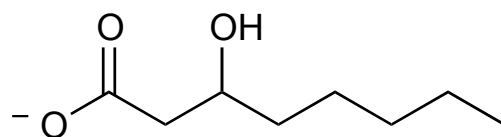
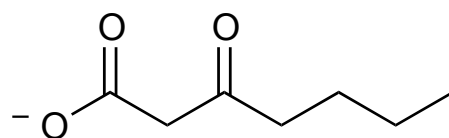


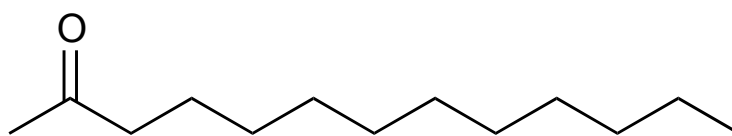
methyl-3-hydroxydodecanoate (M3D)



3-hydroxyoctanoate (3HO)

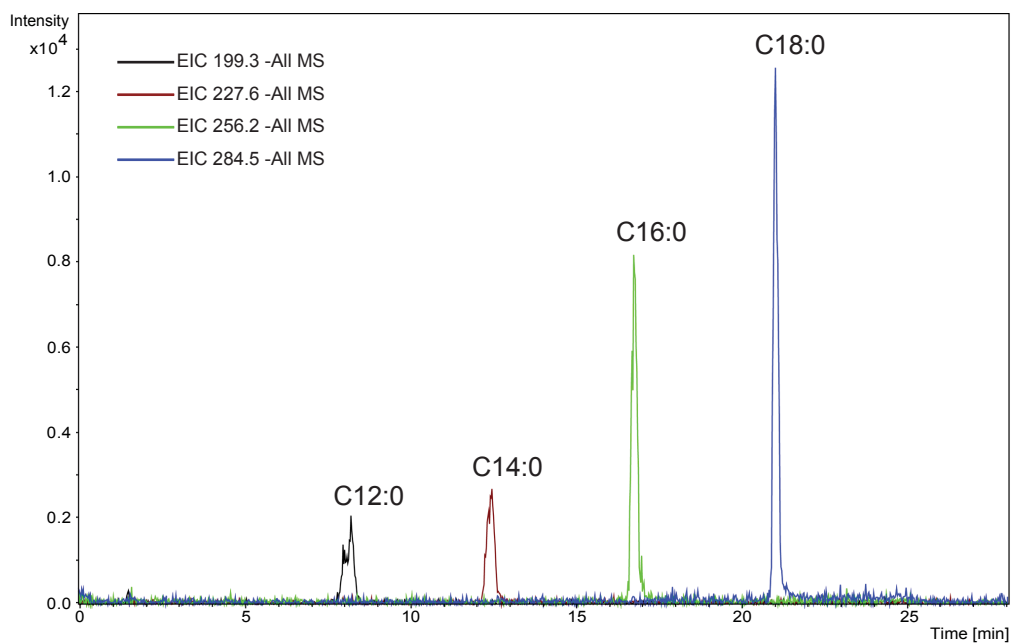


3-ketoheptanoate (3KH)



2-tridecanone (2TD)

**Supplemental Figure 1.** Small molecule ligands used as substrate mimics, with the exception of 2-tridecanone which is a MKS1 natural product. 3-letter code serves as the abbreviation for the respective ligand throughout the text and in the structures deposited to the PDB.

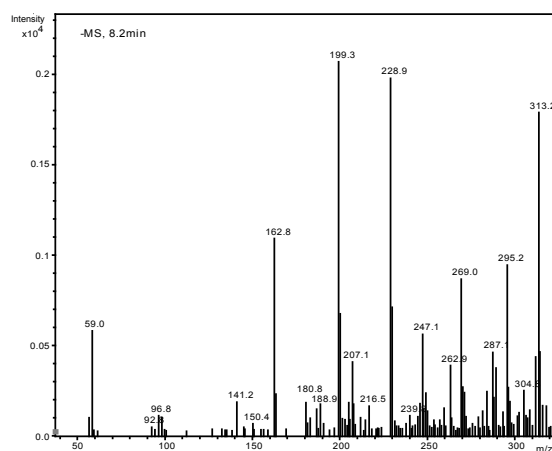


**Supplemental Figure 2.** MKS1 active site serendipitously binds fatty acid during heterologous production in *E. coli*. (A) LC trace showing C12, C14, C16 and C18 fatty acids eluted from purified MKS1. (B) Mass spectra of fatty acid peaks shown in (A).

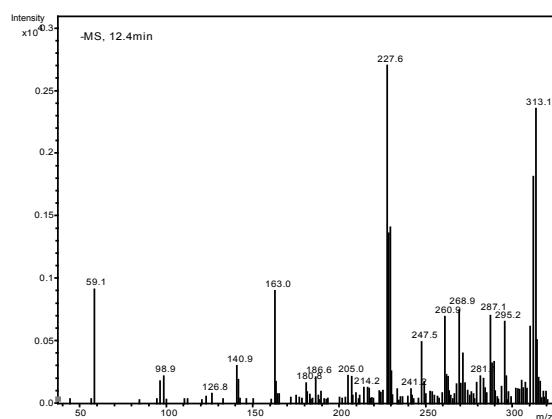
Supplemental Figure 2B

Mass spectra at indicated retention time corresponding to Suppl Fig 1a:

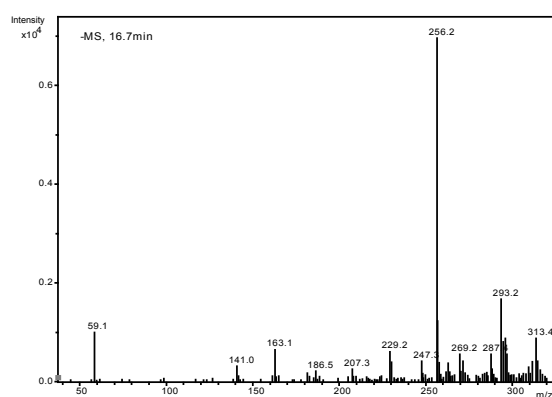
(m/z) of Fatty Acid Standards:



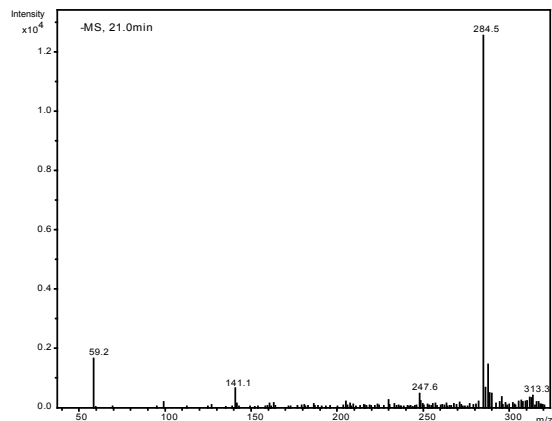
lauric acid (C12:0)  
(m/z)=198.7 ([M -H]-)



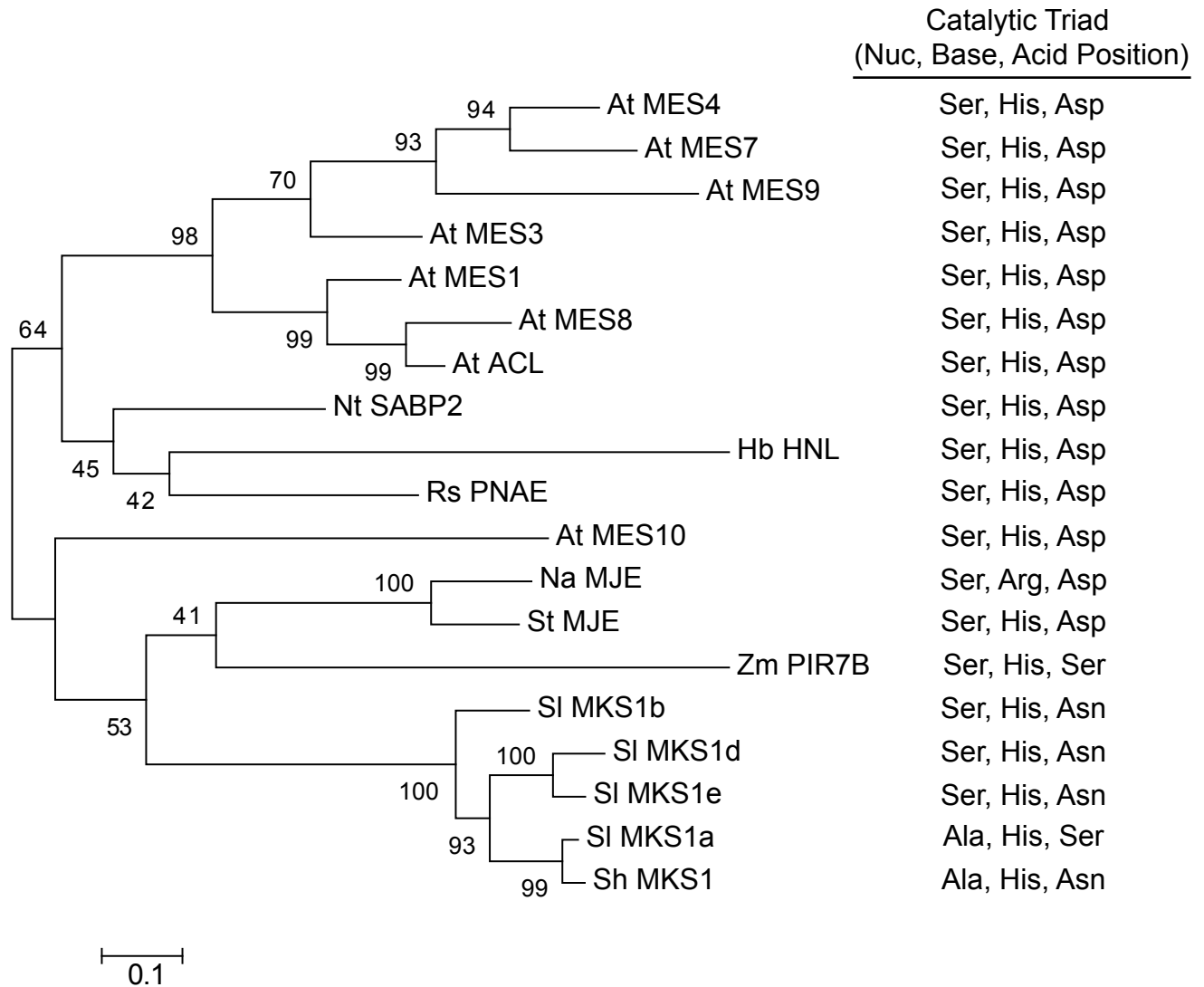
myristic acid (C14:0)  
(m/z)=226.9 ([M -H]-)



palmitic acid (C16:0)  
(m/z)=254.9 ([M -H]-)



stearic acid (C18:0)  
(m/z)=282.9 ([M -H]-)



**Supplemental Figure 3.** Maximum likelihood phylogenetic tree including *S. habrochaites* MKS1 and other closely related proteins. The catalytic triad of each is listed to the right in the order of nucleophile, base and acid. MEGA 5 software was used in tree construction. Sequences included in the phylogeny, with corresponding GI number in parenthesis, are as follows: At MES4 (gi:15227859), At MES7 (gi:15227851), At MES9 (gi:15235445), At MES3 (gi:15227865), At MES1 (gi:15227867), At MES8 (gi:15227861), At ACL (gi:15227863), Nt SABP2 (gi:40549303), Hb HNL (gi:1708278), Rs PNAE (gi:6651393), At MES10 (gi:79439484), Na MJE (gi:164507175), St MJE (gi:56392765), Zm PIR7B (gi:226506656), SI MKS1b (gi:300836821), SI MKS1d (gi:300836824), SI MKS1e (gi:300836826), SI MKS1a (gi:300836819), Sh MKS1 (gi:300836815).

**Supplemental Table 1.** Data Collection and refinement statistics for MKS1 wild-type and variants, with and without analogs.<sup>a</sup>

Complex/Substitution PDB code	Apo 3STT	M3D 3STU	3HO 3STV
Data Collection			
Space group	P2 <sub>1</sub>	P2 <sub>1</sub>	P2 <sub>1</sub>
Cell dimensions:			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	48.2, 105.5, 59.8	48.2, 105.7, 59.9	48.0, 92.4, 60.6
$\alpha$ , $\beta$ , $\gamma$ (°)	90.0, 96.5, 90.0	90.0, 96.5, 90.0	90.0, 97.6, 90.0
# of molec./a.u.	2	2	2
Resolution (Å)	47.9-2.24 (2.28-2.24)	39.5-1.93 (1.98-1.93)	19.6-2.20 (2.30-2.20)
<i>R</i> <sub>sym</sub>	0.041 (0.162)	0.063 (0.499)	0.140 (0.517)
<i>I</i> / $\sigma$ <i>I</i>	33.9 (9.9)	24.7 (3.9)	8.9 (2.5)
Completeness (%)	96.0 (87.4)	97.8 (98.7)	99.1 (99.1)
Redundancy	3.7 (3.5)	3.7 (3.5)	3.1 (3.1)
Refinement			
Resolution (Å)	2.24 (2.28-2.25)	1.93 (1.98-1.93)	2.20 (2.23-2.20)
No. reflections	27180	43407	26450
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.181/0.221 (0.199/0.282)	0.215/0.242 (0.353/0.416)	0.202/0.236 (0.287/0.350)
No. atoms			
Protein	3969	3980	3975
Ligand/ion	36 <sup>b</sup>	28 <sup>c</sup>	12 <sup>d</sup>
Water	260	246	218
<i>B</i> -factors (Å <sup>2</sup> )			
Protein	25.6	28.7	25.5
Ligand/ion	46.9	57.4	66.4
Water	28.7	32.8	25.1
Rmsd			
Bond lengths (Å)	0.006	0.008	0.006
Bond angles (°)	1.40	1.04	1.31
Ramachandran			
% favored	85.8	88.7	86.3
% allowed	13.8	10.4	12.8
% disallowed <sup>d</sup>	0.4	0.9	0.9

Complex/Substitution PDB Code	2TD 3STW	3KH-H243A 3STX	T18A 3STY
Data collection			
Space group	P2 <sub>1</sub>	P2 <sub>1</sub>	P2 <sub>1</sub>
Cell dimensions:			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	48.5, 106.9, 60.1	48.2, 94.4, 60.1	48.3, 106.2, 59.8
$\alpha$ , $\beta$ , $\gamma$ (°)	90.0, 96.5, 90.0	90.0, 97.6, 90.0	90.0, 96.6, 90.0
# of molec./a.u.	2	2	2
Resolution (Å)	20.0-2.31 (2.50-2.31)	19.6-2.37 (2.45-2.37)	37.1-1.70 (1.76-1.70)
<i>R</i> <sub>sym</sub> or <i>R</i> <sub>merge</sub>	0.110 (0.315)	0.109 (0.497)	0.059 (0.499)
<i>I</i> / $\sigma$ <i>I</i>	13.5 (5.8)	12.1 (3.1)	23.8 (2.0)
Completeness (%)	99.6 (99.7)	88.4 (97.3)	98.0 (96.0)
Redundancy	4.6 (4.7)	4.1 (4.3)	3.4 (3.3)
Refinement			
Resolution (Å)	2.31 (2.34-2.31)	2.37 (2.43-2.37)	1.70 (1.71-1.70)
No. reflections	26600	18238	64470
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.198/0.245 (0.219/0.246)	0.187/0.262 (0.251/0.315)	0.196/0.213 (0.317/0.310)
No. atoms			
Protein	3969	3949	3965
Ligand/ion	28	20	24 <sup>b</sup>
Water	222	151	376
<i>B</i> -factors			
Protein	24.8	30.5	23.2
Ligand/ion	41.2	46.6	39.2
Water	24.6	24.1	31.8
Rmsd			
Bond lengths (Å)	0.007	0.009	0.005
Bond angles (°)	1.35	1.36	1.27
Ramachandran			
% favored	85.8	87.2	86.3
% allowed	13.8	12.1	12.8
% disallowed <sup>e</sup>	0.4	0.7	0.9

<sup>a</sup>A single crystal was used for each structure determination. Values shown in parentheses refer to the highest resolution shell.

<sup>b</sup>Ligand present due to serendipitous binding of fatty acid from exogenous growth in *E. coli*.

<sup>c</sup>M3D was observed in monomer B only. Monomer A contained a fatty acid as in MKS1 apo.

<sup>d</sup>3HO was observed in monomer A, while monomer B contained a strong-density peak modeled as Br.

<sup>e</sup>Ala 87 corresponds to the traditional nucleophilic residue of the  $\alpha/\beta$  hydrolases and this residue has previously been reported to have a disallowed conformation in other  $\alpha/\beta$  hydrolase structures (Wagner et al., 1996; Forouhar et al., 2005).

## Supplemental Methods

**Synthesis of 3-keto acids for decarboxylase assays.** The 3-keto acids 3-ketoheptanoate (C7) and 3-ketomyristate (C14) were prepared as previously described (Yu et al., 2010) from methyl 3-oxo esters. Briefly, base-mediated hydrolysis at room temperature, organic extraction to remove the alcohol product, acidification of the resultant aqueous layer and extraction again with methylene chloride was used to obtain the free acids. The resultant organic layers were pooled, washed with saturated aqueous NaCl, dried over sodium sulfate and filtered. Methylene chloride was removed under reduced pressure yielding opaque white powdery products (1 mmoles of 3-ketoheptanoate from 1 mmoles of methyl 3-oxoheptanoate, 99.9% and 1 mmoles of 3-ketomyristate from 1 mmoles of methyl 3-oxotetradecanoate, 95.2%).

TLC of 3-ketoheptanoate (ethyl acetate:hexanes, 1:3 v/v):  $R_f = 0.05$ ; stains with ethanolic  $\text{KMnO}_4$  and heat;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  10.96 (br, 1H), 3.48 (s, 2H), 2.53 (t,  $J = 7.34$  Hz, 2H), 1.54 (m,  $J = 7.34$  Hz, 2H), 1.29 (m,  $J = 7.34$  Hz, 2H), 0.86 (t,  $J = 7.34$  Hz, 3H);  $^{13}\text{C}$  NMR (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  204.1, 172.2, 48.2, 43.0, 25.5, 22.1, 13.8; MS (ESI):  $[\text{M}]^-$  144.4. A working concentration of 20 mM 3-ketoheptanoate in 1,3-bis(tris(hydroxymethyl)methylamino) propane (20 mM, pH 7.0) was made for assays.

TLC of 3-ketomyristate (ethyl acetate:hexanes, 1:3 v/v):  $R_f = 0.12$  stains with ethanolic  $\text{KMnO}_4$  and heat;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  11.96 (br, 1H), 3.523 (s, 2H), 2.559 (m, 2H), 1.618 (m, 2H), 1.257 (m, 16H), 0.880 (t,  $J = 7.34$  Hz, 3H).  $^{13}\text{C}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  206.0, 171.8, 48.5, 43.8, 29.8, 29.6, 29.5, 29.4, 29.3, 29.2, 29.0, 23.8, 22.7, 14.1 ; MS (ESI):  $[\text{M}]^-$  242.4. A working concentration of 20 mM 3-ketomyristate in 1,3-bis(tris(hydroxymethyl)methylamino) propane (10 mM, pH 8.0) was made for assays.

**Fatty acid analysis.** MKS1, purified as in Methods, was analyzed for the presence of fatty acids as follows. Ice-cold HPLC grade ethanol (600  $\mu$ l) was added to purified protein (150  $\mu$ l at 10 mg/ml). The samples were vortexed and incubated at  $-20^{\circ}\text{C}$  for 3 days, following which samples were centrifuged and the supernatant containing the fatty acids was separated in a new glass vial. The solvent was evaporated at  $25^{\circ}\text{C}$  and dried extracts were resuspended in 200  $\mu$ l propanol. The extracts were analyzed by liquid chromatography on an Agilent 1100 HPLC using a Gemini reverse-phase C18 column (4.6  $\mu\text{m}$  x 150mm) at a flow rate of 0.8ml/min, coupled to an electrospray ionization XCT ion trap mass spectrometer in negative-ion mode under the addition of 20mM ammonium acetate (100  $\mu$ l/min). A linear gradient of acetonitrile (30-100%) in 25mM bicarbonate at pH 8 was used. The negative ion-ESI mass spectrum of fatty acid standards were as follows: lauric acid (C12:0) ( $m/z$ )=198.7 ([M  $-$ H] $-$ ), myristic acid (C14:0) ( $m/z$ )=226.9 ([M  $-$ H] $-$ ), palmitic acid (C16:0) ( $m/z$ )=254.9 ([M  $-$ H] $-$ ), and stearic acid (C18:0) ( $m/z$ )=282.9 ([M  $-$ H] $-$ ).

**Phylogenetic analysis.** The MEGA 5 software (Tamura et al., 2011) was used to construct a maximum likelihood tree. Sequences included in the phylogeny, with corresponding GI number in parenthesis, are as follows: At MES4 (gi:15227859), At MES7 (gi:15227851), At MES9 (gi:15235445), At MES3 (gi:15227865), At MES1 (gi:15227867), At MES8 (gi:15227861), At ACL (gi:15227863), Nt SABP2 (gi:40549303), Hb HNL (gi:1708278), Rs PNAE (gi:6651393), At MES10 (gi:79439484), Na MJE (gi:164507175), St MJE (gi:56392765), Zm PIR7B (gi:226506656), SI MKS1b (gi:300836821), SI MKS1d (gi:300836824), SI MKS1e (gi:300836826), SI MKS1a (gi:300836819), Sh MKS1 (gi:300836815).



### **Supplemental References**

- Forouhar, F., Yang, Y., Kumar, D., Chen, Y., Fridman, E., Park, S.W., Chiang, Y., Acton, T.B., Montelione, G.T., Pichersky, E., Klessig, D.F., and Tong, L.** (2005). Structural and biochemical studies identify tobacco SABP2 as a methyl salicylate esterase and implicate it in plant innate immunity. *Proc Natl Acad Sci U S A* **102**, 1773-1778.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S.** (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731-2739.
- Wagner, U.G., Schall, M., Hasslacher, M., Hayn, M., Griengl, H., Schwab, H., and Kratky, C.** (1996). Crystallization and preliminary X-ray diffraction studies of a hydroxynitrile lyase from *Hevea brasiliensis*. *Acta Crystallogr D Biol Crystallogr* **52**, 591-593.
- Yu, G., Nguyen, T.T., Guo, Y., Schauvinhold, I., Auldridge, M.E., Bhuiyan, N., Ben-Israel, I., Iijima, Y., Fridman, E., Noel, J.P., and Pichersky, E.** (2010). Enzymatic functions of wild tomato methylketone synthases 1 and 2. *Plant Physiol* **154**, 67-77.