Received: 9 May 2023 | Revised: 11 September 2023 | Accepted: 19 September 2023

DOI: 10.1002/prot.26599

REVIEW ARTICLE

PROTEINS WILEY

Acyltransferase families that act on thioesters: Sequences, structures, and mechanisms

Caio C. de Carvalho[1] | Ian P. Murray[1] | Hung Nguyen[2] | Tin Nguyen[1,2] | David C. Cantu[1]

[1]Department of Chemical and Materials Engineering, University of Nevada, Reno, Reno, Nevada, USA

[2]Department of Computer Science and Software Engineering, Auburn University, Auburn, Alabama, USA

Correspondence
David C. Cantu, Department of Chemical and Materials Engineering, University of Nevada, Reno, Reno, Nevada 89557, USA.
Email: dcantu@unr.edu

Funding information
National Science Foundation

Abstract

Acyltransferases (AT) are enzymes that catalyze the transfer of acyl group to a receptor molecule. This review focuses on ATs that act on thioester-containing substrates. Although many ATs can recognize a wide variety of substrates, sequence similarity analysis allowed us to classify the ATs into fifteen distinct families. Each AT family is originated from enzymes experimentally characterized to have AT activity, classified according to sequence similarity, and confirmed with tertiary structure similarity for families that have crystallized structures available. All the sequences and structures of the AT families described here are present in the thioester-active enzyme (ThYme) database. The AT sequences and structures classified into families and available in the ThYme database could contribute to enlightening the understanding acyl transfer to thioester-containing substrates, most commonly coenzyme A, which occur in multiple metabolic pathways, mostly with fatty acids.

KEYWORDS
acyltransferases, coenzyme A, enzyme families, enzyme structures

1 | INTRODUCTION

Acyltransferases (AT) are enzymes that catalyze the transfer of an acyl group from a donor molecule to a variety of biomolecules as triacylglycerols, phospholipids, sphingolipids, cholesterol, acetylcholine, prenyl moieties, bile acids, ketone bodies, heme, melatonin, glycosaminoglycans, glycoproteins, gangliosides, proteoglycans, and others.[1,2] The ATs are a part of fatty acid biosynthesis, where they are usually activated for subsequent reactions by the esterification of their carboxyl groups with the thiol group of coenzyme A (CoA) (Figure 1). The chemical structure of the thioester bond CoA-SH allows it to activate carboxylic acids involved in catabolic and anabolic reactions. This review focuses on the ATs that act on thioester bonds, mostly those related to CoA-including substrates. Some examples of reactions involving the thioester bond CoA-SH include lipogenesis, triacylglycerol synthesis, carnitine shuttle, synthesis of cholesterol, and protein acetylation.[3]

Most ATs that act on thioester-containing substrates are CoA-transferases (E.C 2.8.3.-), which have been grouped into three classes based on their catalytic mechanism. Class I CoA-transferases utilize the ping–pong bi–bi mechanism involving the formation of a covalent acyl-enzyme intermediate that can acylate organic nucleophiles like alcohols, amines, and thiols.[4] In this mechanism, for example, a glutamate residue of a glutaconate CoA-transferase from *Acidaminococcus fermentans* works as the acceptor of covalently bound intermediates. The first partial reaction is a nucleophilic attack of the active site glutamate at the CoA-thioester substrate to yield an enzyme-bound acyl-glutamyl.[5,6] Structurally, Class I enzymes belong to the NagB/RpiA/CoA transferase-like superfamily of open α/β-proteins.

Class II CoA-transferases catalyze a partial reaction in citrate or citramalate lyases enzyme complexes. These lyases consist of three subcomplexes, (1) CoA-transferase (α subunit), (2) a lyase (β subunit), and (3) an acyl carrier protein (ACP, γ subunit). They also contain ACP-like γ subunits with 2′-(5″-phosphoribosyl)-3′-dephospho-CoA prosthetic groups, catalytic β subunits, which promote the retro-Claisen cleavage of citrate and citramalate, respectively. During the lyase reaction, the CoA-transferases (EC 2.8.3.10, 2.8.3.11)

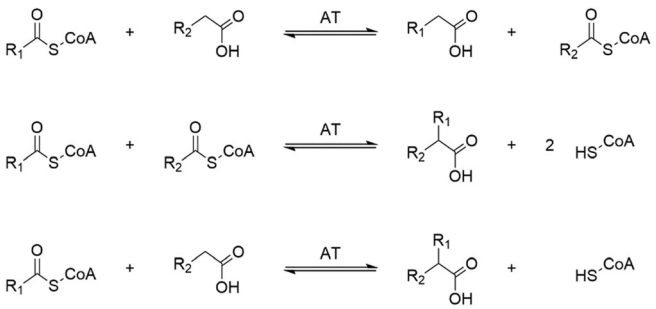

FIGURE 1 General reactions of acyltransferases acting on thioester-containing compounds, most commonly coenzyme A. Table 1 lists the wide variety of R_1 and R_2 chemical functionalities that ATs can catalyze.

catalyze the exchange of acetyl- and citryl (or citramalyl) groups on the thiol of the prosthetic group to ensure the continuity of the catalytic cycle.[6–8]

Class III CoA-transferases act in the metabolism of oxalate, carnitine, bile acids, toluene, and other aromatic compounds, with similar mechanisms to Class I CoA-transferases. In contrast to the classical ping–pong mechanism of Class I enzymes, kinetic studies of formyl-CoA: oxalate CoA-transferases (EC 2.8.3.16) suggest that the carboxylic acid of the donor acyl-CoA does not leave the enzyme complex until the final release of the product, and an apparent reaction mechanism via a ternary complex.[6,7,9]

The ability of ATs to act on a wide diversity of substrates has attracted attention because they efficiently catalyze transacylation reactions to produce esters, thioesters, amides, carbonates, and carbamates.[10] AT have biotechnological applications in the synthesis of modified lipids used in the food, cosmetics, and pharmaceutical industries.[11] The ATs are also used to produce biofuels from plant oils and animal fats.[12] In addition, ATs are being explored as biocatalysts in synthesizing high-value compounds such as fragrances, and flavors.[13] The use of AT in these applications is attractive due to their broad substrate selectivity and efficiency, making them a promising alternative to traditional chemical catalysts.[12,13]

The Class I, II, and III classification of the ATs is based on their catalytic mechanisms and does not consider any sequence or structural similarity. Due to advancements in high-throughput screening for AT and the availability of *in silico* sequence-based methods for predicting acyltransferase activity, a large number of AT have been identified. Since each family is based on at least a sequence experimentally verified to have AT function, classifying the ATs by sequence similarity into families allows to infer the tertiary structure, mechanisms, and catalytic residues for all members in a family, which is useful due to large amount of protein sequences with putative AT function that have no experimental verification.

The enzymes in AT families described here are available in the thioester-active enzyme (ThYme) database, which includes protein and gene identifiers of all the known sequences to date for the families in ThYme, for example, the recently updated thioesterase families.[14] ThYme database also includes links to the Protein Data Bank (PDB) if available, and clearly labels if a sequence has experimental work to date reported in Uniprot.[15,16] The ThYme database is available at https://thyme.engr.unr.edu and has an interactive interface that allows users to narrow down results using various fields, search by identifier and by FASTA sequence, and perform protein BLAST searches.

2 | METHODS

The approach used to group ATs into distinct families followed the same protocol as performed recently with thioesterases.[14] Different sequences appear in the same family when they have ~30% of sequence similarity or a nearly identical tertiary structure confirmed by the spatial superimposition of their structures.

The protocol for creating a new AT family started with extracting from the Swiss-Prot database in UniProt all the sequences with EC 2.3.1.*.[16] Only reviewed sequences with the "Evidence of Protein Level" in the description were considered. Fragments and theoretical proteins were disregarded. These criteria resulted in ~200 sequences experimentally confirmed to have AT activity that were subjected to a BLAST search against the National Center for Biotechnology Information's (NCBI) nr peptide sequence database using the protein–protein algorithm.[17] The BLAST searches were completed using a local instance of blast-2.9.0–2 and the nr database, both downloaded from NCBI on a Unix system. An E-value cutoff of 1×10^{-7} was used to capture as many sequences with the required similarity as possible while minimizing the number of redundant sequences. The highest max target sequences was used to capture all sequences within the E-value cutoff, and the other parameters were left at default settings.

The BLAST results were compared against each other to check for common sequences and identify the query sequences that resulted in the lowest number of BLAST results that included all the AT sequences that had been experimentally verified. The query sequences of unique, non-redundant BLAST results become the representative sequences that will originate new families from all confirmed acyltransferase sequences. The referenced literature in Uniport is checked to confirm experimental acyltransferase activity. The catalytic domain of each of the representative sequences, identified in InterPro,[18] were used to populate each AT family with BLAST as described above.

To confirm each AT family, we obtained all the known tertiary structures for each family from the PDB.[15] To ensure that structural comparisons were accurate, enzyme tertiary structures were reviewed and excluded any fragments, putative proteins, or non-AT domains from multidomain proteins.

For each family, all the monomer structures were extracted and selected a reference structure for each family, which was used as a pivot for superimposing other monomers in each family. To ensure consistency in aligning the core structure and allow for uniform calculation of structural similarity, we selected the shortest monomer in each family as the pivot. We used MultiProt[19] to superimpose all the monomers in each family, setting OnlyRefMol to 1, Scoring to 2, and leaving all other parameters at their default values. To measure the structural similarity of families that contained more than one

structure, we conducted a root mean square distance (RMSD) analysis of the superimposed tertiary structures. We calculated the distances between corresponding alpha carbon atoms (Cα) in the pivot and subject structures, using a cutoff distance determined by the average distance between sequential Cα atoms in the pivot structure. Any pairs of Cα atoms that were more distant than the cutoff were not considered corresponding and were not used in the RMSD calculation. The percentage value (*P*) of Cα atoms used to calculate the RMSD indicates the significance of the calculation. For each family, we superimposed the pivot structure to all other structures, resulting in $n - 1$ calculations, where n is the number of monomers being compared in that family. For families where $n > 2$, we calculated the average RMSD and *P* values ($RMSD_{ave}$ and P_{ave}, respectively).

3 | RESULTS AND DISCUSSION

Based on sequence similarity, following the approach described in the Methods, we identified fifteen acyltransferase families that act on thioester-containing substrates. Their functions, as well as substrate specificity and promiscuity, are discussed in Section 3.1. Their tertiary structures, catalytic residues and mechanisms in Section 3.2.

3.1 | Acyltransferase families and their functions

AT can act in degradative and synthetic pathways. In the degradative direction, ATs catalyze the thiolytic cleavage of ketoacyl-CoA molecules. In the synthetic path, a Claisen condensation reaction is catalyzed, and an acetoacetyl-CoA is synthesized from two molecules of acetyl-CoA.[20] Both forward and backward direction reactions can be found in the AT families,[21–24] except for AT14 and AT15, which include only enzymes that catalyze the forward reaction.

Enzymes in families AT1, AT2, AT4, AT6, AT12, AT14, and AT15 are related to the transfer of CoA from short/medium-chain acyl molecules such as acetyl-CoA, ketoacyl-CoA, malonyl-CoA, succinyl-CoA, butanoyl-CoA, and oxadipyl-CoA, from one carboxylic acid to another.[23,25,26] The families AT7, AT10, and AT13 can act on medium-long acyl-CoA chains (C_8 to C_{20}),[27] including palmitoyl-CoA,[28] and branched-chain CoA,[29] while the families AT8, AT9, and AT11 prefer the aromatic-CoA derivates like caffeoyl-CoA. However, the AT8 family has some members acting on the methylation of 5-hydroxyuridine through the transportation of a carboxyl group from *S*-adenosylmethionine without using an acyl-CoA as substrate.[30]

Members from the same AT families were reported to exhibit highly promiscuous substrate specificity. Some AT13 members can react with various aliphatic and aromatic CoA molecules.[31] There are examples in the AT1 family where the same AT domain could utilize malonate and methylmalonate units, for example, in the polyketide synthase PKS12,[32] resulting in a complex biosynthesis mechanism.

The genes and common protein names, overall function, substrates, EC number, and references related to the AT activity are described in Table 1.

Enzymes in family AT1 are primarily assigned to EC 2.3.1.39 but contain members in EC 2.3.1.161. They transfer the malonyl residue from malonyl-CoA to holo-ACPs and are involved in multiple pathways, for example, in the biosynthesis of lovastatin and antibiotic bacillaene.[33–35] This family contains the acyltransferase domain in many polyketide synthases (PKS), multifunctional enzymes that assemble simple CoA-thioesters (acetyl-CoA, malonyl-CoA) to create complex metabolites as mannosyl-β-1-phosphomycoketide, lovastatin, and anthraquinones.[34,36,37] Some members of AT1 are part of the fatty acid synthase (FAS), found in most eukaryotic organisms. For example, the FAS enzyme in rats has a malonyl transferase domain and the bovine FAS enzymes include an acetyltransferase domain.[35,38]

AT2 members (EC. 2.8.3.5) are involved in ketone-body metabolism, transferring the CoA from either succinyl or butanoyl to acetoacetate to produce acetoacetyl-CoA and succinate or butanoate.[39,40] They are distributed among bacteria and animals, with none reported in plants to date.

The AT3 family (EC 2.8.3.-, 5.1.99.4) includes enzymes with two distinct biological functions but with closer structural similarity. The CoA-transferases encoded by the yfdW gene were reported to enhance the ability of *Escherichia coli* MG1655 to survive under acidic conditions using formyl-CoA to convert the oxalate into oxalyl-CoA.[41] Other enzymes in this family catalyze the racemization of α-methyl-branched CoA esters. Sequence comparisons have shown that this enzyme is a family III CoA transferases member.[42]

Eventually, the ATs are assembled in dimers or tetramers structures, and the catalytic domains may work synergically, promoting either degrative and synthetic reactions in the same protein.[43] The members of AT4 and AT5 are part of dimeric or tetrameric thiolases with low identity between each other (<15%) and were grouped in their respective families. These complexes have distinct substrate channeling paths and substrate preferences.[44] The AT4 family is formed by N-terminal thiolases that are able to catalyze the cleavage of short, medium, and long straight chain 3-oxoacyl-CoAs and medium chain 3-oxoacyl-CoAs.[45] The enzymes in AT5 were reported to be related to the thiolase domain of sterol carrier protein 2/3-oxoacyl-CoA, which are active with medium and long straight chain 3-oxoacyl-CoAs but also with the 2-methyl-branched 3-oxoacyl-CoA and the bile acid intermediate.[45]

Enzymes in the family AT6 are derivates from hydroxymethylglutaryl-CoA synthase (HMG-CoA) (EC 2.3.3.10) and related to acetyl-CoA conversion to mevalonate, using the acetyl-CoA as substrate.[46] The members of AT6 also utilize 3-chloropropionyl-CoA to catalyze the formation of a key intermediate in cholesterogenic and ketogenic pathways.[24,47]

The members from the AT7 family were described using a variety of substrates as acyl-CoA (C_8 to C_{16}), including myristoyl-CoA, acetyl-CoA, iso-, and anteiso-branched chain acyl-CoA in bacterial biosynthesis of long-chain olefins.[48,49] AT7 also includes enzymes that catalyze the condensation reaction of fatty acid synthesis by the addition of an acyl acceptor of two carbons from malonyl-ACP, which are encoded by the FabH gene.[50,51]

TABLE 1 Acyltransferase families, gene and enzyme names, functions, and substrate preference.

Family	Enzyme names (gene names)	Most common AT function	Substrate preference of AT function	Ref.
AT1	Polyketide biosynthesis malonyl CoA-acyl carrier protein (ACP) transacylase (Bae) Mycoketide-CoA synthase (pks12)	Transfer of malonyl from malonyl-CoA to an ACP	Malonyl-CoA	Chopra et al.[36]; Chen et al.[89]
AT2	Succinyl-CoA:3-ketoacid CoA transferase (SCOT) Acetoacetyl CoA:acetate/butyrate: CoA transferase (ctf)	Transfers CoA from succinate to acetoacetate Transfers CoA from acetoacetyl-CoA to acetate, butyrate, and propionate	Succinyl-CoA Butanoyl-CoA	Fischer et al.[39]; Zhang et al.[40]
AT3	CoA-transferase formyl-CoA:oxalate CoA-transferase (FCOCT)	Conversion of acetyl-CoA and oxalate to oxalyl-CoA and acetate	Acyl-CoA Acetyl-CoA Formyl-CoA	Mullins et al.[90]
AT4	Acetoacetyl CoA thiolase (Acat) CoA transferase (ipdAB) 3-Oxoacyl-CoA thiolase	Condensation of two acetyl-CoA molecules into acetoacetyl-CoA	Acetoacetyl-CoA	Antonenkov et al.[45]; Nguyen et al.[91]; Crowe et al.[92]
AT5	Acetoacetyl CoA thiolase (Acat) CoA transferase (ipdAB) 3-Oxoacyl-CoA thiolase	Condensation of two acetyl-CoA molecules into acetoacetyl-CoA	Acetoacetyl-CoA	Antonenkov et al.[45]; Nguyen et al.[91]; Crowe et al.[92]
AT6	3-Hydroxy-3-methylglutaryl CoA synthase (mvaS)	Condensation of acetyl-CoA with acetoacetyl-CoA to form 3-hydroxy-3-methylglutaryl-CoA	Acetyl-CoA	Sutherlin et al.[46]
AT7	Acyl-CoA:acyl-CoA alkyltransferase 3-Oxoacyl-[acyl-carrier-protein] synthase 3	Catalyzes the condensation reaction of fatty acid by the addition of two carbons from malonyl-ACP to an acyl acceptor	Acyl-CoA (C8 to C16) Myristoyl-CoA	Frias et al.[49]; Deckert et al.[93]
AT8	Caffeoyl CoA *O*-methyltransferase (CCoAOMT) 5-Hydroxyuridine methyltransferase (TrmR)	Methylation of caffeoyl-CoA Methylation of 5-hydroxyuridine	Caffeoyl-CoA 5-Methoxyuridine	Rhu et al.[30]; Kilgore et al.[52]; Alazizi et al.[53]
AT9	Coniferyl alcohol acyltransferase (phCFAT) Fatty acid Elongas (ELOV)	Acetylation of coniferyl alcohol Addition of two carbons to long- and very long-chain fatty acids	Acetyl CoA Palmitoyl CoA	Dexter et al.[55]; Sassa et al.[94]
AT10	Acyltransferase PapA3 (papA3)	Transfer of palmitoyl groups to glucose	Palmitoyl-CoA Docosanoyl-CoA	Hatzios et al.[29]
AT11	Bile acid-CoA:amino acid *N*-acyltransferase (hBAT)	Conjugation of bile acid with amino acids glycine or taurine	Cholyl-CoA	Sfakianos et al.[58]
AT12	Succinyl-CoA:acetate CoA transferase (AarC)	Convert acetate to acetyl-CoA and succinate	Succinyl-CoA	Mullins et al.[60]
AT13	Chalcone synthase (CHS1)	Release free CoA from coumaroyl-CoA and malonyl-CoA during the synthesis of chalcones derivates	4-Coumaroyl-CoA 3-Oxoadipyl-CoA	Göbel et al.[26]; Katsuyama et al.[63]
AT14	3-Oxoadipate CoA-transferase	Transfer of CoA from succinate to 3-oxoadipate	3-Oxoadipyl-CoA	Jacob et al.[66]
AT15	3-Oxoadipate CoA-transferase	Transfer of CoA from succinate to 3-oxoadipate	3-Oxoadipyl-CoA	Jacob et al.[66]

The family AT8 has methyl transferases (EC 2.1.1.-) related to the methylation of caffeoyl-CoA in the biosynthesis of lignin monomers, flavonols, and flavones. However, some members of AT8 act on non-thioester acyl transferases that have been characterized in animals and bacteria, catalyzing the methylation of 5-hydroxyuridine to form 5-methoxyuridine. Those enzymes are described as nucleotide-

specific methyltransferase complexed with the cognate tRNA substrate.[30,52,53]

The members of AT9 (E.C 2.3.1.-) can act as anthocyanin acyltransferases (AATs) or coniferyl alcohol derivates and are related to the synthesis of long chain fatty acids.[54] This family also includes enzymes classified as BAHD-AT,[55] which produce small volatile esters, modified anthocyanins, as well as constitutive defense compounds, and phytoalexins.[56] In plants, these enzymes catalyze the acyl transfer from acyl-CoA to the glycosyl moiety of anthocyanins, thus playing an essential role in flower coloration and secondary metabolism.[57]

Family AT10 includes enzymes from gene Pap that utilize palmitoyl-CoA as the preferred substrate in the biosynthesis of polyacyltrealose, a glycolipid linked to mycobacterial pathogenesis, for example, in *Mycobacterium tuberculosis*.[29]

AT11 includes animal enzymes from gene BAAT that catalyze the conjugation of bile acids with the amino acids glycine or taurine using cholyl-CoA as substrate.[58] The enzymes in AT11 are structurally similar TE2 thioesterases with the gene Acot that contribute to hydrolyzing acyl-CoAs to the free fatty acid and CoASH.[59]

Members of the family AT12 include the succinyl-CoA:acetate CoA transferase (EC 2.8.3.18). These enzymes are associated with acetic acid resistance in *Acetobacter aceti*, which utilizes succinyl-CoA to convert toxic acetate to acetyl-CoA and succinate.[60] *Roseburia sp*., a human gut bacterium, plays a crucial role in butyrate formation in the human colon using the butyryl-CoA as substrate.[23]

Enzymes in the family AT13 are mostly found in plants and are related to PKS, a multienzyme complex highly homologous to FAS.[28] The members of this family were experimentally described in the biosynthesis of xanthones (*Hypericum androsaemum*), curcuminoid (*Curcuma longa*), olivetolic acid (*Cannabis sativa*), vancomycin, and teicoplanin antibiotics (*Amycolatopsis orientalis*). Enzymes in AT13 are promiscuous, active on a large variety of substrates. For example, they can use long-chain aliphatic acyl-CoA (C_{12} to C_{20}), feruloyl-CoA, aromatic-CoA (feruloyl-CoA, benzoyl-CoA), and short chain-CoA (malonyl-CoA, butyryl-CoA).[31,61-65]

AT14 and AT15 respectively include the subunit A and subunit B from 3-oxoadipate CoA-transferase (EC 2.8.3.6) found in *Pseudomonas sp*. Subunits A and B do not have enough sequence similarity to be grouped in the same family. Those subunits work synergically to produce acetyl-CoA and succinyl-CoA from 3-oxoadipyl-CoA.[66] The transferase reaction for these families was found to be reversible: the addition of succinate to the assay mixture brought about a decrease in the 3-oxoadipyl-CoA-Mg^{2+} complex.[26] Some AT15 enzymes appear in the pathway for glutamate fermentation in *A. fermentans*.[66]

In addition to the AT function that enzymes in AT families catalyze, some have a wide substrate specificity and/or catalyze different functions. High-throughput screening of AT and an *in silico* sequence-based method for predicting acyltransferase activity provided access to many promiscuous AT. As examples, a human glutathione transferase with the native substrate 2-cyano-1,3-dimethyl-1-nitrosoguanidine has azathioprine as an alternative substrate; or an acyl CoA transferase (YgfH) from *E. coli* transfers the CoA group from succinyl-CoA to succinate, but can also use acetate or butyrate as alternative acceptors.[10,67-69]

Some members in the AT families identified in this work were also identified in a different enzyme group due either to the multiple substrates or the similarity in their structures (Table 2), confirming the promiscuous behavior of AT enzymes. Some members from families AT1, AT7, and AT13 have a ketoacyl synthase activity reported and have enough structural similarity to be also a member of ketoacyl families presented in ThYme.[70] A similar situation was observed for members from families AT11 and AT12, with thioesterase activity also present in thioester families.

This discussion (Section 3.1, Table 1) highlights the main functions and substrate specificities of the enzymes in each AT family that have been experimentally verified. However, each family contains significantly many more sequences that have not been experimentally characterized, which are available in the ThYme database. The classification of these uncharacterized sequences into AT families, with each family based on experimentally verified sequences, allows the selection of sequences likely have a desired function or substrate selectivity from a particular AT family and species.

3.2 | Acyltransferase families and their structures, catalytic residues, and mechanisms

The structural similarity of each AT family was confirmed by the superimposition of their respective tertiary structures that were experimentally resolved and available in the PDB. Table 3 reports the predominant structural domain for each family (confirmed by Pfam and InterPro), as well as the $RMSD_{ave}$, P_{ave} values (see Methods), and a list of the PDB structures that were superimposed. The structural similarity is shown by $RMSD_{ave}$ values <1.03 Å and P_{ave} values of

TABLE 2 Acyltransferase families that have sequences overlapping with other enzyme families.

AT family	Other activity reported or similar structures	Example sequence UniProt identifier	Ref.
AT1	Ketoacyl synthase family KS3	A0A2H3CTK0 and Q0CF73	Engels et al.[95]; Huang et al.[96]
AT7	Ketoacyl synthase family KS1	D0MCQ4 and O67185	Deckert et al.[93]; Hu et al.[97]
AT11	Thioesterase family TE2	Q8N9L9, Q8BGG9, and O55137	Hunt et al.[59]; Huhtinen et al.[98]
AT12	Thioesterase family TE1	P52043, P32316, and P83773	Haller et al.[69]; Lee et al.[99]
AT13	Ketoacyl synthase family KS4	Q9AU09 and B0FYK7	Zheng et al.[100]

TABLE 3 Acyltransferase folds and structure superimposition.

Family	Fold/domain	$RMSD_{ave}$ (Å)	P_{ave} (%)	Structures in the PDB
AT1	Acyl transferase	0.71	35.9	2HG4, 2JFD, 2JFK, 2PFF, 2QC3, 2QJ3, 2QO3, 2UV8, 2VKZ, 3HHD, 3HMJ, 4AMM, 4AMN, 4AMO, 4AMP, 4MZ0, 4RL1, 5BP1, 5YDA, 5YDL, 5YDM, 6C9U, 6FIJ, 6FIK, 6IYR, 6IYT, 6JSH, 6JSI, 6U5T, 6U5U, 6U5V, 6U5W, 7AHB, 7AKC, 7BC4
AT2	CoA transferase	0.97	58.7	1K6D, 1M3E, 1O9L, 1OOY, 1OOZ, 1OPE, 2AHU, 2AHV, 2AHW, 2NRB, 2NRC, 3CDK, 3K6M, 3OXO, 3RRL, 4KGB, 5DBN, 6LP1
AT3	CoA transferase	0.81	62.2	1P5H, 1P5R, 1PQY, 1PT5, 1PT7, 1PT8, 1Q6Y, 1Q7E, 1T3Z, 1T4C, 1VGQ, 1VGR, 1X74, 1XA3, 1XA4, 1XK6, 1XK7, 1XVT, 1XVU, 1XVV, 2G04, 2GCE, 2GCI, 2GD0, 2GD2, 2GD6, 2VJK, 2VJL, 2VJM, 2VJN, 2VJO, 2VJP, 2VJQ, 2YIM, 3UBM, 4ED9, 4HL6, 5YIT, 5YIY, 5YX6
AT4	Thiolase (N-terminal)	1.00	46.7	1AFW, 1DLU, 1DLV, 1DM3, 1M1O, 1M1T, 1M3K, 1M3Z, 1M4S, 1M4T, 1NL7, 1OU6, 1PXT, 1QFL, 1ULQ, 1WDK, 1WDL, 1WDM, 2C7Y, 2C7Z, 2D3T, 2VTZ, 2VU0, 2VU1, 2VU2, 2WKT, 2WKU, 2WKV, 2WL4, 2WL5, 2WL6, 2WU9, 2WUA, 3SS6, 3 W15, 4B3H, 4B3I, 4B3J, 4C2J, 4C2K, 4DD5, 4E1L, 4NZS, 4O99, 4O9A, 4O9C, 4UBT, 4UBU, 4UBV, 4UBW, 4W61, 4WYR, 4WYS, 4XL2, 4XL3, 4XL4, 4YZO, 4ZRC, 5BYV, 5BZ4, 5CBQ, 5F0V, 5F38, 5LP7, 5ONC, 5XYJ, 5XZ5, 5ZQZ, 5ZRV, 6AQP, 6ARE, 6ARF, 6ARG, 6ARL, 6ARR, 6ART, 6BJ9, 6BJA, 6BJB, 6DV2, 6ESQ, 6ET9, 6L2C, 6L2G, 6PCA, 6PCB, 6PCC, 6PCD, 7CW4, 7CW5, 7LBZ, 7LCA, 7LCL, 7LD2, 7LDC, 7LDT, 7LDU, 7LDV, 7LDW, 7O1G, 7O1I, 7O1J, 7O1K, 7O1L, 7O1M, 7O4Q, 7O4R, 7O4S, 7O4T, 7O4V
AT5	Thiolase (C-terminal)	0.62	14.5	1AFW, 1DLU, 1DLV, 1DM3, 1M1O, 1M1T, 1M3K, 1M3Z, 1M4S, 1M4T, 1NL7, 1OU6, 1PXT, 1QFL, 1ULQ, 1WDK, 1WDL, 1WDM, 2C0L, 2C7Y, 2C7Z, 2D3T, 2VTZ, 2VU0, 2VU1, 2VU2, 2WKT, 2WKU, 2WKV, 2WL4, 2WL5, 2WL6, 2WU9, 2WUA, 3SS6, 3 W15, 3ZBG, 3ZBK, 3ZBL, 3ZBN, 4B3H, 4B3I, 4B3J, 4BI9, 4BIA, 4C2J, 4C2K, 4DD5, 4E1L, 4NZS, 4O99, 4O9A, 4O9C, 4UBT, 4UBU, 4UBV, 4UBW, 4W61, 4WYR, 4WYS, 4XL2, 4XL3, 4XL4, 4YZO, 4ZRC, 5BYV, 5BZ4, 5CBQ, 5F0V, 5F38, 5LNQ, 5LOT, 5LP7, 5M3K, 5MG5, 5ONC, 5XYJ, 5XZ5, 5ZQZ, 5ZRV, 6AQP, 6ARE, 6ARF, 6ARG, 6ARL, 6ARR, 6ART, 6BJ9, 6BJA, 6BJB, 6BN2, 6DV2, 6ESQ, 6ET9, 6HRV, 6HSJ, 6HSP, 6L2C, 6L2G, 6PCA, 6PCB, 6PCC, 6PCD, 7CW4, 7CW5, 7LBZ, 7LCA, 7LCL, 7LD2, 7LDC, 7LDT, 7LDU, 7LDV, 7LDW, 7O1G, 7O1I, 7O1J, 7O1K, 7O1L, 7O1M, 7O4Q, 7O4R, 7O4S, 7O4T, 7O4V
AT6	HMG-CoA synthase (homologous to thiolases)	0.89	60.1	1TVZ, 1TXT, 1X9E, 1XPK, 1XPL, 1XPM, 1YSL, 2F82, 2F9A, 2FA0, 2FA3, 2HDB, 2P8U, 3LEH, 3SQZ, 3V4N, 3V4X, 4YXQ, 4YXT, 4YXV, 5HWO, 5HWP, 5HWQ, 5HWR, 5KP5, 5KP6, 5KP7, 5KP8, 7CQT
AT7	ACP synthase III (homologous to thiolase)	0.95	55.0	1EBL, 1HN9, 1HND, 1HNH, 1HNJ, 1HNK, 1MZJ, 1MZS, 1ZOW, 2EBD, 2EFT, 2GYO, 2X3E, 3FK5, 3IL3, 3IL4, 3IL5, 3IL6, 3IL7, 3IL9, 3ROW, 3S1Z, 3S20, 3S21, 3S23, 4DFE, 4KTI, 4KTM, 4KU2, 4KU3, 4KU5, 4NHD, 4RYB, 4WZU, 4X0O, 4X9K, 4X9O, 4YLT, 4YUC, 4YUF, 4Z19, 4Z8D, 5BNM, 5BNR, 5BNS, 5BQS, 5C1J, 5KP2, 5V0P, 5VXD, 5VXE, 5VXF, 5VXG, 5VXH, 5VXI, 6B2R, 6B2S, 6B2T, 6B2U, 6KVS, 6X7R
AT8	*O*-methyltransferase	0.97	48.7	1SUI, 1SUS, 2AVD, 2GPY, 2HNK, 3C3Y, 3CBG, 3DUL, 3DUW, 3R3H, 3TFW, 3TR6, 4OA5, 4OA8, 4PCA, 4PCL, 4QVK, 4YMG, 4YMH, 5KVA, 5LHM, 5LOG, 5X7F, 5ZW3, 5ZW4, 6JCL, 6JCM, 7CVX
AT9	Chloramphenicol acetyltransferase-like domain	1.02	47.5	2BGH, 4G0B, 4G22, 4G2M, 4KE4, 4KEC, 5FAL, 5FAN, 5KJS, 5KJT, 5KJU, 5KJV, 5KJW, 6DD2, 6LPV, 6MK2, 6WAO, 6WCS, 7CYS, 7DTP, 7W26
AT10	Condensation domain (homologous to Chloramphenicol acetyltransferase)	–	–	6AEF
AT11	Bile acid-CoA:amino acid *N*-acetyltransferase (BAAT)	–	–	–
AT12	Acetyl-CoA hydrolase/transferase	1.02	52.8	2G39, 2NVV, 2OAS, 3D3U, 3EH7, 4EU3, 4EU4, 4EU5, 4EU6, 4EU7, 4EU8, 4EU9, 4EUA, 4EUB, 4EUC, 4EUD, 5DDK, 5DW4, 5DW5, 5DW6, 5E5H

TABLE 3 (Continued)

Family	Fold/domain	$RMSD_{ave}$ (Å)	P_{ave} (%)	Structures in the PDB
AT13	Chalcone and stilbene synthases/AMP-binding enzyme	1.02	51.6	1BI5, 1BQ6, 1CGK, 1CGZ, 1CHW, 1CML, 1D6F, 1D6H, 1D6I, 1EE0, 1I86, 1I88, 1I89, 1I8B, 1JWX, 1QLV, 1TED, 1TEE, 1U0M, 1U0U, 1U0V, 1U0W, 1XES, 1XET, 1Z1E, 1Z1F, 2D3M, 2D51, 2D52, 2P0U, 3A5Q, 3A5R, 3A5S, 3ALE, 3AWJ, 3AWK, 3E1H, 3EUO, 3EUQ, 3EUT, 3OIT, 3OV2, 3OV3, 3TSY, 3VS8, 3VS9, 3WD7, 3WD8, 3WXY, 3WXZ, 3WY0, 4B0N, 4WUM, 4YJY, 5UC5, 5UCO, 5W8Q, 5WC4, 5WX3, 5WX4, 5WX5, 5WX6, 5WX7, 5YPT, 6CQB, 6DX7, 6DX8, 6DX9, 6DXA, 6DXC, 6DXE, 6DXF, 6GW3, 6J1M, 6J1N, 6L5U, 6L61, 6L73, 6L78, 6L7J, 6OP5, 7BUR, 7BUS, 7CBF, 7CCT, 7D41, 7DTQ, 7FFA, 7FFC, 7FFG, 7FFH, 7FFI, 7SGY, 7W6G
AT14	CoA transferase	-	-	-
AT15	CoA transferase	-	-	-

>62%, which confirms that enzymes within a family have nearly identical tertiary structures, despite a wide variety of substrate specificities and enzymatic activity for AT families. This is shown in Figure 2, which has two AT2 enzymes superimposed, a succinyl-CoA:3-ketoacid CoA transferase (3CDK) from *Bacillus subtilis* and a *Drosophila melanogaster* succinyl-CoA:3-ketoacid-CoA transferase (SCOT) (4KGB) AT2 enzyme.

An advantage of classifying enzymes into families is that one can make not only structural inferences, but also mechanistic ones. To demonstrate, the catalytic and binding residues of *B. subtilis* succinyl-CoA:3-ketoacid CoA transferase (3CDK) were predicted from the known catalytic and binding residues of the *D. melanogaster* SCOT (4KGB) AT2 enzyme (Figure 2), through structure superimposition and spatial correspondence of the residues. The crystal structure of the *B. subtilis* AT2 enzyme was known; however, no experiments were performed to identify the catalytic residues. Similarly, the three catalytic residues (Cys–His–Cys) of a human acetoacetyl-CoA thiolase (1WL4)[71] were used to predict the catalytic Cys–His–Cys residues in an acetyl-CoA acetyltransferase from *Bacillus anthracis* (3SS6), as shown in Figure 3.

In AT1, the *Saccharomyces cerevisiae* AT domain (2UV8) was described from a fungal fatty acid synthase (FAS I) complex, located in the inner face of the barrel, close to the reaction chamber, facing the ACP domain, formed by heterododecamer in an $\alpha_6\beta_6$ configuration.[72] AT1 enzymes also include the methyltransferase domain in FAS II *M. tuberculosis* (2QC3),[73] as well as enzymes with reversible reactions specific for malonyl-CoA described for *Staphylococcus aureus* and *Streptococcus pneumoniae*.[74] For malonyl CoA-ACP transacylase, the acyl groups are provided by the action of three free-standing AT in a two-step mechanism. First, malonyl-CoA is bound, and the malonyl moiety is transferred to the catalytic serine with CoA released. Second, ACP binds, and the malonyl moiety is transferred to the terminal sulfhydryl of the ACP prosthetic group through the catalytic serine and nearby histidine serving as the catalytic dyad.[74] AT1 also includes a FAS enzyme from *Homo sapiens* (2JFD) that is able to transfer acetyl and malonyl chains and has two subdomains. The large subdomain exhibits the typical α/β hydrolase fold with a multi-stranded parallel β-sheet surrounded by α-helices. The smaller domain is represented by a multi-stranded antiparallel β-sheet packed against two distal α-helices. The active site is located in the cleft between the two subdomains and is composed of residues of His, Ser, and Arg.[75]

The transfer of a CoA moiety from succinate to acetoacetate by a transferase from *D. melanogaster* in the AT2 family (4KGB) is catalyzed by a dimeric succinate-CoA transferase that preserves an α/β/α motif with a central seven-stranded β-sheet. In its reaction mechanism, CoA is covalently bound to the active site by a Glu residue.[40] Mass spectrometric analysis revealed the dimeric enzyme could use a ping–pong mechanism where half of the subunits are catalytically superfluous for the second half reaction, and only one of the subunits transfers the CoA moiety to a carboxylic acid acceptor.[76]

The structure and catalytic mechanism of enzymes in AT3 can be described by a formyl-CoA transferase from *Oxalobacter formigenes* (2VJK). The 2VJK structure contains a covalent β-aspartyl-CoA

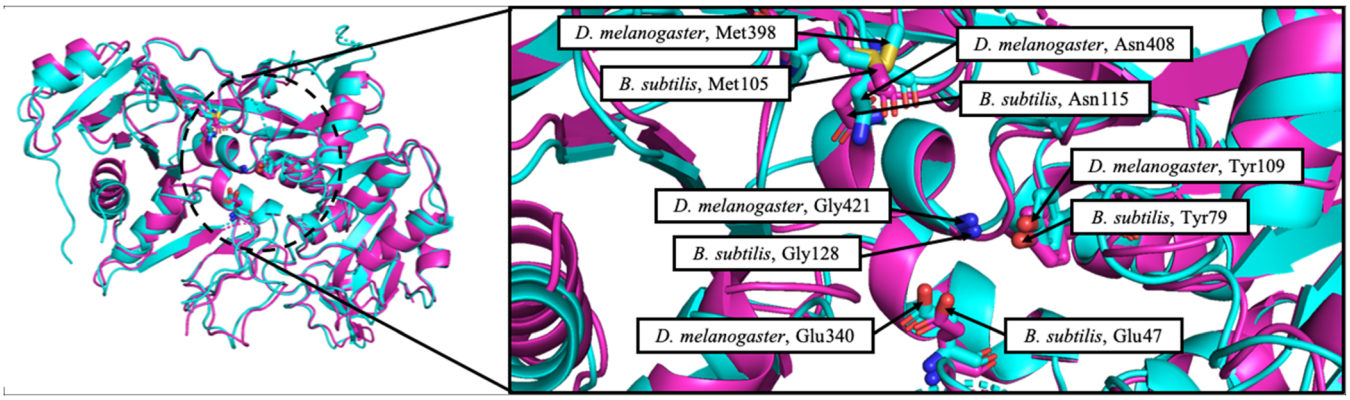

FIGURE 2 AT2: The likely catalytic residues of a succinyl-CoA:3-ketoacid coenzyme A transferase (PDB 3CDK A and B chain) from *Bacillus subtilis* (pink) were predicted based on known residues (PDB 4KGB B chain, blue) from *Drosophila melanogaster*. The superposed Glu340 of 4KGB, known to be the conserved glutamate residue in the active site, was used to predict the Glu47 from 3CDK, and the other residues in stick representation were described for the CoA-binding site.

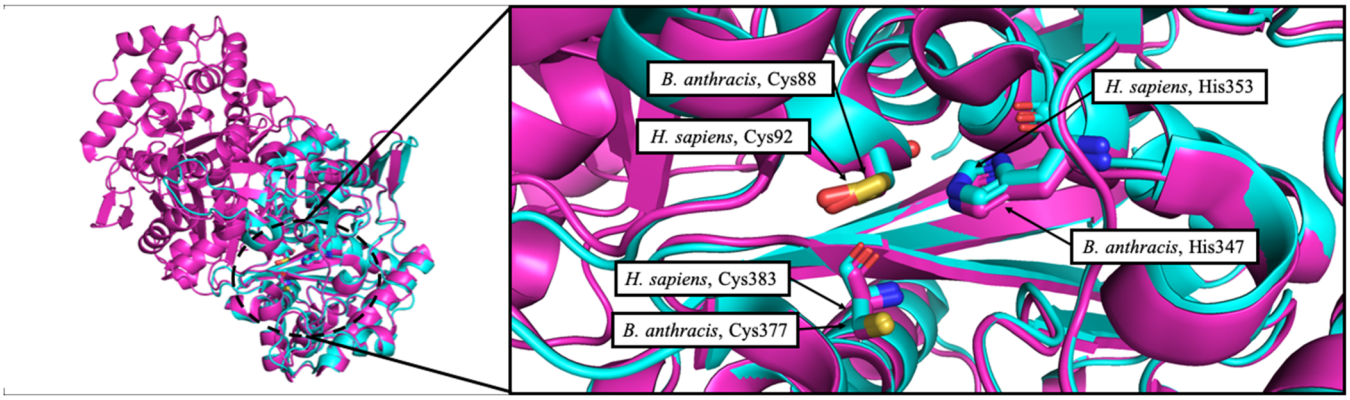

FIGURE 3 AT5: The likely catalytic residues of an acetyl-CoA acyltransferase (PDB 3SS6) from *Bacillus anthracis* (pink) were predicted based on known residues from a *Homo sapiens* acetoacyl-CoA thiolase (PDB 1WL4). The superposed catalytic Cys92, His353, and Cys383 residues of 1WL4 were used to predict the 3SS6 catalytic residues.

thioester, changing the conformation in the active sites of the dimer. A glycine loop binds the formate molecule released during the formyl-CoA catalysis, and a Gln residue prevents the intermediates from early hydrolysis.[77]

AT4 enzymes are homodimeric thiolases with two domains in the core of the structures folded into a mixed five-stranded sheet with the central layer pointed toward the active site. A 3-ketoacyl-CoA thiolase enzyme from *S. cerevisiae* (1PXT) shows the active site with a shallow pocket, shaped by two highly conserved Cys and His residues at the floor of the pocket. Those subunits interact themselves by hydrophobic contacts, hydrogen bonds, and salt bridges.[78]

The general fold for AT5 enzymes, for example in a 3-ketoacyl-CoA thiolase-like protein enzyme from *Leishmania mexicana* (3ZBG), consists of three domains: the N-terminal domain, the C-terminal domain, and the loop domain. The N- and C- domains have the β/α/-β/α/β/α/β/β topology and have the catalytic residues. The function of the loop domain is to orient the CoA moiety through the binding pocket. The active site is located in the interface of the subunits and is formed by residues from each subunit, and is relatively polar. Two Cys and one His are usually involved in the catalytic reaction. One Cys acts as a nucleophile and the other as an acid/base. The His helps the Cys activation and increase its nucleophilicity.[20,43]

The description of reaction mechanisms of enzymes in AT6 is based on HMG-CoA synthases and is similar to the mechanism in thiolases. The HMG-CoA synthase reaction consists of the irreversible condensation of acetyl-CoA with acetoacetyl-CoA to produce HMG-CoA. The acetyl group is transferred from the CoA thioester to a Cys residue forming acetylcysteine, then the acetylcysteine reduces CoA, and the acetoacetyl-CoA binds in the enzyme. Next, acetylcysteine is activated by proton removal, leading to the Claisen condensation, resulting in a free HMG-CoA.[79,80] The structures from family AT6, as shown in a HMG-CoA synthase enzyme from *S. aureus* (1TXT), consist of an N-terminal thiolase-like domain with mixed α-β secondary structural elements dominated by two β-sheets that envelop the two central α-helices forming the active site loop that contains the Cys residue related to the condensation reaction. The C-terminal domain is a structural feature of the HMG-CoA synthases with no experimentally confirmed function, but could be associated with the enzyme

location in the cell or post-translational activity. Beyond the Cys catalytic residue, Asp and Phe also contribute to the catalytic pocket in this enzyme.[79,80]

The dimeric crystal structures (3ROW and 3S23) from *Xanthomonas campestris* OleA are in family AT7. Each monomer contains an active site and uses the ping–pong mechanism to covalently bind the long-chain alkyl substrate to a Cys residue. The active site Cys is a conserved residue, and an active site Glu likely promotes the Claisen condensation.[81] The tertiary structure for the members of this family is similar to the tertiary structure of KS1 ketoacyl synthases,[70] supporting the promiscuous catalytic activity of these enzymes.

The AT8 enzyme structure (1SUI) of an *O*-methyltransferase from *Medicago sativa*, which is involved in the biosynthesis of hydroxycinnamic acid derivatives, can methylate the C_3 position of a lignin monomer and is a bivalent cation dependent enzyme.[82] The catalytic domain exhibits an α/β Rossmann fold, with Lys and Arg residues acting in the orientation of the substrate. The phenolic moiety of the substrate is sequestered by aromatic residues (Tyr and Trp), and Met, Asp, and Asn complete the substrate binding pocket. A divalent cation, such as Ca^{2+}, Mg^{2+}, or Zn^{2+}, was observed in the active site chelated by Thr, Glu, and Asp and revealed to have a critical role in enzyme activity.[83] Three residues (Thr, Glu, and Asp) in the *O*-methyltransferase from *M. sativa* (1SUI) were used to predict that, in the human cathecol-*O*-methyltransferase (2AVD), three similar residues (Thr, Gln, and Asp) appear in the same spatial positions but with a different Thr side chain orientation (same orientation for Glu–Gln and Asp–Asp), which suggests that the human enzyme may not chelate 2^+ ions for the catalytic reaction (Figure 4). However, the different Thr side chain orientation could also result from the fact the 1SUI was crystallized with Ca^{2+} while 2AVD was not.

The structure of enzymes in AT9 are similar to that of BAHD AT. The structure of a shikimate *O*-hydroxycinnamoyl transferase from *Arabidopsis thaliana* (5KJS) consists of two pseudo-symmetric N-terminal and C-terminal domains connected by a long loop. The arrangement comprises a β-sheet core flanked by α-helices with the active site in the interface of the two domains, with a His residue between them.[84] A mechanism was described for a hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase from *A. thaliana* where His153 acts as a general base that deprotonates the 5-hydroxyl of the acyl acceptor substrate shikimate, priming it for nucleophilic attack on the carbonyl carbon of the acyl donor *p*-coumaroyl-CoA.[85]

The only structure available to date in family AT10 (6AEF), a PKS associate enzyme from *M. tuberculosis*, shows an unconventional presence of a zinc finger motif in the N-terminal region, but still presents a classical CoA-dependent acyltransferase fold with a large β sheet flanked by α-helices. The N-terminal subdomain also shows seven mixed-type beta strands, parallel and antiparallel, whereas the C-terminal subdomain contains six mixed beta strands.[86]

AT11 has a single crystallized structure (3K2I) from *Homo sapiens*; however, there is no accompanying literature to describe it. Still, the putative catalytic triad Cys–Asp–His was identified from a human bile acid-CoA:amino acid *N*-acyltransferase by site-direct mutagenesis.[58] The triad Gly–Ser–Gly was assigned for ACOT enzymes in mice, also present in the thioesterase family TE2, with the Ser acting as the nucleophile.[59]

The AT12 crystal structure 2OAS monomer, of an acetyl-CoA hydrolase/transferase enzyme from *Shewanella oneidensis*, is composed of the pseudo-two-fold axis in non-similar N- and C- terminal domains, forming a funnel-shaped active site cleft. The residues Phe, Thr, Ala, Phe, Leu, and Val form the catalytic hydrophobic pocket. This composition and spatial orientation isolate the catalytic nucleophilic glutamate at the bottom of the cleft, allowing the near-ideal angle for the attack as the thioester oxygen with the CoA-derivates.[87]

The tertiary structures of chalcone and stilbene synthases from family AT13 are homologous to the thiolase-like superfamily, and the overall fold for this family resembles the homodimeric β-ketoacyl synthase II and thiolases. The active site for CoA is formed by the conserved residues of Cys, Phe, His, Asn, and leads to the decarboxylation and condensation reaction. The Cys acts as nucleophile in polyketide formation while the Phe orients the substrate to the active site pocket during the elongation process of the polyketide intermediate.[88]

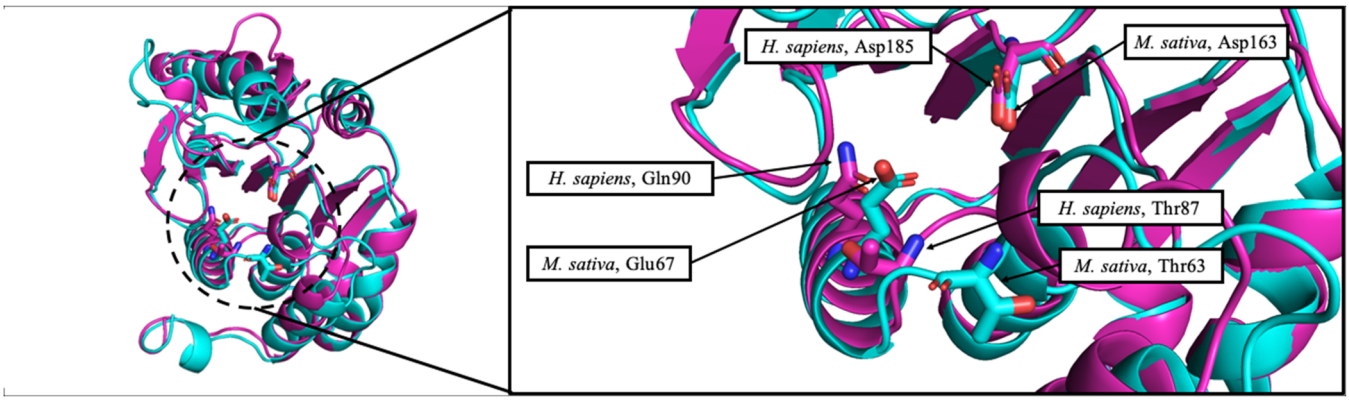

FIGURE 4 AT8: The putative ion chelating residues of a cathecol-*O*-methyltransferase (PDB 2AVD) from *Homo sapiens* (pink) were predicted based on residues from a *Medicago sativa* caffeoyl CoA 3-*O*-methyltransferase (PDB 1SUI) known to chelate 2^+ ions involved in the catalytic reaction.

The families AT14 and AT15 have no known resolved tertiary structures. The AT families described here are mapped to the previous, mechanistic-based Class I, II, III classification: AT1, AT2, AT9, and AT13 families to Class I; AT3 and AT5 to Class II; and AT11, AT14, and AT15 families to Class III. The other AT families do not have their mechanisms resolved in such detail to be able to assign them to Class I, II, or III.

This discussion (Section 3.2, Table 3, and Figures 2–4) highlights the structures and mechanisms of the enzymes in each AT family that have been experimentally verified. Since structures are very well conserved in each family, and since each family contains significantly many more sequences that have not been experimentally characterized, the classification of uncharacterized sequences into AT families enables the selection of sequences with a desired mechanism or structure from a particular AT family and species. Further, the known structures in each family are reliable templates to predict the three dimensional structure of any sequence in the same family.

4 | CONCLUSIONS

All acyl transferases active on thioester-containing substrates, mostly commonly CoA, were classified into fifteen distinct families. Families were established based on sequence similarity and confirmed with tertiary structure superimposition. Each family is based on at least one enzyme experimentally verified to be an acyltransferase. This classification allows to infer the structure, function, catalytic residues, and mechanisms of AT enzymes that have not experimentally characterized. This is of particular importance to the AT enzymes due to their broad substrate specify and promiscuous catalytic activity. All the sequences and structures of the AT families, including external links to protein and gene repositories, are available to the public in the ThYme database (https://thyme.engr.unr.edu).

AUTHOR CONTRIBUTIONS

Caio C. de Carvalho: Investigation; writing – original draft; methodology; validation; writing – review and editing; formal analysis; supervision; data curation; conceptualization; visualization. **Ian P. Murray:** Methodology; validation; visualization; formal analysis; writing – review and editing. **Hung Nguyen:** Data curation; writing – review and editing. **Tin Nguyen:** Funding acquisition; writing – review and editing; supervision. **David C. Cantu:** Conceptualization; investigation; funding acquisition; writing – review and editing; methodology; formal analysis; project administration; supervision.

ACKNOWLEDGMENTS

This work was supported by the U.S. National Science Foundation (Award 2001385).

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/prot.26599.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

David C. Cantu https://orcid.org/0000-0001-9584-5062

REFERENCES

1. van der Sluis R, Erasmus E. Xenobiotic/medium chain fatty acid: CoA ligase–a critical review on its role in fatty acid metabolism and the detoxification of benzoic acid and aspirin. *Expert Opin Drug Metab Toxicol*. 2016;12(10):1169-1179.
2. Röttig A, Steinbüchel A. Acyltransferases in Bacteria. *Microbiol Mol Biol Rev*. 2013;77(2):277-321.
3. Czumaj A, Szrok-Jurga S, Hebanowska A, et al. The pathophysiological role of CoA. *Int J Mol Sci*. 2020;21(23):1-30.
4. Jiang Y, Morley KL, Schrag JD, Kazlauskas RJ. Different Active-Site Loop Orientation in Serine Hydrolases versus Acyltransferases. *Chembiochem*. 2011;12(5):768-776.
5. Selmer T, Buckel W. Oxygen exchange between acetate and the catalytic glutamate residue in glutaconate CoA-transferase from *Acidaminococcus fermentans*. Implications for the mechanism of CoA-ester hydrolysis. *J Biol Chem*. 1999;274(30):20772-20778.
6. Heider J. A new family of CoA-transferases. *FEBS Lett*. 2001;509(3):345-349.
7. Frey PA. *Comprehensive Natural Products II Chemistry and Biology*. 2010. Elsevier Ltd.
8. Xu Z, Zhu L, Chen H. 3.19-Acrylic Acid. In: Moo-Young M, ed. *Comprehensive Biotechnology*. 2nd ed. Academic Press; 2011:201-206. doi:10.1016/B978-0-08-088504-9.00455-4
9. Leutwein C, Heider J. Succinyl-CoA:(R)-benzylsuccinate CoA-transferase: An enzyme of the anaerobic toluene catabolic pathway in denitrifying bacteria. *J Bacteriol*. 2001;183(14):4288-4295.
10. Müller H, Terholsen H, Godehard SP, Badenhorst CPS, Bornscheuer UT. Recent insights and future perspectives on promiscuous hydrolases/acyltransferases. *ACS Catal*. 2021;11(24):14906-14915.
11. King AJ, Judd AS, Souers AJ. Inhibitors of diacylglycerol acyltransferase: a review of 2008 patents. *Expert Opin Ther Pat*. 2010;20(1):19-29.
12. Liu Q, Siloto RMP, Lehner R, Stone SJ, Weselake RJ. Acyl-CoA:diacylglycerol acyltransferase: Molecular biology, biochemistry and biotechnology. *Prog Lipid Res*. 2012;51(4):350-377.
13. Sá AGA, de Meneses AC, de Araújo PHH, de Oliveira D. A review on enzymatic synthesis of aromatic esters used as flavor ingredients for food, cosmetics and pharmaceuticals industries. *Trends Food Sci Technol*. 2017;69:95-105.
14. Caswell BT, de Carvalho CC, Nguyen H, Roy M, Nguyen T, Cantu DC. Thioesterase enzyme families: Functions, structures, and mechanisms. *Protein Sci*. 2022;31(3):652-676.
15. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. *Nucleic Acids Res*. 2000;28(1):235-242.
16. Bateman A, Martin MJ, Orchard S, et al. UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res*. 2021;49(D1):D480-D489.
17. Altschul SF, Madden TL, Schäffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25(17):3389-3402.
18. Paysan-Lafosse T, Blum M, Chuguransky S, et al. InterPro in 2022. *Nucleic Acids Res*. 2023;51(D1):D418-D427.
19. Shatsky M, Nussinov R, Wolfson HJ. A method for simultaneous alignment of multiple protein structures. *Proteins Struct Funct Genet*. 2004;56(1):143-156.

20. Haapalainen AM, Meriläinen G, Pirilä PL, Kondo N, Fukao T, Wierenga RK. Crystallographic and kinetic studies of human mitochondrial acetoacetyl-CoA thiolase: The importance of potassium and chloride ions for its structure and function. *Biochemistry*. 2007; 46(14):4305-4321.
21. Yang SY, Elzinga M. Association of both enoyl coenzyme A hydratase and 3-hydroxyacyl coenzyme A epimerase with an active site in the amino-terminal domain of the multifunctional fatty acid oxidation protein from *Escherichia coli*. *J Biol Chem*. 1993;268(9):6588-6592.
22. Schmitz W, Albers C, Fingerhut R, Conzelmann E. Purification and characterization of an α-methylacyl-CoA racemase from human liver. *Eur J Biochem*. 1995;231(3):815-822.
23. Charrier C, Duncan GJ, Reid MD, et al. A novel class of CoA-transferase involved in short-chain fatty acid metabolism in butyrate-producing human colonic bacteria. *Microbiology*. 2006; 152(1):179-185.
24. Misra I, Charlier HA, Miziorko HM. Avian cytosolic 3-hydroxy-3-methylglutaryl-CoA synthase: evaluation of the role of cysteines in reaction chemistry. *Biochim Biophys Acta*. 1995;1247(2):253-259.
25. Fukao T, Mitchell GA, Song XQ, et al. Succinyl-CoA:3-ketoacid CoA transferase (SCOT): Cloning of the human SCOT gene, tertiary structural modeling of the human SCOT monomer, and characterization of three pathogenic mutations. *Genomics*. 2000;68(2):144-151.
26. Göbel M, Kassel-Cati K, Schmidt E, Reineke W. Degradation of aromatics and chloroaromatics by Pseudomonas sp. strain B13: Cloning, characterization, and analysis of sequences encoding 3-oxoadipate: succinyl-coenzyme A (CoA) transferase and 3-Oxoadipyl-CoA thiolase. *J Bacteriol*. 2002;184(1):216-223.
27. Saxena P, Yadav G, Mohanty D, Gokhale RS. A New Family of Type III Polyketide Synthases in *Mycobacterium tuberculosis*. *J Biol Chem*. 2003;278(45):44780-44790.
28. Kumar P, Schelle MW, Jain M, et al. PapA1 and PapA2 are acyltransferases essential for the biosynthesis of the *Mycobacterium tuberculosis* virulence factor Sulfolipid-1. *Proc Natl Acad Sci U S A*. 2007; 104(27):11221-11226.
29. Hatzios SK, Schelle MW, Holsclaw CM, et al. PapA3 is an acyltransferase required for polyacyltrehalose biosynthesis in *Mycobacterium tuberculosis*. *J Biol Chem*. 2009;284(19):12745-12751.
30. Ryu H, Grove TL, Almo SC, Kim J. Identification of a novel tRNA wobble uridine modifying activity in the biosynthesis of 5-methoxyuridine. *Nucleic Acids Res*. 2018;46(17):9160-9169.
31. Taura F, Tanaka S, Taguchi C, et al. Characterization of olivetol synthase, a polyketide synthase putatively involved in cannabinoid biosynthetic pathway. *FEBS Lett*. 2009;583(12):2061-2066.
32. Trivedi OA, Arora P, Vats A, et al. Dissecting the mechanism and assembly of a complex virulence mycobacterial lipid. *Mol Cell*. 2005; 17(5):631-643.
33. Hendrickson L, Davis CR, Roach C, et al. Lovastatin biosynthesis in Aspergillus terreus: Characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene. *Chem Biol*. 1999;6(7):429-439.
34. Ames BD, Nguyen C, Bruegger J, et al. Crystal structure and biochemical studies of the trans-acting polyketide enoyl reductase LovC from lovastatin biosynthesis. *Proc Natl Acad Sci U S A*. 2012; 109(28):11144-11149.
35. Reed MAC, Schweizer M, Szafranska AE, et al. The type I rat fatty acid synthase ACP shows structural homology and analogous biochemical properties to type II ACPs. *Org Biomol Chem*. 2003;1(3): 463-471.
36. Chopra T, Banerjee S, Gupta S, et al. Novel intermolecular iterative mechanism for biosynthesis of mycoketide catalyzed by a bimodular polyketide synthase. *PLoS Biol*. 2008;6(7):1584-1598.
37. Lin JL, Wheeldon I. Dual N- and C-terminal helices are required for endoplasmic reticulum and lipid droplet association of alcohol acetyltransferases in *Saccharomyces cerevisiae*. *PloS One*. 2014;9(8): e104141.
38. Wright TC, Cant JP, Brenna JT, McBride BW. Acetyl CoA carboxylase shares control of fatty acid synthesis with fatty acid synthase in bovine mammary homogenate. *J Dairy Sci*. 2006;89(7):2552-2558.
39. Fischer RJ, Helms J, Durre P. Cloning, sequencing, and molecular analysis of the sol operon of Clostridium acetobutylicum, a chromosomal locus involved in solventogenesis. *J Bacteriol*. 1993;175(21): 6959-6969.
40. Zhang M, Xu HY, Wang YC, Shi ZB, Zhang NN. Structure of succinyl-CoA:3-ketoacid CoA transferase from *Drosophila melanogaster*. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2013;69(10): 1089-1093.
41. Toyota CG, Berthold CL, Gruez A, et al. Differential substrate specificity and kinetic behavior of *Escherichia coli* YfdW and *Oxalobacter formigenes* formyl coenzyme A transferase. *J Bacteriol*. 2008;190(7): 2556-2564.
42. Savolainen K, Bhaumik P, Schmitz W, et al. α-Methylacyl-CoA racemase from *Mycobacterium tuberculosis*: Mutational and structural characterization of the active site and the fold. *J Biol Chem*. 2005; 280(13):12611-12620.
43. Harijan RK, Kiema TR, Karjalainen MP, et al. Crystal structures of SCP2-thiolases of trypanosomatidae, human pathogens causing widespread tropical diseases: The importance for catalysis of the cysteine of the unique HDCF loop. *Biochem J*. 2013;455(1):119-130.
44. Liang K, Li N, Wang X, et al. Cryo-EM structure of human mitochondrial trifunctional protein. *Proc Natl Acad Sci U S A*. 2018;115(27): 7039-7044.
45. Antonenkov VD, Van Veldhoven PP, Waelkens E, Mannaerts GP. Substrate specificities of 3-oxoacyl-CoA thiolase a and sterol carrier protein 2/3-oxoacyl-coa thiolase purified from normal rat liver peroxisomes. Sterol carrier protein 2/3-oxoacyl-CoA thiolase is involved in the metabolism of 2-methyl-branched fatty ac. *J Biol Chem*. 1997;272(41):26023-26031.
46. Sutherlin A, Hedl M, Sanchez-Neri B, Burgner JW, Stauffacher CV, Rodwell VW. Enterococcus faecalis 3-hydroxy-3-methylglutaryl coenzyme A synthase, an enzyme of isopentenyl diphosphate biosynthesis. *J Bacteriol*. 2002;184(15):4065-4070.
47. Clinkenbeard KD, Reed WD, Mooney RA, Lane MD. Intracellular localization of the 3-hydroxy-3-methylglutaryl coenzme A cycle enzymes in liver. Separate cytoplasmic and mitochondrial 3-hydroxy-3-methylglutaryl coenzyme A generating systems for cholesterogenesis and ketogenesis. *J Biol Chem*. 1975;250(8):3108-3116.
48. Choi KH, Heath RJ, Rock CO. β-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. *J Bacteriol*. 2000;182(2):365-370.
49. Frias JA, Richman JE, Erickson JS, Wackett LP. Purification and characterization of OleA from *Xanthomonas campestris* and demonstration of a non-decarboxylative claisen condensation reaction. *J Biol Chem*. 2011;286(13):10930-10938.
50. Bentley SD, Vernikos GS, Snyder LAS, et al. Meningococcal genetic variation mechanisms viewed through comparative analysis of serogroup C strain FAM18. *PLoS Genet*. 2007;3(2):230-240.
51. Heidelberg JF, Elsen JA, Nelson WC, et al. DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae. *Nature*. 2000;406(6795):477-483.
52. Kilgore MB, Augustin MM, Starks CM, et al. Cloning and characterization of a norbelladine 4′-o-methyltransferase involved in the biosynthesis of the Alzheimer's drug galanthamine in Narcissus sp. aff. pseudonarcissus. *PloS One*. 2014;9(7):e103223.
53. Alazizi A, Liu MY, Williams FE, et al. Identification, characterization, and ontogenic study of a catechol O-methyltransferase from zebrafish. *Aquat Toxicol*. 2011;102(1–2):18-23.
54. Yonekura-Sakakibara K, Tanaka Y, Fukuchi-Mizutani M, et al. Molecular and biochemical characterization of a novel hydroxycinnamoyl-

CoA: anthocyanin 3-O-glucoside-6-O-acyltransferase from Perilla frutescens. *Plant Cell Physiol*. 2000;41(4):495-502.
55. Dexter R, Qualley A, Kish CM, et al. Characterization of a petunia acetyltransferase involved in the biosynthesis of the floral volatile isoeugenol. *Plant J*. 2007;49(2):265-275.
56. Peng M, Gao Y, Chen W, et al. Evolutionarily distinct BAHD N-acyltransferases are responsible for natural variation of aromatic amine conjugates in rice. *Plant Cell*. 2016;28(7):1533-1550.
57. Suzuki H, Sawada S, Watanabe K, et al. Identification and characterization of a novel anthocyanin malonyltransferase from scarlet sage (Salvia splendens) flowers: An enzyme that is phylogenetically separated from other anthocyanin acyltransferases. *Plant J*. 2004;38(6):994-1003.
58. Sfakianos MK, Wilson L, Sakalian M, Falany CN, Barnes S. Conserved Residues in the Putative Catalytic Triad of Human Bile Acid Coenzyme A:Amino Acid N-Acyltransferase. *J Biol Chem*. 2002;277(49):47270-47275.
59. Hunt MC, Rautanen A, Westin MAK, Svensson LT, Alexson SEH. Analysis of the mouse and human acyl-CoA thioesterase (ACOT) gene clusters shows that convergent, functional evolution results in a reduced number of human peroxisomal ACOTs1. *FASEB J*. 2006;20(11):1855-1864.
60. Mullins EA, Francois JA, Kappock TJ. A Specialized Citric Acid Cycle Requiring Succinyl-Coenzyme A (CoA):Acetate CoA-Transferase (AarC) Confers Acetic Acid Resistance on the Acidophile *Acetobacter aceti*. *J Bacteriol*. 2008;190(14):4933-4940.
61. Ridley CP, Khosla C. *Polyketides Defining Statement Biosynthesis Environmental Roles Human Uses Genomes of Prolific Polyketide Producers Search for New Polyketides Heterologous Expression of Polyketides Engineering of PKSs Conclusions Further Reading*. Elsevier; 2009:472-481.
62. Schmidt W, Beerhues L. Alternative pathways of xanthone biosynthesis in cell cultures of *Hypericum androsaemum* L. *FEBS Lett*. 1997;420(2–3):143-146.
63. Katsuyama Y, Kita T, Horinouchi S. Identification and characterization of multiple curcumin synthases from the herb *Curcuma longa*. *FEBS Lett*. 2009;583(17):2799-2803.
64. Chen H, Tseng CC, Hubbard BK, Walsh CT. Glycopeptide antibiotic biosynthesis: Enzymatic assembly of the dedicated amino acid monomer (S)-3,5-dihydroxyphenylglycine. *Proc Natl Acad Sci U S A*. 2001;98(26):14901-14906.
65. Schanz S, Schröder G, Schröder J. Stilbene synthase from Scots pine (Pinus sylvestris). *FEBS Lett*. 1992;313(1):71-74.
66. Jacob U, Mack M, Clausen T, Huber R, Buckel W, Messerschmidt A. Glutaconate CoA-transferase from *Acidaminococcus fermentans*: The crystal structure reveals homology with other CoA-transferases. *Structure*. 1997;5(3):415-426.
67. Gupta RD. Recent advances in enzyme promiscuity. *Sustain Chem Process*. 2016;4(1):1-7.
68. Khersonsky O, Tawfik DS. Enzyme promiscuity: A mechanistic and evolutionary perspective. *Annu Rev Biochem*. 2010;79:471-505.
69. Haller T, Buckel T, Rétey J, Gerlt JA. Discovering new enzymes and metabolic pathways: Conversion of succinate to propionate by *Escherichia coli*. *Biochemistry*. 2000;39(16):4622-4629.
70. Chen Y, Kelly EE, Masluk RP, Nelson CL, Cantu DC, Reilly PJ. Structural classification and properties of ketoacyl synthases. *Protein Sci*. 2011;20(10):1659-1667.
71. Kursula P, Sikkilä H, Fukao T, Kondo N, Wierenga RK. High resolution crystal structures of human cytosolic thiolase (CT): A comparison of the active sites of human CT, bacterial thiolase, and bacterial KAS I. *J Mol Biol*. 2005;347(1):189-201.
72. Lou JW, Iyer KR, Hasan SMN, Cowen LE, Mazhab-Jafari MT. Electron cryomicroscopy observation of acyl carrier protein translocation in type I fungal fatty acid synthase. *Sci Rep*. 2019;9(1):1-8.
73. Li Z, Huang Y, Ge J, et al. The crystal structure of MCAT from *Mycobacterium tuberculosis* reveals three new catalytic models. *J Mol Biol*. 2007;371(4):1075-1083.
74. Hong SK, Kim KH, Park JK, Jeong KW, Kim Y, Kim EEK. New design platform for malonyl-CoA-acyl carrier protein transacylase. *FEBS Lett*. 2010;584(6):1240-1244.
75. Bunkoczi G, Misquitta S, Wu X, et al. Structural basis for different specificities of acyltransferases associated with the human cytosolic and mitochondrial fatty acid synthases. *Chem Biol*. 2009;16(6):667-675.
76. Lloyd AJ, Shoolingin-Jordan PM. Dimeric pig heart succinate-coenzyme A transferase uses only one subunit to support catalysis. *Biochemistry*. 2001;40(8):2455-2467.
77. Berthold CL, Toyota CG, Richards NGJ, Lindqvist Y. Reinvestigation of the catalytic mechanism of formyl-CoA transferase, a class III CoA-transferase. *J Biol Chem*. 2008;283(10):6519-6529.
78. Mathieu M, Jp Z, Pauptit R, Erdmann R, Kunau WH, Wierenga R. The 2.8å Crystal Structure of peroxisomal 3-ketoacyl-CoA thiolase of *Saccharomyces cerevisiae*: a five-layered αβαβα structure constructed from two core domains of identical topology. *Structure*. 1994;2(9):797-808.
79. Steussy CN, Vartia AA, Burgner JW, Sutherlin A, Rodwell VW, Stauffacher CV. X-ray crystal structures of HMG-CoA synthase from Enterococcus faecalis and a complex with its second substrate/inhibitor acetoacetyl-CoA. *Biochemistry*. 2005;44(43):14256-14267.
80. Rodríguez JC, Gil-Gómez G, Hegardt FG, Haro D. Peroxisome proliferator-activated receptor mediates induction of the mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase gene by fatty acids. *J Biol Chem*. 1994;269(29):18767-18772.
81. Goblirsch BR, Frias JA, Wackett LP, Wilmot CM. Crystal structures of *Xanthomonas campestris* OleA reveal features that promote head-to-head condensation of two long-chain fatty acids. *Biochemistry*. 2012;51(20):4138-4146.
82. Fellenberg C, Milkowski C, Hause B, et al. Tapetum-specific location of a cation-dependent O-methyltransferase in *Arabidopsis thaliana*. *Plant J*. 2008;56(1):132-145.
83. Ferrer JL, Zubieta C, Dixon RA, Noel JP. Crystal structures of alfalfa caffeoyl coenzyme A 3-O-methyltransferase. *Plant Physiol*. 2005;137(3):1009-1017.
84. Wang C, Li J, Ma M, et al. Structural and biochemical insights into two BAHD acyltransferases (AtSHT and AtSDT) involved in phenolamide biosynthesis. *Front Plant Sci*. 2021;11:610118.
85. Levsh O, Chiang YC, Tung CF, Noel JP, Wang Y, Weng JK. Dynamic conformational states dictate selectivity toward the native substrate in a substrate-permissive acyltransferase. *Biochemistry*. 2016;55(45):6314-6326.
86. Panchal V, Jatana N, Malik A, et al. A novel mutation alters the stability of PapA2 resulting in the complete abrogation of sulfolipids in clinical mycobacterial strains. *FASEB Bioadv*. 2019;1(5):306-319.
87. Mullins EA, Kappock TJ. Crystal Structures of *Acetobacter aceti* Succinyl-Coenzyme A (CoA):Acetate CoA-Transferase Reveal Specificity Determinants and Illustrate the Mechanism Used by Class I CoA-Transferases. *Biochemistry*. 2012;51(42):8422-8434.
88. Jez JM, Ferrer JL, Bowman ME, Dixon RA, Noel JP. Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. *Biochemistry*. 2000;39(5):890-902.
89. Chen XH, Vater J, Piel J, et al. Structural and functional characterization of three polyketide synthase gene clusters in Bacillus amyloliquefaciens FZB 42. *J Bacteriol*. 2006;188(11):4024-4036.
90. Mullins EA, Sullivan KL, Kappock TJ. Function and X-Ray crystal structure of *Escherichia coli* YfdE. *PloS One*. 2013;8(7):e67901.
91. Nguyen TM, Sawyer JK, Kelley KL, Davis MA, Rudel LL. Cholesterol esterification by ACAT2 is essential for efficient intestinal cholesterol absorption: evidence from thoracic lymph duct cannulation. *J Lipid Res*. 2012;53(1):95-104.
92. Crowe AM, Casabon I, Brown KL, et al. Catabolism of the last two steroid rings in *Mycobacterium tuberculosis* and other bacteria. *MBio*. 2017;8(2). https://doi.org/10.1128/mbio.00321-17

93. Deckert G, Warren PV, Gaasterland T, et al. The complete genome of the hyperthermophilic bacterium Aquifex aeolicus. *Nature*. 1998; 392(6674):353-358.
94. Sassa T, Ohno Y, Suzuki S, et al. Impaired epidermal permeability barrier in mice lacking Elovl1, the gene responsible for very-long-chain fatty acid production. *Mol Cell Biol*. 2013;33(14): 2787-2796.
95. Engels B, Heinig U, Grothe T, Stadler M, Jennewein S. Cloning and characterization of an Armillaria gallica cDNA encoding protoilludene synthase, which catalyzes the first committed step in the synthesis of antimicrobial melleolides. *J Biol Chem*. 2011;286(9): 6871-6878.
96. Huang X, Zhang W, Tang S, Wei S, Lu X. Collaborative biosynthesis of a class of bioactive azaphilones by two separate gene clusters containing four PKS/NRPSs with transcriptional crosstalk in fungi. *Angew Chem Int Ed Engl*. 2020;59(11):4349-4353.
97. Hu Y, Cronan JE. α-Proteobacteria synthesize biotin precursor pimeloyl-ACP using BioZ 3-ketoacyl-ACP synthase and lysine catabolism. *Nat Commun*. 2020;11(1):5598.
98. Huhtinen K, O'Byrne J, Lindquist PJG, Contreras JA, Alexson SEH. The peroxisome proliferator-induced cytosolic type I acyl-CoA thioesterase (CTE-I) is a serine-histidine-aspartic acid alpha /beta hydrolase. *J Biol Chem*. 2002;277(5):3424-3432.
99. Lee FJ, Lin LW, Smith JA. Acetyl-CoA hydrolase involved in acetate utilization in *Saccharomyces cerevisiae*. *Biochim Biophys Acta*. 1996; 1297(1):105-109.
100. Zheng D, Schröder G, Schröder J, Hrazdina G. Molecular and biochemical characterization of three aromatic polyketide synthase genes from Rubus idaeus. *Plant Mol Biol*. 2001;46(1):1-15.

How to cite this article: de Carvalho CC, Murray IP, Nguyen H, Nguyen T, Cantu DC. Acyltransferase families that act on thioesters: Sequences, structures, and mechanisms. *Proteins*. 2024;92(2):157-169. doi:10.1002/prot.26599