2013

Characterization of the Arabidopsis thaliana acetyl-CoA synthetase putative carboxylate binding pocket

Jason H. Hart
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/etd
Part of the Biochemistry Commons

Recommended Citation
http://lib.dr.iastate.edu/etd/13369

This Thesis is brought to you for free and open access by the Graduate College at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Characterization of the *Arabidopsis thaliana* acetyl-CoA synthetase putative carboxylate binding pocket

by

Jason H Hart

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Plant Biology

Program of Study Committee:
Basil J. Nikolau, Major Professor
  David J. Oliver
  Thomas A. Bobik

Iowa State University
Ames, Iowa
2013

Copyright © Jason H Hart, 2013. All rights reserved.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td>CHAPTER 1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>THESIS ORGANIZATION</td>
<td>3</td>
</tr>
<tr>
<td>CHAPTER 2 CHARACTERIZATION OF THE ARABIDOPSIS THALIANA</td>
<td></td>
</tr>
<tr>
<td>ACETYL-COA SYNTHETASE PUTATIVE CARBOXYLATE</td>
<td></td>
</tr>
<tr>
<td>BINDING POCKET</td>
<td>4</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>4</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>5</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>7</td>
</tr>
<tr>
<td>Bacterial Strains, Culture Media, and Growth Conditions</td>
<td>7</td>
</tr>
<tr>
<td>Overexpression and Purification of atACS proteins in <em>E. coli</em></td>
<td>8</td>
</tr>
<tr>
<td>Site-Directed Mutagenesis</td>
<td>9</td>
</tr>
<tr>
<td>Spectrophotometric atACS Activity Assay for Kinetic Characterization</td>
<td>10</td>
</tr>
<tr>
<td>Computational Experiments</td>
<td>11</td>
</tr>
<tr>
<td>Circular Dichroism</td>
<td>12</td>
</tr>
<tr>
<td>RESULTS</td>
<td>12</td>
</tr>
<tr>
<td>Identification of the Residues Forming the Putative atACS Binding Pocket</td>
<td>12</td>
</tr>
<tr>
<td>Mutagenesis of Binding Pocket Residues</td>
<td>14</td>
</tr>
<tr>
<td>Expression and Purification of atACS and Variants</td>
<td>14</td>
</tr>
<tr>
<td>Kinetic Parameters of Wild-type atACS and Val^{399}Ala Variant</td>
<td>15</td>
</tr>
</tbody>
</table>
Kinetic Parameters of the Trp$^{427}$Gly and the Val$^{399}$Ala, Trp$^{427}$Gly Variants................................................................. 15

Utilization of ATP and CoA as Substrates by atACS Wild-type and Variants ........................................................................................................ 17

Determination of atACS Wild-type and Variants’ Secondary Structure by Circular Dichroism ................................................................. 18

DISCUSSION .................................................................................................................................................................................. 19

CHAPTER 3 DISCUSSION AND CONCLUSIONS .............................................. 26

REFERENCES ................................................................................................................................................................................. 29

TABLES AND FIGURES ......................................................................................................................................................... 31
ACKNOWLEDGEMENTS

I would like to take this opportunity to thank all those that have helped me through my research, both scientifically and personally. In particular I must thank my major professor, Dr. Basil Nikolau for his indispensable mentoring, advice, discussion and help throughout my work as well in constructing this manuscript. I would like to express my appreciation for the considerable time and effort that all of the members of the Nikolau laboratory have given to help mentor me and for the numerous helpful and engaging discussions. I would also like to thank my amazing family and friends that have always been supportive and have provided me with joy and encouragement when I have needed it most. I am also grateful for the Center for Biorenewable Chemicals and for the funding provided by the National Science Foundation.
Acetyl-CoA synthetase catalyzes the activation of acetate by the acetylation of the thiol group of Coenzyme A, while hydrolyzing ATP to AMP and pyrophosphate. The Arabidopsis thaliana acetyl-CoA synthetase (atACS) was compared to other acyl-CoA synthetases, and was computationally modeled on the available crystal structures of the Saccharomyces cerevisiae ACS1 and Salmonella enterica ACS. This allowed the identification of the residues that make up the putative carboxylate binding pocket residues. To further understand substrate selectivity and binding within the putative carboxylate binding pocket, selected residues were mutated to resemble the homologous residues in the Pseudomonas chlororaphis isobutyryl-CoA synthetase. Four residues (Ile$^{323}$, Thr$^{324}$, Val$^{399}$, and Trp$^{427}$) were identified that are proposed to form the carboxylate binding pocket. One residue, Trp$^{427}$ was found to be the primary residue in determining the chain length of acceptable carboxylate substrates. By combing two mutations (Val$^{399}$Ala, and Trp$^{427}$Gly) the enzyme was able to utilize butyrate with a catalytic efficiency similar to the wild-type enzyme with acetate. Circular dichroism (CD) was used to evaluate the secondary structure of the wild-type atACS and the mutated variants. The CD spectra showed no difference between the mutated variants and the wild-type and indicated the enzyme is largely composed of α-helices.
CHAPTER 1

INTRODUCTION

Metabolic engineering is the use of modern genetic techniques along with bioreactor, biochemical, and mathematical models to systematically manipulate the metabolic activity of living cells (Bailey, 1996). These techniques have largely been used for the commercial production of compounds that either cannot be produced, or it is economically impossible to produce them via synthetic chemistry (amino acids, proteins, antibiotics, vitamins, etc.) (Stephanopoulos and Vallino, 1991). An important part of most metabolic engineering endeavors is the manipulation of enzymes either genetically or biochemically to contribute new or altered end-products or to alter activity to increase or change chemical yields. The use of metabolic engineering offers a far more rational approach to strain design over the more traditional approach of mutagenesis and screening. These strategies may take advantage of one or several different techniques including heterologous protein expression, widening of available substrate ranges, introduction of new pathways for unique products, improvement of cellular physiology, and improvement in yield or production (Nielson, 2001). Rationally manipulating enzymes, by increasing activity, expanding or changing acceptable substrates, or further modifying function can play an important role in developing useful metabolic pathways.

In this work we describe work toward understanding the Arabidopsis thaliana acetyl-CoA synthetase (atACS), an enzyme that has potential in the developing field of biorenewable chemical production based on a metabolic engineering approach. The Center for Biorenewable Chemicals (CBiRC, http://www.cbirc.iastate.edu/) is a National
Science Foundation-funded engineering research center focused on utilizing a multidisciplinary approach to building a foundation for transforming industrial chemical production from a petroleum-based industry to a renewable resource-based industry (http://www.cbirc.iastate.edu/overview/mission/). This is possible by altering two interconnected pathways that are already common in all organisms: fatty acid synthesis and polyketide synthesis. By slightly modifying the types of chemicals that can be formed from these pathways, economically useful chemicals can be formed that share chemistries with currently used petrochemicals (Nikolau et al., 2008).

Acyl-CoA synthetases, the larger family that atACS belongs to, are of interest because of the role that acyl-CoA synthetases play in the fatty acid/polyketide synthesis pathways. In both pathways the priming molecule is commonly a short (2-4 carbon) CoA-bound acyl molecule. Some organisms can however, utilize other priming substrates, including branched chain-CoA’s (i.e. isobutyryl-CoA) or even aromatic acyl-CoA molecules (i.e. benzoyl-CoA) that can contribute benzyl or phenyl groups into the final product (Rawlings, 2001). These “priming” molecules contribute the initial carbon that will be elongated through reiterative additions of two carbons, and can also be carboxylated themselves to form the “elongating” molecules. Thus, by altering the constituents in the “CoA pool” a variety of molecules may be produced. In this way the atACS enzyme may be capable of providing an important component to the foundation of a biorenewable chemicals industry.

In this work I describe the biochemical characterization of the atACS enzyme and identify the residues that form the carboxylate binding pocket in the enzyme. Mutated variants of the enzyme with one or more of the putative binding pocket residues altered to
resemble the homologous residue in an isobutyryl-CoA synthetase from *Pseudomonas chlororaphis* are used to investigate the role of binding pocket residues in determining the preferred carboxylate substrates as well as enzymatic function. Four residues are identified: Ile$_{323}$, Thr$_{324}$, Val$_{399}$, and Trp$_{427}$, and in particular Val$_{399}$ and Trp$_{427}$ show effects on carboxylate substrate specificity, ATP and CoA utilization, and enzyme turnover number. In these studies I have also investigated the overall secondary structure of the enzyme and have compared data to known acetyl-CoA synthetase structures. These studies provide important information for understanding acetyl-CoA synthetases as well as providing knowledge for enzyme engineering within a metabolic engineering strategy for biorenewable chemical production.

**Thesis Organization**

This thesis is arranged into three chapters. The first chapter is a general introduction to enzyme and metabolic engineering and the role of acetyl-CoA synthetase enzyme engineering in the Center for Biorenewable Chemicals’ (CBiRC) vision for biorenewable chemical production. The second chapter is a manuscript to be submitted to *The Archives of Biochemistry and Biophysics* detailing the kinetic characterization of the *Arabidopsis thaliana* acetyl-CoA synthetase enzyme and describing the putative carboxylate binding pocket. The third chapter summarizes the significant findings of this work and provides a general discussion of the acetyl-CoA synthetase’s place within the engineered metabolic pathway for biorenewable chemical production.
CHAPTER 2
CHARACTERIZATION OF THE ARABIDOPSIS THALIANA ACETYL-COA SYNTHETASE PUTATIVE CARBOXYLATE BINDING POCKET

A manuscript to be submitted to The Archives of Biochemistry and Biophysics

Jason H. Hart\textsuperscript{1}, David J. Oliver\textsuperscript{2}, Basil J. Nikolau\textsuperscript{1,3}

Abstract

Acetyl-CoA synthetase catalyzes the activation of acetate by the acetylation of the thiol group of Coenzyme A, while hydrolyzing ATP to AMP and pyrophosphate. The Arabidopsis thaliana acetyl-CoA synthetase (atACS) was compared to other acyl-CoA synthetases, and was computationally modeled on the available crystal structures of the Saccharomyces cerevisiae ACS1 and Salmonella enterica ACS. This allowed the identification of the residues that make up the putative carboxylate binding pocket residues. To further understand substrate selectivity and binding within the putative carboxylate binding pocket, selected residues were mutated to resemble the homologous residues in the Pseudomonas chlororaphis isobutyryl-CoA synthetase. Four residues (Ile\textsuperscript{323}, Thr\textsuperscript{324}, Val\textsuperscript{399}, and Trp\textsuperscript{427}) were identified that are proposed to form the carboxylate binding pocket.

\textsuperscript{1}Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA 50011, USA
\textsuperscript{2}Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA 50011, US
\textsuperscript{3}Corresponding Author
One residue, Trp$^{427}$ was found to be the primary residue in determining the chain length of acceptable carboxylate substrates. By combing two mutations (Val$^{399}$Ala, and Trp$^{427}$Gly) the enzyme was able to utilize butyrate with a catalytic efficiency similar to the wild-type enzyme with acetate. Circular dichroism (CD) was used to evaluate the secondary structure of the wild-type atACS and the mutated variants. The CD spectra showed no difference between the mutated variants and the wild-type and indicated the enzyme is largely composed of $\alpha$-helices.

**Introduction**

The *Arabidopsis thaliana* AMP-forming acetyl-CoA synthetase (E.C. 6.2.1.1, acetate-CoA ligase, designated as atACS) catalyzes the acylation of the thiol of a Coenzyme A molecule with acetate, to activate the latter to acetyl-CoA. This reaction is a two-step reaction involving the formation of an acetyl-AMP intermediate in the first step, and thioesterification of the acetate to CoA in the second (Berg, 1956). All AMP-forming acetyl-CoA synthetases belong to the adenylate-forming enzyme superfamily containing three subfamilies: the acyl- and aryl-CoA synthetases, nonribosomal peptide synthetases, and firefly luciferase (Conti et al., 1996; Chang et al., 1997; Kleinkauf and Von Döhren, 1996; and Turgay et al., 1992).

Within the acyl-CoA synthetase family, the most common preferred carboxylate substrate is acetate, but there are enzymes within the family that are capable of utilizing different substrates, including propionate, or branched chain carboxylates, such as isobutyrate (Horswill and Escalante-Semerena, 1999; Hashimoto et al., 2005). The
crystal structures of two acetyl-CoA synthetases have been determined; that of the *S. cerevisiae* acetyl-CoA synthetase1 (ACS1, Jogl and Tong, 2004) and *S. enterica* acetyl-CoA synthetase (seACS) (Gulick et al., 2003). The seACS crystal structure was solved in the presence of Coenzyme A and adenosine 5’-propylphosphate, which acts as an inhibitor but is an appropriate analog to the acetyl-adenylate intermediate that is formed in the first half-reaction (Grayson and Westkaemper, 1988). For this reason the adenosine 5’-propylphosphate is useful for approximating the conformation of the carboxylate substrate within the enzyme binding pocket. Ingram-Smith *et al.* (2006) described a hydrophobic cavity formed between the N- and C-terminal domains of the *Methanothermobacter thermoautotrophicus* acetyl-CoA synthetase (MT-ACS) that corresponds to residues found in close proximity to the substrate positions in the *S. enterica* ACS crystal structure and are predicted to form the MT-ACS binding pocket. Mutations of these residues led to altered enzymatic function and in some cases these mutations were able to shift the preferred substrate of the enzyme away from acetate to butyrate (Ingram-Smith *et al.*, 2006).

The characterization of atACS is of particular interest because the only plant acetyl-CoA synthetase to be enzymatically characterized is the *Pisum sativum* acetyl-CoA synthetase (Behal *et al.*, 2002) and the role of this enzyme in plant physiology and metabolism is still unclear. Acetyl-CoA synthetase provides the opportunity to efficiently form acetyl-CoA for fatty acid synthesis in leaves, but the level of acetate within leaves does not appear to be adequate to provide sufficient carbon-flux for fatty acid synthesis (Bao *et al.*, 2000, Behal *et al.*, 2002). It is currently proposed rather, that acetyl-CoA
synthetase is more important in plants for acetate detoxification than fatty acid synthesis (Lin and Oliver, 2008).

In this study we investigate the kinetic characteristics of the wild-type atACS enzyme as well as the changes to these properties as a result of the mutation of four residues predicted to form the carboxylate binding pocket. From these mutagenesis experiments we have found that Ile\textsuperscript{323}, Thr\textsuperscript{324}, Val\textsuperscript{399}, and Trp\textsuperscript{427} are important for enzyme function. Furthermore, Val\textsuperscript{399} and Trp\textsuperscript{427} appear to play important roles in the determination of carboxylate specificity as well as ATP and Coenzyme A binding.

**Materials and Methods**

**Bacterial Strains, Culture Media and Growth Conditions**

*E. coli* Top10 competent cells were obtained from Life Technologies (Carlsbad, CA) and used for the stable propagation of plasmids and long term storage as glycerol frozen stocks. *E. coli* ArcticExpress competent cells were used for overexpression of protein and were obtained from Agilent Technologies (Santa Clara, CA). *E. coli* XL1-Blue supercompetent cells (Agilent Technologies, Santa Clara, CA) were used for transformation of mutagenized plasmids containing the atACS gene. All strains were cultured at 37° C in LB media (Bertani, 1951), unless stated otherwise. Kanamycin and gentamycin were used for selection at 50 μg/ml and 20 μg/ml, respectively, in liquid media and agar plates. All chemicals were obtained from Sigma-Aldrich Co. (St. Louis, MO) unless stated otherwise.
Overexpression and purification of atACS proteins in *E. coli*

High-level expression was performed with the ArcticExpress (DE3) strain transformed with pET24b plasmid vector (Merck KGaA, Darmstadt, Germany), containing either ecACS, ecPCS, pcICS, atACS or one of the 5 mutated versions of atACS. Cells were transformed via the protocol provided by Agilent Technologies. Transformed cells were grown overnight in 5-10 ml LB medium containing 50 μg/ml kanamycin and 20 μg/ml gentamycin at 37°C with shaking at 250 rpm. The overnight culture was used to inoculate 500-1000 ml LB media containing no antibiotics and was grown at 30°C with shaking until an OD<sub>600</sub> of ~0.6 was reached. Expression was then induced by the addition of IPTG at a final concentration of 0.1 mM, and the culture was grown at 13°C with shaking for 24-48 hours.

Cells were harvested by centrifugation at 10,000g for 10 minutes and resuspended in a 20 mM Tris-HCl buffer (pH 7.5) containing 500 mM NaCl, 5 mM imidazole, 0.1% (v/v) Triton X-100, 0.1 mg/ml phenyl methyl sulfonyl fluoride (PMSF) and 10 ul/ml Protease Inhibitor Cocktail (Sigma-Aldich Co., St. Louis, MO) and then disrupted by sonication on ice using 5 bursts of 30 seconds each at 70% intensity to produce a cell-free extract. After centrifugation at 20,000g for 30 minutes the supernatant was filtered through a 0.45 μm filter disc (Corning Inc., Corning, NY) and applied to a column containing 5 ml of PerfectPro Ni-NTA agarose (5 Prime, Inc., Gaithersburg, MD) at 4°C. The column was washed with increasing concentrations of imidazole and the protein was finally eluted using 50 ml of 200 mM imidazole. Samples were immediately dialyzed into 50 mM Tris-HCl pH 7.5 at 4°C. For long-term storage glycerol was added to a final concentration of 20%, and aliquots were flash frozen with liquid nitrogen and stored at -80°C.
The concentration of purified protein was determined by NanoDrop (Thermo Fisher Scientific Inc., Waltham, MA) and the purity was analyzed by SDS-PAGE using 10% polyacrylamide gels and staining by Coomassie Brilliant Blue. Identification of the protein was performed by western blot analysis using 1:1000 dilutions of mouse anti-His antibody followed by 1:3000 dilutions of goat anti-mouse IgG-HRP. Detection was achieved by chemiluminescence with ELC western blotting reagents (GE Healthcare Biosciences, Pittsburgh, PA) and images were obtained and analyzed using a Bio-Rad ChemiDoc™ XRS+ system with Image Lab™ software (Bio-Rad Laboratories Inc., Hercules, CA).

**Site-Directed Mutagenesis**

Initial mutagenesis of putative binding pocket residues was performed by recombination PCR in a two-step fashion as described in Guo and Oliver (2012). pET 24b-atACS was used as template for the first round of PCR, and the PCR products from the first round were used as template for the second round. After the full PCR product was purified it was cloned into a pGEM T-Easy vector and transformed into Top-10 competent *E. coli* cells. The vectors with correct mutations were digested with *HindIII* and *XhoI* and cloned into the corresponding restriction sites in the pET24b vector for expression.

Sequencing was performed to ensure accurate sequence and reading frame within the altered atACS genes in pET24b. Sequencing was performed by the DNA Facility at Iowa State University (http://www.dna.iastate.edu/) using an Applied Biosystems 3730xl DNA analyzer (Life Technologies, Carlsbad, CA).
After sequencing of the ACS genes in the pET24b vector, several unintended mutations were discovered at sites outside of the predicted binding pocket. QuikChange II site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA) was used to return these residues to the original sequence. Mutagenic primers were designed at between 30-45 nucleotides with the altered base at the center of the sequence and amplification and transformation was performed per the provided standard protocol. The mutagenized atACS sequences were confirmed by sequencing at the Iowa State University DNA Facility. All necessary primers were synthesized by the Iowa State University DNA Facility.

**Spectrophotometric atACS Activity Assay for Kinetic Characterization**

Enzymatic activity was observed by coupling the AMP formation of the acyl-CoA synthetase with myokinase, which is capable of forming 2 moles of ADP from 1 mole ATP and 1 mole AMP. The ADP formed can be used by Pyruvate Kinase in the formation of pyruvate from phospho(enol)pyruvate. Lactate Dehydrogenase converts pyruvate to lactate while utilizing NADH as a reducing agent. NADH absorbs at 340 nm and the decrease in absorption is directly correlative to a decrease in NADH with a molar extinction coefficient of 6,220 M$^{-1}$cm$^{-1}$. Assays were performed in 96-well format at 37$^\circ$C with a final well volume of 100 μl. The standard reaction contained 50 mM Tris-HCl at pH 7.5; 5mM Tris(2-carboxylethyl)phosphine hydrochloride (Thermo Fisher Scientific Inc., Waltham, MA), a thiol-free reductant for reducing peptide disulfide bonds (pH 7.5); 5% (v/v) ethanol, to prevent bubble formation that could interfere with spectrophotometric readings; 6mM MgCl$_2$; 5 mM phospho(enol)pyruvate; 0.4 mM NADH; 20 U Pyruvate Kinase/Lactate Dehydrogenase (Sigma-Aldrich Co., St. Louis,
MO); 10U Myokinase from rabbit muscle (Sigma-Aldrich Co., St. Louis, MO); and between 2.5-5 μg of purified ACS per reaction. Concentrations of the three substrates (CoA, ATP-Mg\(^{2+}\), and carboxylic acids) are variable and in the case of determination of apparent kinetic parameters one substrate is varied while the other two are held constant at saturating levels. Assays were performed on a BioTek ELx808\(^{TM}\) Absorbance Microplate Reader using Gen5\(^{TM}\) Data Analysis software (BioTek Instruments, Winooski, VT).

**Computational Experiments**

Structural predictions of the *E. coli* ACS and PCS proteins, *P. chlororaphis* ICS protein, and *A. thaliana* ACS protein were performed at the I-TASSER server (http://zhanglab.ccmb.med.umich.edu/I-TASSER, Zhang, 2008; Roy *et al.* 2010; and Roy *et al.*, 2012). Predicted structures and known crystal structures were visualized and analyzed using PyMOL (http://www.pymol.org, The PyMOL Molecular Graphics System, Version 1.3.0 Schrödinger, LLC.). Multiple sequence alignments were created using Clustal W (Thompson *et al.*, 1994) through Biology Workbench 3.2 (http://workbench.sdsc.edu, Subramaniam, 1998). The chloroplast targeting sequence was identified in the *A. thaliana* ACS by TargetP 1.1 (http://www.cbs.dtu.dk/services/TargetP, Emanuelsson *et al.* 2000 and Nielsen *et al.* 1997). Amino acid similarity and identity values of homologous proteins to the *A. thaliana* ACS protein were obtained using BLASTp (http://blast.ncbi.nlm.nih.gov) with the *A. thaliana* ACS amino acid sequence as the search subject. Secondary structure predictions based on amino acid sequences were performed using PSSPred (http://zhanglab.ccmb.med.umich.edu/PSSpred/)
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 190-250 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 ddH₂O 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 ([θ], degrees cm² dmol⁻¹). Spectra were analyzed using a suite of algorithms collectively called CDPro (Sreerama and Woody, 2000), utilizing algorithms: SELCON3, CDSSTR, and CONTIN/LL.

Results

Identification of the Residues Forming the Putative atACS Binding Pocket

In an effort to understand the binding properties of acyl-CoA synthetases, specifically the *Arabidopsis thaliana* acetyl-CoA synthetase (At5g36880, atACS), a number of acyl-CoA synthetase amino acid sequences were found using Blastp (http://blast.ncbi.nlm.nih.gov) and were aligned using Clustal W (Thompson *et al.*, 1994) (Figure 1). The ACSs that are known to use acetate as the primary substrate all show higher protein similarity (>60%) to the atACS, while those utilizing propionate or
isobutyrate show lower identity and similarity (Table 1). The *Escherichia coli* acetyl-CoA synthetase (ecACS), *Escherichia coli* propionyl-CoA synthetase (ecPCS), and *Pseudomonas chlororaphis* isobutyryl-CoA synthetase (pcICS), along with atACS, were modeled at the I-TASSER server. The templates used in forming the predicted structures were the *S. enterica* ACS (PDB: 1PG4) (Reger et al. 2007) and the *S. cerevisiae* ACS (PDB: 1RY2) (Jogl and Tong, 2004). The predicted models show similar overall structure (Figure 2). The crystal structure of the seACS that was used in these modeling experiments has been solved in the presence of CoA and adenosine 5’-propylphosphate, which mimics the acetyl-AMP intermediate of the ACS-catalyzed reaction and can approximate the location of the acetate molecule within the binding pocket while bound to AMP. A hydrophobic pocket, described by Ingram-Smith *et al.* (2006), surrounds the propyl-AMP intermediate as well as the thiol-group of the CoA molecule. The residues in the atACS protein that line the homologous pocket are Ile$^{323}$, Thr$^{324}$, Val$^{399}$, and Trp$^{427}$. The sequence alignment in Figure 1 shows that these specific residues are highly conserved among the acyl-CoA synthetase proteins that share the same preferred carboxylic acid substrate. It is proposed therefore, that these residues make up the carboxylate binding pocket in atACS. The predicted structures for atACS, ecACS, ecPCS, and pcICS show that these residue positions align well with each other and also with the seACS crystal structure (Figure 3). The propyl group of the propyl-AMP molecule in the seACS crystal structure appears to be coordinated by the wall of the hydrophobic space that is formed by Thr$^{324}$. The presence of a tyrosine residue at that position in the pcICS protein appears to sterically interfere with the propyl group in the predicted structure, suggesting the binding pocket shape is altered considerably from that
of the seACS binding pocket (Figure 4). Similarly, residue Trp$^{427}$ is conserved in all acyl-CoA synthetases other than pcICS, in which it is replaced by a glycine residue. Based on these results, we predict the Trp$^{427}$ residue in atACS plays a role in narrowing the possible substrates by limiting the available space in the binding pocket.

**Mutagenesis of Binding Pocket Residues**

Based on these sequence alignments and the predicted modeled structures the four residues that form the putative binding pocket in atACS were targeted for mutagenesis to resemble the residues that make up the binding pocket in pcICS. The latter is capable of utilizing isobutyrate as a substrate, whereas for the former enzyme isobutyrate is a poor substrate. Thus, the prediction is that these specific changes will alter the substrate specificity of atACS to resemble that of pcICS. Five mutant variants were made to shift the binding pocket residues of atACS to the residues found in the pcICS. These mutations are Val$^{399}$Ala; Trp$^{427}$Gly; the double mutant Ile$^{323}$Ala, Thr$^{324}$Tyr; the double mutant Val$^{399}$Ala, Trp$^{427}$Gly; and the quadruple mutant Ile$^{323}$Ala, Thr$^{324}$Tyr, Val$^{399}$Ala, Trp$^{427}$Gly.

**Expression and Purification of atACS and Variants**

The wild-type atACS and mutated variants were all expressed in ArcticExpress *E. coli* (DE3) using the pET24b vector containing a C-terminal 6x His-tag. After purification using a Ni-NTA column, expression and purity was confirmed using SDS-PAGE (Figure 5). No difference in molecular weight was observed and all proteins showed an acceptable level of purity (>95% pure).
Kinetic Parameters of Wild-type atACS and the Val<sup>399</sup>Ala Variant

All kinetic parameters were determined using the assay diagramed in Figure 6. This is a general assay for AMP production, so the specific enzyme or carboxylate substrate can be varied while all other factors are kept constant for determination of parameters pertaining to carboxylate substrate utilization. The wild-type atACS follows a traditional Michaelis-Menton curve (Figure 7) and was used to confirm that the assay design provides an appropriate amount of accuracy and sensitivity to measure kinetic rates reliably. The wild-type enzyme shows a strong preference for acetate over propionate evidenced by a catalytic efficiency ($k_{cat}/K_m$) approximately 55-fold higher with acetate than propionate (Figure 8c). The Val<sup>399</sup>Ala mutation increases the $K_m$ for acetate as compared to wild-type enzyme by more than 10-fold (Figure 8a), but the binding affinity for propionate is similar between wild-type and Val<sup>399</sup>Ala. The Val<sup>399</sup>Ala variant also shows a minimal increase in turnover number ($k_{cat}$) with both acetate (1.2-fold higher than wild-type) and propionate (4-fold higher than wild-type) as substrates (Figure 8b). The increase in $k_{cat}$ results in an increased catalytic efficiency for Val<sup>399</sup>Ala with propionate as compared to the wild-type enzyme (8-fold higher). The Val<sup>399</sup>Ala variant does not show a discernible preference for either acetate or propionate based on catalytic efficiency, however (Figure 8c). In both the wild-type and Val<sup>399</sup>Ala variant no activity was observed with any branched-chain carboxylate substrates or straight-chain carboxylate substrates longer than propionate (Supplemental Table 1).

Kinetic Parameters of the Trp<sup>427</sup>Gly and the Val<sup>399</sup>Ala, Trp<sup>427</sup>Gly Variants

The Trp<sup>427</sup>Gly mutation was made in atACS to reduce the size of the side chain of the residue that appears to limit the size of the active site pocket, and to resemble the
residue that is in this position for pcICS. This mutation was also made in combination with the Val$^{399}$ Ala mutation to create a double mutant. The Trp$^{427}$Gly single mutant and Val$^{399}$Ala, Trp$^{427}$Gly double mutant variants both show a marked increase in $K_m$ values over wild-type with both acetate and propionate (Figure 9a). The higher $K_m$ is, however, accompanied by an increase in turnover number, especially in Trp$^{427}$Gly mutant with acetate as the substrate (approximately 6-fold increase over wild-type) and Val$^{399}$Ala, Trp$^{427}$Gly with propionate as the substrate (>18-fold increase over wild-type) (Figure 9b). In both variants the catalytic efficiency with acetate is greatly decreased as compared to that of the wild-type enzyme. In the Val$^{399}$Ala, Trp$^{427}$Gly double mutant variant the catalytic efficiency with propionate is modestly increased compared to the wild-type enzyme (2.5-fold increase) (Figure 9c).

The variants containing the Trp$^{427}$Gly mutation are unique in that they gain the ability to utilize straight-chain carboxylate substrates that are longer than propionate. Both the Trp$^{427}$Gly single mutant and Val$^{399}$Ala, Trp$^{427}$Gly double mutant variants are capable of utilizing carboxylate substrates up to eight carbons in length (i.e., octanoate) (Figure 10). With all substrates longer than acetate the Val$^{399}$Ala, Trp$^{427}$Gly double mutant showed a lower $K_m$ value than the Trp$^{427}$Gly single mutant. Both variants showed the lowest $K_m$ with butyrate (2.88 mM ± 0.58 for Trp$^{427}$Gly, 1.5 mM ± 0.14 for Val$^{399}$Ala, Trp$^{427}$Gly) (Figure 10a), but these values are still higher than the $K_m$ for acetate with the wild-type enzyme (0.27 mM ± 0.01). The Trp$^{427}$Gly containing variants did, however, show an increase in turnover number with all substrates as compared to wild-type enzyme with acetate as the substrate.
These findings indicate that the increased size of the binding pocket enabled by the removal of the bulky Trp$^{427}$ side chain is capable of not only allowing the utilization of larger substrates but also increases the rate of turnover in the enzyme catalyzed reaction. The catalytic efficiencies of the Trp$^{427}$Gly single mutant and Val$^{399}$Ala, Trp$^{427}$Gly double mutant variants are both highest when utilizing butyrate (Figure 10c) and the catalytic efficiency of Val$^{399}$Ala, Trp$^{427}$Gly with butyrate is similar to that of the wild-type enzyme with acetate as the substrate (Figure 11).

**Utilization of ATP and CoA as Substrates by atACS Wild-type and Variants**

Enzyme activity assays were performed with the wild-type and all variants of atACS with varying concentrations of ATP and CoA, and from this data $K_m$ and $V_{max}$ values for these substrates were determined. All determinations were performed with saturating amount of the carboxylate substrate that is most preferred for the individual variant enzyme, based on catalytic efficiency. In all experiments 6 mM MgCl$_2$ was used. Based on the structures shown in Figure 4, it is predicted that Val$^{399}$ may be an important residue for establishing the junction between the channels by which ATP and CoA are expected to enter the binding pocket, while the Trp$^{427}$ residue may help coordinate the acetyl-AMP intermediate within the binding pocket. It appears that the shift to a slightly smaller residue at residue Val$^{399}$ is able to decrease the $K_m$ for both ATP and CoA ($>3$-fold and approximately 1.6-fold lower, respectively, as compared to wild-type) (Table 2). The Val$^{399}$Ala single mutant does show a lower $V_{max}$ for both ATP and CoA as compared to the wild-type enzyme, however. The Trp$^{427}$Gly variant shows an even greater impact on ATP binding ($K_m$ approximately 10-fold lower than the wild-type) but has little effect on CoA binding. The Val$^{399}$Ala, Trp$^{427}$Gly double mutant closely resembles the
Val$^{399}$ Ala single mutant in terms of $K_m$ values for both ATP and CoA, but the Val$^{399}$ Ala, Trp$^{427}$ Gly double mutant has an increased maximum rate with both ATP and CoA as compared to the Val$^{399}$ Ala single mutant. The Trp$^{427}$ Gly single mutant and Val$^{399}$ Ala, Trp$^{427}$ Gly double mutant both show higher $V_{\text{max}}$ values than wild-type with ATP and CoA, but the Trp$^{427}$ Gly single mutant shows the highest $V_{\text{max}}$ values with both (approximately 1.5-fold higher than wild-type with ATP, and approximately 2-fold higher than wild-type with CoA).

**Determination of atACS Wild-type and Variants’ Secondary Structure Using Circular Dichroism**

Circular dichroism (CD) was used to investigate if any of the point mutations caused changes in the overall secondary structure of the enzyme. This technique was also of interest to show whether the lack of activity in the Ile$^{323}$ Ala, Thr$^{324}$ Tyr double mutant and the Ile$^{323}$ Ala, Thr$^{324}$ Tyr, Val$^{399}$ Ala, Trp$^{427}$ Gly quadruple mutant was due to a change in overall protein structure or steric and chemical changes in the binding pocket. **Figure 12** shows the spectra of wild-type atACS and all variants. These spectra appear to be very similar, indicating that there is no drastic change in secondary structure in any of the variants as compared to wild-type.

In order to further analyze the spectra shown in **Figure 12**, a suite of analytical algorithms for CD was used, collectively called CDPro (Sreerama and Woody, 2000). Using the three separate algorithms within CDPro allows for an estimation of the error in the computational analysis. **Figure 13** shows the average predicted percentage for four secondary structural components ($\alpha$-helices, $\beta$-sheets, turns, and unordered portions). The error bars indicate the standard error produced by the three different analysis algorithms.
(SELCON3, CDSSTR, and CONTIN/LL). In the wild-type and all variants the average of the three algorithms predict the secondary structure to be >75% α-helices with ≤ 10% of each of the other components. Statistical analysis was done to determine if any of the variant forms differed statistically from the wild-type or from any other variant form. A student’s t-test was performed for each structural component against the same structural component in all the other variants and wild-type. The only difference observed was in the percentage of β-sheet between the Val^{399} Ala single mutant and the quadruple mutant (p-value=0.046). All other variants showed no statistical differences at a p-value of 0.1. The CD spectra indicate that there is insignificant difference in the folding of the enzyme in the mutated variants, and the changes in the kinetic characteristics are indicative of changes in the size and characteristics of the binding pocket.

**Discussion**

In this work we describe the characterization of an acetyl-CoA synthetase from *Arabidopsis thaliana* (atACS). Specifically, we investigated the residues within the active site binding pocket that are important for determining the enzyme’s preference for the carboxylate substrate. Few acyl-CoA synthetases have been biochemically characterized, and those have been, are primarily from bacterial, archael and some mammalian sources (Reger et al., 2007, Ingram-Smith et al., 2006, Fujino et al., 2001). Plant sourced acetyl-CoA synthetases have not been studied extensively in vitro; the exception being the enzyme from pea (*Pisum sativa*) (Behal et al., 2002). The K_m values for the preferred carboxylate substrate of acyl-CoA synthetase tend to be similar, in the
range typically of between 0.04 mM – 5.0 mM) (BRENDA Comprehensive Enzyme Information System, http://www.brenda-enzyme.org). The most comprehensively studied acetyl-CoA synthetases are from bacteria and archaea, S. enterica ACS (seACS) and MT-ACS (Reger et al., 2007; Ingram-Smith et al., 2006). Comparing our data for atACS to these enzymes, we find that the $K_m$ value for acetate in atACS (0.27 mM ± 0.01) is significantly lower than reported values for seACS, 41 mM, (Reger et al., 2007) and 3.5 mM for MT-ACS (Ingram-Smith et al., 2006). Interestingly, the $K_m$ value for the yeast ACS1 with acetate is 0.28 mM (Frenkel and Kitchens, 1981), which is very similar to that of atACS with acetate.

The $K_m$ values for ATP and CoA were also determined (0.2 mM ± .03, with ATP and 0.15 ± 0.02, with CoA). These values are higher than the reported values for seACS (0.07 mM, with ATP and 0.04 mM, with CoA) (Reger et al., 2007) and lower than the values for MT-ACS (3.3 mM with ATP and 0.19 mM with CoA) (Ingram-Smith et al., 2006). The reported values for $k_{cat}$ indicate that the variance between acyl-CoA synthetases from different organisms may be considerable, with the atACS appearing to have a significantly lower turnover number than seACS (Reger et al., 2007) and MT-ACS (Ingram-Smith et al., 2006).

The putative carboxylate binding pocket residues in atACS: Ile$^{323}$, Thr$^{324}$, Val$^{399}$ and Trp$^{427}$ were identified by comparing the atACS amino acid sequence to several other known acyl-CoA synthetase sequences, comparing homologous atACS residues to predicted binding pocket residues from other species (MT-ACS, Ingram-Smith et al., 2006), and computational modeling of atACS using available crystal structures as templates (seACS, Reger et al., 2007, and yeast ACS1, Jogl and Tong, 2004).
identified residues are likely to be important for enzymatic function based on the fact that they are conserved in all acyl-CoA synthetases that utilize acetate as the preferred carboxylate substrate. In this study we tested these hypotheses by site directed mutagenesis, in which each of these residues were individually, or in different combinations, changed to residues without a significant side-chain (i.e., Gly and Ala) or changed to residues that occur in the Pseudomonas chlororaphis isobutyryl-CoA synthetase, which shows preference for isobutyrate over other carboxylate substrates (Hashimoto et al., 2002).

The Val^{399} Ala mutation impacts enzyme function in a number of ways. It appears to have an effect on the acceptance of carboxylate substrates, increasing the K_m value with acetate as compared to wild-type, but causing no discernible change in the K_m with propionate. The Val^{399} Ala mutated variant has an obvious effect on the K_m values of both ATP and CoA, though. The K_m for ATP is decreased dramatically (0.2 mM ±0.03 in wild type and 0.06 mM ±0.01 in Val^{399} Ala variant). With CoA, the decrease in Km value is more modest but still apparent (0.15 mM ±0.02 in wild-type and 0.09 mM ±0.02 in Val^{399} Ala variant). The other major difference from wild-type is an increased turnover number with both acetate and propionate, including an increased catalytic efficiency for propionate as compared to wild-type (0.16 s^{-1} mM^{-1} ±0.05 in wild-type and 1.08 s^{-1} mM^{-1} ±0.10 in Val399Ala variant).

Trp^{427} appears to play the largest role in the determination of acceptable carboxylate substrates. This is consistent with similar findings in the MT-ACS by Ingram-Smith et al. (2006). The Trp^{427}Gly variant of atACS protein displays a greatly expanded library of carboxylate substrates as compared to the wild-type enzyme which
indicates that the \( \text{Trp}^{427} \text{Gly} \) mutation does indeed enlarge the binding pocket significantly and that steric interference is likely the strongest factor in determining substrate specificity. The \( \text{Val}^{399} \text{Ala, Trp}^{427} \text{Gly} \) double mutant was also able to use longer, straight-chain carboxylates, and could do so with higher catalytic efficiency. This variant was still not unable to use any branched-chain carboxylates at a level high enough to perform any determination of kinetic parameters, though. Straight-chain carboxylates longer than eight carbons were not attempted and it is possible that, especially in the \( \text{Val}^{399} \text{Ala, Trp}^{427} \text{Gly} \) double mutant variant, there may be some activity with these longer chain carboxylates. It is clear from these results that butyrate is the preferred substrate for both the \( \text{Trp}^{427} \text{Gly} \) single mutant and \( \text{Val}^{399} \text{Ala, Trp}^{427} \text{Gly} \) double mutant variants, based on \( K_m \) and \( k_{\text{cat}}/K_m \) values. The \( \text{Val}^{399} \text{Ala, Trp}^{427} \text{Gly} \) double mutation is even able to utilize butyrate with a catalytic efficiency very similar to wild-type with acetate. This is accomplished because not only does the \( \text{Trp}^{427} \text{Gly} \) mutation result in the acceptance of larger substrates but also increases the turnover number of the enzyme greatly (>5-fold higher than wild-type, both with acetate). The \( \text{Trp}^{427} \text{Gly} \) mutation also reduces the \( K_m \) value for ATP by approximately 10-fold, but has little effect on the \( K_m \) value for CoA indicating that the mutation likely opens the area in which ATP enters the binding pocket, but not the channel that CoA enters.

The combination of the \( \text{Val}^{399} \text{Ala} \) mutation and the \( \text{Trp}^{427} \text{Gly} \) mutation decreases the \( k_{\text{cat}} \) with all substrates except propionate and butyrate indicating that the \( \text{Val}^{399} \) residue may be important for high enzymatic rates. The \( \text{Val}^{399} \text{Ala, Trp}^{427} \text{Gly} \) double mutant also shows lower \( K_m \) values than the \( \text{Trp}^{427} \text{Gly} \) single mutant, with all substrates except acetate. From the position of \( \text{Val}^{399} \) in the predicted models and the seACS crystal
structure, the smaller Ala residue may allow more space within the binding pocket for acceptance of larger carboxylate substrates. This may be helpful in accepting longer-chain carboxylates and may also open the channels by which ATP and CoA enter the binding pocket, resulting in lower $K_m$ values for those substrates as well.

The Ile$^{323}$Ala, Thr$^{324}$Tyr double mutant and the Ile$^{323}$Ala, Thr$^{324}$Tyr, Val$^{399}$Ala, Trp$^{427}$Gly quadruple mutant both showed no activity with any carboxylate substrate, indicating that the residues Ile$^{323}$ and Thr$^{324}$ are essential residues required for catalysis, but more mutated variants would need to be made to determine if both residues are necessary for proper enzyme function and if other residue could function at this position. Additionally, the position of the Tyr in the Thr$^{324}$Try mutation is predicted to interfere with the position of the carboxylate while bound to the AMP molecule after the first half-reaction. More mutations of the Thr$^{324}$ residue to different or smaller amino acids may allow for enzyme catalysis and provide a larger binding pocket for larger straight-chain or bulky carboxylates (branched, hydroxylated, etc.). The lack of activity in the quadruple mutant indicates that the atACS and pcICS binding pockets must differ structurally in ways other than the four residues described here. The atACS and pcICS enzymes only share 30% identity and approximately 44% similarity though, so overall differences in the enzyme architecture, as well as the binding pockets residues, are likely contributing to substrate preference.

Circular dichroism (CD) is an important tool used in this study to investigate the overall secondary structure of the wild-type enzyme and the mutated variants. Only two experimentally determined crystal structures are known of acetyl-CoA synthetase, the *S. enterica* acetyl-CoA synthetase (Gulick *et al.*, 2002 and Reger *et al.*, 2007) and the *S.
cerevisiae acetyl-CoA synthetase1 (Jogl and Tong, 2004). These enzymes share 55% and 47% identity to atACS, respectively, and 47% identity between them, and these two structures were used for the computational modeling studies of the atACS. By performing CD we were able to compare the secondary structure of our expressed atACS, and mutated variants with the known catalogued structures. The CD spectra of wild-type atACS indicate that over 75% of the protein has a secondary structure that is α-helix and less than 10% is in β-sheet, turns, and unordered regions. This is in stark contrast to the structures found in the S. eneterica ACS (33% α-helix, 28% β-sheet, 2% turns, and 37% unordered regions) and the S. cerevisiae ACS1 (25% α-helix, 19% β-sheet, 2% turns, and 54% unordered regions). Although this may indicate that the overall structure between acyl-CoA synthetases may often be quite different, our alignment and mutagenesis studies show that the binding pocket residues are likely conserved among ACSs from different organisms. CD was used to compare the wild-type atACS and all of the mutated variants. This was helpful in determining whether any changes that were observed in the kinetic assays were the effect of discreet structural changes associated with each mutation in the binding pocket or in changes to the overall secondary structure of the enzyme. These analyses indicate that the overall secondary structure was not significantly altered in any of the mutated variants as compared to wild-type. This reassures one that the lack of activity in the Ile\textsuperscript{323}Ala, Thr\textsuperscript{324}Tyr double mutant and the Ile\textsuperscript{323}Ala, Thr\textsuperscript{324}Tyr, Val\textsuperscript{399}Ala, Trp\textsuperscript{427}Gly quadruple mutant is not due to misfolding, but rather must be a result of the Ile\textsuperscript{323}Ala and Thr\textsuperscript{324}Tyr mutations removing structural components that are needed for catalysis.
The understanding of how acyl-CoA synthetases’ substrate preferences are determined could have potential in engineering enzymatic pathways for the production of unique fatty acid products for use in biorenewable chemical production. By producing unique precursor acyl-CoA molecules, unique downstream products may be formed by polyketide/fatty acid synthesis as described by Nikolau et al. (2008). Further understanding of the binding pocket and enzymatic function may allow for the development of unique acyl-CoA synthetases that may be useful for biorenewable chemical productions.

In summary, the atACS enzyme has been characterized and shows a strong preference for acetate over all other carboxylates that were attempted. By identifying the residues that form the putative binding pocket we have made site-directed mutations to these residues to resemble the putative binding pocket of the P. chlororaphis isobutyryl-CoA synthetase, which has a strong preference for isobutyrate. Two keys residues were found, Val\textsuperscript{399} and Trp\textsuperscript{427}. When both are altered to smaller residues, it allowed for a greatly increased library of carboxylates that may be used as a substrate, and shifted the enzymes substrate preference from acetate to butyrate. Two other residues have been identified that are predicted to be important for carboxylate binding (Ile\textsuperscript{323} and Thr\textsuperscript{324}). When mutated together, these result in the loss of enzymatic activity. This work has provided a biochemical characterization of the Arabidopsis thaliana acetyl-CoA synthetase, important knowledge on the structure of the enzyme, and the identification of four residues that form the putative binding pocket and play an important role in carboxylate substrate specificity.
CHAPTER 3
DISCUSSION AND CONCLUSIONS

In this work I describe the characterization of the Arabidopsis thaliana acetyl-CoA synthetase (atACS) and the identification of residues that form the putative carboxylate binding pocket. This is important for understanding how this enzyme determines the carboxylate substrate it prefers, which is a major concern for the potential use of acyl-CoA synthetases in a metabolic engineering strategy for biorenewable chemical production. The atACS enzyme is of particular interest because it is plant sourced and there is little biochemical data about plant sourced acetyl-CoA synthetases, with the only available data from the Pisum sativa acetyl-CoA synthetase (Behal et al., 2002).

In this thesis I describe the four residues: Ile\textsuperscript{323}, Thr\textsuperscript{324}, Val\textsuperscript{399}, and Trp\textsuperscript{427} that are predicted to form the carboxylate binding pocket, and show that alterations of these residues can have great effects on the substrate specificity as well as catalytic efficiency of the enzyme. These results may be useful in future work to integrate acyl-CoA synthetases into a metabolic pathway engineered for production of molecules that may have economic importance.

The mutated variants of the atACS enzyme are interesting in their universal loss of preference for acetate, based on catalytic efficiency. This indicates that the identified binding pocket residues, which have been conserved in all enzymes that share a common preference for acetate, are likely to be necessary for the strong preference for acetate observed in the wild-type enzyme(s). However, when all four residues are mutated to
resemble the homologous residues from *Pseudomonas chlororaphis* isobutyryl-CoA synthetase, which prefers isobutyrate, activity with isobutyrate is not detected. This indicates that there are either other residues contributing to the carboxylate binding pocket, or that larger structural changes contribute to substrate utilization.

The mutated variants did show activity with novel carboxylate substrates, though. The Trp$^{427}$Gly single mutant gained the ability to use four to eight carbon straight-chain carboxylate substrates. The Val$^{399}$Ala, Trp$^{427}$Gly double mutant was able to use these substrates as well, and with greater catalytic efficiency. The Val$^{399}$Ala, Trp$^{427}$Gly double mutant was even able to achieve a catalytic efficiency with butyrate that is similar to that of the wild-type enzyme with acetate as the substrate. The primary target for acyl-CoA synthetase engineering for biorenewable chemical production is for the introduction of novel chemistries in the acyl-CoA pool. These “primer” acyl-CoA molecules may be extended via fatty acid synthase or polyketide synthase resulting in fatty acids/polyketides with more than one functional group on the product. In order to be integrated into a metabolic engineering strategy further research would be needed to determine if bulkier carboxylates could be utilized by these enzymes.

The research presented provides essential information about the atACS enzyme and confirms past research describing the carboxylate binding pocket in the *Methanothermobacter thermoautotrophicus* ACS (Ingram-Smith *et al*., 2006). These works, along with the conservation of these residues in other acetyl-CoA synthetases (Figure 1), indicates that the carboxylate binding pocket residues described here are common to all acetyl-CoA synthetases. Future work should focus on mutations to the two residues that resulted in the loss of activity in this study (Ile$^{323}$Ala and Thr$^{324}$Tyr).
Mutation to smaller residues or individual mutations at these positions may result in an active enzyme. Further investigation of the putative binding pocket by traditional structural biology techniques (i.e. X-ray crystallography) may also be useful to identify more residues that may be a factor in carboxylate substrate determination. Structural studies may at least be able to confirm or deny the difference in secondary structure between the atACS enzyme and the crystal structures available for other acetyl-CoA synteh/ses (Reger et al., 2007 and Jogl and Tong, 2004) that is predicted by computational algorithms, and confirmed by circular dichroism spectra presented here. Additional research could also be focused on determining the factors that contribute to the enzyme’s turnover number. Both the Val\textsuperscript{399}Ala and Trp\textsuperscript{427}Gly mutations resulted in an increased turnover number with all carboxylates, as compared to wild-type, and other residues may play a role in regulating turnover number as well. Improving turnover number as well as shifting carboxylate substrate specificity to a more economically useful carboxylate may allow for the useful integration of atACS into a metabolically engineered pathway for biorenewable chemical production.
REFERENCES


B.J. Nikolau, M.A.D.N. Perera, L. Brachova, B. Shanks, Plant J. 54 (2008), 536–545


K. Turgay, M. Krause, M.A. Marahiel, Mol. Microbial. 6 (1992) 529-546


M. Lin, D.J. Oliver, Plant Physiol. 147 (2008) 1822-1829

G. Bertani, J. Bacteriol. 62 (1951) 293-300


J.D. Thompson, D.G. Higgins, T.J. Gibson Nucl. Acid Res. 22 (1994) 4673-4680


S. Subramaniam, Protein 32 (1998) 1-2


H. Nielsen, J. Engelbrecht, S. Brunak, G. von Heijne Protein Eng. 10 (1997) 1-6


Figure 1: Partial sequence alignment showing the predicted carboxylate binding pocket and surrounding residues of aTACS and several acyl-CoA synthetases that utilize acetate, propionate, or isobutyrate as the preferred substrates. Alignment of amino acids was performed using Clustal W (Thompson et al., 1994). The asterisks indicate the residues proposed to be forming the carboxylate binding pocket and only the portion of sequences containing putative binding pocket residues is shown. White letters on a black background indicate identity. Gray letters on a black background indicate similarity. Black letters on a white background indicate no conservation.
Table 1: Identity and Similarity between several acyl-CoA synthetase sequences and the *A. thaliana* ACS sequence

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Identity</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. cerevisiae</em> ACS1</td>
<td>47%</td>
<td>62%</td>
</tr>
<tr>
<td><em>S. cerevisiae</em> ACS2</td>
<td>47%</td>
<td>63%</td>
</tr>
<tr>
<td><em>S. enterica</em> ACS</td>
<td>55%</td>
<td>69%</td>
</tr>
<tr>
<td><em>E. coli</em> ACS</td>
<td>55%</td>
<td>69%</td>
</tr>
<tr>
<td><em>M. thermoautotrophicus</em> ACS</td>
<td>45%</td>
<td>62%</td>
</tr>
<tr>
<td><em>E. coli</em> PCS</td>
<td>37%</td>
<td>56%</td>
</tr>
<tr>
<td><em>S. enterica</em> PCS</td>
<td>37%</td>
<td>56%</td>
</tr>
<tr>
<td><em>P. chlororaphis</em> ICS</td>
<td>30%</td>
<td>44%</td>
</tr>
</tbody>
</table>
Figure 2: Predicted protein structure of atACS (purple), ecACS (blue), ecPCS (pink), and pcICS (green). Structures predicted using the I-TASSER server and aligned and visualized with PyMOL.
Figure 3: Predicted position of the residues that make up the putative carboxylate binding pocket in atACS (purple), ecACS (blue), ecPCS (pink), and pcICS (green) aligned with the experimentally determined crystal structure of the seACS (beige). Numbers used represent the position of the homologous residue in the atACS protein.
Figure 4: Putative binding pocket residues from atACS (purple), ecACS (blue), seACS (beige), ecPCS (pink), and pcICS (green) aligned with Coenzyme A and the acetyl-AMP intermediate. SeACS residues and substrate positioning are from the experimentally determined crystal structures derived by Reger et al. (2007).
Figure 5: Purification of atACS and mutated variants. Predicted size of atACS and all variants is 72.6 kDa. All proteins preparations show >95% purity.
Figure 6: Diagram of the assay components. The conversion of NADH to NAD$^+$ occurs in the same molar ratio as the condensation of carboxylate to acyl-CoA.
Figure 7: AtACS activity is dependent on acetate concentration in a typical hyperbolic Michaelis-Menton response. The y-axis shows μmoles of NADH converted to NAD$^+$ per minute, which occurs in equal molar ratio to the conversion of acetate to acetyl-CoA. Each point is the average of three replications and the error bars are present in all data points, but in some cases the error value is too small to show graphically. All data points are normalized to a blank assay containing all assay components except acetate. Curve was fitted using GraphPad Prism ver. 6.01 for Windows using the kcat (Michaelis-Menton kinetics) function (GraphPad Software, La Jolla, CA, www.graphpad.com).
Figure 8:
A) $K_m$ values for atACS wild-type and mutated variants with acetate or propionate as substrates.

B) $k_{cat}$ values for atACS wild-type and mutated variants with acetate or propionate as substrates.

C) $k_{cat}/K_m$ values for atACS wild-type and mutated variants with acetate or propionate as substrates.

No activity was detected in the I$^{323}$A, T$^{324}$Y double mutant or I$^{323}$A, T$^{324}$Y, V$^{399}$A, W$^{427}$G quadruple mutant variants. All values were determined in assays containing saturating concentrations of the co-substrates.
Figure 9:

A) $K_m$ values for the wild-type, W$^{427}$G single mutant and V$^{399}$A, W$^{427}$G double mutant variants with acetate or propionate as substrates.

B) $K_{cat}$ values for the W$^{427}$G single mutant and V$^{399}$A, W$^{427}$G double mutant variants with acetate or propionate as substrates.

C) $K_{cat}/K_m$ values for the W$^{427}$G single mutant and V$^{399}$A, W$^{427}$G double mutant variants with acetate or propionate as substrates.

No activity was detected in the I$^{323}$A, T$^{324}$Y double mutant; or I$^{323}$A, T324Y, V399A, W427G quadruple mutant variants. All values were determined in assays containing saturating concentrations of the co-substrates.
Figure 10:

D) $K_m$ values for the W$^{427}$G single mutant and V$^{399}$A, W$^{427}$G double mutant variants with acetate, propionate, butyrate, valerate, hexanoate, heptanoate, or octanoate as substrates.

E) $K_{cat}$ values for the W$^{427}$G single mutant and V$^{399}$A, W$^{427}$G double mutant variants with acetate, propionate, butyrate, valerate, hexanoate, heptanoate, or octanoate as substrates.

F) $K_{cat}/K_m$ values for the W$^{427}$G single mutant and V$^{399}$A, W$^{427}$G double mutant variants with acetate, propionate, butyrate, valerate, hexanoate, heptanoate, or octanoate as substrates.

No activity was detected with butyrate, valerate, hexanoate, heptanoate, or octanoate in the wild-type or V$^{399}$A single mutant. No activity was detected with any substrate in the I$^{323}$A, T$^{324}$Y double mutant or I$^{323}$A, T324Y, V399A, W427G quadruple mutant variants. All values were determined in assays containing saturating concentrations of the co-substrates.
Figure 11: A comparison of the catalytic efficiency in wild-type utilizing acetate and the V^{399}A, W^{427}G double mutant utilizing butyrate.
Table 2: $K_m$ and $V_{max}$ values for ATP and CoA for atACS and the variant forms

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$K_m$ (MgATP) (mM)$^a$</th>
<th>$V_{max}$ (MgATP) (mM)$^a$</th>
<th>$K_m$ (CoA) (mM)$^a$</th>
<th>$V_{max}$ (CoA) (mM)$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild-type</td>
<td>0.20 ±0.03</td>
<td>1.47 ±0.05</td>
<td>0.15 ±0.02</td>
<td>1.44 ±0.06</td>
</tr>
<tr>
<td>V$^{399}$A</td>
<td>0.06 ±0.01</td>
<td>0.84 ±0.02</td>
<td>0.09 ±0.02</td>
<td>1.18 ±0.06</td>
</tr>
<tr>
<td>W$^{427}$G</td>
<td>0.02 ±0.00</td>
<td>2.17 ±0.11</td>
<td>0.14 ±0.05</td>
<td>3.17 ±0.41</td>
</tr>
<tr>
<td>V$^{399}$A, W$^{427}$G</td>
<td>0.06 ±0.02</td>
<td>1.92 ±0.20</td>
<td>0.09 ±0.02</td>
<td>2.50 ±0.24</td>
</tr>
</tbody>
</table>

$^a$ Kinetic values were obtained in assays containing saturating concentrations of the cosubstrates.
Figure 12: Comparison of CD spectra from atACS and the mutated variants. Measurements are an average of three scans with a reading every 0.2 nm. Units are in molar ellipticity $[\theta]$, to correct for small changes in protein concentration during spectra acquisition.
Figure 13: Determination of secondary structure components by CDPro analysis of CD data. Three different algorithms were used to predict the percentage of secondary structure based on CD spectra. Error bars indicate the standard error between algorithms.
## Supplemental Table 1: Kinetic parameters of atACS and the mutated variants

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate</th>
<th>$K_m$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_m$ (s$^{-1}$ mM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>atACS WT</td>
<td>Acetate</td>
<td>0.27 ±0.01</td>
<td>2.41 ±0.03</td>
<td>8.91 ±0.48</td>
</tr>
<tr>
<td></td>
<td>Propionate</td>
<td>3.99 ±1.16</td>
<td>0.66 ±0.07</td>
<td>0.16 ±0.05</td>
</tr>
<tr>
<td></td>
<td>Isobutyrate</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>atACS V$^{399}$A</td>
<td>Acetate</td>
<td>3.07 ±0.23</td>
<td>3.04 ±0.05</td>
<td>0.99 ±0.07</td>
</tr>
<tr>
<td></td>
<td>Propionate</td>
<td>3.12 ±0.36</td>
<td>2.65 ±0.07</td>
<td>0.85 ±0.10</td>
</tr>
<tr>
<td></td>
<td>Isobutyrate</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>atACS W$^{427}$G</td>
<td>Acetate</td>
<td>10.79 ±1.80</td>
<td>14.29 ±0.88</td>
<td>1.32 ±0.24</td>
</tr>
<tr>
<td></td>
<td>Propionate</td>
<td>37.51 ±5.02</td>
<td>6.18 ±0.45</td>
<td>0.16 ±0.03</td>
</tr>
<tr>
<td></td>
<td>Butyrate</td>
<td>2.88 ±0.58</td>
<td>9.42 ±0.54</td>
<td>3.27 ±0.68</td>
</tr>
<tr>
<td></td>
<td>Valerate</td>
<td>6.47 ±0.73</td>
<td>13.79 ±0.42</td>
<td>2.13 ±0.25</td>
</tr>
<tr>
<td></td>
<td>Hexanoate</td>
<td>12.61 ±1.42</td>
<td>13.03 ±0.47</td>
<td>1.03 ±0.12</td>
</tr>
<tr>
<td></td>
<td>Heptanoate</td>
<td>10.01 ±1.76</td>
<td>11.0 ±0.58</td>
<td>1.10 ±0.20</td>
</tr>
<tr>
<td></td>
<td>Octanoate</td>
<td>48.24 ±11.79</td>
<td>5.78 ±0.84</td>
<td>0.12 ±0.03</td>
</tr>
<tr>
<td></td>
<td>Isobutyrate</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>atACS I$^{323}$A, T$^{324}$Y</td>
<td>Acetate</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Propionate</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Isobutyrate</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>atACS V$^{399}$A, W$^{427}$G</td>
<td>Acetate</td>
<td>44.23 ±5.91</td>
<td>5.07 ±0.31</td>
<td>0.11 ±0.02</td>
</tr>
<tr>
<td></td>
<td>Propionate</td>
<td>31.21 ±4.97</td>
<td>12.94 ±1.07</td>
<td>0.41 ±0.07</td>
</tr>
<tr>
<td></td>
<td>Butyrate</td>
<td>1.50 ±0.14</td>
<td>14.87 ±0.57</td>
<td>9.93 ±1.01</td>
</tr>
<tr>
<td></td>
<td>Valerate</td>
<td>3.98 ±0.56</td>
<td>12.25 ±0.47</td>
<td>3.08 ±0.45</td>
</tr>
<tr>
<td></td>
<td>Hexanoate</td>
<td>7.25 ±0.87</td>
<td>9.20 ±0.31</td>
<td>1.27 ±0.16</td>
</tr>
<tr>
<td></td>
<td>Heptanoate</td>
<td>5.80 ±0.86</td>
<td>9.12 ±0.43</td>
<td>1.57 ±0.24</td>
</tr>
<tr>
<td></td>
<td>Octanoate</td>
<td>6.32 ±1.99</td>
<td>3.04 ±0.30</td>
<td>0.48 ±0.16</td>
</tr>
<tr>
<td></td>
<td>Isobutyrate</td>
<td>b</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Isovalerate</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-Methylvalerate</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>atACS I$^{323}$A, T$^{324}$Y, V$^{399}$A, W$^{427}$G</td>
<td>Acetate</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Propionate</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Isobutyrate</td>
<td>ND</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a* Kinetic values were obtained in saturating concentrations of cosubstrates.

*b* Activity was observed at high concentrations of substrate but was too low for determination of kinetic parameters.

ND, Activity not detectable.