiramide structural diversity in marine bacteria. *J. Am. Chem. Soc.* **2013,** 135, 1155-1162.
6. Zhang, J. J.; Tang, X.; Huan, T.; Ross, A. C.; Moore, B. S. Pass-back chain extension expands multimodular assembly line biosynthesis. *Nat. Chem. Biol.* **2020,** 16, 42-49.

7. Lundy, T. A.; Mori, S.; Thamban Chandrika, N.; Garneau-Tsodikova, S. Characterization of a unique interrupted adenylation domain that can catalyze three reactions. *ACS Chem. Biol.* **2020,** 15, 282-289.
8. Xu, Y.; Kersten, R. D.; Nam, S. J.; Lu, L.; Al-Suwailem, A. M.; Zheng, H.; Fenical, W.; Dorrestein, P. C.; Moore, B. S.; Qian, P. Y. Bacterial biosynthesis and maturation of the didemnin anti-cancer agents. *J. Am. Chem. Soc.* **2012,** 134, 8625-8632.
9. Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004,** 32, 1792-1797.
10. Edgar, R. C. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **2004,** 5, 113.
11. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987,** 4, 406-425.
12. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019,** 47, W256-W259.
13. Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **1988,** 16, 10881-10890.
14. Shrestha, S. K.; Garneau-Tsodikova, S. Expanding substrate promiscuity by engineering a novel adenylating-methylating NRPS bifunctional enzyme. *ChemBioChem* **2016,** 17, 1328-1332.
15. Stachelhaus, T.; Marahiel, M. A. Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. *J. Biol. Chem.* **1995,** 270, 6163-6169.
16. Watanabe, K.; Hotta, K.; Praseuth, A. P.; Koketsu, K.; Migita, A.; Boddy, C. N.; Wang, C. C.; Oguri, H.; Oikawa, H. Total biosynthesis of antitumor nonribosomal peptides in *Escherichia coli*. *Nat. Chem. Biol.* **2006,** 2, 423-428.
17. Lombó, F.; Velasco, A.; Castro, A.; de la Calle, F.; Brana, A. F.; Sanchez-Puelles, J. M.; Mendez, C.; Salas, J. A. Deciphering the biosynthesis pathway of the antitumor thiocoraline from a marine actinomycete and its expression in two *Streptomyces* species. *ChemBioChem* **2006,** 7, 366-376.
18. Rouhiainen, L.; Paulin, L.; Suomalainen, S.; Hyytiainen, H.; Buikema, W.; Haselkorn, R.; Sivonen, K. Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in *Anabaena* strain 90. *Mol. Microbiol.* **2000,** 37, 156-167.

19. Ansari, M. Z.; Sharma, J.; Gokhale, R. S.; Mohanty, D. *In silico* analysis of methyltransferase domains involved in biosynthesis of secondary metabolites. *BMC Bioinformatics* **2008,** 9, 454.