

Supplementary Information

Computationally-guided exchange of substrate selectivity motifs in a modular polyketide synthase acyltransferase

Edward Kalkreuter^{1,2,3}, Kyle S Bingham^{1,4}, Aaron M Keeler¹, Andrew N Lowell^{5,7}, David H Sherman^{5,6}, and Gavin J Williams^{1,2*}

¹ Department of Chemistry, NC State University, Raleigh, North Carolina 27695, United States

² Comparative Medicine Institute, NC State University, Raleigh, North Carolina 27695, United States

³ Present address: Department of Chemistry, The Scripps Research Institute, Jupiter, Florida 33458, United States

⁴ Present address: UNC Chapel Hill School of Medicine, Chapel Hill, North Carolina 27516, United States

⁵ Life Sciences Institute, Department of Medicinal Chemistry, University of Michigan, Ann Arbor, Michigan 48109, United States

⁶ Department of Chemistry, Department of Microbiology & Immunology, University of Michigan, Ann Arbor, Michigan 48109, United States

⁷ Present address: Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, United States

* Corresponding author, gjwillia@ncsu.edu

Table of contents

Supplementary Tables

Supplementary Table 1. Wild-type EryAT6 amino acid sequences for MD simulations.

Supplementary Table 2. High-Resolution LC-MS retention times, calculated masses, and observed masses for DEBS PKS-catalyzed reaction products.

Supplementary Table 3. High-resolution LC-MS peak areas for PKS-catalyzed reaction products with five extenders.

Supplementary Table 4. Mutations in Ery6TE that did not result in a shift from the wild-type product distribution.

Supplementary Table 5. High-resolution LC-MS peak areas for PKS-catalyzed reaction products with two extenders.

Supplementary Table 6. Statistics for high-resolution LC-MS peak areas for PKS-catalyzed reaction products with two extenders.

Supplementary Table 7. DEBS PKS DNA FASTA sequences.

Supplementary Table 8. Primers for DEBS PKS construction and mutagenesis.

Supplementary Table 9. High-resolution LC-MS parameters and gradient.

Supplementary Figures

Supplementary Figure 1. Conserved motifs of PKS acyltransferases.

Supplementary Figure 2. Evaluation of EryAT6 models.

Supplementary Figure 3. Per-residue protein secondary structure of EryAT6 models.

Supplementary Figure 4. RMSD of EryAT6 MD simulations.

Supplementary Figure 5. Time-averaged structures of EryAT6 variants.

Supplementary Figure 6. RMSF plots for EryAT6 simulations.

Supplementary Figure 7. Distance plots for EryAT6 simulations.

Supplementary Figure 8. Representative HPLC traces for MatB-synthesized malonyl-CoAs.

Supplementary Figure 9. Motif sequences from natural acyltransferases.

Supplementary Figure 10. Representative mass spectra of 10-deoxymethynolide B compounds from lysate module reactions.

Supplementary Figure 11. Representative mass spectra of keto-10-deoxymethynolide B compounds from lysate module reactions.

Supplementary Methods

Site-directed mutagenesis of Ery6TE

Construction of Ery6TE motif chimeras

Expression and purification of wild-type and mutant MatB

Synthesis of acyl-CoAs by MatB

Supplementary References

Supplementary Tables

Supplemental Table 1. Wild-type EryAT6 amino acid sequences for MD simulations. Linkers are in bold.

EryAT6 Construct	Amino Acid FASTA Sequences for MD Simulations
Wild-Type with Both Linkers	AEPPEPEPLPEPGPVGV LAAANSVPVLLSARTETALAAQARLLES AVDDSVPLTA LASALATGRAHLPRRAALLAGDHEQLRGQLRAVAEGVAAPGATTGTASAGGVV FVFPQGGAQWEGMARGLLSVPVFAESIAECDAVLSEVAGFSASEVLEQRPDAPSL ERVDVVQPVLF SVMVSLARLWGACGVSPSAVIGHSQGEIAAAVVAGVLSLEDGVR VVALRAKALRALAGKGGMVSLAAPGERARALIAPWEDRISVAAVNSPSSVVVSGD PEALAE LVARCEDEGVRAKTL PVDYASHSRHVEEIRETILADLDGISARRAAIPLYST LHGERRDGADMGP RYWDNLRSQVRFDEAVSAAVADGHATFVEMSPHPVLTA A VQEIAADAVAIGSLHRDTAEEHLIAELARAHVHGVAVDWRNVFPAA
Wild-Type with N-Terminal Linker	AEPPEPEPLPEPGPVGV LAAANSVPVLLSARTETALAAQARLLES AVDDSVPLTA LASALATGRAHLPRRAALLAGDHEQLRGQLRAVAEGVAAPGATTGTASAGGVV FVFPQGGAQWEGMARGLLSVPVFAESIAECDAVLSEVAGFSASEVLEQRPDAPSL ERVDVVQPVLF SVMVSLARLWGACGVSPSAVIGHSQGEIAAAVVAGVLSLEDGVR VVALRAKALRALAGKGGMVSLAAPGERARALIAPWEDRISVAAVNSPSSVVVSGD PEALAE LVARCEDEGVRAKTL PVDYASHSRHVEEIRETILADLDGISARRAAIPLYST LHGERRDGADMGP RYWDNLRSQVRFDEAVSAAVADGHATFVEMSPHPVLTA A VQE
Wild-Type with No Linkers	VFPQGGAQWEGMARGLLSVPVFAESIAECDAVLSEVAGFSASEVLEQRPDAPSLE RVDVVQPVLF SVMVSLARLWGACGVSPSAVIGHSQGEIAAAVVAGVLSLEDGVRV VALRAKALRALAGKGGMVSLAAPGERARALIAPWEDRISVAAVNSPSSVVVSGDP EALAE LVARCEDEGVRAKTL PVDYASHSRHVEEIRETILADLDGISARRAAIPLYSTL HGERRDGADMGP RYWDNLRSQVRFDEAVSAAVADGHATFVEMSPHPVLTA AV QE

Supplemental Table 2. High-Resolution LC-MS retention times, calculated masses, and observed masses for DEBS PKS-catalyzed reaction products. N.D. = Not Detected.

Compound & Retention Time (min)		Calculated Mass ([M-H ₂ O+H] ⁺)	Observed Mass ([M-H ₂ O+H] ⁺)	Calculated Mass ([M+H] ⁺)	Observed Mass ([M+H] ⁺)	Calculated Mass ([M+Na] ⁺)	Observed Mass ([M+Na] ⁺)
4	8.03	279.1955	279.1962	297.2060	297.2069	319.1880	319.1904
5a	9.68	293.2111	293.2119	311.2217	311.2227	333.2036	333.2044
5b	8.96	303.1955	303.1963	321.2060	321.2068	343.1880	343.1890
5c	10.63	307.2268	307.2274	325.2373	325.2378	347.2193	347.2198
5d	N.D.	321.2424	N.D.	339.2530	N.D.	361.2349	N.D.
5e	10.12	305.2111	305.2105	323.2217	323.2201	345.2036	345.2035
5f	N.D.	335.2581	N.D.	353.2686	N.D.	375.2506	N.D.
6a	12.39	--	--	309.2060	309.2067	331.1880	331.1886
6b	11.80	--	--	319.1904	319.1911	341.1723	341.1732
6c	11.79	--	--	323.2217	323.2224	345.2036	345.2044
6d	14.56	--	--	337.2373	337.2380	359.2193	359.2200
6e	12.67	--	--	321.2060	321.2055	343.1880	343.1876
6f	15.07	--	--	351.2530	351.2525	373.2349	373.2345

Supplemental Table 3. High-resolution LC-MS peak areas for PKS-catalyzed reaction products with five extenders. Extracted ion count peak areas from one replicate are shown for each reaction condition, prior to correction with enzyme concentration. N.D. = Not Detected.

Ery6TE Variant	Representative EIC Peak Areas				
	4	5a	6a	5b	6b
WT	211,060,257	8,350,707	3,201,023	60,428,975	33,336,885
V742A	1,074,530,536	74,181,724	8,289,132	621,019,227	62,604,070
Y744R	201,560,426	18,273,398	10,450,848	425,370,308	221,736,187
L673G	412,033,607	11,485,854	3,946,441	96,211,067	47,618,973
A672S	918,117,379	17,567,468	16,886,399	174,246,461	190,537,319
L676V	913,134,214	20,437,965	21,610,476	202,357,895	254,127,057
G681A	258,612,993	7,211,911	4,379,076	42,211,949	38,450,516
D613E	1,112,681,260	25,901,052	23,610,928	246,571,818	263,851,451
R674A	1,137,937,000	22,945,009	25,007,975	178,486,582	202,512,483
H643S	244,645,247	12,743,255	8,023,759	70,681,023	47,042,985
H643F	215,279,432	1,629,754	960,271	2,131,336	2,185,209
I648G	365,230,755	869,632	475,403	1,770,386	1,262,385
L797E	565,835,459	782,613	2,352,056	829,163	1,769,043
Cin1	1,122,551	14,392	29,947	1,093,275	2,750,716
Tha13	29,754,912	1,565,605	928,229	31,038,395	17,687,295
Ery6TE Variant	Representative EIC Peak Areas				
	5c	6c	5d	6d	
WT	11,975,213	22,923,302	N.D.	3,401,711	
V742A	41,645,025	4,688,419	N.D.	32,702,810	
Y744R	87,834,327	60,330,553	N.D.	110,340,604	
L673G	12,692,412	27,903,903	N.D.	15,313,652	
A672S	7,799,155	195,172,467	N.D.	18,325,167	
L676V	2,065,221	161,120,479	N.D.	28,957,619	
G681A	13,320,019	63,892,775	N.D.	6,301,166	
D613E	20,124,173	244,569,998	N.D.	37,219,146	
R674A	16,724,087	256,290,037	N.D.	19,347,354	
H643S	25,063,933	35,489,999	N.D.	18,784,167	
H643F	74,197	3,727,611	N.D.	589,128	
I648G	595,456	6,006,129	N.D.	1,376,049	
L797E	128,196	1,623,863	N.D.	1,573,907	
Cin1	39,917	439,077	N.D.	1,402,533	
Tha13	22,492,383	39,231,371	N.D.	123,779,145	

Supplemental Table 4. Mutations in Ery6TE that did not result in a shift from the wild-type product distribution. All mutants that were constructed and tested in vitro with **1** and non-native extenders but showed no difference in product distribution (other than total activity) from wild-type Ery6TE were excluded from the main text but are included in the following table.

Ery6TE Mutants Indistinguishable from Wild-Type		
A596L	A672S	L740V
E598D	R674A	P741Δ
V612D	A675L	P741G
E647D	G689A	P790A
G662A	D701G	L797S
A670T	A707G	
A670V	S717A	

Supplemental Table 5. High-resolution LC-MS peak areas for PKS-catalyzed reaction products with two extenders. Extracted ion count peak areas from the three replicates are shown for each reaction condition, prior to correction with enzyme concentration. N.D. = Not Detected.

Substrates	Product	EIC Peak Areas				
		WT	I648G	Y744R	Cin1	Tha13
1+2b	4	532,578,041	799,908,718	78,684,251	29,848,810	26,068,751
	5b	31,090,929	2,583,102	320,918,411	20,904,574	45,737,083
	6b	75,186,842	N.D.	109,088,198	44,368,738	37,497,560
	4	716,257,624	1,281,730,797	288,363,018	27,967,173	110,869,247
	5b	10,491,816	3,456,956	93,514,897	20,078,581	129,467,348
	6b	3363144	N.D.	367,954,412	45,478,751	66,654,955
	4	663,536,505	851,319,123	454,754,654	25,447,589	27,668,082
	5b	90,668,644	3,371,564	1,416,351,093	14,412,667	18,014,853
	6b	37,810,814	N.D.	563,471,339	44,512,521	49,507,603
1+2d	4	1,241,364,749	650,111,454	230,399,192	478,245	29,616,734
	5d	N.D.	N.D.	N.D.	N.D.	N.D.
	6d	6,953,819	N.D.	88,178,801	4,152,337	193,602,896
	4	442,846,997	701,947,321	40,688,455	18,686,181	13,798,643
	5d	N.D.	N.D.	N.D.	N.D.	N.D.
	6d	3,252,033	N.D.	13,599,517	64,279,072	172,045,949
	4	1,109,471,085	1,318,378,357	353,628,391	17,770,595	56,310,379
	5d	N.D.	N.D.	N.D.	N.D.	N.D.
	6d	6,247,017	N.D.	35,401,742	134,608,436	499,303,732
1+2e	4	1,263,268,254	1,467,946,423	122,464,111	45,392,299	43,610,053
	5e	7,647,855	N.D.	28,872,748	964,634	10,197,042
	6e	21,784,302	N.D.	53,689,472	17,802,789	33,024,254
	4	1,347,621,460	1,166,990,061	231,683,367	41,470,130	54,256,708
	5e	8,642,746	N.D.	31,211,506	5,053,195	12,186,663
	6e	27,600,123	N.D.	61,477,390	13,093,086	22,590,082
	4	749,218,220	787,105,102	395,020,747	40,718,563	60,618,920
	5e	3,627,008	N.D.	55,760,678	6,361,402	18,086,700
	6e	14,433,977	N.D.	157,017,607	10,680,958	30,154,791
1+2f	4	1,144,743,869	1,115,568,669	622,633,725	29,298,295	30,377,170
	5f	N.D.	N.D.	N.D.	N.D.	N.D.
	6f	1,583,892	N.D.	217,518,820	26,459,590	473,073,518
	4	905,943,636	1,217,656,413	98,396,869	33,943,300	20,218,413
	5f	N.D.	N.D.	N.D.	N.D.	N.D.
	6f	3,663,759	N.D.	27,012,724	40,092,569	271,527,843
	4	1,153,814,036	813,348,137	880,467,719	34,365,774	39,255,832
	5f	N.D.	N.D.	N.D.	N.D.	N.D.
	6f	4,307,982	N.D.	187,905,304	47,369,923	602,913,839

Supplemental Table 6. Statistics for high-resolution LC-MS peak areas for PKS-catalyzed reaction products with two extenders. Statistics of the extracted ion count peak areas from **Table S5** are shown for each reaction condition. Results also shown in **Figure 7**. N.D. = Not Detected.

Substrates	Product	Average Peak Area Distribution					Standard Deviation of Peak Area Distribution				
		WT	I648G	Y744R	Cin1	Tha13	WT	I648G	Y744R	Cin1	Tha13
1 + 2b	4	88.4%	99.7%	24.2%	30.5%	29.7%	8.4%	0.1%	12.4%	0.8%	6.2%
	5b + 6b	11.6%	0.3%	75.8%	69.5%	70.3%	8.4%	0.1%	12.4%	0.8%	6.2%
1 + 2d	4	99.4%	100.0%	79.4%	14.8%	10.3%	0.1%	0.0%	10.1%	6.7%	2.9%
	5d + 6d	0.6%	0.0%	20.6%	85.2%	89.7%	0.1%	0.0%	10.1%	6.7%	2.9%
1 + 2e	4	97.6%	100.0%	65.4%	70.3%	55.6%	0.2%	0.0%	5.9%	0.6%	5.4%
	5e + 6e	2.4%	0.0%	34.6%	29.7%	44.4%	0.2%	0.0%	5.9%	0.6%	5.4%
1 + 2f	4	99.7%	100.0%	78.3%	46.8%	6.4%	0.1%	0.0%	4.2%	5.3%	0.5%
	5f + 6f	0.3%	0.0%	21.7%	53.2%	93.6%	0.1%	0.0%	4.2%	5.3%	0.5%

Supplemental Table 7. DEBS PKS DNA FASTA sequences. AT domains in modules are in bold. Pik docking domain is italicized.

Construct	Nucleotide Sequences
ThaAT13 Motifs	GACCCAATACGCGCCGTCGCTGGAGCGTCTGGATGTGAATCAGCCGGTTCTGTTTAGT GTTATGGTGAGTCTGGCCCGTCTGTGGGGTGCCGTGTGGTGTGAGTCCGAGTGCCGTTA TTGGCCATAGCCAGGGTGAAATTGCAGCAGCAGTTGTTGCAGGCGTGCTGAGTCTGGA AGATGGTGTGCGCGTGGTTGCACTGCGTGCCAAAGCACTGCGTGCGCTGGCAGGCAAA GGTGGCATGGTTAGCCTGGCCGCACCGGGCGAACGCGCTAGAGCTCTGATTGCCCGT GGGAAGATCGTATTAGCGTTGCCGCAGTTAATAGTCCGAGTAGCGTGGTGGTGAGTGG TGACCCGGAAGCCCTGGCCGAAGTGGTGGCACGTTGTGAAGATGAAGGTGTTTCGCGCA AAAGTTCTGCCGGGTGCCGATGCCGCAGGCCACTCCCGCCACGTCGAGACCCAAT
CinAT1 Motifs	ACGCGCCGTCGCTGGAGGAAGGCGATATTCAGCAGCCGGTTCTGTTTAGTGTGATGGT TAGTCTGGCACGCCTGTGGGGCGCCTGCGGTGTTAGTCCGAGTGCAAGTTATTGGCCAT AGCCAGGGCGAAATTGCAGCCGCCGTTGTTGCAGGCGTCTGAGCCTGGAAGATGGCG TTCGTGTTGTGGCCCTGCGCGCCAAAGCACTGCGCGCACTGGCAGGCAAAGGTGGTAT GGTGAGTCTGGCAGCACCGGGTGAACGCGCCCGTGCTCTGATTGCACCGTGGGAAGAT CGCATTAGCGTGGCCGCAGTGAATAGCCCCGAGCAGCGTTGTTGTTAGTGGCGATCCGG AAGCCCTGGCAGAACTGGTGGCCCGTTGCGAAGATGAAGGCGTTCGTGCCAAAGCCCT GCGCGTGGAACGCGCAGGCCACTCCCGCCACGTCGA
Ery6 WT	ATGGTTCGGCGCAGCAGAGGCGGAGCAAGCCCCGGCGCTCGTGCGCGAGGTGCCGAAGG ATGCCGACGACCCGATCGCGATCGTCGGCATGGCCTGCCGCTTCCCCGGCGGCGTGCA CAACCCCGGTGAGCTGTGGGAGTTCATCGTCGGCGGCGGAGACGCCGTGACGGAGATG CCCACCGACCGCGGCTGGGACCTCGACGCGCTGTTTCGACCCCGACCCGCAGCGCCACG GAACCAGCTACTCGCGACACGGCGCGTTCCTCGACGGGGCCCGCCGACTTCGACGCGGC GTTCTTCGGGATCTCGCCGCGCGAGGCGCTGGCGATGGACCCGCAGCAGCGCCAGGTC CTGGAACGACGTGGGAGCTGTTTCGAGAACGCCGGCATCGACCCGCACTCGCTGCGGG GCAGCGACACCGGCGTCTTCCTCGGCGCCGCGTACCAGGGCTACGGCCAGGACGCGGT GGTGGCCGAGGACAGCGAGGGCTACCTGCTCACCGGCAACTCCTCCGCCGTGGTGTCC GGCCGGGTGCGCTACGTGCTGGGGCTGGAAGGCCCCGCGGTACGGTGGACACGGCGT GTTTCGTGCTCGTTGGTGGCCTTGCAATTCGGCGTGTGGGTGCTTGCCTGACGGTGA CGGTCTTGCGGTGGCCGGTGGTGTGTGCGGTGATGGCGGGCCCGGAGGTGTTACCGAG TTCTCCCGCCAGGGCGGCTTGCCCGTGGACGGGCGCTGCAAGGCGTTCTCCGCGGAGG CCGACGGCTTCGGTTTCGCCGAGGGCGTTCGCGGTGGTTCCTGCTCCAGCGGTTGTCCGA CGCCCGCAGGGCGGGTCGCCAGGTGCTCGGCGTGGTTCGCGGGCTCGGCGATCAACCAG GACGGCGCGAGCAACGGTCTCGCGGCGCCGAGCGGCGTTCGCCAGCAGCGCGTGATCC GCAAGGCGTGGGCGCGTGCAGGGGATCACGGGCGCGGATGTGGCCGTGGTGGAGGCGCA TGGGACCGGTACGCGGCTGGGCGATCCGGTGGAGGCGTTCGGCGTTGCTGGCTACTTAC GGCAAGTCGCGCGGGTCTGCGGGCCCGGTGCTGCTGGGTTCGGTGAAGTCGAACATCG GTCACGCGCAGGCGGCCGCGGGTGTGCGGGCGTGATCAAGGTGGTTCCTGGGGTTGAA CCGCGGCCCTGGTGCCGCCGATGCTCTGCCGCGGCGAGCGGTTCGCCGCTGATCGAATGG TCCTCGGGTGGTGTGGAACCTGCCGAGGCGGTGAGCCCGTGGCCTCCGGCCGCGGACG GGGTGCGCCGGGCGGTGTGTCGGCGTTCGGGGTGAGCGGGACGAACGCGCACGTGAT CATCGCCGAGCCCCCGGAGCCCGAGCCGCTGCCGGAACCCGGACCGGTGGGCGTGCTG GCCGCTGCGAACTCGGTGCCCGTACTGCTGTGCGCCAGGACCGAGACCGCGTTGGCAG CGCAGGCGCGGCTCCTGGAGTCCGCAGTGGACGACTCGGTTCCGTTGACGGCATTGGC TTCCGCGCTGGCCACCGGACGCGCCACCTGCCGCGTCTGCGGCGTTGCTGGCAGGC GACCACGAACAGCTCCGCGGGCAGTTGCGAGCGGTTCGCCAGGGCGTTGCGGCTCCCG GTGCCACCACCGGAACCGCCTCCGCCGGCGGCGTGGTTTTCTGCTTTCCAGGTCAGGG TGCTCAGTGGGAGGGCATGGCCCCGGGGCTTGCTCTCGGTCCCCGTCTTCGCCGAGTCG

	ATCGCCGAGTGCATGCGGTGTTGTTCGGAGGTGGCCGGGTCTCTCGGCCTCCGAAGTGC TGGAGCAGCGTCCGGACGCGCCGTCTGCTGGAGCGGGTCGACGTCGTACAGCCGGTGT GTTCTCCGTGATGGTGTCTGCTGGCGCGGCTGTGGGGCGCTTTCGGAGTCAGCCCCTCG GCCGTATCGGCCATTTCGAGGGCGAGATCGCCGCCGCGGTGGTGGCCGGGGTGTGT CGCTGGAGGACGGCGTGCCTGCTGGCCCTGCGCGCGAAGGCGTTGCGTGCCTGGC GGGCAAGGGCGGCATGGTCTCGTTGGCGGCTCCCGGTGAACGCGCCCGCGCGCTGATC GCACCGTGGGAGGACCGGATCTCCGTGCGGGCGGTCAACTCCCCGTCTCGGTCTGG TCTCCGGCGATCCGGAGGCGCTGGCCGAACCTCGTCGCACGTTGCGAGGACGAGGGCGT GCGCGCCAAGACGCTCCCGGTGGACTACGCTCGCACTCCCGCCACGTCGAGGAGATC CGCGAGACGATCTCGCCGACCTCGACGGCATCTCCGCGCGGCGTGCCGCCATCCCGC TCTACTCCACGCTGCACGGCGAACGGCGGACGGCGCCGACATGGGTCCGCGGTACTG GTACGACAACCTGCGCTCCCAGGTGCGCTTCGACGAGGCGGTCTCGGCCGCCGTCTGCC GACGGTCACGCCACCTTCGTGAGATGAGCCCGCACCCGGTGCTCACCGCGGCGGTGC AGGAGATCGCCGCGGACGCGGTGGCCATCGGGTCTGCTGCACCGCGACACCGCGGAGGA GCACCTGATCGCCGAGCTCGCCCGGGCGCACGTGCACGGCGTGCCCGTGGACTGGCGG AACGTCTTCCCGGCGGCACCTCCGGTGGCGCTGCCCAACTACCCGTTTCGAGCCCCAGC GGTACTGGCTCGCGCCCGGAGGTGTCCGACCAGCTCGCCGACAGCCGCTACCGCGTCTGA CTGGCGACCGCTGGCCACCACGCCGTGGACCTGGAAGGCGGCTTCCTGGTCCACGGG TCCGCACCGGAGTCGCTGACCAGCGCAGTCGAGAAGGCCGGAGGCCGCGTCTGCGCGG TCGCCTCGGCCGACCGCGAAGCGCTCGCGGCGGGCCCTGCGGGAGGTGCCGGGCGAGGT CGCCGGCGTGCTCTCGGTCCACACCGGCGCCGCAACGCACCTCGCCCTGCACCAGTCG CTGGGTGAGGCCGGCGTGCGGGCCCCGCTCTGGCTGGTACACAGCCGAGCGGTCTGCGC TCGGGGAGTCCGAGCCGTCGATCCCGAGCAGGCGATGGTGTGGGGTCTCGGGCGCGT CATGGGCCCTGGAGACCCCGGAACGGTGGGGCGGTCTGGTGGACCTGCCCGCCGAACCC GCGCCGGGGGACGGCGAGGCGTTCTGCTGCCTGCCTCGGCGCGGACGGCCACGAGGACC AGGTCGCGATCCGTGACCACGCCCGCTACGGCCGCCGCTCGTCCGCGCCCCGCTGGG CACCCGCGAGTCGAGCTGGGAGCCGGCGGGCACGGCGCTGGTACCGGCGGCACCGGT GCGCTCGGCGGCCACGTGCCCCGCCACCTCGCCAGGTGCGGGGTGGAGGACCTGGTGC TGGTCAGCAGGCGCGGCGTCTGACGCTCCCGGCGCGGCCGAGCTGGAAGCCGAACCTGGT CGCCCTCGGCGCGAAGACGACCATCACCGCTGCGACGTGGCCGACCGCGAGCAGCTC TCCAAGCTGCTGGAAGAACTGCGCGGGCAGGGACGTCCGGTTCGGACCGTCTGTCACA CCGCCGGGGTGCCCGAATCGAGGCGGCTGCACGAGATCGGCGAGCTGGAGTCGGTCTG CGCGGCGAAGGTGACCGGGGCCCGGTGCTCGACGAGCTGTGCCCGGACGCCGAGACC TTCGTCTGTTCTCGTCCGGAGCGGGGTGTGGGGCAGTGCGAACCTCGGCGCCTACT CCGCGGCCAACGCCTACCTCGACGCGCTGGCCCACCGCCGCCGTGCGGAAGGCCGTGC GGCGACGTCCGTGCGTGGGGCGCTGGGCGGGCGAGGGCATGGCCACCGGCGACCTC GAGGGGCTACCCGGCGCGGCTGCGCCCGATGGCGCCCGAGCGCGCGATCCGCGCGC TGCACCAGGCGCTGGACAACGGCGACACGTGCGTTTCGATCGCCGACGTCGACTGGGA GCGCTTCGCGGTGCGCTTCACCGCCGCCCGGCGCGTCCGCTGCTGGACGAGCTCGTC ACGCCGGCGGTGGGGGCCGTCCCCGCGGTGCAGGCGGCCCCGCGCGGGAGATGACGT CGCAGGAGTTGCTGGAGTTCACGCACTCGCACGTGCGGCGATCCTCGGGCATTCAG CCCGGACGCGGTTCGGGACGACAGCCGTTACCGAGCTCGGCTTCGACTCGCTGACC GCGGTTCGGGCTGCGCAACCAGCTCCAGCAGGCCACCGGGCTCGCGCTGCCCGCGACCC TGGTGTTCGAGCACCCACGGTCCGCAGGTTGGCCGACCACATAGGACAGCAGCTCTG A
PDDEry6TE WT	ATGACGAGTTCCAACGAACAGTTGGTGGACGCTCTGCGCGCCTCTCTCAAGGAGAACG AAGAACTCCGGAAAGAGAGCCGTGCGCGGGCCGACCGTTCGGCAGGAGGAGATCGCGAT CGTCGGCATGGCCTGCCGTTCCCCGGCGGCGTGACAAACCCGGTGAGCTGTGGGAG TTCATCGTCGGCGGCGGAGACCGGTGACGGAGATGCCACCGACCGCGGCTGGGACC TCGACGCGCTGTTGACCCCCGACCCGACGCGCCACGGAACAGCTACTCGCGACACGG

CGCGTTCTCGACGGGGCCGCCGACTTCGACGCGGCGTTCTTCGGGATCTCGCCGCGC
 GAGGCGCTGGCGATGGACCCGACGAGCGCCAGGTCTTGAAACGACGTGGGAGCTGT
 TCGAGAACGCCGCGCATCGACCCGCACTCGCTGCGGGGACGACACCGGCGTCTTCCT
 CGGCGCCGCGTACCAGGGCTACGGCCAGGACGCGGTGGTGCCCGAGGACAGCGAGGGC
 TACCTGCTCACC GGCAACTCCTCCGCCGTGGTGTCGGGCCGGGTTCGCTACGTGCTGG
 GGCTGGAAGGCCCGCGGTACAGGTGGACACGGCGTGTTCGTCGTCGTTGGTGCCCTT
 GCATTTCGGCGTGTGGGTTCGTTGCGTGACGGTGACTGCGGTCTTGCGGTGGCCGGTGGT
 GTGTCGGTGATGGCGGGCCCGGAGGTGTTACCGAGTTCTCCCGCCAGGGCGGCTTGG
 CCGTGGAACGGGCGCTGCAAGGCGTTCTCCGCGGAGGCCGACGGCTTCGGTTTCGCCGA
 GGGCGTCGCGGTGGTTCCTGCTCCAGCGTTGTCCGACGCCCGCAGGGCGGGTCGCCAG
 GTGCTCGGCGTGGTCGCGGGCTCGGCGATCAACCAGGACGGCGCGAGCAACGGTCTCG
 CGGCGCCGAGCGGCGTCCGCCAGCAGCGCGTGATCCGCAAGGCGTGGGCGCGTGCGGG
 GATCACGGGCGCGGATGTGGCCGTGGTGGAGGCGCATGGGACCGGTACGCGGCTGGGC
 GATCCGGTGGAGGCGTCGGCGTTGCTGGCTACTTACGGCAAGTCGCGCGGGTTCGTTCG
 GCCCGGTGCTGCTGGGTTCGGTGAAGTCGAACATCGGTACGCGCAGGCGGCCGCGGG
 TGTCGCGGGCGTGATCAAGGTGGTCTGGGGTTGAACGCGGCCTGGTGCCGCCGATG
 CTCTGCCGCGGCGAGCGGTCCGCCGTGATCGAATGGTCCTCGGGTGGTGTGGAACCTG
 CCGAGGCCGTGAGCCCGTGGCCTCCGGCCGCGGACGGGGTGCGCCGGGCCGGTGTGTC
 GCGTTTCGGGGTGAGC**GGGACGAACGCGCACGTGATCATCGCCGAGCCCCGGAGCCC**
GAGCCGCTGCCGGAACCCGGACCGGTGGGCGTGCTGGCCGCTGCGAACTCGGTGCCCG
TACTGCTGTCGGCCAGGACCGAGACCGCGTTGGCAGCGCAGGCGCGGCTCCTGGAGTC
CGCAGTGGACGACTCGGTTCGGTTGACGGCATTGGCTTCCGCGCTGGCCACCGGACGC
GCCACCTGCCGCGTCGTGCGGCGTTGCTGGCAGGCGACCACGAACAGCTCCGCGGGC
AGTTGCGAGCGGTCCCGAGGGCGTTGCGGCTCCCGGTGCCACCACCGGAACCGCCTC
CGCCGGCGGCGTGTTTTCTGCTCTTCCAGGTACGGGTGCTCAGTGGGAGGGCATGGCC
CGGGGCTTGCTCTCGGTCCCCGTCTTCGCCGAGTCGATCGCCGAGTGCGATGCGGTGT
TGTCGGAGGTGGCCGGGTCTCGGCCTCCGAAGTGCTGGAGCAGCGTCCGGACGCGCC
GTCGCTGGAGCGGGTCGACGTCGTACAGCCGGTGTTGTTCTCCGTGATGGTGTGCTG
GCGCGGCTGTGGGGCGCTTGCGGAGTCAGCCCCTCGGCCGTCATCGGCCATTGCGAGG
GCGAGATCGCCGCCGCGGTGGTGGCCGGGGTGTTGTCGCTGGAGGACGGCGTGCGCGT
CGTGCCCTGCGCGCGAAGGCGTTGCGTGCGCTGGCGGGCAAGGGCGGCATGGTCTCG
TTGGCGGCTCCCGGTGAACGCGCCCCGCGCGCTGATCGCACCGTGAGGACCGGATCT
CCGTCGCGGCGGTCAACTCCCCGTCTCGGTGCTGGTCTCCGGCGATCCGGAGGCGCT
GGCCGAACTCGTGCGACGTTGCGAGGACGAGGGCGTGCGCGCCAAGACGCTCCCGGTG
GACTACGCCTCGCACTCCCGCCACGTCGAGGAGATCCGCGAGACGATCCTCGCCGACC
TCGACGGCATCTCCGCGCGGCGTGCCGCCATCCCGCTCTACTCCACGCTGCACGGCGA
ACGGCGCGACGGCGCCGACATGGGTCCGCGGTACTGGTACGACAACCTGCGCTCCCAG
GTGCGCTTCGACGAGGCGGTCTCGGCCGCCGTGCCGACGGTCACGCCACCTTCGTG
AGATGAGCCCGCACCCGGTGCTCACCGCGGCGGTGCAGGAGATCGCCGCGGACGCCGT
GGCCATCGGGTCGCTGCACCGCGACACCGCGGAGGACACCTGATCGCCGAGCTCGCC
CGGGCGCACGTGCACGGCGTGCCGTGGACTGGCGGAACGTCTTCCCGCGGCACCTC
CGGTGGCGCTGCCCAACTACCCGTTTCGAGCCCCAGCGGTACTGGCTCGCGCCGGAGGT
 GTCCGACCAGCTCGCCGACAGCCGCTACCGCGTCGACTGGCGACCGCTGGCCACCACG
 CCGGTGGACCTGGAAGGCGGCTTCCTGGTCCACGGGTCCGCACCGGAGTCGCTGACCA
 GCGCAGTCGAGAAGGCCGGAGGCCGCGTCGTGCCGGTCGCCTCGGCCGACCGCGAAGC
 GCTCGCGGGCGGCCCTGCGGGAGGTGCCGGGCGAGGTGCCGGCGTGCTCTCGGTCCAC
 ACCGGCGCCGCAACGCACCTCGCCCTGCACCAGTCGCTGGGTGAGGCCGGCGTGCGGG
 CCCCCTCTGGCTGGTACACAGCCGAGCGGTGCGCTCGGGGAGTCCGAGCCGGTTCGA
 TCCCGAGCAGGCGATGGTGTGGGGTCTCGGGCGCGTCATGGGCCTGGAGACCCGGAA
 CCGTGGGGCGGTCTGGTGGACCTGCCCGCCGAACCCGCGCCGGGGGACGGCGAGGCGT

TCGTGCGCCTGCCTCGGCGCGGACGGCCACGAGGACCAGGTGCGCGATCCGTGACCACGC
CCGCTACGGCCCGCCGCTCGTCCGCGCCCCGCTGGGCACCCGCGAGTCGAGCTGGGAG
CCGGCGGGCACGGCGCTGGTCACCGGCGGCACCGGTGCGCTCGGCGGCCACGTGCCCC
GCCACCTCGCCAGGTGCGGGGTGGAGGACCTGGTGCTGGTCAGCAGGCGCGGCGTGA
CGCTCCCGGCGCGGCCGAGCTGGAAGCCGAACCTGGTCGCCCTCGGCGCGAAGACGACC
ATCACCGCCTGCGACGTGGCCGACCGCGAGCAGCTCTCCAAGCTGCTGGAAGAACTGC
GCGGGCAGGGACGTCCGGTGCGGACCGTCTGTGCACACCGCCGGGTGCCCCGAATCGAG
GCCGCTGCACGAGATCGGCGAGCTGGAGTCGGTCTGCGCGGCGAAGGTGACCGGGGCC
CGGCTGCTCGACGAGCTGTGCCCGGACGCCGAGACCTTCGTCTGTCTCGTCCGGAG
CGGGGGTGTGGGGCAGTGCGAACCTCGGCGCCTACTCCGCGGCCAACGCCTACCTCGA
CGCGCTGGCCCACCGCCGCGGTGCGGAAGGCCGTGCGGCGACGTCCGTGCGGTGGGGC
GCCTGGGCGGGCGAGGGCATGGCCACCGGCGACCTCGAGGGGCTCACCGGCGCGGCC
TGCGCCCGATGGCGCCCGAGCGCGCGATCCGCGCGCTGCACCAGGCGCTGGACAACGG
CGACACGTGCGTTTCGATCGCCGACGTGCGACTGGGAGCGCTTCGCGGTTCGGCTTCACC
GCCGCCCGGCCGCGTCCGCTGCTGGACGAGCTCGTCACGCCGGCGGTGGGGGCCGTCC
CCGCGGTGCAGGCGGCCCCGGCGCGGGAGATGACGTGCGAGGAGTTGCTGGAGTTCAC
GCACTCGCACGTGCGGGCGATCCTCGGGCATTCCAGCCCGGACGCGGTTCGGGCAGGAC
CAGCCGTTACCGAGCTCGGCTTCGACTCGCTGACCGCGGTTCGGGTGCGCAACCAGC
TCCAGCAGGCCACCGGGCTCGCGCTGCCCGCGACCCTGGTGTTCGAGCACCCACGGT
CCGCAGGTGGCCGACCACATAGGACAGCAGCTCGACAGCGGGACTCCCGCCCCGGGAA
GCGAGCAGCGCTCTTCGCGACGGCTACCGGCAGGCGGGCGTGTGCGGCAGGGTCCGGT
CCTACCTCGACCTGCTGGCGGGGCTGTGCGGACTTCGCGGAGCACTTCGACGGCTCCGA
CGGGTTCTCCCTCGATCTCGTGGACATGGCCGACGGTCCCGGAGAGGTACGGTGATC
TGCTGCGCGGGAACGGCGGCGATCTCCGGTCCGCACGAGTTCACCCGGCTCGCCGGGG
CGCTGCGCGGAATCGCTCCGGTTCGGGCCGTGCCCCAGCCCGGCTACGAGGAGGGCGA
ACCTCTGCCGTCGTCGATGGCGGCGGTGGCGGCGGTGCAGGCCGATGCGGTTCATCAGG
ACACAGGGGGACAAGCCGTTTCGTGGTGGCCGGTCACTCCGCGGGGGCACTGATGGCCT
ACGCGCTGGCGACCGAACTGCTCGATCGCGGGCACCCGCCACGCGGTGTCGTCCTGAT
CGACGTCTACCCGCCCGGTACCAGGACGCGATGAACGCCTGGCTGGAGGAGCTGACC
GCCACGCTGTTTCGACCGCGAGACGGTTCGGATGGACGACACCAGGCTCACCGCCCTGG
GCGCCTACGACCGCCTCACCGGTCACTGGCGACCCCGGGAAACCGGGCTGCCGACGCT
GCTGGTTCAGCGCCGGCGAGCCGATGGGTCCGTGGCCCGACGACAGCTGGAAGCCGACG
TGCCCCCTTCGAGCACGACACCGTTCGCCGTCCCCGGCGACCACTTCACGATGGTGCAGG
AACACGCCGACGCGATCGCGCGGCACATCGACGCCTGGCTGGGCGGAGGGAATTCAGG
A

Supplemental Table 8. Primers for DEBS PKS construction and mutagenesis. GA = Gibson Assembly. RTH = 'Round-the-horn mutagenesis'.¹ RTH primers have 5' phosphates.

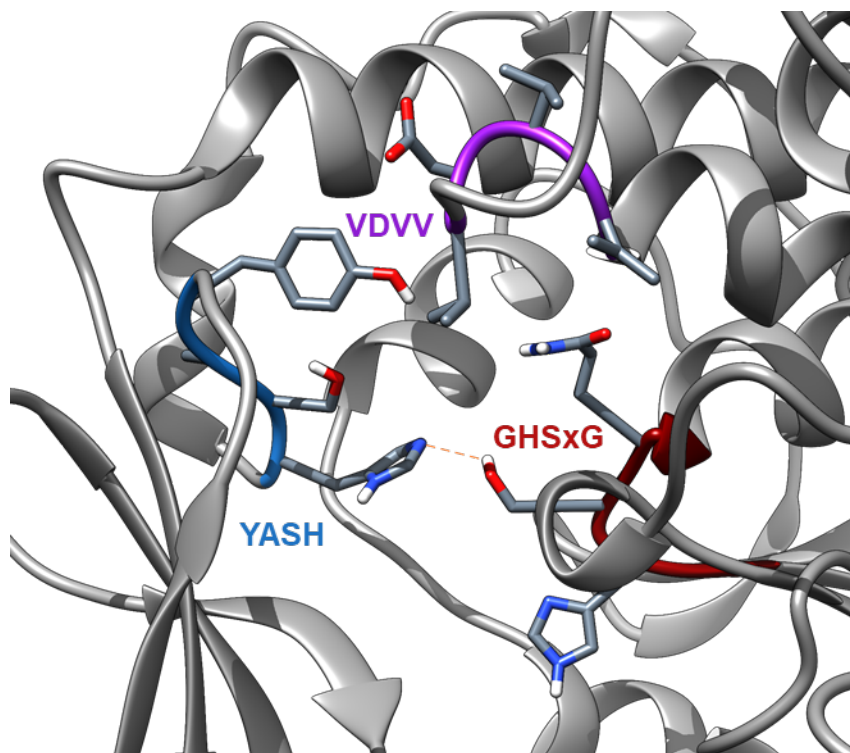
Primer Name	5'→3' Primer Sequence	Primer Function
AT_Motif.Gib1	CACTCCCGCCACGTCGAG	Amplifies Ery6TE without AT motifs for GA
AT_Motif.Gib2	CTCCAGCGACGGCGCG	
AT_Motif.FOR	ACGCGCCGTCGCTGG	Amplifies motif fragments for GA
AT_Motif.REV	TCGACGTGGCGGGAGTG	
Ery6V742A.FOR	GCGGACTACGCCTCGCACTCCC	Ery6 V742A RTH
Ery6V742A.REV	CGGGAGCGTCTTGCGCG	
Ery6Y744RGV.RTH1	SKGGCCTCGCACTCCCGC	Ery6TE Y744R/G RTH
Ery6Y744RGV.RTH2	GTCCACCGGGAGCGTCTTG	
Ery6L676V.FOR	GTGGCGGGCAAGGGCG	Ery6 L676V RTH
Ery6L676V.REV	CGCACGCAACGCCTTCG	
Ery6L673G.FOR	GGACGTGCGCTGGCGGG	Ery6 L673G RTH
Ery6L673G.REV	CGCCTTCGCGCGCAG	
Ery6E647D.FOR	GATATCGCCGCCGCGG	Ery6 E647D RTH
Ery6E647D.REV	GCCCTGCGAATGGCCGAT	
Ery6L797X.FOR	NNKCGCTCCCAGGTGCGC	Ery6 L797X RTH
Ery6L797X.REV	GTTGTCGTACCAGTACCGCGGA	
Ery6I648A.FOR	GCCGCCGCGGTGGTGG	Ery6 I648G/A/V RTH
Ery6I648X.REV	GNSCTCGCCCTGCGAATGGC	
Ery6H643F.FOR	TTTTCGCAGGGCGAGATCGC	Ery6 H643F RTH
Ery6H643F.REV	GCCGATGACGGCCGAGG	
Ery6H643S.FOR	TCTTCGCAGGGCGAGATCGC	Ery6 H643S RTH
Ery6H643F.REV	GCCGATGACGGCCGAGG	
Ery6L740V.FOR	GTCCCGGTGGACTACGCCTC	Ery6 L740V RTH
Ery6L740V.REV	CGTCTTGCGCGCACG	
Ery6A707G.FOR	GGGGTCAACTCCCCGTCCTCG	Ery6 A707G RTH
Ery6A707G.REV	CGCGACGGAGATCCGGTC	
Ery6G681A.FOR	GCAATGGTCTCGTTGGCGGC	Ery6 G681A RTH
Ery6G681A.REV	GCCCTTGCCCGCCAGC	
Ery6V751A.FOR	GCCGAGGAGATCCGCGAGAC	Ery6 V751A RTH
Ery6V751A.REV	GTGGCGGGAGTGCGAGG	
Ery6S644A.FOR	GTCAGCCGTTATCGGTCATGCT	Ery6 S644A RTH
Ery6S644A.REV	CCGCAATTTGCCCTG	
Ery6D613E.FOR	GAAGTCGTACAGCCGGTGTTG	Ery6 D613E RTH
Ery6D613E.REV	GACCCGCTCCAGCGACG	
Ery6V751A.FOR	GCCGAGGAGATCCGCGAGAC	Ery6 V751A RTH
Ery6V751A.REV	GTGGCGGGAGTGCGAGG	

Supplemental Table 9. High-resolution LC-MS parameters and gradient.

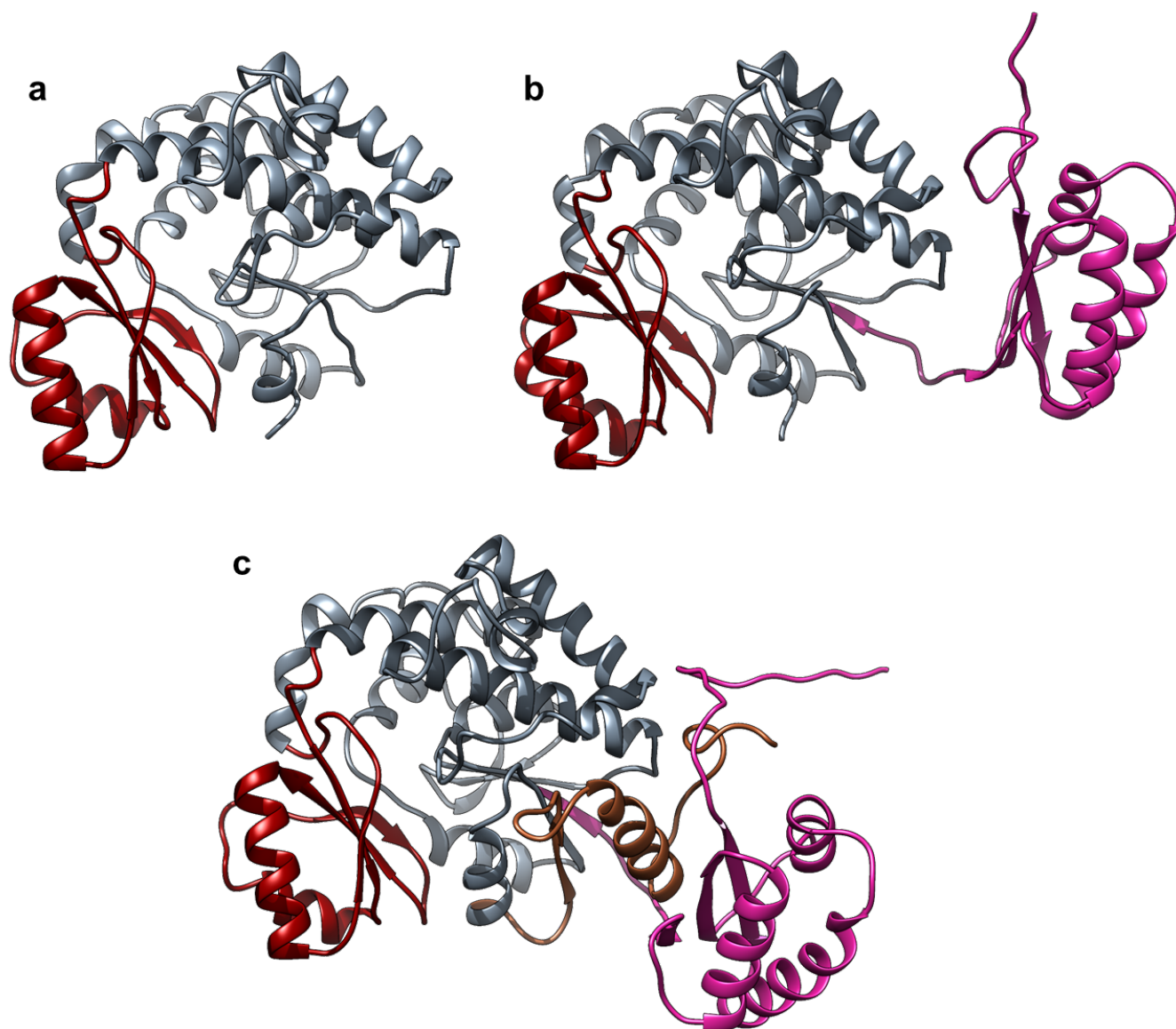
HESI Source Parameters	
Spray Voltage	3.5 kV
Capillary Temperature	350 °C
Heater Temperature	300 °C
S Lens RF Level	70 V
Sheath Gas Flow Rate	60 au
Resolution	70,000 FWHM
Scan Range	100-1000 m/z

LC Gradient	
Time (min)	% B
0.0	25.0
1.0	25.0
10.0	59.0
11.0	85.0
12.0	85.0
12.5	25.0
16.5	25.0

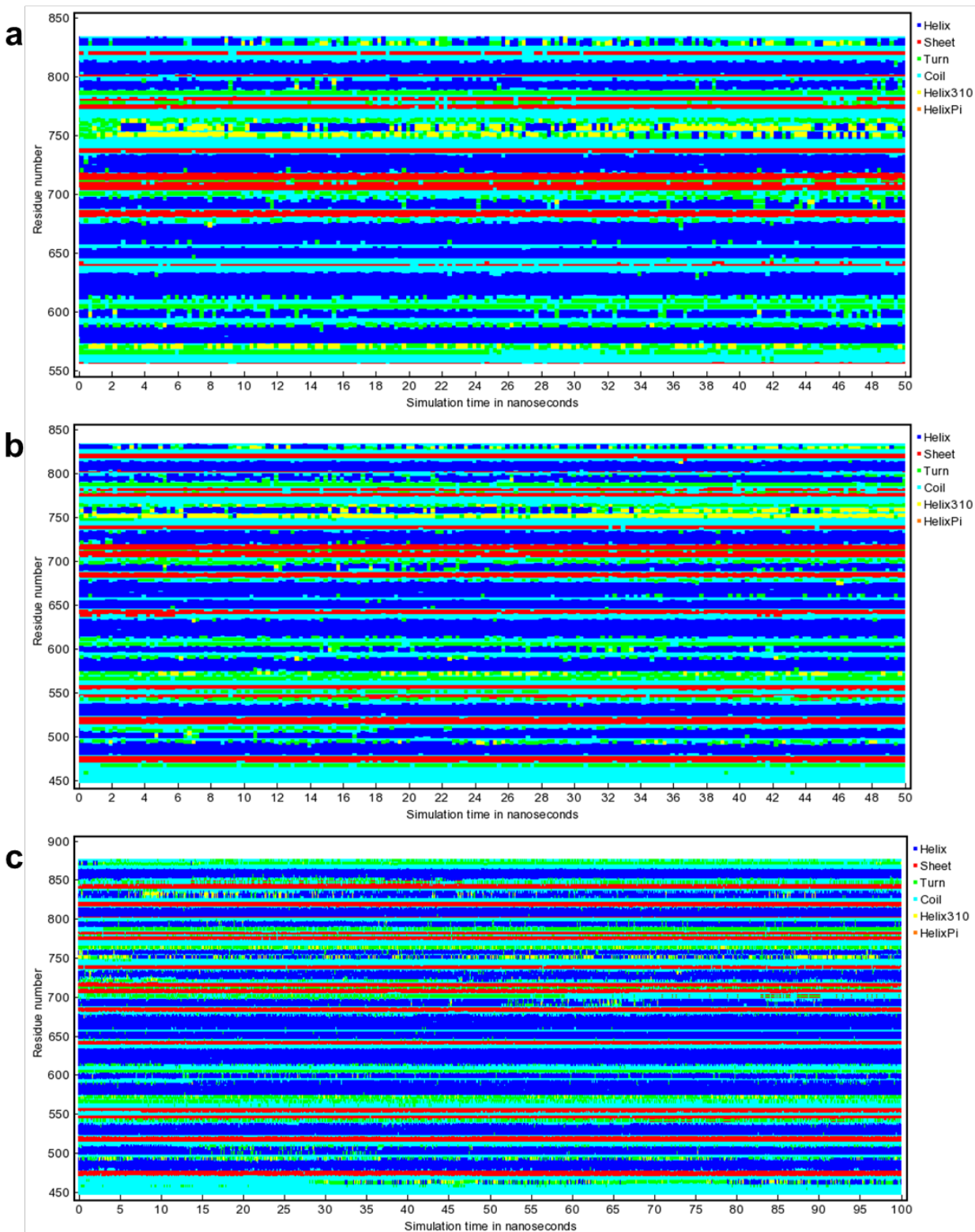
Supplementary Figures



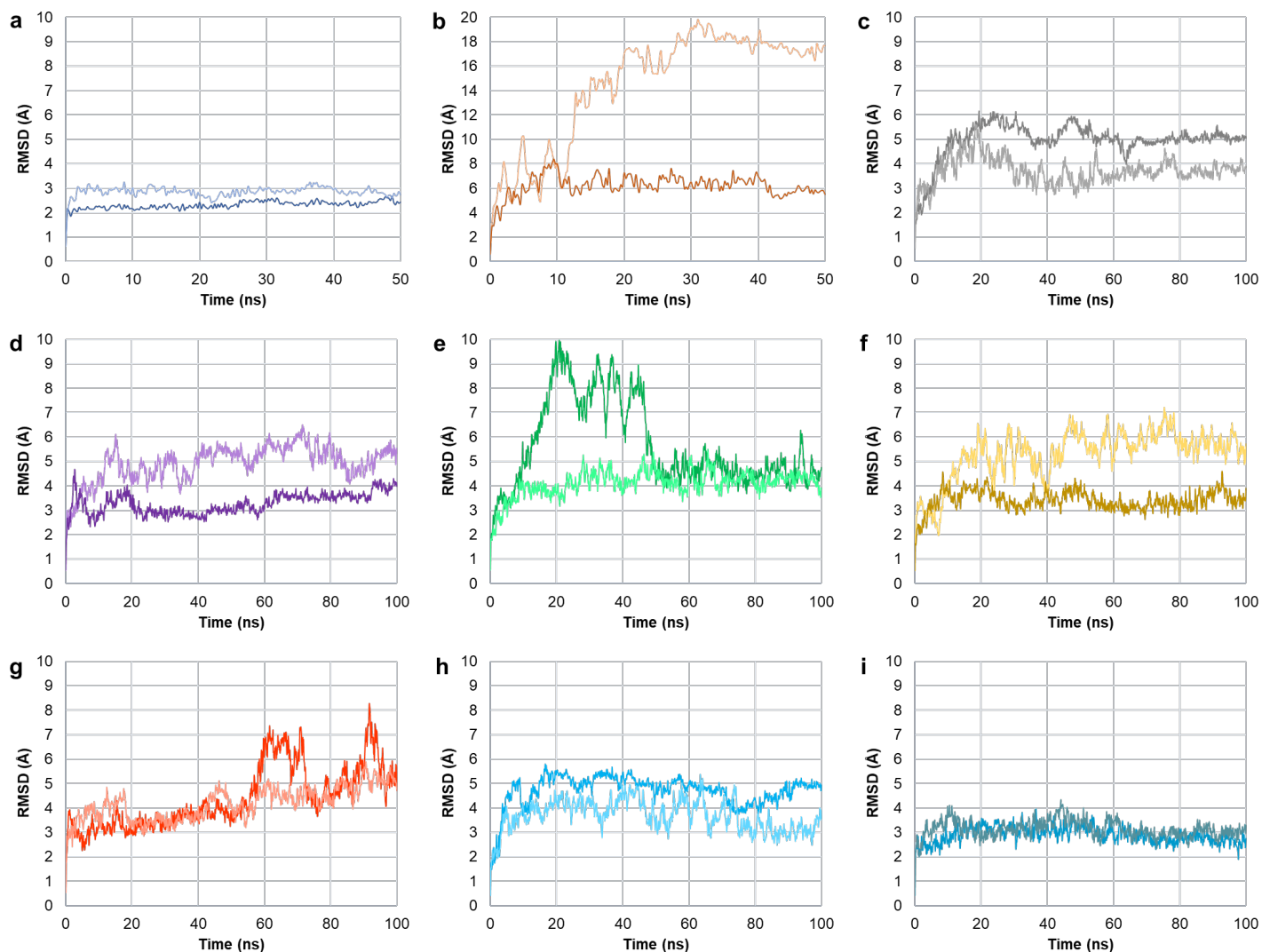
Supplemental Figure 1. Conserved motifs of PKS acyltransferases. Active site of EryAT6 with well-conserved motifs from methylmalonyl-CoA-specific ATs depicted as sticks (SeqLogos in **Supplemental Figure 9**).



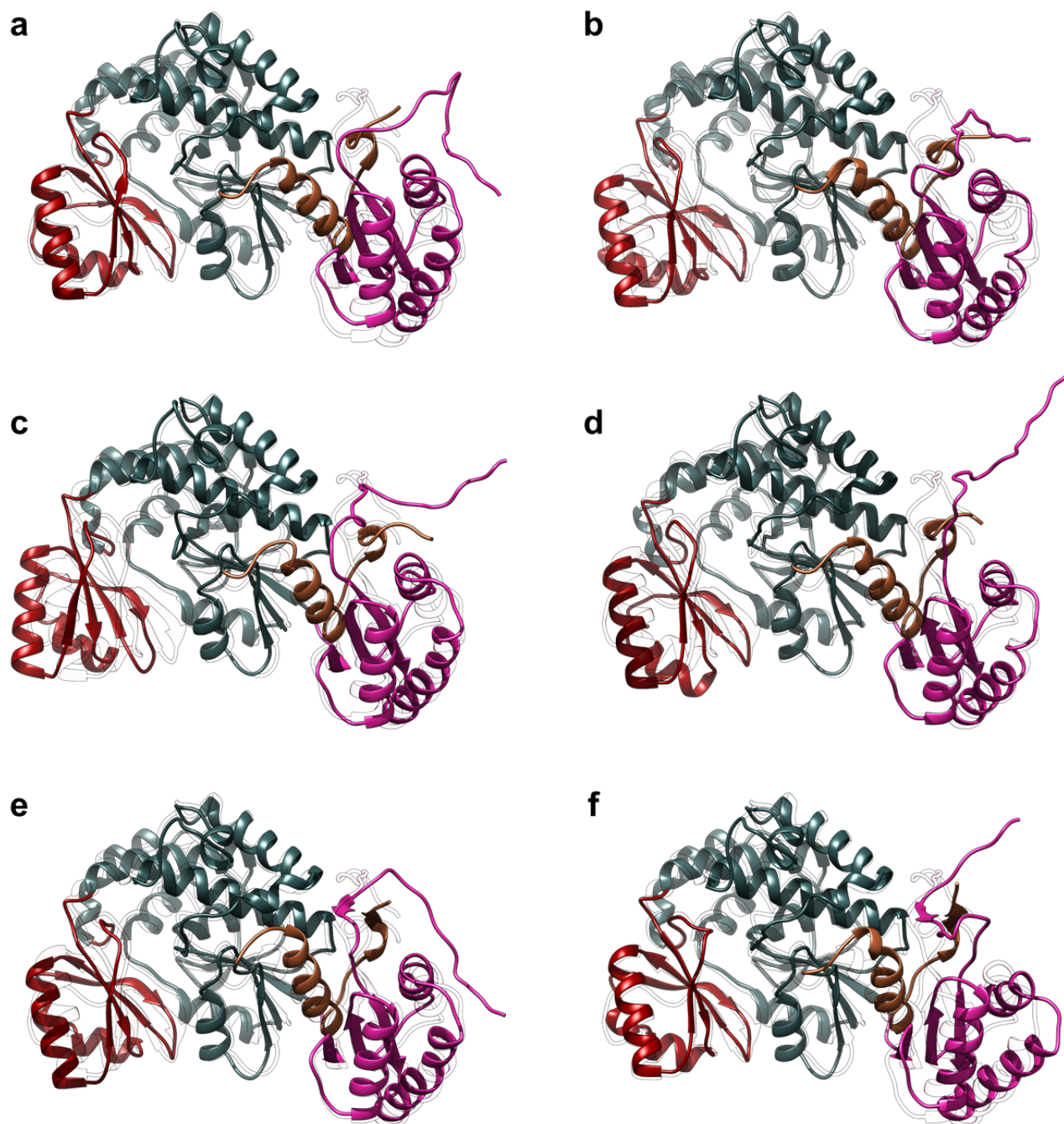
Supplemental Figure 2. Evaluation of EryAT6 models. Time-averaged wild-type EryAT6 homology model structures from MD simulations with (a) no linkers and a total of 278 residues, (b) an N-terminal linker (magenta) and a total of 387 residues, and (c) an N-terminal linker (magenta), a C-terminal linker (brown), and a total of 430 residues. This final model was selected for all remaining simulations, and the per-residue secondary structure over time is shown in Figure S3. The small (red) and large (grey) subunits are included in all models.



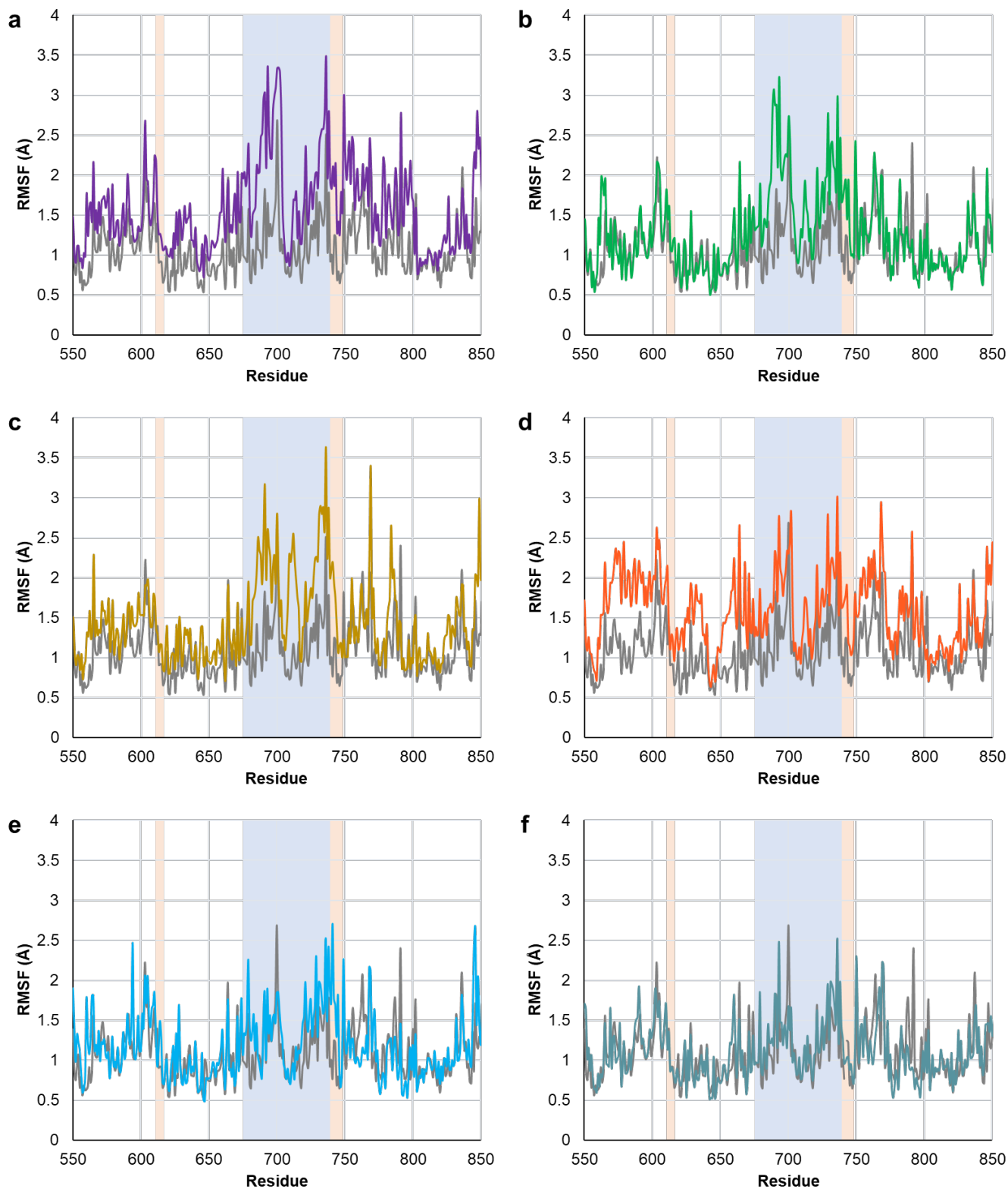
Supplemental Figure 3. Per-residue protein secondary structure of EryAT6 models. Wild-type EryAT6 homology models simulated with (a) no linkers and a total of 278 residues, (b) an N-terminal linker (magenta) and a total of 387 residues, and (c) an N-terminal linker (magenta), a C-terminal linker (brown), and a total of 430 residues each showed different levels of secondary structure retention over time. In the two models lacking a C-terminal linker, several secondary structures were less stable than in the two-linker model, especially in the helices near residue 750, which are important for the integrity of the active site.



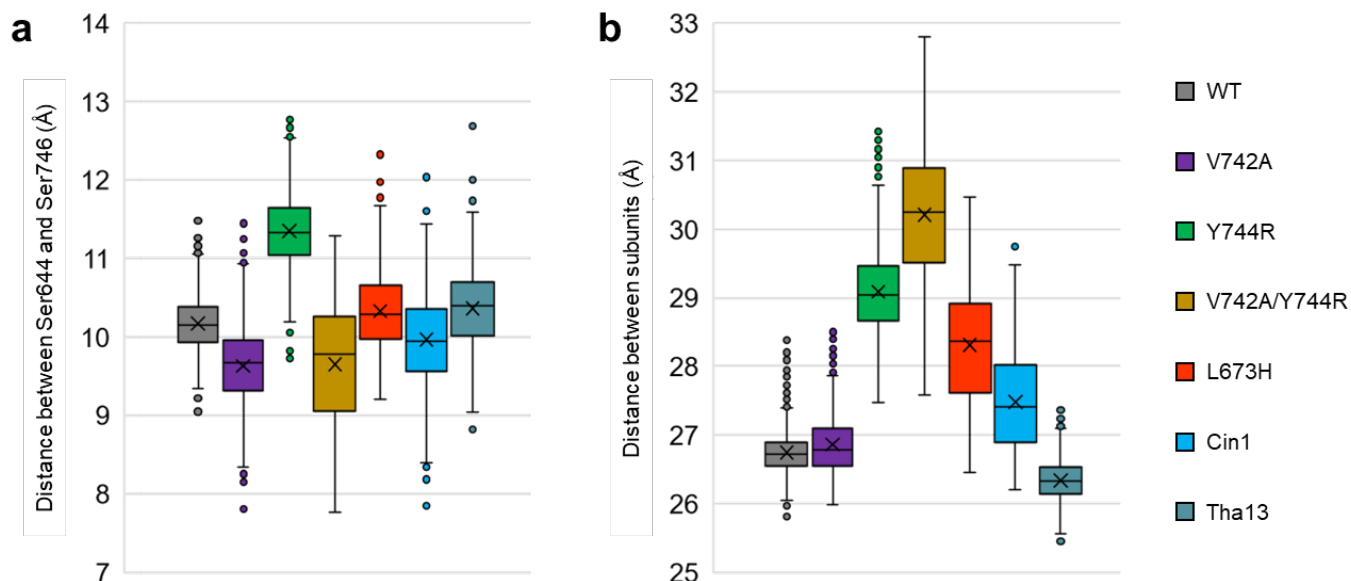
Supplemental Figure 4. RMSD of EryAT6 MD simulations. RMSD of the C α from all residues within the AT for each of the EryAT6 models: (a) wild-type with no linkers. (b) wild-type with only the N-terminal linker. (c) wild-type. (d) V742A. (e) Y744R. (f) V742A/Y744R. (g) L673H. (h) Cin1. (i) Tha13. Each simulation was run in duplicate and RMSDs are shown for each simulation. Rotation of the N- and/or C-terminal linkers resulted in large RMSD shifts in some simulations.



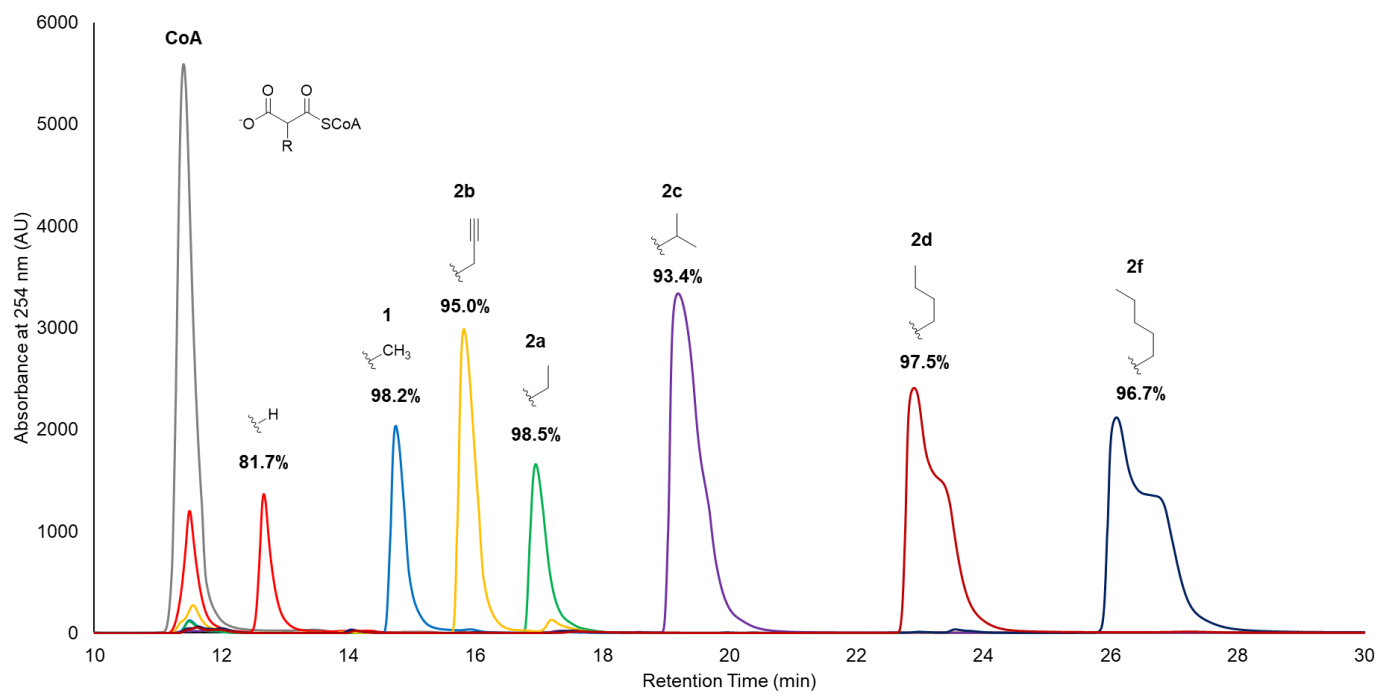
Supplemental Figure 5. Time-averaged structures of EryAT6 variants. Time-averaged structures of EryAT6 variants derived from the final 80 ns of the MD simulations are overlaid on the time-averaged wild-type EryAT6 structure (transparent). (a) The V742A mutant does not show a major rearrangement of the protein backbone relative to wild-type. (b) The Y744R mutant shows a significant increase in the width of the active site, especially in the proximity of the small subunit motif. (c) The double mutant V742A/Y744R shows significant perturbation of the small subunit, with little to no interaction between the catalytic residues. (d) The L673H mutant only shows a slight shift in the position of the small subunit relative to the large subunit. Both motif-swapped mutants, Cin1 (e) and Tha13 (f), show little difference in the average position of the two subunits, especially compared with the other two mutants containing Y744R.



Supplemental Figure 6. RMSF plots for EryAT6 simulations. Per-residue root-mean-square fluctuation (RMSF) measurements were calculated for the final 80 ns of each 100 ns MD simulation of the EryAT6 models to allow for an equilibration period. In each panel, the grey trace represents the movement seen in the wild-type simulation, overlaid by a second trace for each mutant: (a) V742A, (b) Y744R, (c) V742A/Y744R, (d) L673H, (e) Cin1, and (f) Tha13. Higher values correspond to increased movement of a residue over the time frame. The blue box represents the residues found within the small subunit of the AT. The orange boxes highlight the residues located within the large subunit motif (LSM; 611-617) and the small subunit motif (SSM; 739-747). Linkers were excluded from the analyses for clarity.

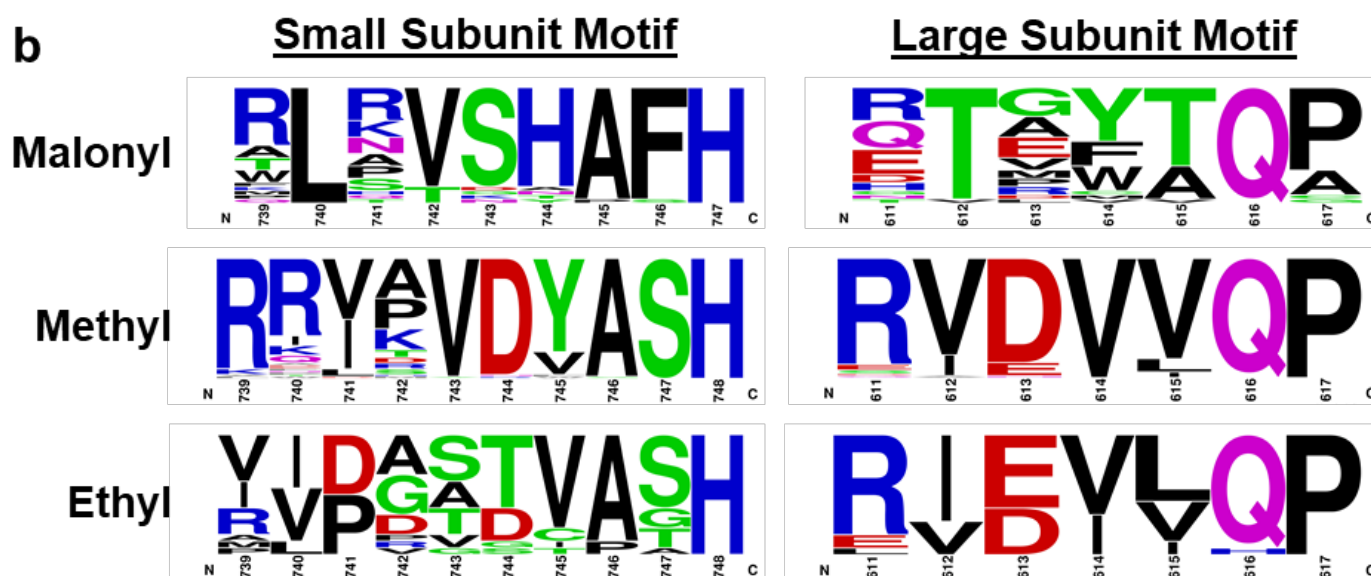


Supplemental Figure 7. Distance plots for EryAT6 simulations. Distance measurements were calculated for the final 80 ns of each 100 ns MD simulation of the EryAT6 models to allow for an equilibration period. **(a)** The distances between the catalytic Ser644 and the third residue of the YASH motif (Ser746 in wild-type) shows little variation except for a wider active site in the Y744R mutant. **(b)** The distance between the large and small subunits (by center of mass) shows a much shorter distance for the two motif-swapped mutants relative to most of the point mutants or even wild-type. The X indicates the mean. The boxes represent the range from the first to third quartiles with the center line representing the median. The whiskers represent the minima and maxima or, in the presence of outliers indicated by points, 1.5 times the interquartile range. These distances were measured every 0.1 ns over the final 80 ns ($n = 800$) of two independent simulations.



Supplemental Figure 8. Representative HPLC traces for MatB-synthesized malonyl-CoAs. Representative HPLC traces for MatB-synthesized malonyl-CoAs. Percent conversions were determined by remaining CoA peak. Individual reaction conditions are described in the methods.

a	(2S)-X-malonyl-CoA	Acyltransferase	Small Subunit Motif	Large Subunit Motif
	H	Avermectin AT5	TLP-TNHAFH	QTPYAQP
	Methyl	Erythromycin AT6	TLP-VDYASH	RVDVVQP
	Methyl	Rifamycin AT3	RVA-VDYASH	RVDVVQP
	Ethyl/Methyl	Monensin AT5	AVA-SDVAGH	RIDVVQP
	Ethyl	Niddamycin AT5	PIPGVDTAGH	RVDVVQP
	Allyl/Ethyl	FK506 AT4	RIA-VDCPTH	RVDVVHP
	Chloroethyl	Salinosporamide AT1	AVR-VDRPGH	ATDIQQP
	Butyl	Thailandin AT13	VLPGADAAGH	RLDVNQP
	Hexyl	Cinnabaramide AT1	ALR-VERAGH	EGDIQQP
	Benzyl	Splenocin AT1	RLR-MPAAAH	--TAAFM



Supplemental Figure 9. Motif sequences from natural acyltransferases. (a) Selected examples of SSM and LSM sequences for ATs with different CoA-linked substrates. (b) Consensus sequences of AT domains with substrate specificities for malonyl-CoA (100 sequences), methylmalonyl-CoA (100 sequences), and ethylmalonyl-CoA (19 sequences) are shown. Sequence logos were created using WebLogo.⁶

a

Relative Abundance

RT: 8.01
AA: 157987498

RT: 8.26
AA: 46103

RT: 8.93
AA: 450149

RT: 9.61
AA: 1121712

RT: 10.73
AA: 1037340

RT: 12.37
AA: 110650

RT: 13.39
AA: 2674806

RT: 14.24
AA: 4004082

RT: 14.92
AA: 1135633

RT: 0.84
AA: 5232697

RT: 1.01
AA: 3636743

RT: 2.76
AA: 773880

RT: 3.94
AA: 181924

RT: 5.02
AA: 293271

RT: 6.41
AA: 26808

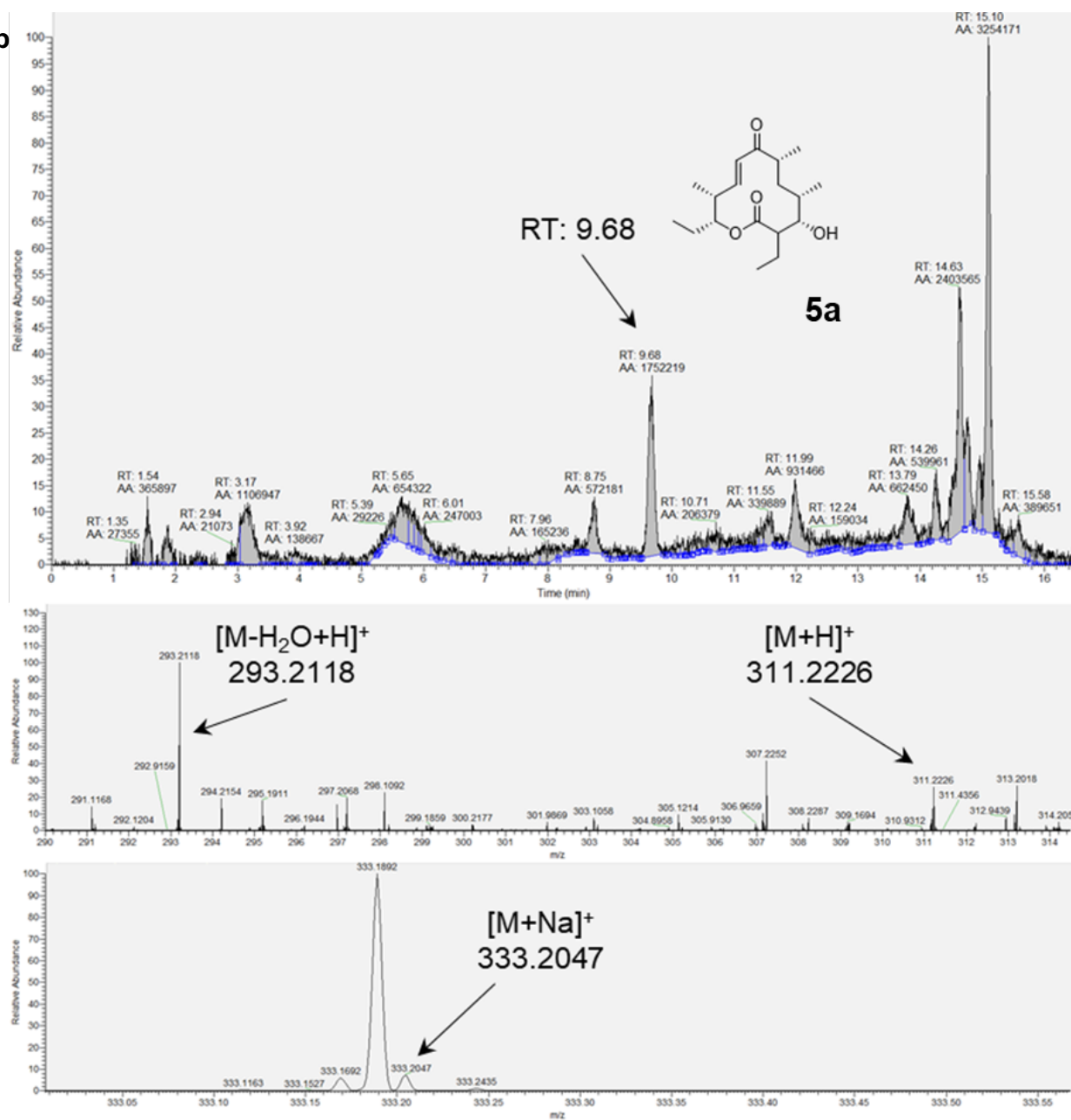
RT: 7.22
AA: 185562

Time (min)

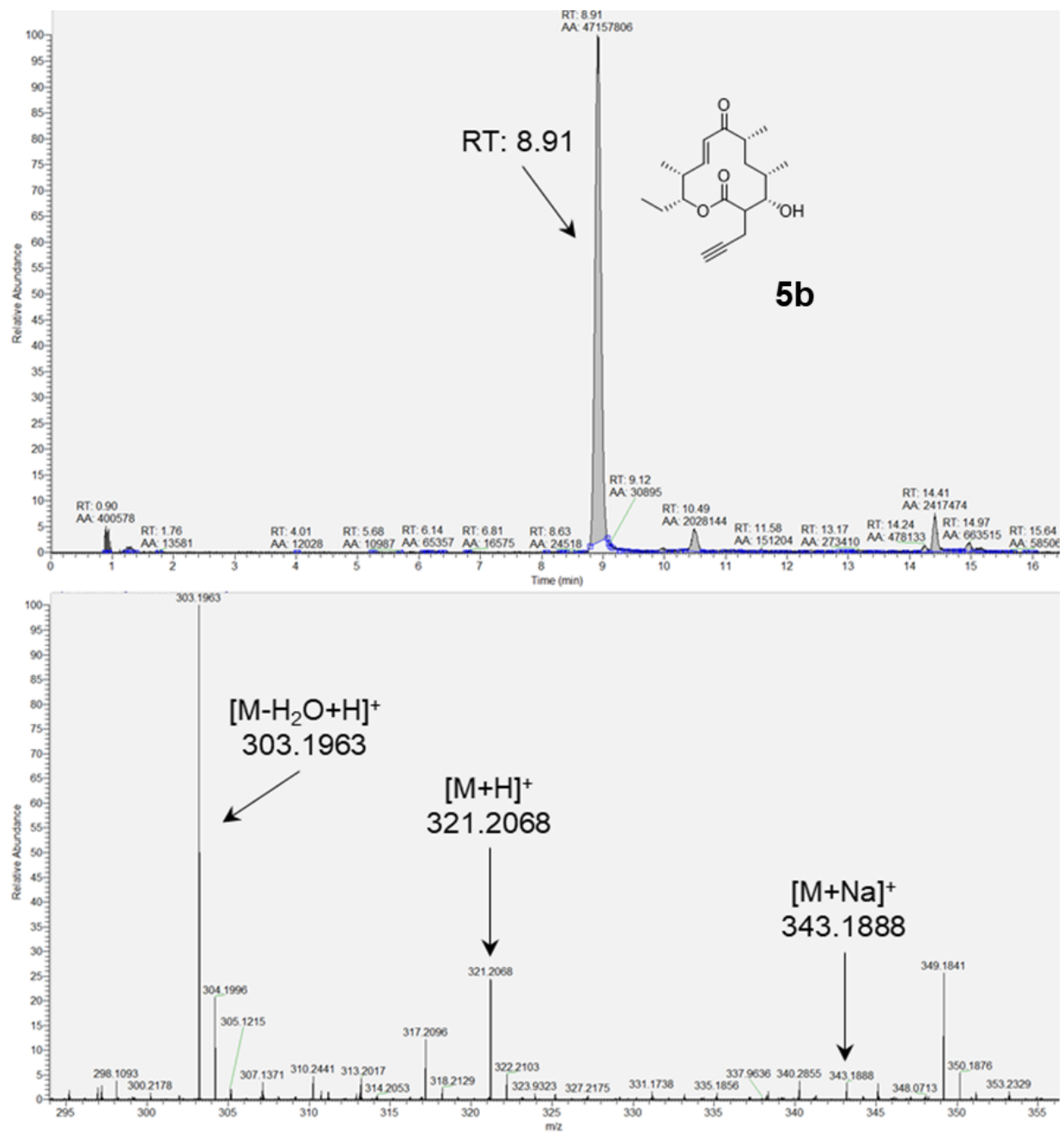
4

CC1(C)C(=O)C(=C)C2(C)C(C1OC2=O)C(C)(C)C(C)O

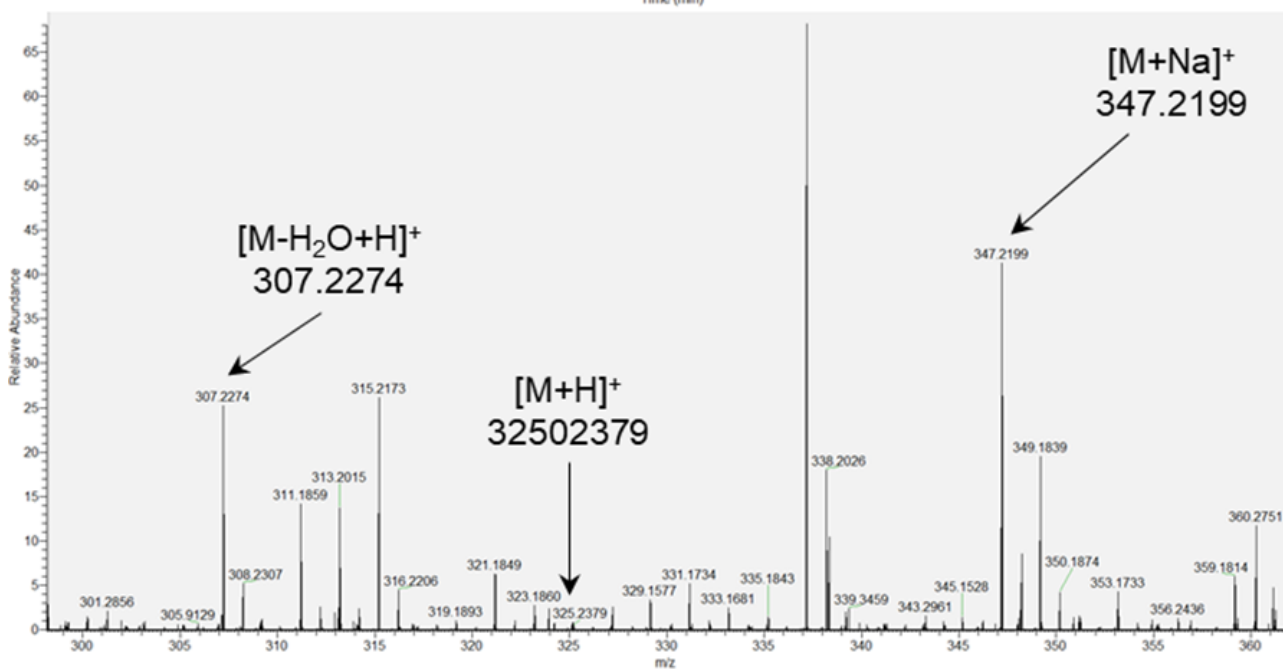
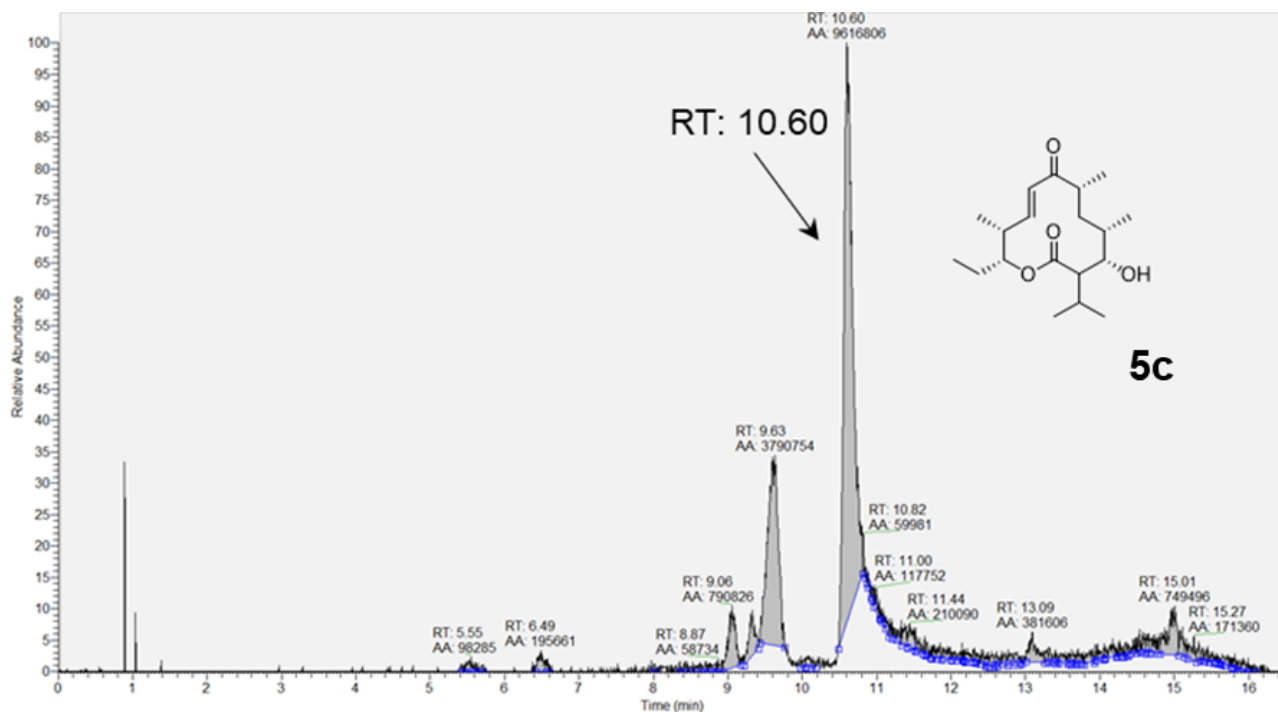
b

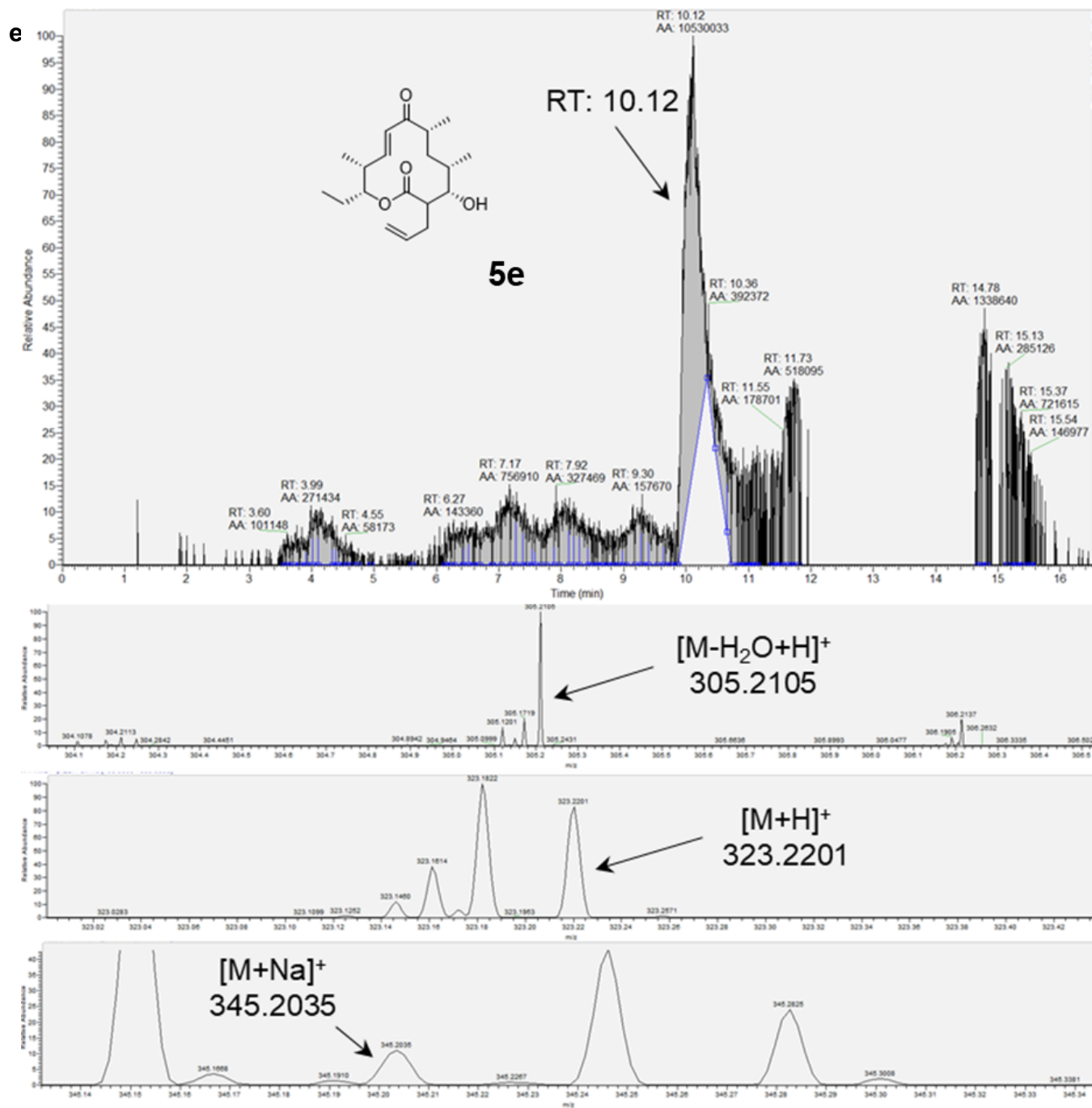


c

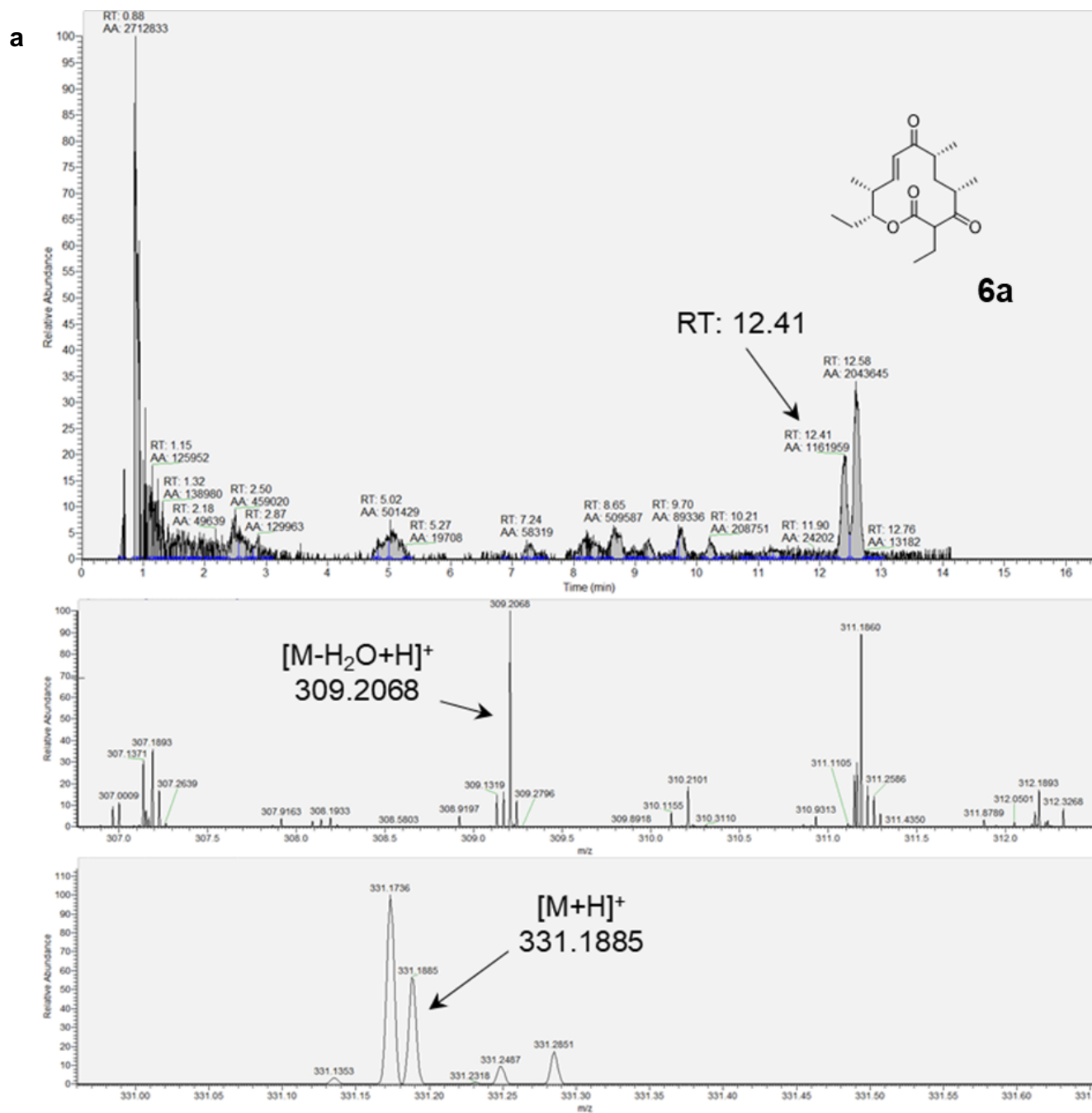


d

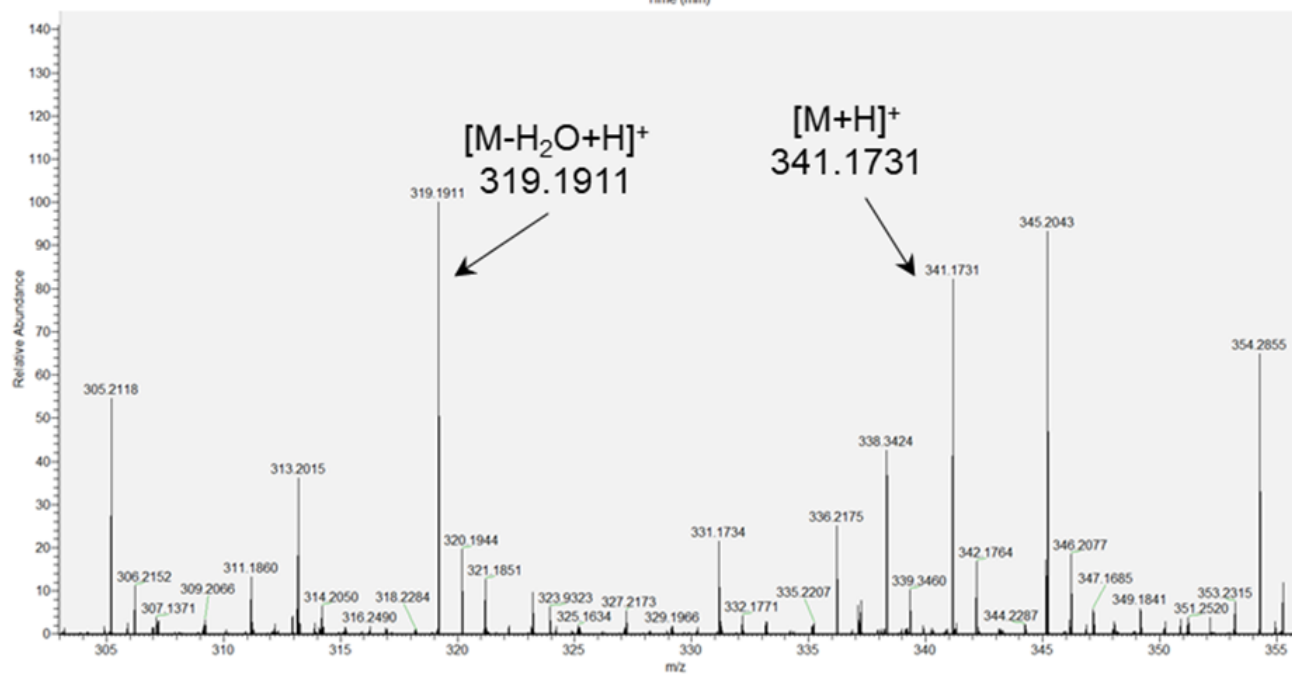
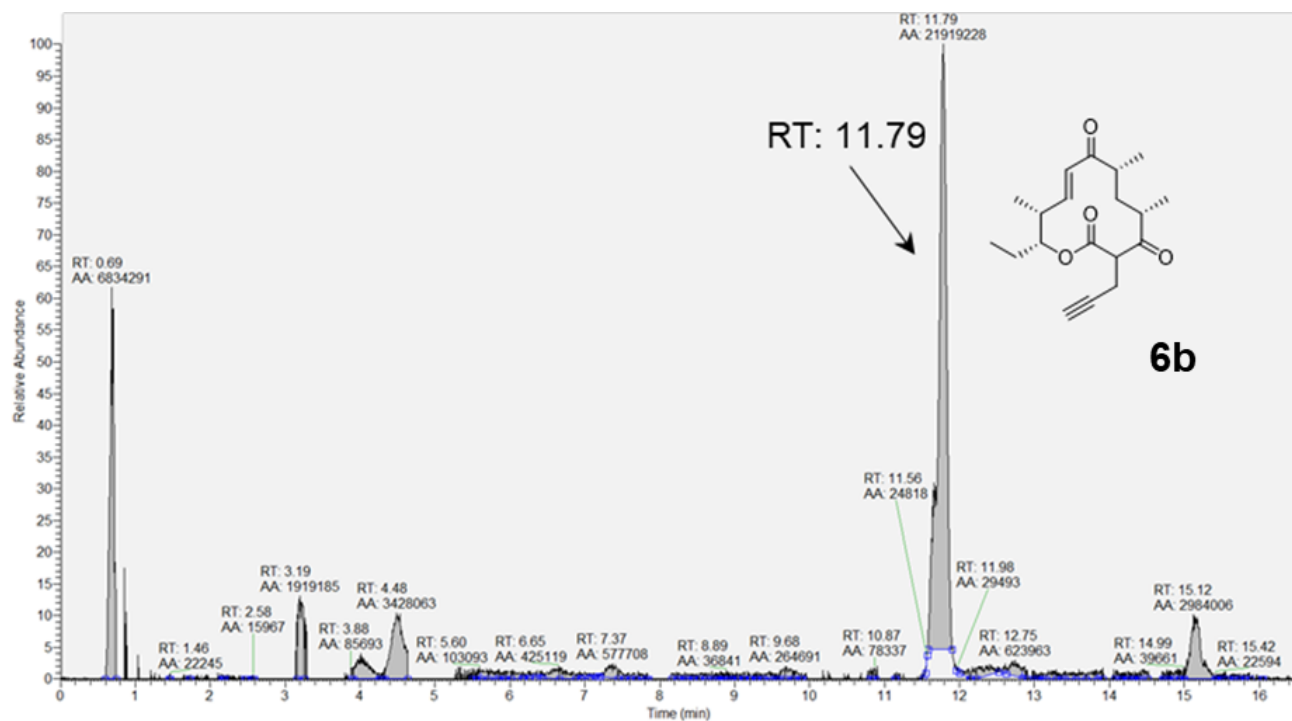


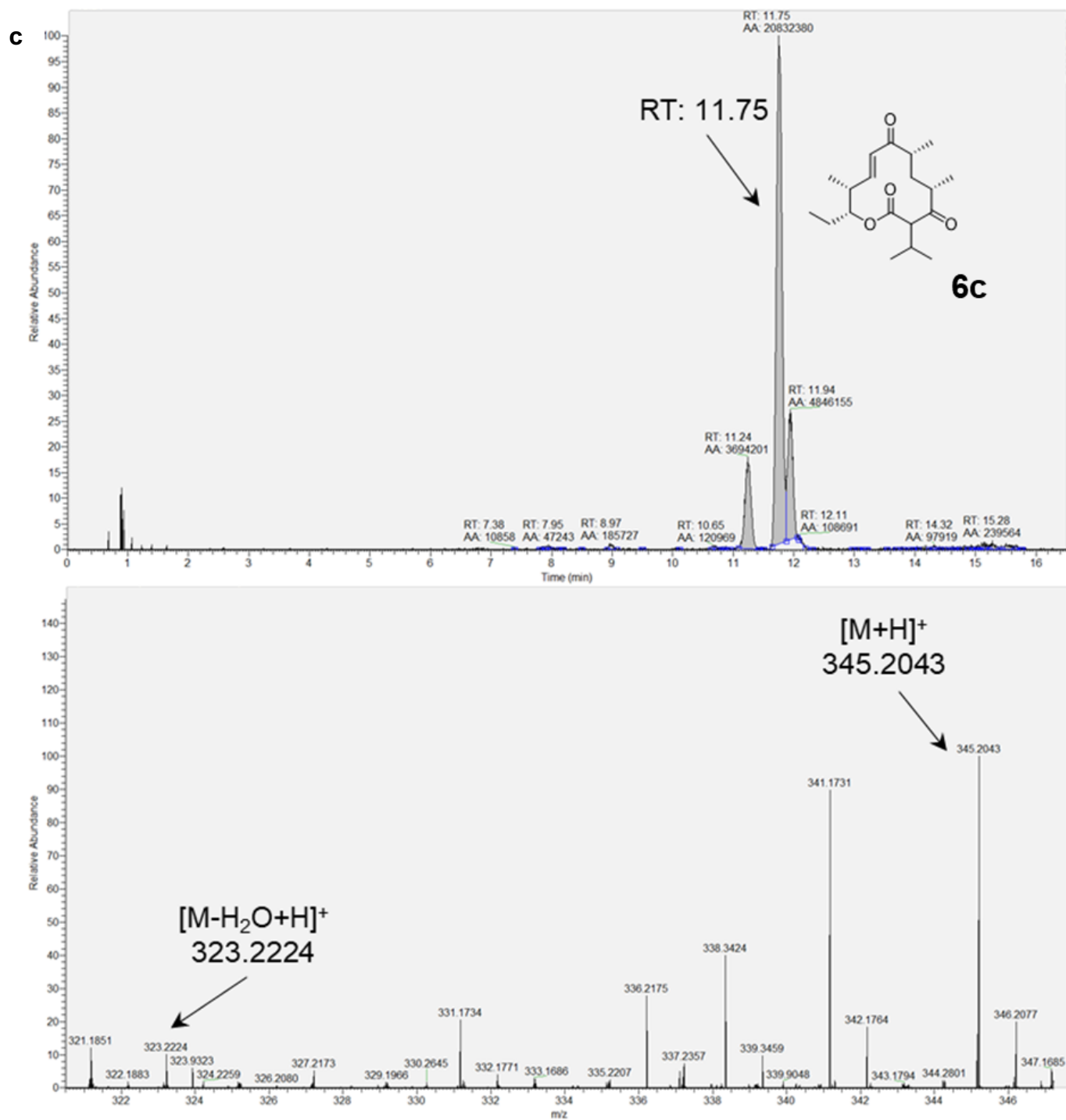


Supplemental Figure 11. Representative mass spectra of keto-10-deoxymethynolide B compounds from lysate module reactions. Top panel for each compound shows the extracted ion chromatogram. Bottom panel(s) show the total ion spectra.

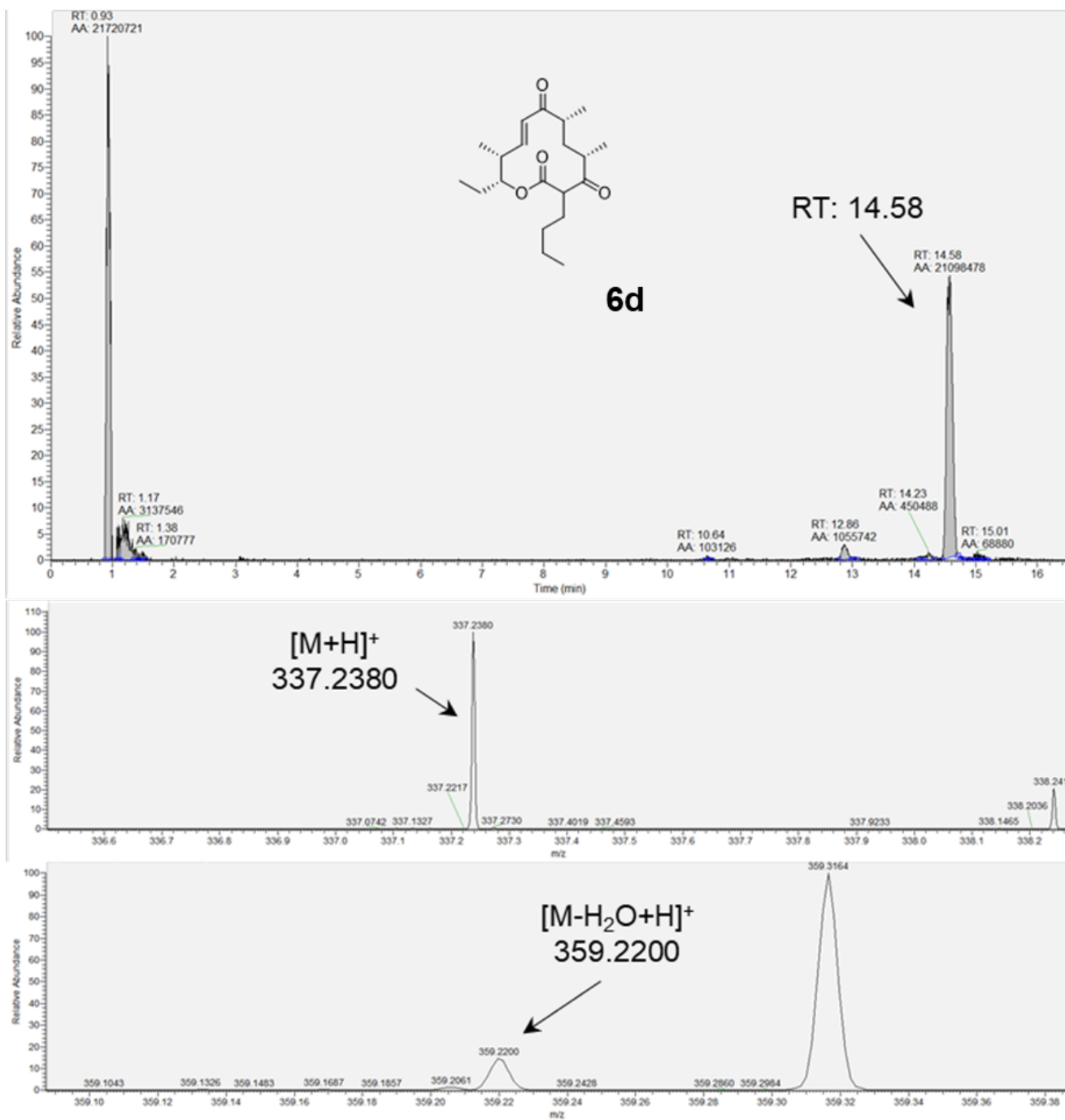


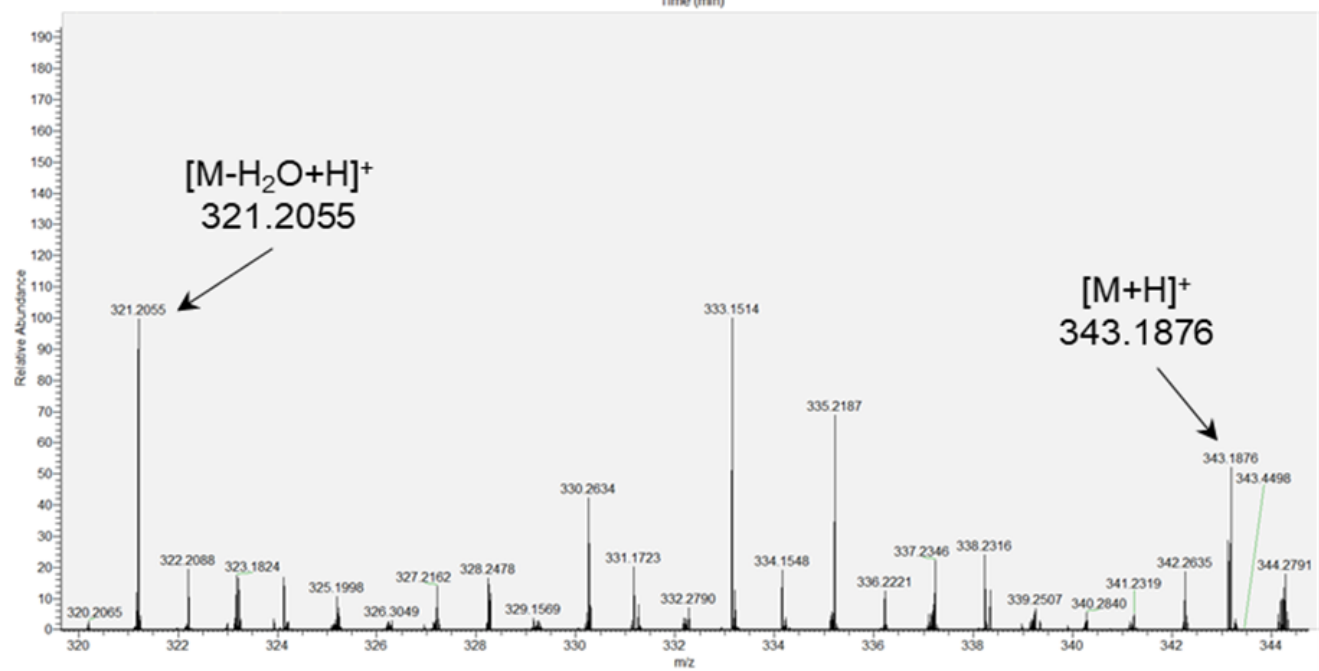
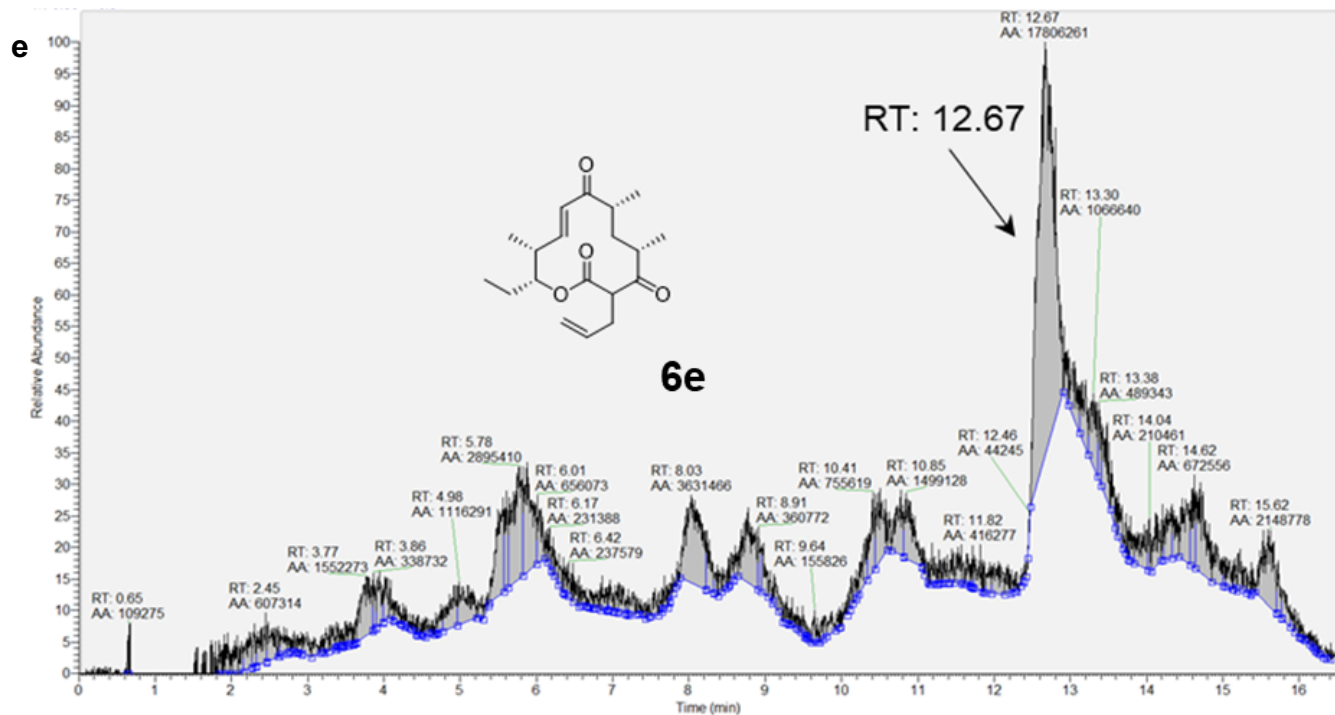
b

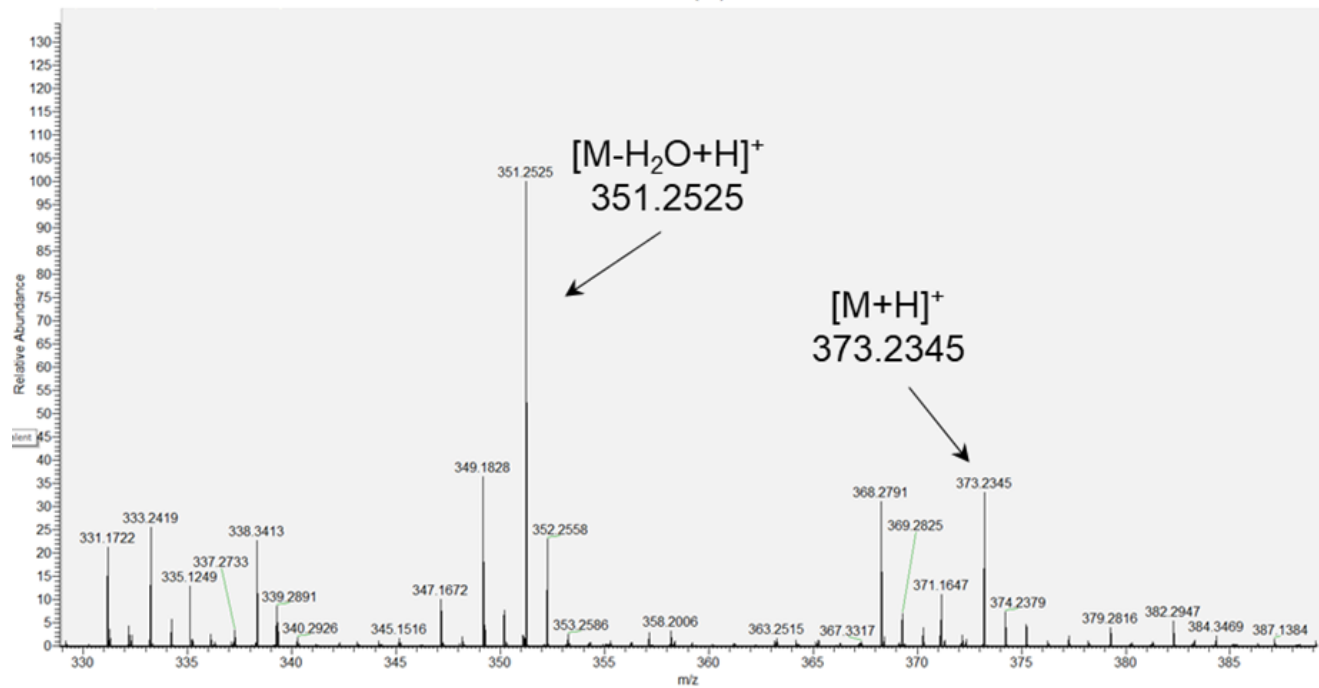
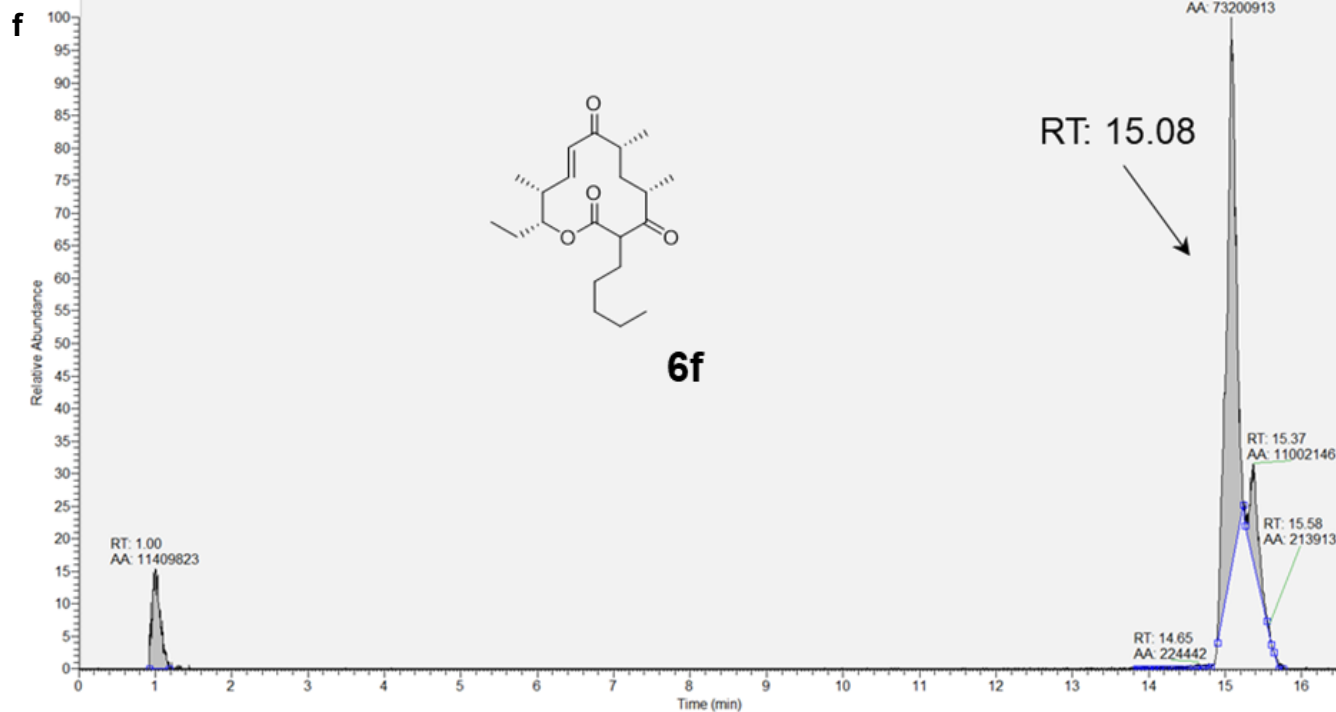




d







Supplementary Methods

Site-directed mutagenesis of Ery6TE

All Ery6 and Er6TE site-directed mutants were constructed by 'round-the-horn' mutagenesis using the templates from **Supplemental Table 7** and the oligonucleotide sequences described **Supplemental Table 8**.¹ Each PCR contained: 5X GC Phusion DNA polymerase buffer (4 μ L), DNAase free water (11.8 μ L), forward/reverse primer mix (2 μ L, 10 μ M), template DNA (1 μ L, 50 ng/ μ L), dNTPs (0.4 μ L, 2.5 mM each dNTP), DMSO (0.6 μ L), and Phusion High Fidelity DNA Polymerase (0.2 μ L) (New England Biolabs, NEB). PCR cycling parameters: step 1) 98 °C, 60 s; step 2) 29x [a) 98 °C, 10 s; b) 72 °C, 5 min]; step 3) 72 °C, 5 min. Next, the PCR mixture was subjected to restriction enzyme digest with *DpnI* to remove any remaining template DNA. The *DpnI* reaction mixture contained: 10X CutSmart Buffer (2 μ L), the PCR mixture (17 μ L), and *DpnI* (1 μ L). The mixture was vortexed, centrifuged, and incubated for 3 h at 37 °C. The digested reaction mixture was then purified by agarose gel electrophoresis, with a total of 8 μ L DNA eluted at the final step. The PCR product was then ligated using T4 DNA ligase in a reaction containing: 1 μ L of the 10X T4 DNA ligase buffer, 8 μ L of the *DpnI*-treated and purified PCR product, and 1 μ L of the T4 DNA ligase (overnight incubation at 16 °C). Each subsequent ligation reaction was transformed directly into *E. coli* 10G electrocompetent cells (Lucigen). Individual transformants were sequenced to identify incorporation of mutant codons.

Construction of Ery6TE motif chimeras

Motif swaps of Ery6TE AT were constructed using Gibson assembly (NEB) and boundaries defined by the Keasling group.² The Gibson assembly mixtures were used to transform *E. coli* DH5 α competent cells. Successful incorporation of mutations was confirmed by DNA sequencing of purified plasmids from single transformants. Each mutant plasmid was transformed into *E. coli* K207-3 competent cells and plated onto LB agar plates (50 μ g mL⁻¹ kanamycin and 100 μ g mL⁻¹ spectinomycin).

Expression and purification of wild-type and mutant MatB

The expression and purification of MatB wild-type and T207G/M306I have been previously described.³⁻⁵ Briefly, *E. coli* BL21(DE3) pLysS competent cells were transformed with plasmid, and positive transformants were selected on LB agar supplemented with 50 μ g/mL kanamycin. A single colony was transferred to LB (3 mL) supplemented with kanamycin (50 μ g/mL) and grown at 37 °C and 250 rpm overnight. The culture was used to inoculate LB media (1 L) supplemented with kanamycin (50 μ g/mL). One liter culture was incubated at 37 °C and 250 rpm to an OD₆₀₀ of 0.6, at which time protein synthesis was induced by the addition of IPTG to a final concentration of 1 mM. After incubation at 18 °C and 200 rpm for 18 h, cells were collected by centrifugation at 5,000 *g* for 20 min and resuspended in 100 mM Tris-HCl pH 8.0 (20 mL) containing NaCl (300 mM) and then lysed by sonication. Following centrifugation at 10,000 *g*, the soluble extract was loaded onto a 1 mL HisTrap HP column (GE Healthcare, Piscataway, NJ) and purified by fast protein liquid chromatography using the

following buffers: wash buffer [20 mM phosphate (pH 7.4) containing 0.5 M NaCl and 20 mM imidazole] and elution buffer [20 mM phosphate (pH 7.4) containing 0.5 M NaCl and 200 mM imidazole]. The purified protein was concentrated using an Amicon Ultra 30 kDa MWCO centrifugal filter (Millipore Corp., Billerica, MA) and stored as 10% glycerol stocks at -80 °C. Protein purity was verified by SDS-PAGE. Protein quantification was carried out using the Bradford Protein Assay Kit from Bio-Rad.

Synthesis of acyl-CoAs by MatB

The MatB-catalyzed synthesis of extender units **1** (wild-type MatB) and **2a-f** (MatB T207G/M306I) has been previously described.³⁻⁵ Briefly, reactions were performed in a 50 µL reaction mixture containing 100 mM sodium phosphate (pH 7), MgCl₂ (2 mM), ATP (12 mM), coenzyme A (8 mM), malonate or corresponding analog (16 mM) and wild-type or mutant MatB (10 µg) at 25 °C. Aliquots were removed after 3 h incubation, and quenched with an equal volume of ice-cold methanol, centrifuged at 10,000 *g* for 10 min, and cleared supernatants used for HPLC analysis on a Varian ProStar HPLC system. A series of linear gradients was developed from 0.1% TFA (A) in water to methanol (HPLC grade, B) using the following protocol: 0-32 min, 80% B; 32-35 min, 100% A. The flow rate was 1 mL/min, and the absorbance was monitored at 254 nm using Pursuit XRs C18 column (250 x 4.6 mm, Varian Inc.). To ensure complete conversion, the malonate analogue and the acyl-CoA product HPLC peak areas were integrated, and the conversion (%) calculated as a percent of the total peak area. Product elution times and LC-MS data were in complete agreement with that previous described.^{3,4}

Supplementary References

- 1 Moore, S. 'Round-the-horn site-directed mutagenesis', <https://openwetware.org/wiki/Round-the-horn_site-directed_mutagenesis>
- 2 Yuzawa, S. *et al.* Comprehensive in vitro analysis of acyltransferase domain exchanges in modular polyketide synthases and its application for short-chain ketone production. *ACS Synth. Biol.* **6**, 139-147 (2017).
- 3 Koryakina, I. *et al.* Poly specific trans-acyltransferase machinery revealed via engineered acyl-CoA synthetases. *ACS Chem. Biol.* **8**, 200-208 (2013).
- 4 Koryakina, I., McArthur, J. B., Draelos, M. M. & Williams, G. J. Promiscuity of a modular polyketide synthase towards natural and non-natural extender units. *Org. Biomol. Chem.* **11**, 4449-4458 (2013).
- 5 Koryakina, I. & Williams, G. J. Mutant malonyl-CoA synthetases with altered specificity for polyketide synthase extender unit generation. *ChemBioChem* **12**, 2289-2293 (2011).
- 6 Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188-1190 (2004).