ACS Synthetic Biology
Research Article
Cite This: ACS Synth. Biol. XXXX, XXX, XXX−XXX
pubs.acs.org/synthbio

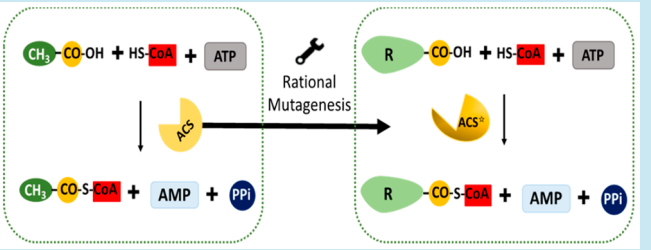
Altering the Substrate Specificity of Acetyl-CoA Synthetase by Rational Mutagenesis of the Carboxylate Binding Pocket

Naazneen Sofeo,[†,§,∥] Jason H. Hart,[†,§] Brandon Butler,[§] David J. Oliver,[‡,§] Marna D. Yandeau-Nelson,[†,‡,§,∥] and Basil J. Nikolau*[,†,§,∥]

[†]Roy J. Carver Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011, United States
[‡]Department of Genetics, Development, and Cell Biology, Iowa State University, Ames, Iowa 50011, United States
[§]Engineering Research Center for Biorenewable Chemicals, Iowa State University, Ames, Iowa 50011, United States
[∥]Center for Metabolic Biology, Iowa State University, Ames, Iowa 50011, United States

Supporting Information

ABSTRACT: Acetyl-CoA synthetase (ACS) is a member of a large superfamily of enzymes that display diverse substrate specificities, with a common mechanism of catalyzing the formation of a thioester bond between Coenzyme A and a carboxylic acid, while hydrolyzing ATP to AMP and pyrophosphate. As an activated form of acetate, acetyl-CoA is a key metabolic intermediate that links many metabolic processes, including the TCA cycle, amino acid metabolism, fatty acid metabolism and biosynthetic processes that generate many polyketides and some terpenes. We explored the structural basis of the specificity of ACS for only activating acetate, whereas other members of this superfamily utilize a broad range of other carboxylate substrates. By computationally modeling the structure of the *Arabidopsis* ACS and the *Pseudomonas chlororaphis* isobutyryl-CoA synthetase using the experimentally determined tertiary structures of homologous ACS enzymes as templates, we identified residues that potentially comprise the carboxylate binding pocket. These predictions were systematically tested by mutagenesis of four specific residues. The resulting rationally redesigned carboxylate binding pocket modified the size and chemo-physical properties of the carboxylate binding pocket. This redesign successfully switched a highly specific enzyme from using only acetate, to be equally specific for using longer linear (up to hexanoate) or branched chain (methylvalerate) carboxylate substrates. The significance of this achievement is that it sets a precedent for understanding the structure−function relationship of an enzyme without the need for an experimentally determined tertiary structure of that target enzyme, and rationally generates new biocatalysts for metabolic engineering of a broad range of metabolic processes.



KEYWORDS: *acetyl-CoA synthetase, targeted mutagenesis, enzyme engineering, substrate specificity, homologous enzymes, Arabidopsis*

The promise of synthetic biology is based on the premise that the systematic application of engineering principles to redesign individual molecular entities can outcompete natural evolutionary-based optimized design of complex biological processes. Furthermore, these individual genetically encoded redesigned entities can be fabricated into assemblages of novel systems (chassis), which do not already exist in the natural world.[1,2] Enzyme catalysts that provide specificity and enhanced rates of chemical conversions are key biological entities that are targets of such human-driven optimization. A limiting factor in such optimization redesigns is the fact that enzymatic functionality is based on tertiary structural design principles that cannot *a priori* be deduced from primary structural data, which are the most readily available data sets concerning structures of enzymes. This study illustrates a research path for redesigning a specific enzyme whose tertiary structure is unknown, but could be computationally modeled based on tertiary structures of functionally related enzymes.[3,4] The systematic targeted engineering of an AMP-generating acetyl-CoA synthetase from *Arabidopsis thaliana* (atACS) led to the redesign of this enzyme to generate new synthetic enzymes that display distinct catalytic capabilities not present in any of the enzymes that served as templates for the redesign.

The atACS enzyme is a member of the AMP-forming acyl-CoA synthetase family, which is part of the adenylate forming superfamily of enzymes.[5−7] The initial half reaction catalyzed by this superfamily of enzymes activates a carboxylate substrate by converting the hydroxyl leaving group of the substrate to an adenosine monophosphate (adenylate intermediate) using the hydrolysis of PP_i to drive the reaction in the forward direction[7−9] (Figure 1). In the second half reaction, the

Received: January 10, 2019

DOI: 10.1021/acssynbio.9b00008
ACS Synth. Biol. XXXX, XXX, XXX−XXX

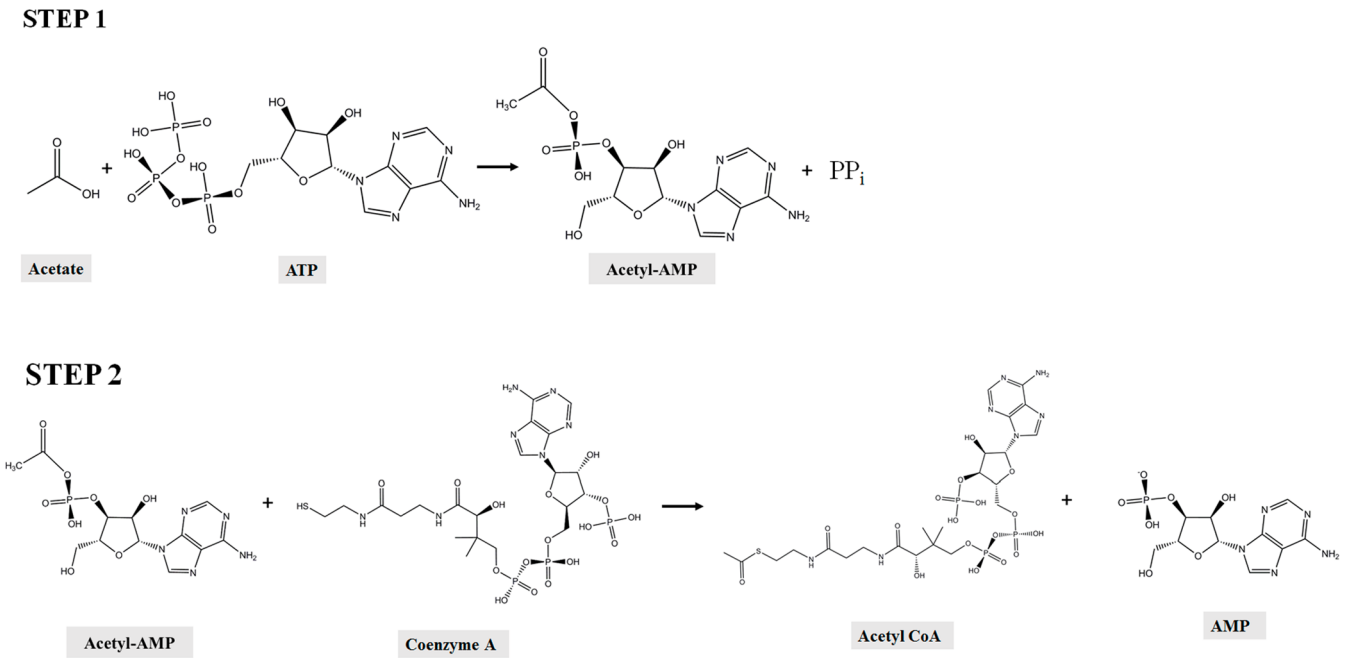


Figure 1. Enzymatic mechanism of AMP-forming acetyl-CoA synthetase. The reaction occurs in two steps, where an acetyl-adenylate intermediate is formed in the first half-reaction. This intermediate then reacts with Coenzyme A to form acetyl-CoA and release AMP.

Table 1. Sequence Identity (%) and RMSD (Å) Values of All Atom Structural Alignment of 3D Structures of atACS and Several Acyl-CoA Synthetases[a]

	1PG4[c]		1RY2[c]		pcICS[b]		atACS[b]		3NYR[c]		1MD9[c]		mtACS[b]		atMCS[b]		sePCS[b]	
	%	Å	%	Å	%	Å	%	Å	%	Å	%	Å	%	Å	%	Å	%	Å
1PG4[c]			45.5	10.5	30	13.7	53.9	1.8	23.4	6.7	19.9	13.1	48	1.5	22.9	8.6	37	2.4
1RY2[c]					30.5	8.8	45.9	10.5	21	16.9	18	6.6	43	10.6	21.8	11.7	37.3	10.8
pcICS[b]							29.4	14	24.5	14.7	19.0	10.4	31.4	14	21.6	12.6	26.0	13.5
atACS[b]									22	8.3	21	13.7	46.4	3.7	22	6.8	35.8	3.2
3NYR[c]											27.6	12.5	23.3	6.5	39.8	4.6	21	9.2
1MD9[c]													22	13.1	23.9	11.8	21.5	12.4
mtACS[b]															22.7	11.6	37.3	2.8
atMCS[b]																	18	10.2
sePCS[b]																		

[a]**1PG4**: *Salmonella enterica* acetyl-CoA synthetase; **1RY2**: *Saccharomyces cerevisiae* acetyl-CoA synthetase 1; **pcICS:** *Pseudomonas chlororaphis* isobutyryl-CoA synthetase; **atACS**: *Arabidopsis thaliana* acetyl-CoA synthetase; **3NYR**: *Streptomyces coelicolor* malonyl-CoA ligase; **1MD9**: *Bacillus subtilis* adenylation domain of NRPS to activate 2,3-dihydroxybenzoate; **mtACS**: *Methanobacter thermoautotrophicus* acetyl-CoA synthetase; **atMCS**: *Arabidopsis thaliana* malonyl-CoA synthetase; **sePCS**: *Salmonella enterica* propionyl-CoA synthetase. [b]Indicates proteins whose tertiary structures were computationally generated. [c]Indicates proteins whose tertiary structures were experimentally determined.

adenylate intermediate reacts with the thiol group of the phosphopantothenate group of Coenzyme A, generating an acyl-CoA product, which can be considered as the "activated" form of the carboxylate substrate.[7,10] Different acyl-CoA synthetases in this family act upon different carboxylate substrates, and these include for example, acetate, propionate, malonate and chloro-benzoate,[4,11−15] which vary in the nature of the alkyl group of the carboxylate substrate.

Thioester bonds, as occurs in acyl-CoA, are crucial to a large number of anabolic and catabolic processes. These include fatty acid and polyketide metabolism, assembly of complex lipids, amino acid metabolism, assembly of mevalonate for the biosynthesis of isoprenoids, and porphyrin biosynthesis.[10,16] Chemically the formation of the thioester bond facilitates the subsequent reactions of the acyl moiety because of the high energy content of the bond, and because the thiol group is a better leaving group than the hydroxyl group of the initial carboxylate substrate for the nucleophilic reactions, that lead to the formation of different carbonyl-containing products (*e.g.*, esters, amides and ketones). Indeed, because of the high energy content of the thioester bond, and its chemical flexibility, it has been theorized that it could have been a precursor to life, preceding the role of ATP as the common energy carrier in a so-called "thioester world".[17]

The most prevalent thioester-containing intermediate of metabolism is acetyl-CoA, and biological systems utilize multiple enzymatic mechanisms to generate this central metabolite. Most prevalent of these reactions is likely the oxidative decarboxylation of pyruvate, catalyzed by the pyruvate dehydrogenase complex,[18] and the retro-Claisen cleavage of citrate, catalyzed by ATP-citrate lyase.[19] However, ACS is part of a large family of enzymes that is widely prevalent in biological systems, and in some cases, such as in plants, the physiological relevance of this enzyme to

DOI: 10.1021/acssynbio.9b00008
ACS Synth. Biol. XXXX, XXX, XXX−XXX

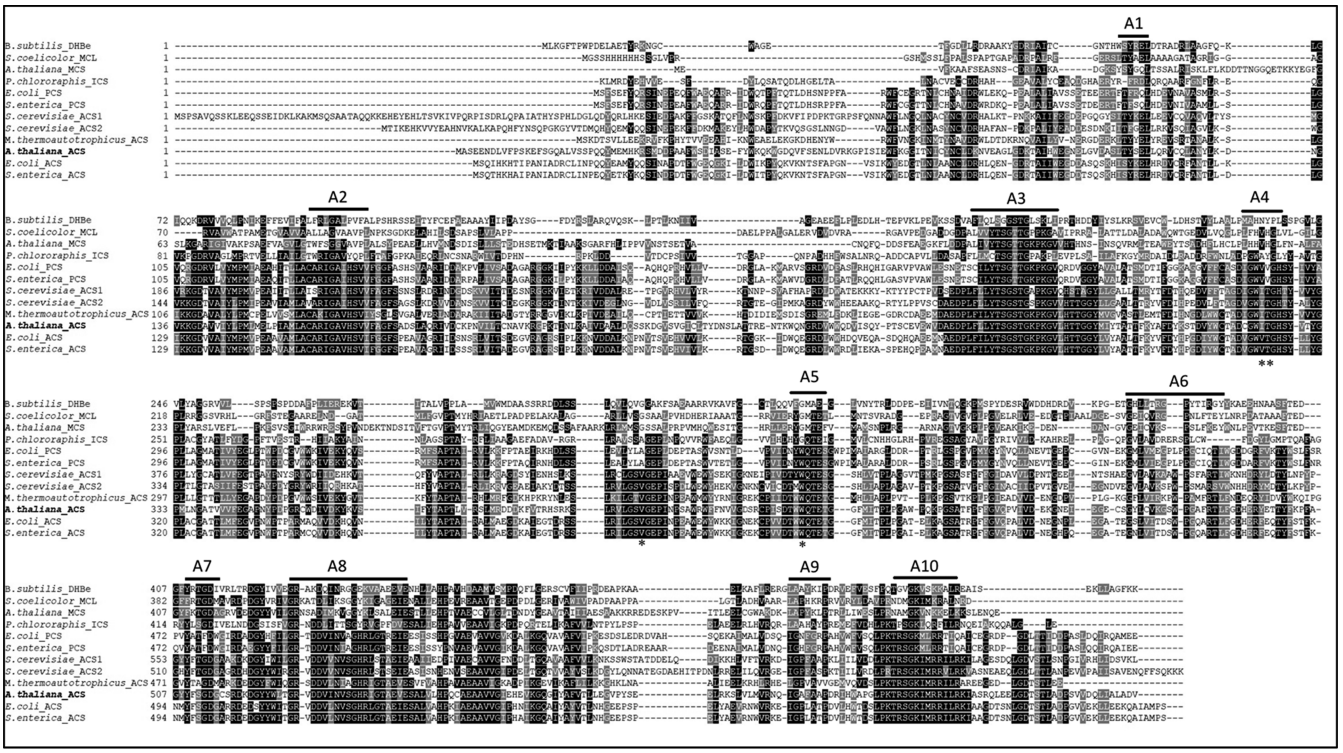


Figure 2. Multiple sequence alignment of the amino acid sequences of acyl-CoA synthetases that utilize acetate (ACS), propionate (PCS), isobutyrate (ICS), malonate (MCS or MCL) and 2,3-dihyroxybenzoate (DHBe) as the preferred substrate. Alignment was performed using Clustal Omega. Asterisks indicate the residues proposed to be forming the carboxylate binding pocket. White letters on a black background indicate identity, white letters on a gray background indicate similarity, and black letters on a white background indicate no conservation. Ten conserved motifs of the AMP-forming family of acyl-CoA synthetases are indicated (A1-A10). UniProt/GenBank Accession numbers are *S. enterica* ACS (A0A0F7JF17), *E. coli* ACS (P27550), *A. thaliana* ACS (B9DGD6), *M. thermoautotrophicus* ACS (DQ274062.1), *S. cerevisiae* ACS1 (Q01574), *S. cerevisiae* ACS2 (P52910), *S. enterica* PCS (A0A0F7J6L9), *E. coli* PCS (A0A222QH41), *P. chlororaphis* ICS (WP_069076265), *A. thaliana* MCS (Q8H151), *S. coelicolor* MCL (Q9L0A2) and *B. subtilis* DHBe (P40871).

metabolism is not clear.[18,20−23] Due to this potential metabolic redundancy, engineering ACS to utilize different carboxylate substrates has considerable potential in impacting the end-products of metabolism, particularly in plant systems. Acetyl-CoA is a key precursor of fatty acid biosynthesis and many polyketide biosynthetic processes, and in both complex processes it is the precursor of the malonyl-CoA substrate that is used to elongate the fatty acids or polyketides (*e.g.*, chalcone synthase).[24−26] Hence, engineering ACS to utilize novel carboxylate substrates would generate alternative acyl-CoA substrates for fatty acid or polyketide biosynthesis would provide the means to explore the enzymological flexibilities of these biosynthetic machineries and build a technology platform to produce a diversity of chemical products with wide ranging applications.[27] As a prerequisite to such redesign of biological chassis, in this study we have been successful in not only broadening the substrate specificity of a eukaryotic ACS, but shifting its substrate specificity, using a systematic targeted mutagenesis strategy.

RESULTS AND DISCUSSION

Identification of the Residues Forming the Putative Carboxylate Binding Pocket of atACS. Alignment of sequences of multiple acyl-CoA synthetases, which utilize different carboxylate substrates, reveal that atACS shares higher sequence identity (>40%) with other acyl-CoA synthetases that utilize acetate as the preferred substrate, rather than those that utilize propionate, isobutyrate or malonate (Table 1). This conservation in sequence is focused in ten motifs (A1 to A10) (Figure 2), three of which (motifs A3, A4 and A5) have been associated with (a) the carboxylate substrate binding pocket of acyl-CoA synthetases; (b) the amino acid binding pocket of nonribosomal peptide synthases; and (c) in interacting with the pyrophosphate product of these enzymes.[7,10,28] We explored the basis for the high substrate specificity of atACS for acetate, by comparing and contrasting the structure of this enzyme with that of the *Pseudomonas chlororaphis* isobutyryl-CoA synthetase (pcICS), which is equally specific, but for isobutyrate. Higher order structural insights of atACS and pcICS were obtained by using 3D structural models generated with I-TASSER,[29−31] using as a guide the experimentally determined structures of the acetate-utilizing homologous ACS from *Salmonella enterica* (PDB: 1PG4)[32] and *Saccharomyces cerevisiae* (PDB: 1RY2).[33]

Despite the fact that atACS and pcICS differ in primary structure (sharing only 30% sequence identity), both computationally generated 3D models of atACS and pcICS share a high degree of tertiary structural alignment with each other and with both *S. enterica* ACS and *S. cerevisiae* ACS. The 3D structural alignment of atACS and pcICS structures has an all-atom root mean square deviation (RMSD) of 14 Å for 3159 atoms; a numerical evaluation of the structural alignment between two tertiary structures. Comparison of the computationally generated structures of atACS and pcICS with the experimentally determined structures of *S. enterica* ACS (1PG4) and *S. cerevisiae* ACS (1RY2) indicates a higher

DOI: 10.1021/acssynbio.9b00008
ACS Synth. Biol. XXXX, XXX, XXX−XXX

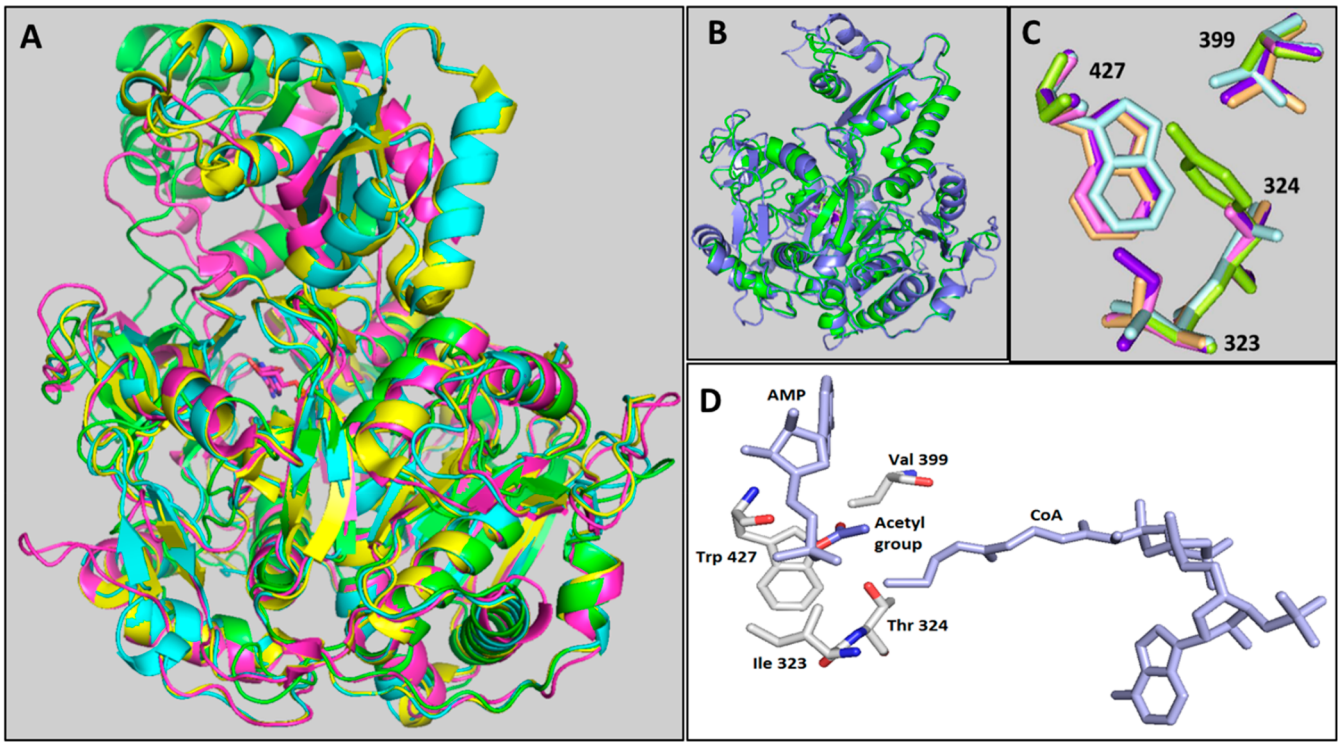


Figure 3. (A) Comparisons of the 3D structural homology models of atACS (cyan ribbon diagram) and pcICS (green ribbon diagram) predicted using I-TASSER, and the experimentally determined structures of scACS (1RY2) (pink ribbon diagram) and seACS (1PG4) (yellow ribbon diagram). These structures were aligned using PyMOL. (B) Comparisons of the 3D structural homology models of atACS (purple ribbon diagram) and pcICS (green ribbon diagram) computationally predicted using I-TASSER and aligned using PyMOL. (C) Predicted position of the residues that constitute the putative carboxylate binding pocket of atACS (purple), ecACS (blue), ecPCS (pink), and pcICS (green). These alignments were guided with the experimentally determined tertiary structure of the seACS (beige). Numbers represent the position of the homologous residue in the atACS protein. (D) Putative binding pocket residues of atACS from the homology structure predicted using I-TASSER, aligned with the derived structures of Coenzyme A, AMP and the acetyl group from the experimentally determined structure of these substrates in 1RY2 and 1PG4.

degree of 3D structural alignment among the enzymes that share the common substrate, acetate. Specifically, the all-atom RMSD comparison between atACS and *S. enterica* ACS or *S. cerevisiae* ACS is 1.8 and 10.5 Å, respectively, whereas the homologous comparisons between pcICS and *S. enterica* ACS or *S. cerevisiae* ACS is 13.7 and 8.8 Å, respectively (Table 1 and Figure 3A,B). Therefore, the atACS model has smaller RMSD values when compared to enzymes that act on smaller carboxylate substrates (*e.g.*, acetate, malonate and propionate) as compared to enzymes that act on larger carboxylate substrates (*e.g.*, isobutyrate or benzoate) (Table 1). These insights are consistent with prior studies of this class of enzymes, which indicate that substrate specificity cannot be assigned solely based upon primary sequence similarity but can more accurately be deduced based upon higher order structural templates.[34] Therefore, these 3D structural comparisons provide more valuable and accurate insights into the design principles of the substrate binding pocket that can be experimentally validated.

These computational higher structural order comparisons identified four residues that may comprise the carboxylate binding pocket of atACS (*i.e.*, residues Ile^{323}, Thr^{324}, Val^{399}, and Trp^{427}), and of pcICS (*i.e.*, residues Ala^{278}, Tyr^{279}, Ala^{313}, and Gly^{338}) (Figure 3C,D). Confidence that these identified residues are important for enzymatic function is further enhanced by the fact that the former set are conserved in all acyl-CoA synthetases that utilize acetate as the preferred carboxylate substrate (Figure 2). Furthermore, these residues were previously identified by the comparison of ACS from *Methanobacter thermautotrophicus* and *Archaeoglobus fulgidus*,[35,36] and they also coincide with the 15 residues previously predicted by the *in silico* analysis of this superfamily of enzymes.[34,37]

The accuracy of our predictions were evaluated by singularly and in combination mutating the atACS residues (*i.e.*, Ile^{323}, Thr^{324}, Val^{399}, and Trp^{427}) to the corresponding residues that comprise the binding pocket of pcICS (Figure 2). The expectation is that these mutations would affect catalysis and/or alter the carboxylate substrate specificity of the atACS enzyme. The kinetic parameters (V_{max}, K_m, and k_{cat}) of the resulting mutant proteins were determined with a range of different carboxylate substrates, including a series of straight chain carboxylates (acetate to octanoate) and branched chain carboxylates (isobutyrate, 3-methylvalerate, 4-methylvalerate, and 4-methylhexanoate), and these data were compared to the parallel data generated with the wild-type atACS.

Mutations That Identify an Important Residue for Catalysis. The atACS enzyme follows traditional Michaelis–Menten kinetics, with a strong preference for acetate as the carboxylate substrate. Specifically, the catalytic efficiency (k_{cat}/K_m) is approximately 55-fold higher with acetate than propionate, and there is no detectable activity with longer chain carboxylates (Table 2). The singular replacement of the Thr^{324} residue with the Tyr residue that occurs at this position in the pcICS structure inactivates atACS with all tested substrates. Furthermore, each multisite mutant that contained

DOI: 10.1021/acssynbio.9b00008
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Table 2. Kinetic Parameters of atACS Variants[a]

atACS variant	carboxylate substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})
wild-type	acetate	0.27 ± 0.01	2.41 ± 0.03	8.9 ± 0.5
	propionate	4.0 ± 1.2	0.66 ± 0.07	0.16 ± 0.05
$Val^{399}Ala$	acetate	3.1 ± 0.2	3.04 ± 0.05	0.99 ± 0.07
	propionate	3.1 ± 0.4	2.65 ± 0.07	0.85 ± 0.10
$Trp^{427}Gly$	acetate	10.8 ± 1.8	14.3 ± 0.9	1.32 ± 0.24
	propionate	37.5 ± 5.0	6.2 ± 0.5	0.16 ± 0.03
	butyrate	2.9 ± 0.6	9.4 ± 0.5	3.3 ± 0.7
	valerate	6.5 ± 0.7	13.8 ± 0.4	2.13 ± 0.25
	hexanoate	12.6 ± 1.4	13.0 ± 0.5	1.03 ± 0.12
	heptanoate	10.0 ± 1.8	11.0 ± 0.6	1.10 ± 0.20
	octanoate	48 ± 12	5.8 ± 0.8	0.12 ± 0.03
$Val^{399}Ala$, $Trp^{427}Gly$	acetate	44 ± 6	5.07 ± 0.31	0.11 ± 0.02
	propionate	31 ± 5	12.94 ± 1.07	0.41 ± 0.07
	butyrate	1.5 ± 0.1	14.87 ± 0.57	9.9 ± 1.0
	valerate	4.0 ± 0.6	12.25 ± 0.47	3.1 ± 0.5
	hexanoate	7.3 ± 0.9	9.20 ± 0.31	1.3 ± 0.2
	heptanoate	5.8 ± 0.9	9.12 ± 0.43	1.6 ± 0.2
	octanoate	6.3 ± 2	3.04 ± 0.30	0.48 ± 0.16
$Thr^{324}Val$	acetate	0.042 ± 0.008	0.99 ± 0.02	24 ± 5
	propionate	0.62 ± 0.23	1.1 ± 0.1	1.8 ± 0.8
	isobutyrate	2.1 ± 0.8	0.33 ± 0.05	0.15 ± 0.08
	3-methylvalerate	125 ± 416	3.9 ± 1	0.03 ± 0.2
$Thr^{324}Gly$	acetate	0.18 ± 0.04	1.91 ± 0.08	10.4 ± 2.9
	propionate	3.2 ± 1.1	1.78 ± 0.26	0.56 ± 0.28
	butyrate	0.19 ± 0.09	0.13 ± 0.01	0.68 ± 0.34
	isobutyrate	0.15 ± 0.30	0.23 ± 0.08	1.6 ± 2.4
	3-methylvalerate	0.04 ± 0.06	0.077 ± 0.009	1.9 ± 2.9
	4-methylvalerate	0.07 ± 0.05	0.17 ± 0.01	2.4 ± 2.1
wild-type*	acetate	0.12 ± 0.04	5.7 ± 0.3	48 ± 7
$Thr^{324}Gly$, $Val^{399}Ala$, $Trp^{427}Gly$*	acetate	7.8 ± 17	0.6 ± 0.7	0.07 ± 0.2
	propionate	98 ± 45	3.3 ± 0.9	0.03 ± 0.02
	butyrate	20 ± 2	6.7 ± 0.2	0.33 ± 0.04
	valerate	0.4 ± 0.1	8.8 ± 0.5	22 ± 6
	hexanoate	0.13 ± 0.05	5.7 ± 0.4	43 ± 18
	heptanoate	0.71 ± 0.45	11.0 ± 1.9	15 ± 9.8
	octanoate	2.1 ± 0.7	5.2 ± 0.6	2.4 ± 0.9
	isobutyrate	219 ± 69	4.0 ± 0.9	0.018 ± 0.007
	3-methylvalerate	2.7 ± 0.6	7.7 ± 0.6	2.9 ± 0.6
	4-methylvalerate	0.22 ± 0.07	6.6 ± 0.4	29 ± 10
	4-methylhexanoate	0.25 ± 0.05	5.9 ± 0.2	2 ± 4
$Ile^{323}Ala$, $Val^{399}Ala$, $Trp^{427}Gly$*	propionate	98 ± 42	3.3 ± 0.8	0.03 ± 0.02
	valerate	2.3 ± 0.7	2.7 ± 0.3	1.20 ± 0.38
	hexanoate	0.58 ± 0.05	2.67 ± 0.06	4.62 ± 0.40
	heptanoate	1.05 ± 0.13	2.22 ± 0.08	2.11 ± 0.27
	octanoate	4.1 ± 2.4	6.1 ± 1.6	1.47 ± 0.95
	isobutyrate	110 ± 42	3.2 ± 0.8	0.03 ± 0.01
	3-methylvalerate	1.5 ± 0.4	1.9 ± 0.2	1.22 ± 0.37
	4-methylvalerate	4.3 ± 0.5	6.7 ± 0.3	1.57 ± 0.19
	4-methylhexanoate	1.0 ± 0.3	1.5 ± 0.1	1.6 ± 0.51
$Ile^{323}Ala$, $Val^{399}Gly$, $Trp^{427}Gly$*	butyrate	4.5 ± 0.9	3.8 ± 0.4	0.9 ± 0.2
	valerate	2.1 ± 1.8	2.2 ± 0.7	1.0 ± 1.0
	hexanoate	7.8 ± 1.0	4.8 ± 0.4	0.62 ± 0.09
	heptanoate	5.0 ± 1.3	5.4 ± 0.7	1.1 ± 0.3
	octanoate	2.4 ± 0.7	4.5 ± 0.5	1.9 ± 0.5
	3-methylvalerate	3.01 ± 0.4	2.5 ± 0.1	0.8 ± 0.1
	4-methylvalerate	1.2 ± 0.4	3.0 ± 0.3	2.4 ± 0.8
$Thr^{324}Ser$, $Val^{399}Ala$, $Trp^{427}Gly$*	butyrate	16 ± 5	19 ± 4	1.2 ± 0.5
	valerate	1.5 ± 0.5	7.2 ± 0.9	4.9 ± 1.8
	4-methylvalerate	1.2 ± 0.3	8.2 ± 0.6	6.7 ± 1.6
$Trp^{427}Ala$*	acetate	1.9 ± 0.4	11.0 ± 0.8	6.0 ± 1.2
	propionate	23 ± 11	21 ± 7	0.9 ± 0.5

DOI: 10.1021/acssynbio.9b00008
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Table 2. continued

atACS variant	carboxylate substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})
	butyrate	5.1 ± 0.7	10.8 ± 0.8	2.1 ± 0.3
$Trp^{427}Ser$*	acetate	8.9 ± 3.8	7.2 ± 1.8	0.81 ± 0.4
	isobutyrate	1.6 ± 1.4	1.5 ± 0.5	1.0 ± 0.9
$Trp^{427}His$*	acetate	14 ± 11	15 ± 7	1.1 ± 0.9
$Trp^{427}Val$*	acetate	10 ± 2	10.6 ± 1.4	1.0 ± 0.3
	propionate	38 ± 27	2.5 ± 1.5	0.06 ± 0.06
$Trp^{427}Phe$*	acetate	1.2 ± 0.4	11.7 ± 1.1	15.2 ± 4.7

[a]The variants marked with asterisk (*) were extracted with the purification protocol using HEPES buffer, while all other enzymes were purified with Tris-HCl buffer (see Methods). For reasons that are currently unknown, the enzyme purified with the HEPES buffer displayed higher activity than the preparations made with Tris-HCl buffer. All enzyme variants were assayed with all 11 carboxylate substrates listed in the Methods, but data are only presented with the carboxylate substrates that exhibited catalytic activity.

this $Thr^{324}Tyr$ mutation (*i.e.*, the double mutant $Ile^{323}Ala$, $Thr^{324}Tyr$ and the quadruple mutant $Ile^{323}Ala$, $Thr^{324}Tyr$, $Val^{399}Ala$, $Trp^{427}Gly$) also fails to exhibit enzymatic activity with any of the tested carboxylate substrates (Table 2). The CD spectra of these mutants indicate near normal folding of the mutant proteins (Figure 4). However, homology modeling

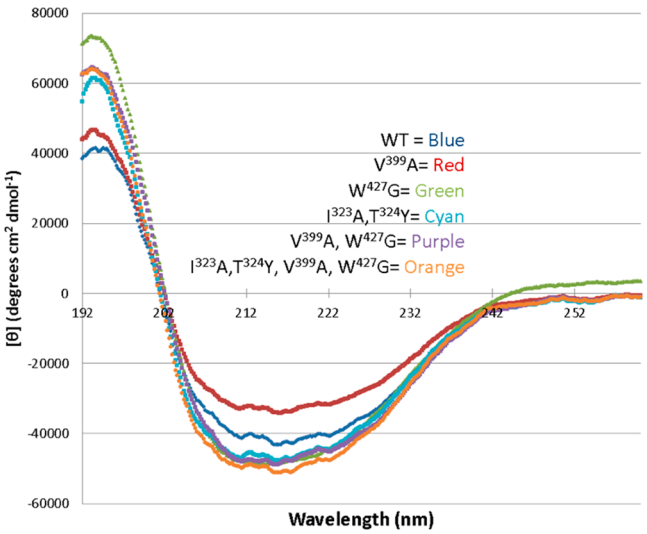


Figure 4. Comparison of CD spectra of wild-type atACS and the mutated variants. Measurements are an average of three scans with a reading every 0.2 nm. Units are in molar ellipticity [θ].

of the $Thr^{324}Tyr$ mutant protein revealed a change in the surface topology of the predicted substrate binding pocket resulting in a decreased volume for binding of the carboxylate substrate (Figure 5). This is consistent therefore with steric hindrance hampering the ability of the carboxylate substrates to access the catalytic site and thereby inactivating the enzyme.

Mutations That Switch Substrate Specificity to Longer, Linear Chain Carboxylate Substrates. Two single site mutants, $Trp^{427}Gly$ and $Val^{399}Ala$, resulted in altered substrate specificity of atACS. The former mutation causes the most dramatic alteration, switching the substrate preference from acetate to longer, straight-chain carboxylate substrates. Namely, this mutant enzyme is able to utilize "new" carboxylate substrates that range between 3- to 8-carbons in length (Table 2). Furthermore, this switch is accompanied by a 6.5-fold decrease in catalytic efficiency (k_{cat}/K_m) with acetate, the native substrate of the wild-type enzyme (Figure 6A).

The other single-site mutant that affected catalytic capabilities of atACS is the $Val^{399}Ala$ switch, which broadened the substrate specificity of the enzyme so that it could equally utilize both acetate and propionate. However, unlike the $Trp^{427}Gly$ mutant, the $Val^{399}Ala$ mutant enzyme could not utilize longer chain carboxylate substrates of 4−8 carbon atoms (Table 2).

On the basis of the fact that these two single site mutations at Val^{399} and Trp^{427} individually generate somewhat distinct changes in catalytic activity of atACS, we evaluated the effect of combining the two mutations in the double mutant, $Val^{399}Ala$, $Trp^{427}Gly$. This double mutant further enhanced the switch in substrate specificity that was observed with the $Trp^{427}Gly$ single mutant (Figure 6A). Specifically, as compared to the wild-type enzyme, the $Val^{399}Ala$, $Trp^{427}Gly$ double mutant shows an 80-fold decrease in catalytic efficiency with the acetate-substrate, and a concomitant increase in catalytic efficiency with the longer, straight chain carboxylate substrates (*i.e.*, propionate to octanoate). With propionate as the substrate, this double mutant enzyme shows a 2.5-fold increase in catalytic efficiency as compared to either the wild-type enzyme or the $Trp^{427}Gly$ single mutant (Figure 6A). The catalytic efficiency with butyrate, the best substrate for this mutant enzyme, matched the catalytic efficiency expressed by the wild-type enzyme with acetate. With these "new" longer chain-length substrates, the double mutant enzyme displayed a lower K_m value (ranging from 1.5 to 7.5-fold lower values) than the $Trp^{427}Gly$ single mutant (Table 2). Analysis of the structural homology models indicate that both the $Trp^{427}Gly$ and $Val^{399}Ala$ single mutations enlarge the carboxylate binding pocket significantly, and further enlargement is apparent in the $Val^{399}Ala$, $Trp^{427}Gly$ double mutant. Thus, these models are consistent with the ability of these variant enzymes to use the larger substrates, and in general it is likely that steric hindrance is a major factor in narrowing the substrate specificity for the native atACS enzyme (Figure 5).

Mutations That Switch Substrate Specificity to Branched Chain Carboxylate Substrates. The Thr^{324} residue of atACS, which we showed to be critical for catalytic competence because its replacement with the equivalent residue that occurs in pcICS (*i.e.*, Tyr) resulted in an inactive enzyme, was subsequently replaced with residues that have smaller side chains. These mutations were systematically implemented to explore whether maintaining either the hydroxyl-group functionality (*i.e.*, a serine switch) or the branched alkyl chain functionality (*i.e.*, a valine switch) or completely eliminating the side chain (*i.e.*, a glycine switch) of the existing Thr residue is important for catalysis.

The $Thr^{324}Ser$ variant did not show any activity with acetate or longer straight chain substrates (Table 2); however, the

DOI: 10.1021/acssynbio.9b00008
ACS Synth. Biol. XXXX, XXX, XXX−XXX

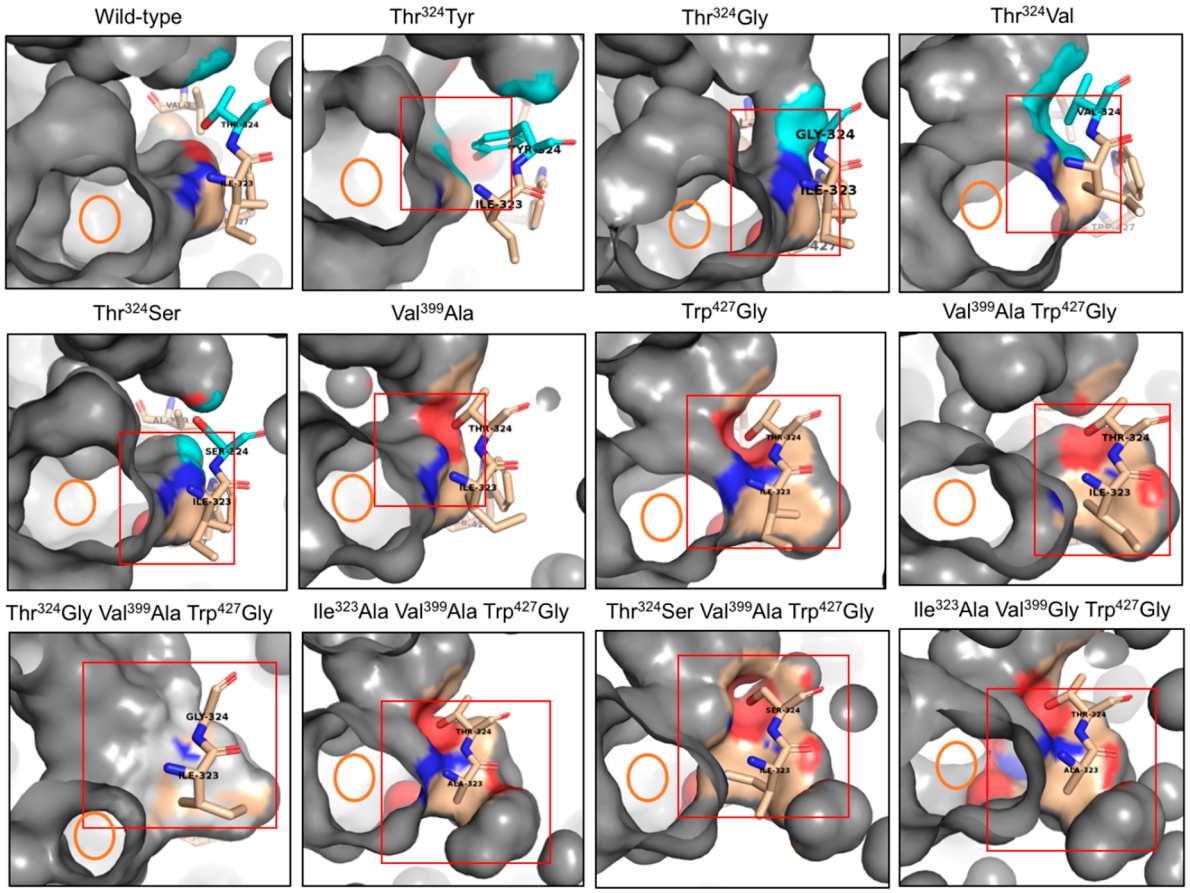


Figure 5. Comparison of the carboxylate binding pocket surface computationally generated from 3D structure homology models of the wild-type and mutants of atACS. The mutated residues are colored in tan. The change in binding pocket size and shape is highlighted with a red box. The nitrogen and oxygen atoms of the targeted residues are shown in blue and red, respectively. The orange circle indicates the cut away opening to reveal the inside of the carboxylate binding pocket. The chemo-physical nature of the cavity surface is color-coded as tan for hydrophobic, red for negatively changed polarity, and blue for positively charged polarity. The binding pocket model was generated using PyMOL. The Ala^{399} and Gly^{427} residues in the double and triple mutant images are not in the view as they are situated behind the expanded carboxylate binding cavity.

Thr^{324}Val and Thr^{324}Gly variants maintain the ability to utilize acetate as a substrate, but with quantitative differences between them (Figure 6B; Table 2). Namely, the catalytic efficiency of the Thr^{324}Val mutant is about 2-fold higher than either the wild-type or Thr^{324}Gly variant with acetate, and this was primarily due to a lower K_m value for this substrate. Moreover, the Thr^{324}Val and Thr^{324}Gly single mutants also showed qualitative differences in substrate specificity, gaining the ability to utilize branched chain carboxylate substrates, such as isobutyrate and 4-methylvalerate, a capability that is undetectable with the wild-type atACS (Figure 6B; Table 2).

The 3D homology models of these Thr^{324} single mutants reveal the potential explanations of these effects on catalytic ability of the variant enzymes. In contrast to the Thr^{324}Tyr single mutant, which inactivates the enzyme due to an apparent reduction in the size of the carboxylate binding pocket (Figure 5), the Thr^{324}Gly single mutant is predicted to have a larger carboxylate binding pocket as compared to the wild-type enzyme, thus explaining the ability of this mutant to utilize the larger, branched chain carboxylate substrates (Figure 5). Similarly, the Thr^{324}Val single mutant is also predicted to have a larger carboxylate binding pocket, although smaller than the Thr^{324}Gly single mutant, which explains the ability of this single mutant to utilize straight-chain substrates, up to three carbon atoms in length (*i.e.*, propionate) and also branched chain substrates, albeit with lower efficiency as compared to the Thr^{324}Gly single mutant. The size of the binding pocket of the Thr^{324}Ser mutant is similar to the wild-type enzyme (Figure 5), thus the inactivity of this single site mutant cannot be explained solely based on steric interference between the substrate and pocket-size, but instead may be attributable to the difference in the polarity of the pocket affected by the orientation of the hydroxyl group of the Ser side chain as compared to Thr in wild-type (Figure 5). These findings therefore indicate that both steric effects and the chemo-physical properties (polarity *versus* hydrophobicity) of the binding pocket affect the carboxylate specificity of the enzyme.

On the basis of these observations, we systematically combined the singular Thr^{324}Gly or Thr^{324}Ser mutations with the double mutant, Val^{399}Ala, Trp^{427}Gly, to evaluate the interrelationships among these residues to affect catalytic capabilities of atACS. The enzymatic capabilities of the resulting Thr^{324}Gly, Val^{399}Ala, Trp^{427}Gly and Thr^{324}Ser, Val^{399}Ala, Trp^{427}Gly triple mutants were compared to the wild-type and the parental double mutant enzymes.

These triple mutants express enhanced catalytic capabilities in utilizing longer chain linear carboxylate substrates than the Val^{399}Ala, Trp^{427}Gly double mutant enzyme (Figure 6C). Thus, whereas the double mutant can utilize butyrate with the highest efficiency (Figure 6A), the Thr^{324}Gly, Val^{399}Ala,

DOI: 10.1021/acssynbio.9b00008
ACS Synth. Biol. XXXX, XXX, XXX−XXX

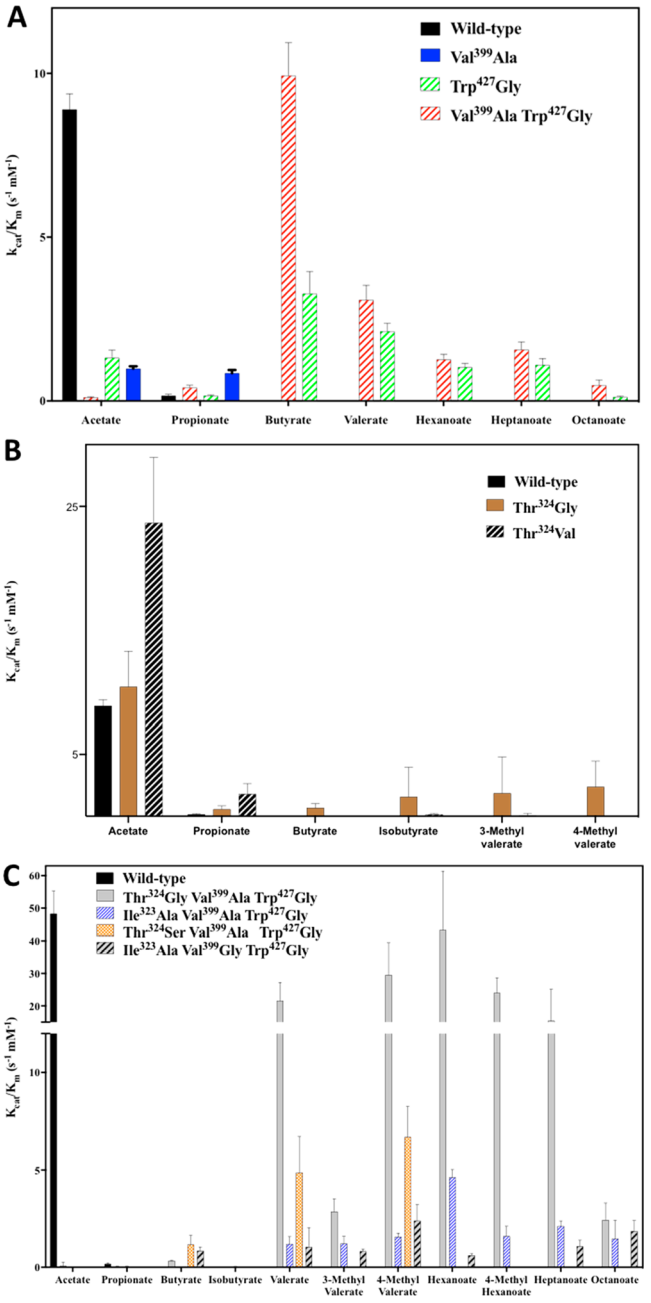


Figure 6. Comparison of the catalytic efficiency (k_{cat}/K_m) of the wild-type atACS with single and double site mutants (A, B), and triple site mutants (C), assayed with straight- and branched-chain carboxylate substrates. The variant enzymes in panel A were purified using the Tris-HCl extraction buffer, and variants in panel C were purified using the HEPES extraction buffer (see Methods). Data are the average of triplicate determination ± standard error.

Trp^{427}Gly triple mutant can utilize hexanoate with a catalytic efficiency that is comparable to that of the wild-type enzyme with its native acetate substrate (Figure 6C). Moreover, these triple mutants also utilize branched chain substrates, such as 3-methylvalerate, 4-methylvalerate and 4-methylhexanoate. The catalytic efficiency of this triple mutant with branched chain carboxylate substrates is comparable to the activity expressed by the wild-type enzyme with acetate as the substrate. Compared to the Val^{399}Ala, Trp^{427}Gly double mutant, this additional switch in the preferred carboxylate substrate by the triple mutants is associated with decreased K_m values for the longer chain length carboxylate substrates of more than 5-carbon atoms, and increased K_m values for the shorter chain substrates, such as propionate, butyrate and isobutyrate (Table 2).

Thus, comparing the modeled structure of the carboxylate binding pockets of the Val^{399}Ala, Trp^{427}Gly double mutant, with the Thr^{324}Gly, Val^{399}Ala, Trp^{427}Gly triple mutant indicates that the replacement of Thr^{324} with a smaller side chain residue, *i.e.*, glycine, affects both the shape and size of the binding pocket, and its polarity (Figure 5), resulting in a shift in substrate specificity from a 2-carbon substrate, to more than 4-carbon long, straight chain and branched substrates, which are also less polar. These findings further strengthen the conclusion that both steric interference and chemo-physical properties of the binding pocket plays a role in determining the carboxylate substrate specificity of atACS.

Even more dramatic changes in substrate specificity were generated when the Ile^{323}Ala mutation was combined with the Val^{399}Ala, Trp^{427}Gly or Val^{399}Gly, Trp^{427}Gly double mutations. Thus, Figure 6C shows that both the Ile^{323}Ala, Val^{399}Gly, Trp^{427}Gly and Ile^{323}Ala, Val^{399}Ala, Trp^{427}Gly triple mutants are inactive with acetate and only show barely detectable activity with propionate, both of these triple mutants can utilize 4- to 8-carbon straight chain carboxylates as substrates, as well as the branched chain carboxylates, 3-methylvalerate and 4-methylvalerate. Detailed examination of the kinetic properties of all the triple mutants indicate that they prefer 4-methyl branched chain carboxylates (*i.e.*, 4-methylvalerate and 4-methylhexanoate) over the 3-methyl branched carboxylate (*i.e.*, 3-methylvalerate), manifested by lower K_m values and higher catalytic efficiencies.

The structural homology models of the triple mutants indicate that these mutations enlarge the carboxylate binding pocket, thus explaining the capability of these mutants to utilize longer chain substrates (Figure 5). The triple mutant Thr^{324}Gly, Val^{399}Ala, Trp^{427}Gly which displays the largest shift in substrate specificity (Figure 6C), shows the most nonpolar binding pocket cavity (Figure 5), which provides the potential explanation of why this variant is more active with the most hydrophobic carboxylate substrates. This explanation was further explored and confirmed by computational docking experiments that compared the binding of acetate to the wild-type atACS and the binding of 4-methyl valerate and hexanoate to the substrate pocket of this triple mutant (Supporting Information Figure S2).

Because the Trp^{427}Gly mutation has a large effect on substrate specificity and promiscuity (Table 2), we explored the consequence of changing the Trp^{427} residue to all other possible amino acids. These variants thereby evaluated how the change in the chemo-physical property of this amino acid side chain affects substrate specificity (Table 2). Only five of the 18 additional variants were catalytically active (Val, Ala, Phe, Ser or His), and the other 13 were inactive with all tested carboxylates (the consequence of changing Trp^{427} to Gly had already been evaluated; Figure 6). The five active single site mutants utilize acetate as a substrate, though with lower catalytic efficiency (Table 2). The mutations that incorporate small hydrophobic side chain residues (*i.e.*, Trp^{427}Val or Trp^{427}Ala) enable the enzyme to better utilize straight chain carboxylate substrates, of 3- and 4-carbon atoms. In addition, the mutations that conserve the ring structure of the native Trp^{427} (*i.e.*, Trp^{427}Phe or Trp^{427}His) maintain enzymatic activity with acetate. Therefore, these findings indicate that

DOI: 10.1021/acssynbio.9b00008
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Table 3. Michaelis–Menten Kinetic Parameters of ATP and CoA for Wild-Type and Variant atACS Enzymes[a]

		ATP		CoA	
enzyme	carboxylate substrate	K_m (mM)	k_{cat} (s^{-1})	K_m (mM)	k_{cat} (s^{-1})
wild-type	acetate	0.20 ± 0.03	3.56 ± 0.12	0.15 ± 0.02	3.47 ± 0.14
$Val^{399}Ala$	acetate	0.06 ± 0.01	2.28 ± 0.04	0.09 ± 0.02	3.20 ± 0.15
$Trp^{427}Gly$	butyrate	0.02 ± 0.003	6.41 ± 0.31	0.14 ± 0.05	9.36 ± 1.21
$Val^{399}Ala$, $Trp^{427}Gly$	butyrate	0.06 ± 0.02	6.17 ± 0.66	0.09 ± 0.02	8.04 ± 0.78
wild-type*	acetate	0.050 ± 0.002	16.04 ± 0.18	0.11 ± 0.03	16.03 ± 0.80
$Thr^{324}Gly$, $Val^{399}Ala$, $Trp^{427}Gly$*	hexanoate	0.26 ± 0.05	5.04 ± 0.26	0.14 ± 0.03	5.07 ± 0.20
$Ile^{323}Ala$, $Val^{399}Ala$, $Trp^{427}Gly$*	hexanoate	0.11 ± 0.03	1.60 ± 0.09	0.04 ± 0.02	1.45 ± 0.11

[a]The variants marked with an asterisk (*) were purified using HEPES buffer, as described in the Methods.

Trp^{427} is an important residue in the carboxylate binding pocket, and in addition to its aromaticity and/or polarity associated with the indole ring, the size of the side chain has an effect on restricting the specificity of the wild-type enzyme to acetate.

Collectively therefore, using this rational redesign strategy, we switched an enzyme's substrate preference from acetate to be equally active with medium-length, straight and branched chain carboxylate substrates. This engineered enzyme displays catalytic activity that is comparable to the native enzyme with its natural substrate, acetate (k_{cat} ~ 5 s^{-1}). Additional improvements can be envisioned, based on the fact that naturally existing medium-chain acyl-CoA synthetases have been characterized that display k_{cat} ~ 60 s^{-1} (*e.g.*, CSAAE3 isolated from *Cannabis sativa*[38]).

Such improvements could be achieved by combining this rational redesign with a randomized saturation mutagenesis of the four residues that were targeted in this study. Although we partially implemented this strategy by specifically targeting one of the four residues (*i.e.*, Trp^{427}), combining this set of mutations with a similar randomized saturation mutagenesis strategy at the other 3 residues would generate 160 000 (=20^4) variants. Assaying such a large collection of variants would require a high-throughput assay, and the enzyme-coupled spectroscopic assay that we used herein could be adapted for such a high-throughput parallel assessment. Hence, combining all possible mutant combinations at the 4 targeted residues has the potential of uncovering synergistic combinations to improve and alter the carboxylate substrate preference of this class of enzymes.

Kinetics of ATP and CoA Dependence. The wild-type atACS and selected mutants showing altered preference for the carboxylate substrate were also characterized to evaluate the effects of these mutations on the kinetic parameters associated with the other two substrates of the reaction, ATP and CoA (Table 3). The assays used to determine the K_m and k_{cat} values for these substrates were conducted with saturating concentrations of the most preferred carboxylate substrate for the individual variant enzymes. These characterizations demonstrate that the mutations that were designed to affect the specificity of the carboxylate substrate also affected the enzyme's dependence on these two additional substrates, although the magnitude of these changes are considerably smaller than the changes in the kinetic parameters for the carboxylate substrates.

The most dramatic of these changes is associated with the kinetic parameters for ATP, with up to 10-fold decrease in K_m in mutants that enable the enzyme to better utilize longer chain linear carboxylate substrates (*e.g.*, $Trp^{427}Gly$). In contrast, there are relatively minor effects on the kinetic parameters associated with CoA. While in the triple-site mutant (*i.e.*, $Thr^{324}Gly$ $Val^{399}Ala$ $Trp^{427}Gly$), which switches the carboxylate substrate preference from acetate to the branched chain substrates, there is increase in K_m and decrease in k_{cat} values for both ATP and CoA. The alterations in these kinetic parameters are consistent with the mechanism of the ACS-catalyzed reaction; namely ATP and the carboxylate substrates first react to form the adenylate-intermediate, which subsequently reacts with CoA. Thus, changing the carboxylate binding-site would be expected to have effects on ATP kinetics as these two substrates have to come together at an active site that is juxtaposed next to each substrate binding pocket.

CONCLUSION

Despite the conservation of ACS enzymes among a wide range of evolutionary phyla, this enzyme has maintained a restricted substrate specificity.[32] For example, the plant ACS that was the focus of this study only utilizes acetate efficiently as a substrate, and propionate to a much lesser extent. In this study we extrapolated from experimentally determined 3D structural data of two phylogenetically distantly related ACS enzymes, and modeled the potential carboxylate substrate binding pockets of two enzymes that also assemble thioester bonds but display different specificities toward the carboxylate substrate (acetate or isobutyrate). Rational mutagenesis experiments guided by this modeling were successful in initially introducing substrate promiscuity to the enzyme, and ultimately altering its substrate specificity. We have been particularly successful in shifting the substrate specificity of this eukaryotic ACS away from its native substrate (acetate), to be equally efficient in utilizing novel branched chain carboxylates that are not substrates of any of the guiding model enzymes. Similar studies have targeted microbial ACS enzymes,[35,36,39,40] which expanded the carboxylate substrates that these enzymes can utilize, thus converting a very specific enzyme to a promiscuous enzyme. Furthermore, apart from the four residues targeted in this study, additional residues have been identified (*i.e.*, Lys^{627}, Trp^{426} *etc.*) that appear to be important in determining substrate selectivity and catalysis of this class of enzymes.[34,37] The variant enzymes that we have generated are distinct in that they are not promiscuous, but they show shifted substrate specificity, being equally efficient with the new substrates, as the wild-type enzyme is with its native substrate.

Hence, this is an exemplary study of a research path to generate novel enzymes. Moreover, the novel acyl-CoA synthetase enzymes that were generated herein can serve as genetically encoded catalytic elements for bioengineering multiple metabolic pathways and potentially producing useful novel chemicals.[27,41,42] For example, bioengineering organisms with these novel acyl-CoA synthases may enable the

DOI: 10.1021/acssynbio.9b00008
ACS Synth. Biol. XXXX, XXX, XXX–XXX

biosynthesis of novel fatty acids or polyketides which would have internal branched structures, thus generating bioproducts with distinct chemo-physical properties, and can potentially serve as bioprivileged molecules.[43]

METHODS

Generation of atACS Variants. The atACS ORF was cloned into the pET24b vector (Merck KGaA, Darmstadt, Germany) between the *Hin*dIII and *Xho*I sites. atACS was also subcloned into the pET30f vector at the *Bam*HI restriction site. pET30f vector is a modified pET30 vector (Takara, Clontech, Mountain View, CA) encoding an N-terminal poly histidine coding sequence, followed by a TEV cleavage site. Mutagenesis of specific residues was performed either using the QuikChange Lightning site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA) *via* manufacturer instructions or by recombination PCR in a two-step fashion.[13] The atACS ORF cloned into the plasmid pET24b, was used as template for QuikChange or recombination PCR. The full-length atACS ORF PCR product containing the mutation was purified and cloned into a pGEM T-Easy vector, and then transformed into Top-10 competent *E. coli* cells. Sequence-confirmed atACS variants in pGEM T-Easy were digested with *Hin*dIII and *Xho*I to release the atACS ORF and were subcloned into the corresponding restriction sites in the pET24b vector. Sequence-confirmed pET30f and pET24b constructs were subsequently transformed into *E. coli* ArcticExpress (DE3) cells (Agilent Technologies, Santa Clara, CA) for protein expression.

Overexpression and Purification of Wild-Type and Variant atACS Proteins. Arctic Express strains harboring atACS variants were grown overnight at 37 °C with shaking at 250 rpm, in 5−10 mL LB medium containing the antibiotics kanamycin (50 μg/mL) and gentamycin (20 μg/mL). The overnight culture was used to inoculate 0.5−1.0 L LB media containing no antibiotics, and these cultures were grown at 30 °C with shaking at 250 rpm, until an OD_{600} of ∼0.6. Expression was then induced by the addition of IPTG to a final concentration of 0.1 mM, and the culture was grown at 13 °C with shaking at 250 rpm for 24−48 h.

Cells were harvested by centrifugation at 10 000*x*g for 10 min and resuspended in a buffer containing 0.5 M NaCl, 5 mM imidazole, 0.1% (v/v) Triton X-100, 0.1 mg/mL phenyl methyl sulfonyl fluoride (PMSF), and 10 μL/ml Protease Inhibitor Cocktail (Sigma-Aldrich Co., St. Louis, MO) and disrupted by sonication on ice. Two different buffer systems were used in this cell disruption step, either 10 mM HEPES-KOH, pH 7.5, or 20 mM Tris-HCl, pH 7.5. The resulting extract was centrifuged at 20 000*x*g for 30 min, and the supernatant was retained, filtered through a 0.45 μm filter disc (Corning Inc., Corning, NY) and applied to a column containing 2−5 mL PerfectPro Ni-NTA agarose (5 Prime, Inc., Gaithersburg, MD) at 4 °C. The column was washed with increasing concentrations of imidazole, and the atACS protein was eluted with 0.2 M imidazole. Samples were immediately dialyzed at 4 °C into either 10 mM HEPES-KOH, pH 7.5, 10 mM KCl or 50 mM Tris-HCl pH 7.5; these two methods yield enzyme preparations that expressed different specific activities, and these differences are specifically identified in the Results. In all comparative studies of mutants, both the wild-type and all variant enzymes were prepared identically, so as to avoid this buffer-induced variation in catalytic activity, whose origins are as of yet unknown.

Spectrophotometric atACS Activity Assay. ACS activity was measured by coupling the carboxylate-dependent formation of the AMP product to the oxidation of NADH (Supporting Information Figure S1), using the enzymes myokinase, pyruvate kinase and lactate dehydrogenase[4] (Sigma-Aldrich Co., St. Louis, MO). The oxidation of NADH was monitored by the rate of decrease in absorbance at 340 nm. Each enzyme variant was assayed with 11 different carboxylate substrates that differed in the alkyl chain-length (2−8 carbons) or whether the alkyl chain was linear or branched, and their concentrations in assays ranged between 0 and 10 mM. The specific substrates were: acetate, propionate, butyrate, valerate, hexanoate, heptanoate, octanoate, isobutyrate, 3-methyl valerate, 4-methyl valerate and 4-methyl hexanoate. Assays were performed in 96-well microtiter dishes at 37 °C with a final well-volume of 100 μL, and A_{340} of each well was monitored with a BioTek ELx808TM Absorbance Microplate Reader, using Gen5TM Data Analysis software (BioTek Instruments, Winooski, VT). The standard reaction contained 50 mM Tris-HCl pH 7.5, and 5 mM Tris(2-carboxylethyl) phosphine hydrochloride (Thermo Fisher Scientific Inc., Waltham, MA); 5% (v/v) ethanol; 5 mM ATP; 6 mM $MgCl_2$; 5 mM phospho(enol)pyruvate; 0.4 mM NADH; 2−20 U Pyruvate Kinase/Lactate Dehydrogenase mix (Sigma-Aldrich Co., St. Louis, MO); 10U Myokinase from rabbit muscle or 1U Myokinase from chicken muscle (Sigma-Aldrich Co., St. Louis, MO); and between 2.5 and 5 μg of purified recombinant atACS protein.

Computational Experiments. The chloroplast targeting sequence was identified in the atACS sequence using TargetP 1.1[47,48] (http://www.cbs.dtu.dk/services/TargetP). Amino acid similarity and identity values of homologous proteins were obtained using BLASTp (http://blast.ncbi.nlm.nih.gov). Structural predictions were performed using the I-TASSER server.[28,41,44] Predicted structures and known crystal structures were visualized and analyzed using PyMOL (The PyMOL Molecular Graphics System, Version 1.3.0 Schrödinger, LLC). Multiple sequence alignments were generated using Clustal Omega.[45] Docking experiments were performed using AutoDockVina[46] and results visualized using PyMOL.

Circular Dichroism. CD spectra were obtained at the Iowa State University Protein Facility (http://www.protein.iastate.edu/) using a J700 spectropolarimeter (JASCO Inc., Easton, MD). Spectra were taken from 190 to 250 nm with readings every 0.2 nm. Samples were measured in a 1 mm dichroically neutral quartz cuvette and the average of 3−5 scans was used for analysis. Protein samples in 50 mM Tris-HCl were diluted into ddH_2O to a final concentration of 0.15−0.25 mg/mL for spectra acquisition. A baseline spectrum for 5 mM Tris-HCl was obtained in the same manner and subtracted from the experimental protein spectra. After data collection, spectra were converted from millidegrees to molar ellipticity ($[\theta]$, degrees cm^2 $dmol^{-1}$). Spectra were analyzed using a suite of algorithms collectively called CDPro,[44] utilizing algorithms: SELCON3, CDSSTR, and CONTIN/LL.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.9b00008.

DOI: 10.1021/acssynbio.9b00008
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Figure S1: Diagram of the assay components; Figure S2: Docking of substrates to atACS variants (PDF)

AUTHOR INFORMATION

Corresponding Author

*Tel: 515-294-9423. E-mail: dimmas@iastate.edu.

ORCID

Naazneen Sofeo: 0000-0001-5724-3394

Author Contributions

B.J.N., M.D.Y.-N., and D.J.O. conceived the study. N.S., J.H.H., M.D.Y.-N., and B.J.N. designed the experiments. N.S., J.H.H., and B.B. conducted the experiments. N.S., J.H.H., M.D.Y.-N., and B.J.N. wrote the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was partially supported by the State of Iowa, through the Center of Metabolic Biology, and by the National Science Foundation, which supported the Engineering Research Center for Biorenewable Chemicals (Award No. EEC-0813570). Brandon Butler acknowledges support from the Research Experience for Undergraduates program of the National Science Foundation, provided through the Engineering Research Center for Biorenewable Chemicals. The authors thank Dr. Fuyuan Jing of Iowa State University for gifting the pET30f vector used for protein expression, and Dr. Alan T. Culbertson, who evaluated the phylogenetic-based interrelationships among ACS enzymes.

ABBREVIATIONS

atACS, *Arabidopsis thaliana* acetyl CoA synthetase; ACS, acetyl CoA synthetase; pcICS, *Pseudomonas chlororaphis* isobutyryl CoA synthetase; ATP, adenosine triphosphate; CoA, coenzyme A.

REFERENCES

(1) McCarty, N. S., and Ledesma-Amaro, R. (2019) Synthetic Biology Tools to Engineer Microbial Communities for Biotechnology. *Trends Biotechnol. 37* (2), 181−197.

(2) Liu, Y., Liu, L., Li, J., Du, G., and Chen, J. (2019) Synthetic Biology Toolbox and Chassis Development in Bacillus Subtilis. *Trends Biotechnol. 37* (5), 548−562.

(3) Hashimoto, Y., Hosaka, H., Oinuma, K.-I., Goda, M., Higashibata, H., and Kobayashi, M. (2005) Nitrile Pathway Involving Acyl-CoA Synthetase: Overall Metabolic Gene Organization and Purification and Characterization of the Enzyme. *J. Biol. Chem. 280* (10), 8660−8667.

(4) Horswill, A. R., and Escalante-Semerena, J. C. (1999) The PrpE Gene of Salmonella Typhimurium LT2 Encodes Propionyl-CoA Synthetase. *Microbiology 145* (6), 1381−1388.

(5) Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997) Structural Basis for the Activation of Phenylalanine in the Non-Ribosomal Biosynthesis of Gramicidin S. *EMBO J. 16* (14), 4174−4183.

(6) Gulick, A. M. (2009) Conformational Dynamics in the Acyl-CoA Synthetases, Adenylation Domains of Non-Ribosomal Peptide Synthetases, and Firefly Luciferase. *ACS Chem. Biol. 4* (10), 811−827.

(7) Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997) Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chem. Rev. 97* (7), 2651−2674.

(8) Berg, P., and Newton, W. (1956) The Technical Assistance of G. Acyl Adenylates: An Enzymatic Mechanism of Acetate Activation. *J. Biol. Chem. 222* (2), 991−1013.

(9) Schmelz, S., and Naismith, J. H. (2009) Adenylate-Forming Enzymes. *Curr. Opin. Struct. Biol. 19* (6), 666−671.

(10) Starai, V. J., and Escalante-Semerena, J. C. (2004) Acetyl-Coenzyme A Synthetase (AMP Forming). *Cell. Mol. Life Sci. 61* (16), 2020−2030.

(11) Gulick, A. M., Lu, X., and Dunaway-Mariano, D. (2004) Crystal Structure of 4-Chlorobenzoate:CoA Ligase/Synthetase in the Unliganded and Aryl Substrate-Bound Statest. *Biochemistry 43* (27), 8670−8679.

(12) Guan, X., and Nikolau, B. J. (2016) AAE13 Encodes a Dual-Localized Malonyl-CoA Synthetase That Is Crucial for Mitochondrial Fatty Acid Biosynthesis. *Plant J. 85* (5), 581−593.

(13) Guo, Y., and Oliver, D. J. (2012) E. Coli Propionyl-CoA Synthetase Is Regulated in Vitro by an Intramolecular Disulfide Bond. *Prikl. Biokhim. Mikrobiol. 48* (3), 289−293.

(14) Hughes, A. J., and Keatinge-Clay, A. (2011) Enzymatic Extender Unit Generation for In Vitro Polyketide Synthase Reactions: Structural and Func-Tional Showcasing of Streptomyces Coelicolor MatB. *Chem. Biol. 18* (2), 165−176.

(15) Chang, K. H., Xiang, H., and Dunaway-Mariano, D. (1997) Acyl-Adenylate Motif of the Acyl-Adenylate/Thioester-Forming Enzyme Superfamily: A Site-Directed Mutagenesis Study with the Pseudomonas Sp. Strain CBS3 4-Chlorobenzoate:Coenzyme A Ligase. *Biochemistry 36* (50), 15650−15659.

(16) Pandey, S. K., Yadav, S., Temre, M. K., and Singh, S. M. (2018) Tracking Acetate through a Journey of Living World: Evolution as Alternative Cellular Fuel with Potential for Application in Cancer Therapeutics. *Life Sci. 215*, 86−95.

(17) Duve, C. d. (1995) The Beginnings of Life on Earth. *Am. Sci. 83* (5), 428−437.

(18) Ke, J., Behal, R. H., Back, S. L., Nikolau, B. J., Wurtele, E. S., and Oliver, D. J. (2000) The Role of Pyruvate Dehydrogenase and Acetyl-Coenzyme A Synthetase in Fatty Acid Synthesis in Developing Arabidopsis Seeds. *Plant Physiol. 123* (2), 497−508.

(19) Fatland, B. L., Nikolau, B. J., and Wurtele, E. S. (2005) Reverse Genetic Characterization of Cytosolic Acetyl-CoA Generation by ATP-Citrate Lyase in Arabidopsis. *Plant Cell 17* (1), 182−203.

(20) Bao, X., Pollard, M., and Ohlrogge, J. (1998) The Biosynthesis of Erucic Acid in Developing Embryos of Brassica Rapa. *Plant Physiol. 118* (1), 183−190.

(21) Behal, R. H., Lin, M., Back, S., and Oliver, D. J. (2002) Role of Acetyl-Coenzyme A Synthetase in Leaves of Arabidopsis Thaliana. *Arch. Biochem. Biophys. 402* (2), 259−267.

(22) Lin, M., and Oliver, D. J. (2008) The Role of Acetyl-Coenzyme A Synthetase in Arabidopsis. *Plant Physiol. 147* (4), 1822−1829.

(23) Zeiher, C. A., and Randall, D. D. (1990) Identification and Characterization of Mitochondrial Acetyl-Coenzyme A Hydrolase from Pisum Sativum L. Seedlings. *Plant Physiol. 94* (1), 20−27.

(24) Austin, M. B., and Noel, J. P. (2003) The Chalcone Synthase Superfamily of Type III Polyketide Synthases. *Nat. Prod. Rep. 20* (1), 79−110.

(25) Liedvogel, B. (1986) Acetyl Coenzyme A and Isopentenylpyrophosphate as Lipid Precursors in Plant Cells - Biosynthesis and Compartmentation. *J. Plant Physiol. 124* (3−4), 211−222.

(26) Brooks, J. L., and Stumpf, P. K. (1966) Fat Metabolism in Higher Plants. XXXIX. Properties of a Soluble Fatty Acid Synthesizing System from Lettuce Chloroplasts. *Arch. Biochem. Biophys. 116* (1), 108−116.

(27) Nikolau, B. J., Perera, M. A. D. N., Brachova, L., and Shanks, B. (2008) Platform Biochemicals for a Biorenewable Chemical Industry. *Plant J. 54* (4), 536−545.

(28) Gulick, A. M., Starai, V. J., Horswill, A. R., Homick, K. M., and Escalante-Semerena, J. C. (2003) The 1.75 Å Crystal Structure of Acetyl-CoA Synthetase Bound to Adenosine-5???-Propylphosphate and Coenzyme A. *Biochemistry 42* (10), 2866−2873.

(29) Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: A Unified Platform for Automated Protein Structure and Function Prediction. *Nat. Protoc. 5* (4), 725−738.

DOI: 10.1021/acssynbio.9b00008
ACS Synth. Biol. XXXX, XXX, XXX−XXX

(30) Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015) The I-TASSER Suite: Protein Structure and Function Prediction. *Nat. Methods 12* (1), 7−8.
(31) Zhang, Y. (2008) I-TASSER Server for Protein 3D Structure Prediction. *BMC Bioinf. 9* (1), 40.
(32) Reger, A. S., Carney, J. M., and Gulick, A. M. (2007) Biochemical and Crystallographic Analysis of Substrate Binding and Conformational Changes in Acetyl-CoA Synthetase. *Biochemistry 46* (22), 6536−6546.
(33) Jogl, G., and Tong, L. (2004) Crystal Structure of Yeast Acetyl-Coenzyme A Synthetase in Complex with AMP. *Biochemistry 43* (6), 1425−1431.
(34) Khurana, P., Gokhale, R. S., and Mohanty, D. (2010) Genome Scale Prediction of Substrate Specificity for Acyl Adenylate Superfamily of Enzymes Based on Active Site Residue Profiles. *BMC Bioinf. 11* (1), 57.
(35) Ingram-Smith, C., Woods, B. I., and Smith, K. S. (2006) Characterization of the Acyl Substrate Binding Pocket of Acetyl-CoA Synthetase. *Biochemistry 45* (38), 11482−11490.
(36) Ingram-Smith, C., and Smith, K. S. (2006) AMP-Forming Acetyl-CoA Synthetases in Archaea Show Unexpected Diversity in Substrate Utilization. *Archaea 2* (2), 95−107.
(37) Soumya, N., Tandan, H., Damre, M. V., Gangwal, R. P., Sangamwar, A. T., and Singh, S. (2016) Leucine-684: A Conserved Residue of an AMP-Acetyl CoA Synthetase (AceCS) from Leishmania Donovani Is Involved in Substrate Recognition, Catalysis and Acetylation. *Gene 580* (2), 125−133.
(38) Stout, J. M., Boubakir, Z., Ambrose, S. J., Purves, R. W., and Page, J. E. (2012) The Hexanoyl-CoA Precursor for Cannabinoid Biosynthesis Is Formed by an Acyl-Activating Enzyme in Cannabis Sativa Trichomes. *Plant J. 71* (3), 353−365.
(39) Khersonsky, O., Lipsh, R., Avizemer, Z., Ashani, Y., Goldsmith, M., Leader, H., Dym, O., Rogotner, S., Trudeau, D. L., Prilusky, J., et al. (2018) Automated Design of Efficient and Functionally Diverse Enzyme Repertoires. *Mol. Cell 72* (1), 178−186.
(40) Trudeau, D. L., Edlich-Muth, C., Zarzycki, J., Scheffen, M., Goldsmith, M., Khersonsky, O., Avizemer, Z., Fleishman, S. J., Cotton, C. A. R., Erb, T. J., et al. (2018) Design and in Vitro Realization of Carbon-Conserving Photorespiration. *Proc. Natl. Acad. Sci. U. S. A. 115* (49), E11455−E11464.
(41) Steen, E. J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., del Cardayre, S. B., and Keasling, J. D. (2010) Microbial Production of Fatty-Acid-Derived Fuels and Chemicals from Plant Biomass. *Nature 463* (7280), 559−562.
(42) Schwartz, T. J., O'Neill, B. J., Shanks, B. H., and Dumesic, J. A. (2014) Bridging the Chemical and Biological Catalysis Gap: Challenges and Outlooks for Producing Sustainable Chemicals. *ACS Catal. 4* (6), 2060−2069.
(43) Shanks, B. H., and Broadbelt, L. J. (2019) A Robust Strategy for Sustainable Organic Chemicals Utilizing Bioprivileged Molecules. *ChemSusChem*, cssc.201900323.
(44) Sreerama, N., and Woody, R. W. (2000) Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. *Anal. Biochem. 287* (2), 252−260.
(45) Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res. 22* (22), 4673−4680.
(46) Trott, O., and Olson, A. J. (2009) AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem. 31* (2), 455−461.
(47) Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol. 300*, 1005−1016.
(48) Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng., Des. Sel. 10*, 1−6.

DOI: 10.1021/acssynbio.9b00008
ACS Synth. Biol. XXXX, XXX, XXX−XXX