true or highly likely substrate	method of verification	assay type	reference	doi	fc correct	ss correct	organism	predicted functional class fc	fc prediction probability	predicted substrate specificity ss	ss prediction probability	accession	perc aa identity to closest training set
2,3-dihydroxybenzoate	purified protein in heterologous host	malachite green pyrophosphate assay	Zane et al. (2014) *JACS.* 136(15), 5615-5618	doi: 10.1021/ja5019942.	1	1	*Vibrio campbellii ATCC BAA-1116*	Aryl-CoA ligase	0.98	aryl and biaryl derivatives	0.95	ABU70375.1	59.7
proline	biosynthetic logic based on gene knockouts and natural product isolation	biosynthetic logic sequence-based prediction	Fu et al. (2019) *ChemBioChem* 20(6), 764-769	doi: 10.1002/cbic.201800791.	1	1	*Streptomyces armeniacus*	Nonribosomal peptide synthetase	0.98	cyclic aliphatic	0.66	AZY92000.1	64.4
glycine	purified protein from heterologous host, in vitro biochemical assay	fluorogenic pyrophosphate production assay	Shi et al. (2019) *Chemical Science* 10(10), 3042-3048	doi: 10.1039/c8sc05670f	1	1	*Streptomyces lividans SBT19*	Nonribosomal peptide synthetase	0.98	tiny	0.42	QCE20608.1	36.2
tyrosine	product abolished after genetic knockout, stable isotope labeling to show tyrosine incorporation	gene deletions	Lin et al. (2019) *Molecules* 24, 2267	doi: 10.3390/molecules24122267	1	1	*Neonectria* sp. DH2	Nonribosomal peptide synthetase	0.96	bulky mainly phenyl derivatives	0.75	BGC0002035|c1|342|	58.1
C12 - C16 (preferred) C10-C18 tested	product purification from native host, MS-based assessment of structural variability of fatty acid moiety	product determination by ultra-high performance liquid chromatography plus (HPLC+)	Mares et al. (2019) *AEM*. 85(4), 17	doi: 10.1128/aem.02675-18	1	1	*Cylindrospermum moravicum CCALA 993*	Fatty-acyl AMP ligase	0.91	C13 through C17	0.72	AXN93578.1	94.1
3-formamidosalicylate	BAC cloning and expression of biosynthetic cluster in heterologous host, biosynthetic logic based on natural product isolation	HPLC-MS analysis of natural product intermediates, sequence-based prediction	Awakawa et al. (2018) *Nat. Comm.* 9, 3534	doi: 10.1038/s41467-018-05877-z.	1	1	*Streptomyces orinoci*	Aryl-CoA ligase	0.91	cinnamate and succinylbenzoate derivatives	0.63	BBD17766.1	79.3
C12 - C16 (preferred) C10-C18 tested	product purification from native host, MS-based assessment of structural variability of fatty acid moiety	product determination by ultra-high performance liquid chromatography plus (HPLC+)	Mares et al. (2019) *AEM*. 85(4), 17	doi: 10.1128/aem.02675-19	1	1	*Anabaena minutissima UTEX B 1613*	Fatty-acyl AMP ligase	0.91	C13 through C17	0.44	AXN93595.1	68.7
3-formamidosalicylate	BAC cloning and expression of biosynthetic cluster in heterologous host, biosynthetic logic based on natural product isolation	HPLC-MS analysis of natural product intermediates, sequence-based prediction	Awakawa et al. (2018) *Nat. Comm.* 9, 3534	doi: 10.1038/s41467-018-05877-z.	1	1	*Streptomyces sp.*	Aryl-CoA ligase	0.9	cinnamate and succinylbenzoate derivatives	0.6	BBD17745.1	79.2
C12 - C16 (preferred) C10-C18 tested	product purification from native host, MS-based assessment of structural variability of fatty acid moiety	product determination by ultra-high performance liquid chromatography plus (HPLC+)	Mares et al. (2019) *AEM*. 85(4), 17	doi: 10.1128/aem.02675-20	1	1	*Anabaena sp. UHCC-0399*	Fatty-acyl AMP ligase	0.89	C13 through C17	0.47	AXN93608.1	67.9
glycine, serine, alanine	purified protein from heterologous host, in vitro biochemical assay	fluorogenic pyrophosphate production assay	Shi et al. (2019) *Chemical Science* 10(10), 3042-3048	doi: 10.1039/c8sc05670f	1	0	*Streptomyces lividans SBT20*	Nonribosomal peptide synthetase	0.88	cysteine	0.43	QCE20598.1	45.8
N-formyl glycine	biosynthetic logic-based on gene knockouts and natural product isolation	biosynthetic logic sequence-based prediction	Bauman et al. (2019) *Cell Chem. Biol.* 26, 724	doi: 10.1038/s41467-018-05877-z.	1	1	*Streptomyces sp. CNB091*	Nonribosomal peptide synthetase	0.85	small hydrophilic	0.3	WP_063738219.1	47.0
decanoic acid	biosynthetic logic-based on gene knockouts and natural product isolation	biosynthetic logic sequence-based prediction	Koomsiri et al. (2019) *J. Nat. Prod.* 82(8), 2144-2151	doi: 10.1021/acs.jnatprod.9b00074	1	0	*Streptomyces sp.*	Fatty-acyl AMP ligase	0.83	C13 through C17	0.51	QED55419.1	47.4
(2*Z*,4*E*)-8-methyldeca-2,4-dienoic acid	TAR cloning and expression of biosynthetic cluster in heterologous host, biosynthetic logic	biosynthetic logic sequence-based prediction	Wu et al. (2019) *JACS* 141(9), 3910-3919	doi: 10.1021/jacs.8b12087	1	0	uncultured bacterium metagenome	Fatty-acyl AMP ligase	0.83	C13 through C17	0.43	QBC75017.1	51.8
phenylacetate	purified protein from heterologous host	NADH-consumption continuous spectrophotometric assay	Burckhardt et al. (2020) *Mol. Microbiol.* 113, 253-269	doi: 10.1111/mmi.14414	1	1	*Streptomyces lividans TK25*	Aryl-CoA ligase	0.82	aryl and biaryl derivatives	0.74	EFD64737.1	66.4

C12 - C16 (preferred) C10-C18 tested	product purification from native host, MS-based assessment of structural variability of fatty acid moiety	product determination by ultra-high performance liquid chromatography plus (HPLC+)	Mares et al. (2019) *AEM* 85(4), 17	doi: 10.1128/aem.02675-21	1	1	*Anabaena minutissima UTEX B 1613*	Fatty-acyl AMP ligase	0.81	C6 through C12	0.55	AXN93598.1	67.7
C12 - C16 (preferred) C10-C18 tested	product purification from native host, MS-based assessment of structural variability of fatty acid moiety	product determination by ultra-high performance liquid chromatography plus (HPLC+)	Mares et al. (2019) *AEM* 85(4), 17	doi: 10.1128/aem.02675-22	1	1	*Anabaena sp. UHCC-0399*	Fatty-acyl AMP ligase	0.81	C6 through C12	0.55	AXN93611.1	66.0
2,3-dihydroxybenzoate	purified protein from heterologous host	NADH-consumption continuous spectrophotometric assay	Burckhardt et al. (2020) *Mol. Microbiol.* 113, 253-269	doi: 10.1111/mmi.14415	1	1	*Streptomyces lividans TK24*	Aryl-CoA ligase	0.79	aryl and biaryl derivatives	0.76	EFD64524.1	54.9
myristoleic acid	biosynthetic logic-based on isolation of natural product and intermediates	biosynthetic logic sequence-based prediction	Moosmann et al. (2018) *Org. Letters* . 20(17), 5238-5241	doi: 10.1021/acs.orglett.8b02193	1	1	*Fischerella sp. PCC 9339*	Fatty-acyl AMP ligase	0.78	C13 through C17	0.49	DAB41914.1	54.4
alanine	purified protein from heterologous host, in vitro biochemical assay	fluorogenic pyrophosphate production assay	Shi et al. (2019) *Chemical Science* 10(10), 3042-3048	doi: 10.1039/c8sc05670f	1	1	*Streptomyces lividans SBT21*	Nonribosomal peptide synthetase	0.77	small hydrophilic	0.3	QCE20609.1	39.8
luciferin	purified protein from heterologous host, structure determination, steady-state kinetics	spectrophotometric assay	Carrasco-Lopez et al. (2018) *Life Sci Alliance* 1(4), e201800072	doi: 10.26508/lsa.201800072	1	1	*Amydetes vivianii*	Luciferase	0.76	luciferin	0.79	6AAA_A	86.4
2-aminobenzoate	purified protein from heterologous host	NADH-consumption continuous spectrophotometric assay	Burckhardt et al. (2020) *Mol. Microbiol.* 113, 253-269	doi: 10.1111/mmi.14416	1	1	*Streptomyces lividans TK27*	Aryl-CoA ligase	0.7	aryl and biaryl derivatives	0.57	EFD66106.1	33.7
4-acetamidopyrrole-2-carboxylate	biosynthetic logic-based on gene knockouts and natural product isolation	biosynthetic logic sequence-based prediction	Vingadassalon et al. (2015) *ACS Chem. Biol.* 10(2), 601-610	doi: 10.1021/cb500652n	1	1	*Streptomyces netropsis*	Aryl-CoA ligase	0.62	cinnamate and succinylbenzoate derivatives	0.28	AIS24844.1	31.9
4-coumarate	purified protein from heterologous host	NADH-consumption continuous spectrophotometric assay	Burckhardt et al. (2020) *Mol. Microbiol.* 113, 253-269	doi: 10.1111/mmi.14417	1	1	*Streptomyces lividans TK28*	Aryl-CoA ligase	0.61	cinnamate and succinylbenzoate derivatives	0.42	EFD67678.1	44.0
octanoate; hexanoate	substrate synthesis and culture feeding experiments, biosynthetic logic	HPLC-MS analysis of natural product intermediates, sequence-based prediction	Rachid et al. (2011) *ChemBioChem* . 12(6), 922-931	doi: 10.1002/cbic.201100024.	1	0	*Streptomyces cinnabarigriseus*	Fatty-acyl AMP ligase	0.6	C13 through C17	0.29	CBW54660.1	61.7
(1-(3-(carboxymethyl)-1,8,9-trihydroxyanthracen-2-yl)-1-hydroxyethyl)-L-serine	biosynthetic logic-based on gene knockouts and natural product isolation	biosynthetic logic sequence-based prediction	Panter et al. (2018) *Chemical Science* 9(21), 4898-4908.	doi: 10.1039/c8sc01325j	1	1	*Pyxidicoccus fallax*	Aryl-CoA ligase	0.58	cinnamate and succinylbenzoate derivatives	0.3	AXM42922.1	36.2
3-oxo-3,4-dihydro-2*H*-benzo[*b*][1,4]oxazine-2,5-dicarboxylic acid	purified protein from heterologous host, in vitro biochemical assay	fluorogenic pyrophosphate production assay	Shi et al. (2019) *Chemical Science.* 10(10), 3042-3048	doi: 10.1039/c8sc05670f	1	1	*Streptomyces sp.*	Aryl-CoA ligase	0.56	aryl and biaryl derivatives	0.34	QCE20613.1	48.4
3-methylmercaptopropionate; also active with short chain fatty acids up to C6	purified protein from heterologous host	NADH-consumption continuous spectrophotometric assay	Bullock et al. (2014) *J. Biol. Chem.* 196, 1275-1285	doi: 10.1016/j.chembiol.2019.02.004	0	0	*Ruegeria pomeroyi*	Aryl-CoA ligase	0.47	cinnamate and succinylbenzoate derivatives	0.23	WP_011047771.1	44.5
5-hydroxy-6-methoxy-1*H*-indole-2-carboxylic acid	biosynthetic logic based on gene knockouts, protein purification of other biosynthetic enzymes, and isolation of natural products and intermediates	biosynthetic logic sequence-based prediction	Huang et al. (2012) *JACS* 64(1),163–166	doi: 10.1038/ja.2010.150	1	1	*Streptomyces sp. TP-A2060*	Aryl-CoA ligase	0.46	cinnamate and succinylbenzoate derivatives	0.27	ADZ13551.1	35.3

benzoxazolinate	purified protein from heterologous host, in vitro biochemical assay	fluorogenic pyrophosphate production assay	Shi et al. (2019) *Chemical Science.* 10(10), 3042-3048	doi: 10.1039/c8sc05670f	1	1	*Streptomyces lividans SBT18*	Aryl-CoA ligase	0.46	aryl and biaryl derivatives	0.23	QCE20614.1	34.5
3-dimethylallyl-4-hydroxybenzoic acid	purified protein in heterologous host	radioactive ATP-pyrophosphate exchange assay	Steffensky et al. (2000) *JBC* . 275(28), 21754-21760	doi: 10.1074/jbc.M003066200	1	1	*Streptomyces niveus*	Aryl-CoA ligase	0.46	cinnamate and succinylbenzoate derivatives	0.22	AAF67505.1	35.3
1-carboxy-2-cyanopropane	biosynthetic logic-based on gene knockouts and natural product isolation	biosynthetic logic sequence-based prediction	Wang et al. (2017) *ACS Chem. Biol.* 12, 3067−3075	doi: 10.1021/cb500652n	0	0	*Streptomyces thioluteus*	Long chain acyl-CoA synthetase	0.45	C13 through C17	0.55	ATY72527.1	58.6
(2*S*,3*R*)-2-hydroxy-3-(1*H*-indol-3-yl)butanoic acid	purified protein from heterologous host, in vitro biochemical assays	HPLC/LC-MS analysis	Du et al. (2015) *PNAS.* 112(9), 2717-2722	doi: 10.1073/pnas.1419964112	1	1	*Streptomyces griseus*	Aryl-CoA ligase	0.45	aryl and biaryl derivatives	0.23	AJT38684.1	31.2
3-methylmercaptopropionate	purified protein from heterologous host, structure determination, steady-state kinetics	HPLC analysis	Shao et al. (2019) *Mol. Microbiol.* 111(4), 1057–1073.	doi: 10.1002/cbic.201100024.	0	0	*Ruegeria lacuscaerulensis ITI-1157*	Aryl-CoA ligase	0.44	cinnamate and succinylbenzoate derivatives	0.24	6IJB_A	44.1
caproate, valerate, heptanoate	purified protein from heterologous host	NADH-consumption continuous spectrophotometric assay	Burckhardt et al. (2020) *Mol. Microbiol.* 113, 253-269	doi: 10.1111/mmi.14418	0	0	*Streptomyces lividans TK29*	Aryl-CoA ligase	0.37	C13 through C17	0.25	EFD68037.1	54.0
3-hydroxypropionate	purified protein from heterologous host, structure determination, steady-state kinetics	NADH-consumption continuous spectrophotometric assay	Bernhardsgrütter et al. (2018) *Nat. Chem. Biol.* 14(12), 1127-1132	doi: 10.1021/ja5019942.	1	1	*Erthythrobacter* sp. *NAP1*	Short chain acyl-CoA synthetase	0.36	C2 through C5	0.24	6EQO_A	27.4
2-hydroxyisobutyric acid	purified protein from heterologous host, structure determination, steady-state kinetics	HPLC analysis	Zahn et al. (2019) *J Mol. Biol.* 431(15), 2747–2761	doi: 10.1016/j.jmb.2019.05.027	0	0	*Aquincola tertiaricarbonis*	Aryl-CoA ligase	0.36	cinnamate and succinylbenzoate derivatives	0.17	6HDW_A	34.0
myristate	biosynthetic logic based on gene knockouts, protein purification of other biosynthetic enzymes, and isolation of natural products and intermediates	biosynthetic logic sequence-based prediction	Duitman et al. (1999) *PNAS.* 96(23), 13294–13299	doi: 10.1073/pnas.96.23.13294	0	1	*Bacillus subtilis subsp. spizizenii ATCC 6633*	Aryl-CoA ligase	0.32	C13 through C17	0.15	AAF08801.1	92.9
valerate, butyrate, caproate	purified protein from heterologous host	NADH-consumption continuous spectrophotometric assay	Burckhardt et al. (2020) *Mol. Microbiol.* 113, 253-269	doi: 10.1111/mmi.14419	0	0	*Streptomyces lividans TK26*	Aryl-CoA ligase	0.31	C13 through C17	0.21	EFD64965.1	39.1
3-amino-2-hydroxy-4-methoxybenzoic acid	crude lysate assays, in vitro biochemical assays on partially purified protein	LC-MS-based analysis	Waldman et al. (2018) *J. Org Chem.* 83(14), 7539-46	doi: 10.1021/acs.joc.8b00367	1	0	*Streptomyces cremeus*	Aryl-CoA ligase	0.29	C13 through C17	0.18	ALA99210.1	43.3
1-acetyl-9*H*-pyrido[3,4-*b*]indole-3-carboxylic acid	purified protein from heterologous host, structure determination	HPLC analysis with substrate panel	Petchey et al. (2018) *Angew. Chem. Int. Ed.*	doi: 10.1002/anie.201804592	1	1	*Marinactinospora thermotolerans*	Aryl-CoA ligase	0.26	cinnamate and succinylbenzoate derivatives	0.16	6H1B_A	33.2
				Average	0.83	0.73						**Average**	52.5