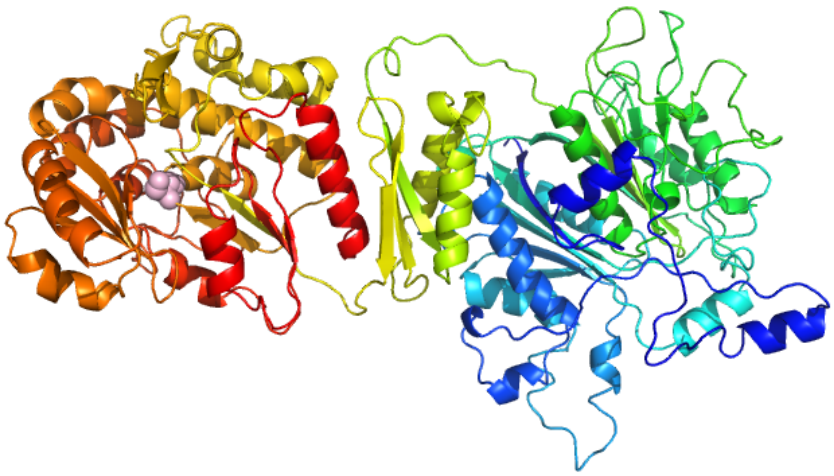
Change in Substrate Binding Specificity of the Tandem Acyl Carrier Protein Domains of Polyunsaturated Fatty Acid Synthesis



Carla Vela Iglesias

Master in Molecular Biology and Biomedicine

Department of Molecular Biology, UC

Department of Microbiology and Genomics, IBBTEC

Academic course 2017-2018

Supervised by Dr. Gabriel Moncalian

Acknowledgments

It is a great pleasure to acknowledge my deepest thanks and gratitude to Dr. Gabriel Moncalian for his kind guidance. It is a great honor to work under his supervision. Also, I would like to express my thanks and sincere appreciation to all the members of the 02.08 laboratory, Laura, Omar, Lorena, Raul, Arantza, Jorge, and Martin. Thank you for all your help, support and patience. I want to thank as well to all my coworkers of the IBBTEC.

I am most grateful to my family for all their emotional support and their patients during this last year, especially to my beloved mother. A year without you feels like an eternity.

Index

Abbreviations

ACP	Acyl carrier protein
ALA	α-Linolenic acid
Amp	Ampicillin
ARA	Arachidonic acid
AT	Acyl transferase
CLF	Chain length factor
CoA	Coenzyme-A
DAG	Diacylglycerol
DEBS	6-Deoxyerythronolide B synthase
DGAT	Diacylglycerol acyltransferase
DH	Dehydrogenase
DHA	Docosahexaenoic acid
DNA	Deoxyribonucleic acid
EPA	Eicosapentaenoic acid
ER	Enoyl reductase
FA	Fatty acid
FAS	Fatty acid synthase
IPTG	Isopropyl-β-D-1tiogalactopiranósido
kDa	Kilodalton
Kn	Kanamycin
KR	Keto reductase
KS	Keto synthase
LC	Long chain
LD	Lipid droplet
MAT	Malonyl acyltransferase
NADPH	Nicotinamideadenine dinucleotide phosphate
PHA	Polyhydroxyalkanoate
PKS	Polyketide synthase
PMSF	Phenylmethylsulfonyl fluoride
PPT	4-phosphopantetheine
PPTase	Phosphopantetheinyl transferase
PUFA	Polyunsaturated fatty acids
SDS	Short-chain dehydrogenase/reductase
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

1. Abstract

Polyunsaturated fatty acids synthases are large enzymatic complexes that produce PUFAs, such as eicosapentaenoic acid (EPA, 20:5ω3) and docosahexaenoic acid (DHA, 22:6ω3). This omega-3 fatty acids are essential for the human health and nutrition. In general, ω-3 fatty acids are found in fish, seeds, and some unicellular organisms like marine gammaproteobacteria, myxobacteria or schizochytrium. PUFA synthases are analogous to polyketide synthases (PKS), which are proteins involve in the bacterial antibiotic production mechanism. In marine proteobacteria, PUFA synthases enzymology and biochemistry understanding continues to be rudimentary. The system contains a main conserved long protein known as PfaA, which features multiple acyl carrier protein (ACP) domains. In this project, PfaA domain, contained in the pfa cluster of Moritella marina, has been studied to understand the mechanism in which PfaA initiates the synthesis by the selection of malonyl units. Also, the objective of this study was to observe the change in substrate binding specificity of the five tandem acyl carrier proteins when a one-point mutation is introduced in the serine 703 of the active center of the AT domain. The results performed by radioactivity assays showed that when serine 703 is mutated, malonyl-coA cannot bind to the protein and so the synthesis cannot be performed. However, when the radiolabeled acetyl-coA was introduced with the mutated serine 703, it was observed that the acetyl-coA was capable of binding to the protein. Moreover, ATS703A was able to acetylate PfaA ACP domain in vitro. At the same time, other one-point mutations were tried in other two near serine residues (serine 607 and 863) to identify with accuracy the binding site of acetyl-coA. Further experiments must be carried on in order to find why acetyl-coA binds the active center of AT mutated.

2. Introduction

2.1. Chemical nature of Omega-3 PUFAs

Omega-3 fatty acids are polyunsaturated fatty acids (PUFAs), which starting from the FA methyl end have a double bond that is located between the third and fourth carbons. All of the polyunsaturated fatty acids have an aliphatic chain. They are all in a cis conjugation structure and vary in the length of their chains, number and position of double bonds. The most significant PUFAs are omega-6 fatty acids such as arachidonic acid (20:4 ARA), and omega-3 fatty acids, such as eicosapentaenoic acid (20:5 EPA) and docosahexaenoic acid (22:6 DHA). For being the longest fatty acids, ARA, EPA, and DHA fatty acids are classified as long-chain polyunsaturated fatty acids (LC-PUFAs). Figure I-1 shows the chemical structure of all LC-PUFAs fatty acids and the omega-6 fatty acids.

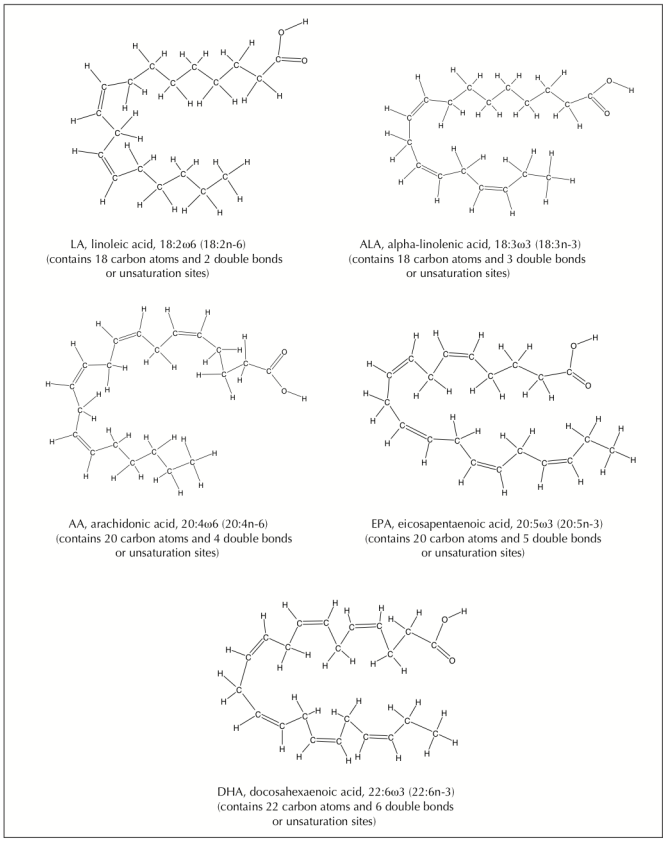


Figure I-1: the figure shows the chemical structures of LA, ALA, AA, EPA, and DHA. The figure was obtained from Holub, B. J. 2002.

In the past decades, most doctors and health organizations recommend to take omega-3 fatty acids supplements for pregnant woman and the ill Omega-3 LC-PUFAs such as EPA and DHA are have a fundamental role, since they are precursors of signaling molecules, such as prostaglandins, thromboxane, and leukotrienes (Tapiero et al. 2002), molecules involved in many processes. These molecules have an important role on the stimulation of growth, a protective property of the cardiovascular system, and the reduction of inflammatory symptoms (Wall et al. 2010). Researchers have also studied the benefits of omega-3 FA intake in various diseases, such as Alzheimer's disease, allergies, and cancer (Cole et al. 2005; Rose and Connolly 1999). It is recommended a balanced diet that includes the same ratio of intake of omega-6 and omega-3 fatty acids through one's life.

Most eukaryote organisms produce PUFAs in nature by desaturation and elongation of short saturated fatty acids. In this pathway, desaturases introduce cis double bonds into the saturated fatty acids to synthesize PUFAs (Pereira, Leonard, and Mukerji 2003). In the case of human metabolism, we are incapable of synthesizing ALA since higher animals do not poses LC-PUFA desaturases. Therefore, we must obtain our sources of omega-3 FA by consuming produce in high quantities of ALA, EPA and DHA.

The current mayor sources of omega-3 FA in nature are the marine products such as fish and fish oils. Nevertheless, there are other products like microalgae and some proteobacteria from the deep sea that are able to produce important amounts of LC-PUFAs. The terrestrial organisms and the lipids from marine organisms contain a wide variety of fatty acids. Usually, marine oil products consist of C14 to C22 fatty acids.

Also, some of the known marine bacteria that are capable of producing omega-3 fatty acids are from the genus Colwellia, Moritella, Shewanella, and Photobacterium. These bacteria use an alternative biosynthetic pathway which synthesizes LC-PUFAs from short substrates. They have common genes known as pfa clusters, which code for the protein complex involved in the synthesis of LC-PUFAs. The pfa cluster has a similar domain configuration as bacterial polyketide synthases (PKS) or fatty acid synthases (FAS).

Currently, most supplements in the market are made with fish, which are not actually recommended as some species of fish contain a high level of contaminants and fisheries are overexploited. Supplements of EPA and DHA are presently used in the pharmaceutical industry,

to produce food for animals, and other comestibles such as infant products. However, there is not enough production of fish oils to sustain the massive demand of omega-3 fatty acids. Therefore, alternative sources of LC-PUFAs are required in order to fulfill the demand. Bacteria and eukaryotic microorganisms have been proposed as new sources of LC-PUFAs. Some of the advantages of using microorganisms are their constant and stable supply of the final product, as well as, the lack of contaminants and allergies. Nowadays, there are a number of eukaryotic microorganisms that are used to produce LC-PUFAs. However, the production of LC-PUFAs by bacterial microorganisms is still being investigated as the mechanism of synthesis of fatty acids is almost unknown and has a low productivity rate compared with other sources.

2.2. Health Benefits of ω-3 Fatty Acids in Humans

Over the last two decades, omega-3 long-chain PUFA, such as EPA and DHA, and omega-6 fatty acids, such as AA, have been intensely studied due to their impact on human health. Omega-3 long-chain PUFA play a key role in anti-inflammatory processes, as well as, in the viscosity of the cell membranes. In the case of omega-3 fatty acids, they have anti-inflammatory, antiarrhythmic, and anti-thrombotic properties. Moreover, DHA plays an essential role in fetal development and aging, since, DHA is an important component of cell membranes from the brain and retina. Additionally, EPA and DHA are essential precursors of several metabolites, which are fundamental for lipid mediators that could be beneficial for the prevention or treatment of severe diseases (Swanson et al. 2012).

However, as previously said, it can be arduous to get a balance intake of EPA and DHA through our diet alone, even though EPA and DHA are mainly produced by marine plant like algae and other marine animals, such as grease fish. For humans, it is not possible to synthesize another type of omega-3 fatty acid with a shorten chain named α-linolenic acid (ALA), which is commonly found in land plants and seeds. We can convert ALA into EPA and DHA by enlongase and desaturase enzymes but unfortunately just small amounts of omega-3 can be synthesize by this process.

On the other hand, over the last century, the human diet has changed drastically by increasing the amount of ω-6 fatty acids intake and lowering the intake of ω-3 PUFA. Omega-6 fatty acids have proinflammatory and prothrombotic properties (Cowington, M. B. 2004). An increase in the consumption of products that contain high amounts of ω-6 fatty acids, such as vegetable oils, has led to an increase in cardiovascular disease, poor fetal development, increase in inflammatory processes, and an increase in developing Alzheimer's disease, as well as, increasing the risk of sudden death and play a key role in the development of behavioral diseases.

In summary, it is beneficial for the human health to have a balanced diet that includes, as well as other omega-3 fatty acids. Researchers have proved a correlation between the intake of essential omega-3 PUFA and the prevention or treatment of some diseases. Epidemiological studies concluded that sick as well as healthy people should increase the consumption of fish or fish oil products, which are important sources of omega-3 fatty acids (Holub, B. J. 2002). Nowadays, in the market, there are omega-3 fatty acids supplements that are highly recommended for people that suffers cardiopathies, and atherosclerosis. Also, it has been proved that omega-3 could have beneficial effects in people at risk of cerebrovascular events (Chowdhury, R. et al, 2012). The supplements are a positive approach to increase the intake of ω-3 PUFA, as they don't contain any contaminant like some fishes do (Kris-Etherton et al 2002). Similarly, there are studies, which indicate that omega-3 polyunsaturated fatty acids are potential treatment targets for a variety of neurodegenerative and neurological disorders (Dyall, S. C. 2015). Further investigations are needed in order to determine the full potential of omega-3 PUFA in the human health.

2.3. Overall mechanism of fatty acid synthesis

The general biosynthesis of fatty acids begins with the incorporation of an acetyl-CoA unit to form a new fatty acid molecule. Then, a consecutive series of carboxylation rounds incorporates the malonyl-CoA to the growing chain. Consequently, malonyl-CoA will be needed in all elongation steps and will be synthesized in previous reactions, which will be catalyzed by acetyl-CoA carboxylase that uses bicarbonate and acetate as precursors (Alberts and Vagelos 1968). After, the new FA molecule will be reduced at the end of each condensation step until the fatty acid molecule matures for being use by the cell. In almost all animal FA synthesis, there is an even number of carbon atoms due to the assembly of segments of two carbons length together (Wakil, Stoops, and Joshi 1983).

There are two types of fatty acid synthesis pathways known as type I and type II FAS. Both type I and type II FAS require a small protein named acyl carrier protein (ACP) to transport the fatty acid chain from one domain to the other so the condensation and reduction steps can be done (Chan and Vogel 2010). Via post-translational modification the ACP synthesis is activated by the phosphopantetheinyl transferase (PPTase) enzyme, which adds a coenzyme A derived from a PPT, indispensable to form the tether upon which the growing fatty acid is bound (Flugel et al 2000). Conclusively, the central role of the ACP domain is to keep the fatty acid chain attached and exposed to the rest of the protein domains.

The function of the malonyl-CoA acyl transferase domain (AT or MAT if specific for malonyl-CoA) is to transfer the malonyl-CoA units to the ACPs. In E. coli fabD is responsible for catalyze the loading of fatty acid building blocks by the conversion of malonyl-CoA to malonyl-CoA (Rock and Cronan 1996). Generally, the AT domain is first malonylated at the active site serine residue. Then, the malonyl groups are transported to the ACP PPT, activating them in the process for the successive synthesis reactions. The synthesis is always initiated by a molecule of acetyl-CoA or acetyl-ACP, even though malonyl is always the basic extender unit (John E. Cronan and Rock 2008), this initial acetyl molecule is condensed with malonyl-ACP to form acetoacetyl-ACP in the first synthesis cycle. In E. coli this process is due in the domain FabH. The acetate unit would become the methyl end of the future fatty acid. Conclusively, the AT domains are the starters and the leading biosynthesis proteins, which select the fundamental blocks with which the fatty acid would be constructed. AT domains have preferences for malonyl-CoA but they can also accept multiple CoA analogs in vitro (Ghayourmanesh and Kumar 1981; Joshi, Witkowski, and Smith 1997).

Once the ACPs are charged with the malonyl-CoA, the ketoacyl synthase (KS) domains located adjacent to the AT domains will extent the fatty acid chain. In E. coli the KS domains are represented as FabH, FabB or FabF, either one of them. The KS domains catalyze, through an iterative process, the chain extension via successive attachments of the α-carbon of the malonyl-thioester onto the acyl-thioester, which intermediates with decarboxylation of the malonyl group, producing β-keto-thioesters as the resulting product (Johan Gotthardt Olsen et al. 2001). Consequently, the fatty acid molecule is extended by the KS domains by channeling the malonyl groups previously selected and bound to the ACP domain by the acyl transferase. The characteristic bonding pattern of the fatty acids is produced by the β-keto groups that are generated after each condensation and subsequently reduced.

In this phase of the synthesis process, the domains keto reductase (KR), dehydratase (DH), and enoyl reductase (ER) are essential and they are perfectly coordinated between them. The previously generated and condensed keto-group is first reduced by the KR domains (FabG in E. coli), that is typically an NADPH-dependent 3-ketoacyl-ACP reductase (Heath and Rock 1995). After, DH domain, a 3-hydroxyacyl-ACP dehydratase (FabZ and FabA in E. coli) that generates a trans or cis double bond, removes a water molecule (Heath and Rock 1996). The final reduction step is catalyzed by the ER domain or in the case of E. coli FabI, which gives an acyl-ACP that can be used as additional substrate if the expected chain length hasn't been reached. To be able to generate the fatty acids with unsaturations at different position, it isn't necessary to

make reductions in all the cycles. The ER domains are able to remove the double bonds but it is conditioned by its isomeric form. It normally catalyzes the reduction of the trans-2-acyl-ACP isomer (Massengo-Tiassé and Cronan 2009).

The final fatty acid chain is released by a thioesterase (TE) domain when the fatty acid has reached its definitive size. The chain-termination step of fatty acid synthesis is due by the TE domain via the hydrolysis of the acyl-S-Ppant thioester bound to the ACP domain (Chakravarty et al. 2004). Figure I-2 shows the complete summary of the palmitate synthesis.

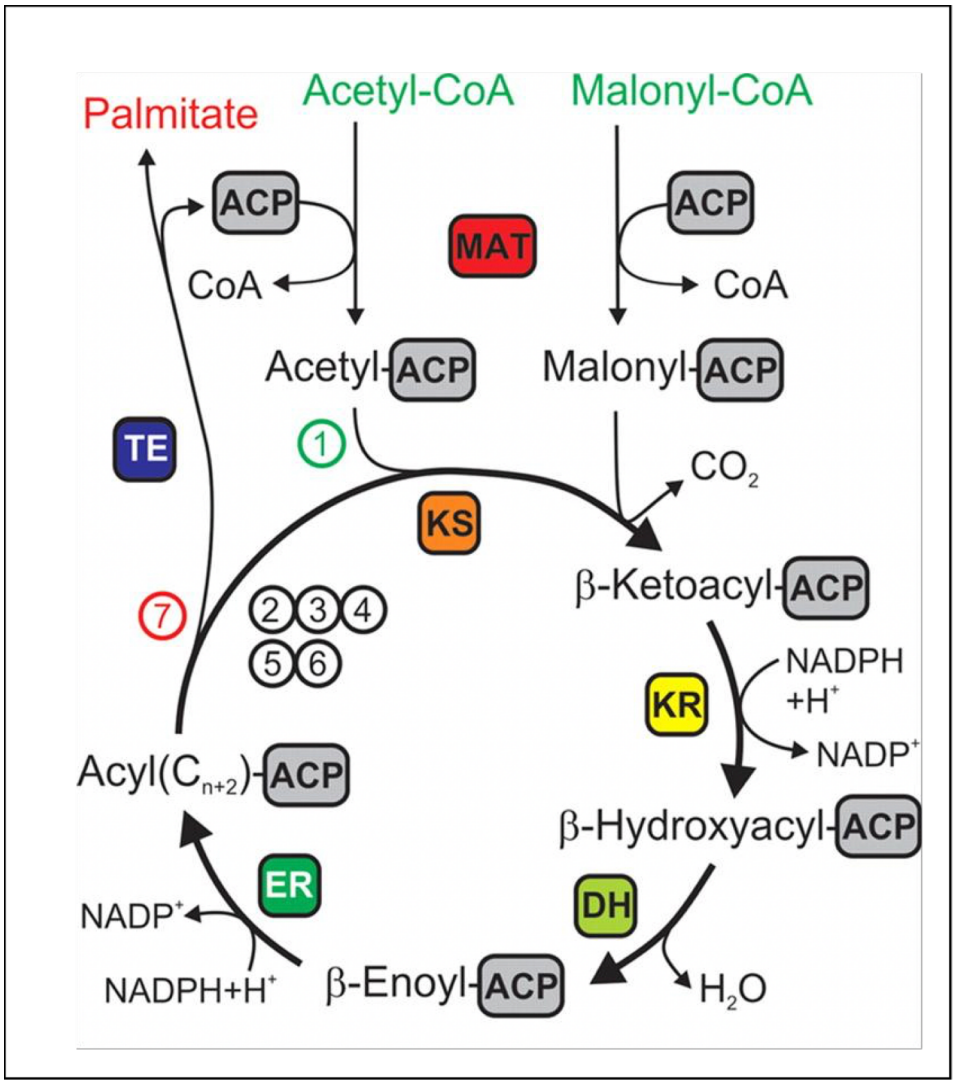


Figure I-2: The figure shows the catalytic cycle model for palmitate synthesis. The FAS synthesis is initiated by the transfer of the acyl moiety of acetyl-CoA to the ACP domain, which is then catalyzed by malonyl-CoA/acetyl-CoA-ACP-transacylase (MAT). The KS domain catalyzes the decarboxylative condensation of the acyl intermediated with malonyl-ACP to a β-ketoacyl-ACP intermediate. The KR domain then processes the β-carbon, resulting in β-hydroxyacyl-ACP dehydration by the DH domain to a β-enoyl. Then, it is reduced by a NADPH-dependent ER domain to form a four-carbon acyl substrate for further elongation processes until the substrate length is between C_{16} to C_{18} is reached. The final step consists in the release of the product from the ACP by the TE domain. The figure was taken from (Maier, Jenni, & Ban, 2006).

2.4. Omega-3 fatty acid synthesis

Traditionally, omega-3 fatty acids have been classified as secondary lipids since they are not included in the classical metabolism pathways of synthesis of PUFAs. Depending on the fatty acid, they have a chain formed by 18 carbons or more in a cis position. Omega-3 fatty acids take their name because of its proximity to the functional carboxylic acid group. Omega-3 polyunsaturated fatty acids have many double bonds starting from the third carbon atom when counting from the methyl end of the FA molecule (Holub, B. J. 2002).

PUFA synthesis in plants and some animals is achieved by special enzymes known as desaturases and elongases. To produce longer fatty acids, they use the FAS synthase routes to elongate their chain and introduce more double bonds, which generates the unsaturation of fatty acids. The function of the elongase enzymes is to incorporate malonyl units into the previous fatty acid, to elongate the chain. At the same time, desaturase enzymes introduce double bonds on the acyl chain at specific positions. These enzymes are necessary in order to produce long chain polyunsaturated fatty acids, such as EPA and DHA. The principal problem in mammalian cells is that they cannot express the 12 and 15 double bonds activities required in order to produce PUFAs by the de novo pathway. Conclusively, because mammals have a deficient production of ALA, consequently it must be consumed in the diet or by taking supplements.

The aerobic route of PUFA synthases uses desaturases as oxygenase enzymes, which removes the hydrogens from the acyl chain by using the activated molecular oxygen. The entire aerobic route requires more than 30 distinct enzymes and approximately 70 biochemical reactions (Metz et al. 2001). Its complexity makes it difficult to design genetic constructions to execute bioengineering experiments and to produce PUFAs in the laboratory.

2.4.1 Omega-3 FA aerobic pathway

PUFA synthesis in plants and some animals is achieved by special enzymes known as desaturases and elongases. To produce longer fatty acids, they use the FAS synthases routes by elongating their chain and introducing more double bonds, which generates fatty acids unsaturations. The function of the elongase enzymes is to incorporate malonyl units into the previous fatty acid, to elongate the chain. At the same time, desaturase enzymes introduce double

bonds on the acyl chain at specific positions. These enzymes are necessary in order to produce long chain polyunsaturated fatty acids, such as EPA and DHA. The principal problem in mammalian cells is that they cannot express the 12 and 15 double bond activities required in order to produce PUFAs by the de novo pathway. Conclusively, because mammals have a deficient production of EPA and DHA, these must be consumed in the diet or by taking supplements.

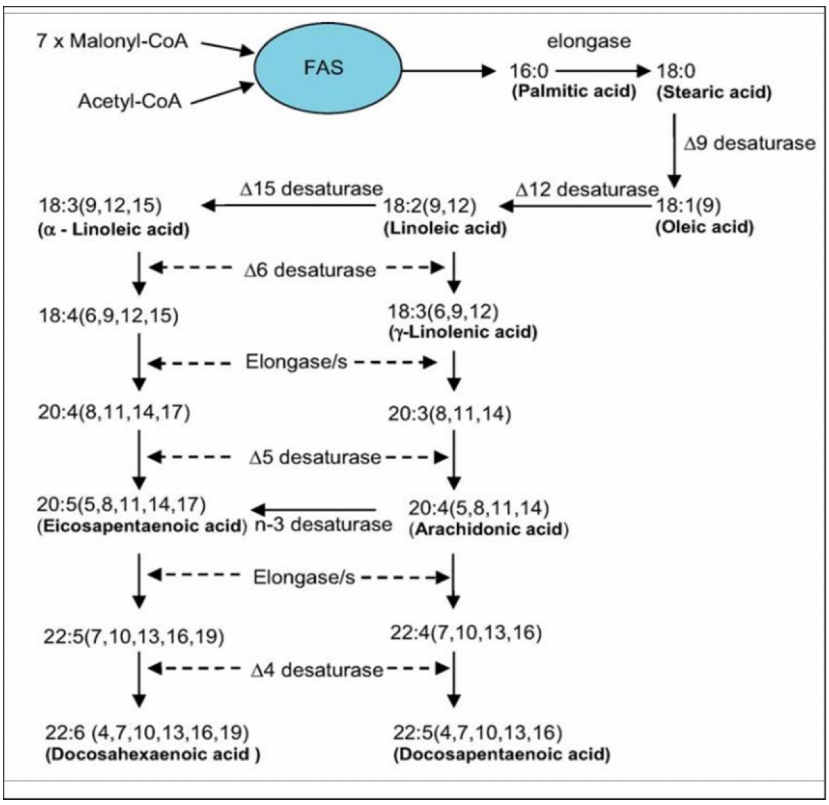


Figure I-3: Summary of the aerobic FAS route for the formation of polyunsaturated fatty acids (PUFAs). The FAS complex of enzymes utilizes acetyl-CoA and malonyl-CoA to start the synthesis of fatty acids. Then, the saturated fatty acid stearic acid C18:0 is successively desaturated and elongated through a series of reactions leading to the formation of various PUFAs (the n-3 and n-6 series). The figure was taken from Ratledge, 2004.

The aerobic route of PUFA synthases uses desaturases as oxygenase enzymes, which removes the hydrogens from the acyl chain by using the activated molecular oxygen. The entire aerobic route requires more than 30 distinct enzymes and approximately 70 biochemical reactions (Metz et al. 2001). Its humongous complexity makes it difficult to design genetic constructions to execute bioengineering experiments and to produce PUFAs in the laboratory.

2.4.2. Anaerobic pathway of pfa synthases

In nature, there is more than one pathway to synthesize PUFAs. Previously, it was described the aerobic pathway of polyunsaturated fatty acids synthesis in which elongases and aerobic desaturases are used in saturated products of FAS enzymes. In recent times, other pathway of fatty acids synthesis is the anaerobic pathway, which is able to produce omega-3 fatty acids from basic molecules. Researchers have discovered marine gamma-proteobacteria such as S. pealeana and M. marina, which are able to produce LC-PUFAs by using the anaerobic route of fatty acids synthesis. They have gene clusters independent from the aerobic pathway of PUFA synthases (Nichols et al. 1999). The first of these gene clusters type to be identify was in S. pealeana. It had a high homology with some enzymes involved in fatty acid synthesis. These gene clusters possessed five open reading frames (ORFs) that were fundamental and adequate for the EPA production in E. coli (Yazawa 1996[a]). Some years ago, it was discovered that Moritella marina has a homolog gene cluster, which produces DHA in E. coli (Orikasa et al. 2006). These new pathways share common characteristics with iterative PKSs type II and FAS systems domains, organization of gene clusters, sequence homology, and its biological structure. The de novo bacterial PUFA biosynthesis pathway is formed by these gene clusters that are known as the "pfa" cluster (Nishida et al. 2006; Orikasa et al. 2004). Figure I-3 shows the domain architecture of pfa genes and their homologs in different organisms.

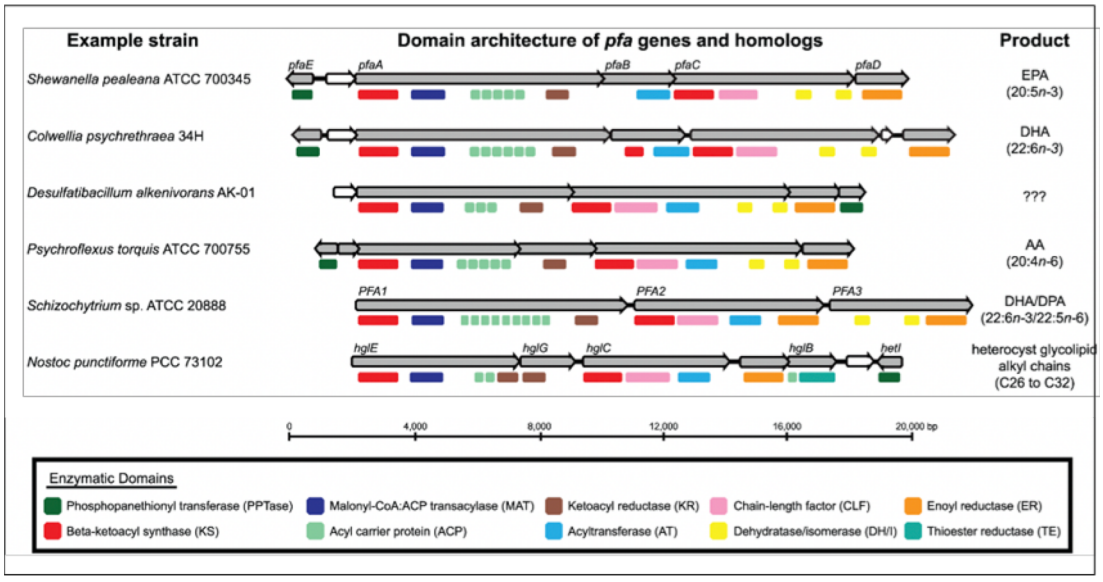


Figure I-4: The figure shows microbial genomes with the pfa cluster or a homolog. Also, it contains information about the structural organization of the clusters, as well as, the product that each organism can elaborate. The figure was obtained from Shulse & Allen, 2011.

These pfa synthases are extensive multifunctional enzyme complexes conformed by three to five subunits, depending on the organism, that are named as pfaA, pfaB, pfaC, pfaD, and pfaE. These protein complexes are indispensable for the synthesis of fatty acids. The common domains that can be found in these proteins as well as in PKSs are ketoacyl synthase (KS), acyltransferase (AT), acyl carrier protein (ACP), ketoacyl reductase (KR), dehydratase (DH), enoyl reductase (ER), and chain length factor (CFL). PfaE is a phosphopantetheine transferase (PPTase) which function is to start the synthesis of fatty acids by activating the ACPs with the PPT arm. In a regular FAS or PKS system, there is a single ACP domain but in the case of LC-PUFA synthases, there are multiple ACP domains (Jiang, Rajski, and Shen 2009). Also, some other domains such as KS, AT or DH are duplicated in some marine organisms. However, the reason for its duplication in a pfa cluster is still under investigation.

The biochemical reactions of PUFA anaerobic biosynthesis are established by the structural domains of the independent proteins of the pfa gene cluster and the similarities with them of the known FAS and PKS systems. Nonetheless, researchers have not been able to determine the individual enzymatic reactions of the genes responsible for the production of EPA or DHA (Yoshida et al. 2016). The initial biosynthesis reaction in LC-PUFA synthesis is likely to be the activation of ACPs of PfaA from their inactive apo-form to their active holo-form by the PPTase of the protein PfaE (Chan and Vogel 2010). In E. coli this biosynthesis reaction can be replaced by a heterologous gene that is not specific (Sugihara, Orikasa, and Okuyama 2010). The next step in the synthesis of fatty acids after the activation of the ACP domains, is the acquisition of acetyl and malonyl groups from acetyl-CoA and malonyl-CoA, which is catalyzed by the AT domain. Then, KS domain catalyzes the Claisen condensation reaction of the two molecules on the ACPs of PfaA with the release molecule of CO_2. Consequently, the keto group would be reduced by a keto reductase domain, which would form a hydroxyl moiety. Then, the hydroxyl moiety is dehydrated by a dehydratase domain that forms a trans-double bond, followed by a reduction to form a saturated C-C bond or an isomerization by and enoyl reductase and a dehydratase that finally forms a cis-double bond, which would linger in the final fatty acid chain (Yoshida et al. 2016).

The cluster of pfa genes is the responsible for the synthesis of EPA, DHA, and ARA. However, the mechanism in which PUFA synthases determine the final length of the carbon chains is not fully explained. The organisms S. pneumatophori and M. marina produce EPA and DHA respectively as was showed in figure I-3. The second gene of the pfa cluster, PfaB has a

KS-AT domain which seems to determine the final length of the carbon chain. Conclusively, this gene determines the production of EPA or DHA depending on the length of the carbon chain (Orikasa et al. 2009[a]). However, researchers do not completely understand why and how this protein have these effects in the final carbon chain length or if any other complex is involved in the process.

In recent years, researchers have tried to analyze the structure of the proteins that conform the pfa synthesis cluster by X-ray crystallography. Nevertheless, the crystallization of flexible proteins like the tandem ACPs is extremely challenging. There have been prediction models of the possible structures of some proteins but a substantially amount of work still remaining for the scientists.

2.5. Objectives of the project

In this work, we planned to study the structural biology of the PfaA cluster of Moritella marina that is one of the key proteins involved in the synthesis of omega-3 fatty acids, in particular DHA. Protein engineering could allow us to understand the function of the different domains of PfaA and to analyze the substrate binding specificity of the tandem ACP domains. Our aims for this project were to understand the mechanism by which ACP domains change in substrate binding specificity when a one-point mutation is introduced in the active center serine (Ser703) of the AT domain. To do this, we planned to solve the structure of the mutated protein ATS703A by using X-ray crystallography techniques. Also, we focus the study on trying biochemical and radioactivity assays to determine the enzymatic contribution of each domain in the complex and to analyze the changes in substrate specificity by using site-directed mutagenesis.

3. Materials & Methods

3.1. Bacterial Strains

The bacterial strain used for molecular cloning throughout this project was E. coli DH5α. Because of its lack of the natural E. coli recombinase system, the stability of the plasmids increases drastically, providing an improved system to insert plasmids. Similarly, the bacterial strain used to express proteins was E. coli BL21 (DE3), which contains the phage T7 RNA polymerase gene regulated by IPTG.

3.2. Microbiological Techniques

3.2.1 General Culture Conditions

The general culture medium used for bacterial growth was LB culture medium (10g/L of tryptone, 5g/L of yeast extract and 5g/L of NaCl) provided by Pronadisa Spain. For standard bacterial growth in solid medium plates LB agar was used. The LB medium was supplemented with 1.5% agar, which was also provided by Pronadisa Spain. All the general culture medium was sterilized in an autoclave at 120°C during twenty minutes. For bacterial strains containing plasmids with specific antibiotic resistances, the culture medium was supplemented with their corresponded antibiotic concentration: Ampicillin at 100 μg/ml, or Kanamycin at 50μg/ml.

3.2.2. Electroporation Protocol

The standard electroporation transformation protocol to transform E. coli cells establishes a previous microdialysis step, in which the DNA is desalted. This essential step consists of depositing a drop of the DNA on a 0.05μM nitrocellulose filter in a petri dish filled with sterile milliQ water for approximately 20 minutes. Then, the drop is collected and stored.

In a regular transformation, when the DNA sample concentration is approximately 100ng/μl, the amount of DNA required is between 1 and 5μl, which after it's mixed with 50μl of competent cells. Then, the mixture is pipetted to a previously cooled 0.2 cm cuvette Gene Pulser. The cuvette is then placed in a MicroPulser TM Electroporator that produces an electric pulse of 2.5kV. Immediately after the electric pulse, 1ml of sterile LB, previously heated at 37°C, was added to the cuvette, mixing it up and transferring it to a 1.5ml Eppendorf. Then, the cells were incubated at 37°C for one hour. Afterwards the cells were plated on a solid medium of LB agar with the proper antibiotic, according with its plasmid resistance.

3.3. Oligonucleotides

All the primers were purchased from Sigma and designed using VectorNTI. The primers, constructions, and sequences are summarized in the table M-1.

Table M-1. Description of the oligonucleotides used in each construction

Primer	Construction	Sequence
T7_primer	Multi cloning site in pET plasmids	taatacgaactcactataggg
pT7_primer		gctagttattgctcagcgg
ATS607A	Site-directed mutagenesis	cgaaccttgccctgcgaatagcgcaaccact agtggttgcgctattcgcagggcaaggttcg
ATS863A		tttcttgattgaaatgaacagcttccagcatgtggtttttcagg cctgaaaaaccacatgctggaagctgttcatttcaatcaagaaa
KSAT-FW1	Sequencing pET plasmids	ttccaccgtcgtgcaggtatc
KSAT-FW2		tggtcgtttcaaatctatttac
KSAT-RV		aaatttagcgctatcaaccg

3.4. Plasmids

The plasmids used during the course of the project were designed by using the technique of cloning by PCR, which generate fragments that were inserted in the following vectors pET29c with a resistance to kanamycin or pET3a with resistance to ampicillin.

Table M-2. List of vectors used for the different constructions.

Plasmid	Construct
pET29c (kn)	pDHA domains
pET3a (Ap)	pfaE

3.5. DNA Engineering

3.5.1. DNA Extraction and Concentration Measuring

The method used to extract the plasmid DNA was the extraction kit from ThermoScientific Gene JET Plasmid Miniprep Kit, which protocol can be found in their website. Then, the

concentration of DNA was measured in a Nano-Drop ND-1000 spectrophotometer apparatus by measuring the absorbance at 260nm.

3.5.2. Site Directed Mutagenesis

Site directed mutagenesis was used to replace amino acids by changing the one or two base pairs. The mechanism of action of the site directed mutagenesis is explained in figure M-1 (QuikChange II Site-Directed Mutagenesis Kit). Polymerase chain reaction technique was used to amplify all the DNA fragments for site-directed mutagenesis using Phusion polymerase enzyme from ThermoScientific.

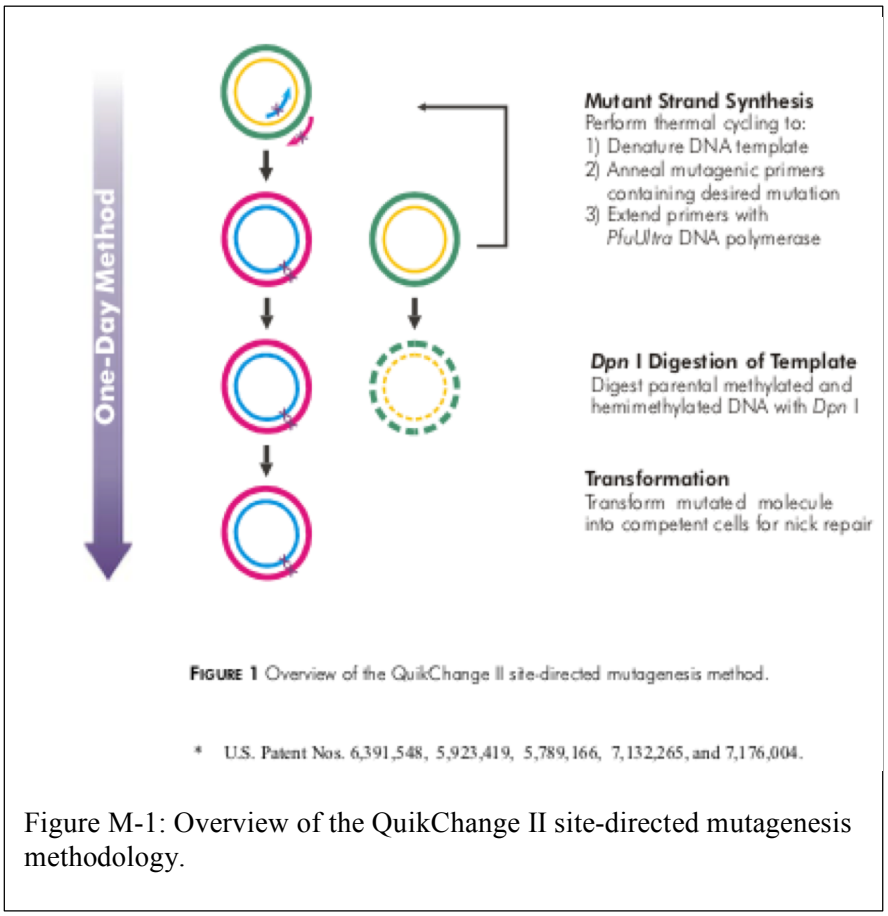


Figure M-1: Overview of the QuikChange II site-directed mutagenesis methodology.

The standard protocol for a site directed mutagenesis PCR initiates by designing the mutagenic sense and antisense primers using the Agilent webpage. The mutation should be placed in the middle of the sequence flanked by 10 to 15 base pairs at both sides, and the primers should have a melting point not higher than 78°C. Then, a mutant strand synthesis reaction is made by adding 10µl of 5x Buffer Phusion HF, 1µl of dNTPs, 2,5µl of sense primer, 2,5µl of antisense primer, 10ng of DNA template, 31µl of ddH_2O, and 0,5µl of Phusion DNA polymerase, which is

added the last. Next the PCR tubes were placed in the C-1000 Touch™ Thermo Cycler from Bio-Rad following the subsequently parameters: 98°C for 1 minute, 98°C for 30 seconds, 65°C for 50 second, 72°C for 2 minutes, 72°C for 10 minutes, and 4°C for 5 minutes.

For the digestion reaction of the PCR product, 1μl of DpnI was added. Then, the reaction tubes were placed again in the Thermo Cycler in a program where the reactions were incubated at 37°C for one hour and then for 20 minutes at 80°C. Finally, xx ul of the reaction mixture were electroporated onto DH5α cells, as explained above. After 1h at 37C, cells were plated and incubated overnight at 37 °C.

3.5.3. DNA Sequencing

All of the DNA sequences of the cloned PCR fragments and plasmids were sent to the company StabVida in Portugal where they were analyzed by Sanger automated sequencing protocol. For sequencing, 1000ng of the template DNA was mixed with 25pmol of the primers (3μl).

3.6. Protein Purification Protocol

3.6.1. Protein overexpression in E. coli

The overexpression of all the proteins used in this project were produced by the IPTG-inducible system. To begin with the overexpression of the proteins, a 50 ml flask containing LB media was added with 50μl of the correspondent antibiotic (kanamycin or ampicillin). Then, the flask was inoculated with the E. coli BL21 strain, carrying the targeted protein construction. Next, it was overexpressed overnight at 37°C with shaking at 160 rpm. Later, in a 1L flasks that contained a selective medium of LB, supplemented 1000μl of the appropriate antibiotics was added to the media. A dilution 1/20 was prepared. Next, the 1L flasks were incubated at 37°C for 1:30hr until its OD was between 0,6-0,8. When the culture reached the required OD, the flasks were placed in a bath of ice for 10 minutes, immediately after the isopropyl b-D-thiogalactoside (IPTG) with a concentration of 0.5mM was added to the culture. After, the cultures were placed in an incubator at 15°C overnight. Finally, to be able to purify the protein, the 1L flasks cultures were centrifugated in a Beckman Coulter Avanti j-26 centrifuge at 5.500rpm at 12°C for 30 minutes in a JA10 rotor (Beckman Coulter, USA). The pellets from the cultures were collected in falcons and frozen at -80°C.

3.6.2. Protein Purification by Affinity Chromatography

The method used for protein purification throughout this work was affinity chromatography nickel column. It's a general technique in which the polyhistidine tail that the pET29 vector contain, adhere the histidine tail to the nickel, so the other compounds can pass through and be eliminated as we obtain at least a purification of the 90% in some elusion fractions.

First, the equipment must be prepared. The ÄKTA system and the column were cleaned with MilliQ water using the manual run with a flow rate of 1 ml/min (for a column of 1ml) and a concentration of Buffer B to 50%. It was used between 20 to 30 ml of MilliQ to verify that the cables and the column were properly cleaned. Then this step was repeated with Buffer A (300mM of NaCl, 50mM Tris with pH 7,5 and 20mM of Imidazole) and Buffer B (300mM of NaCl, 50mM Tris with pH 7,5 and 500mM of Imidazole) each one at a concentration of 50 %. To finish cleaning the apparatus, the previous step was repeated with just Buffer A. While cleaning the apparatus, the sample protein that previously was frozen at -80 °C, was defrosted and resuspended in Buffer A with a total volume of 50 ml, to which we added 50 µl of PMSF 1%. Then, the sample was sonicated four times during one minute in a cold room at -20°C. After, the sample was centrifuged during 30 minutes at 4°C and 40.000 rpm in a Sorvall WX Ultra series ultracentrifuge from Thermo Fisher. When centrifuged, 10 µl of the supernatant and a piece of the pellet were placed in 1,5 ml eppendorfs to be used later for verification. In a cold room, a pump was used to process the protein to the column so the protein can go through the column and stay attached to it, while the waste pass through it. Previous to its use, the pump's cables were cleaned with MilliQ and Buffer A. Ones the protein is in the column, the column is placed in the ÄKTA apparatus again. The final step of the protein purification protocol is to collect the fractions in which our protein is incorporated. To leak the protein from the column, it is necessary to run the ÄKTA in a program were the concentration of Buffer B increases at a constant rate. When we had the fractions that supposedly contains our protein a verification gel of acrylamide at 12% was made in order to verify in which fractions there was protein.

3.6.3. Protein Electrophoresis

Protein electrophoresis is a testing method use to separate proteins based on their net charges, size, and shape. Our experiments used 12% polyacrylamide SDS-Page gels and a loading buffer

containing 50 mM Tris-HCl pH 6,8, 4 % SDS, 4 % glycerol and 0,02 % bromophenol blue. The molecular weight marker used was PageRuler Plus Prestained Protein Ladder from ThermoFisher. To stain and destain the gels it was used a solution containing Coomassie Brilliant Blue R-250 (Merck, Germany) in methanol, acetic acid and deionized water, and another solution of methanol, acetic acid, and deionized water, respectively. The electrophoresis gels were run at 190V for 50 minutes.

3.6.4. Preparation of Native Gel for Electrophoresis

In this work, it was necessary to prepare different types of native urea-Page gels and native gels without urea, at concentrations of 17,5 and 20% of acrylamide. The urea-PAGE gels at 17,5% and 20% were prepare by adding 0.5M urea, 40% acrylamide/bisacrilamide with a radio of 19:1, TBE 10x, Ammonium persulfate, TEMED, and distilled water (ddH_2O) up to 15ml for three gels. Then, the loading buffer was prepared by adding blue bromophenol, Tris 1M pH 6.8, sucrose, and ddH_2O. The loading buffer was mixed with DTT 2mM or DTT 50mM depending of the concentration for each experiment. 10μl of the loading buffer and the protein were added and loaded into the gels. Finally, all the samples were run at 100V for three hours in an electrophoresis apparatus with a TBE 1%. All the gels, including the SDS-PAGE gels, were stained with Coomassie Brilliant Blue R250, 100% methanol, 100% acetic acid glacial, and ddH_2O. Similarly, the destain used was made of methanol, glacial acetic, and ddH_2O.

The protocol used to prepare the native-PAGE without urea was. First, it was necessary to prepare a stacking gel at 6% containing 40% acrylamide/bisacrilamide (radio 19:1), 0.5M Tris pH 6.8, 10%APS, TEMED, and ddH_2O. Then, it was prepared a separating gel at 20% with 40% acrylamide/bisacrilamide (radio 19:1), 1,5M Tris pH 8.8, 10% APS, TEMED, and ddH_2O. The loading buffer used in this experiment was made of 100mM Tris pH 7.5, 300mM NaCl, and ddH_2O. In this case, for running the electrophoresis, the buffer used was made of 25mM of Tris-HCl, and 192mM of glycine at a final pH of 8.3. As in previous electrophoresis, the samples were run for three hours at 100V and using the same stain and destain solutions.

3.6.5. Buffer Exchange

To be able to crystallize a protein, it was necessary to change the buffer to a tampon with a lower tris and salt concentration. The loading buffer and crystallization buffer used for these proteins contained 20mM of Tris with a pH of 7,5, 50mM of NaCl, and 1mM of EDTA. The new

buffer was used to dilute the fractions of protein to reach 1/10 of the initial protein concentration. The purified proteins were concentrated for crystallization to 10ng/ml in a centricon tube from Millipore. This concentration method was used two to three times to remove the initial buffer A and reach the sought concentration for farther experiments. The concentration of protein was measured in each case using a Nano-Drop ND-1000 spectrophotometer, which measured the absorbance at 280nm.

3.7. Radioactivity Assays

Radioactivity assays were used to study the enzymatic activity of different pfaA domains in vitro. Different combinations of proteins were incubated with radiolabeled substrates such as Malonyl-CoA or Acetyl-CoA, which were analyzed by 12% radio-SDS-Page gels. In each combination of proteins, it was tested different protein and substrate concentrations. When the radiolabeled substrate was added to the protein mixture, the reaction was incubated at room temperature for at least 30 minutes. Then, the protein loading buffer was added and inserted directly onto the 12% SDS gels. Finally, the gels were placed in a drier for two hours and putted to autoradiography for one day.

3.8. X-ray Crystallization

To obtain high quality diffraction crystals is required to have an elevated purity protein samples. An adequate candidate for crystallization must be a protein capable of producing soluble and stable crystals. To find the precise crystallization condition in which the protein sample crystallizes correctly, screening kits from Hampton research crystal screen one and two were used. The screening kits contained different reagents to form a broad range of precipitants, pH, and organic compounds. Each reagent was pipetted into reservoirs of 96 wells plate in 50µl aliquots. Using the sitting drop system, 1µl of the protein sample was placed in its correspondent chamber and immediately after 1µl of the precipitant was added. Next, the 96 wells plate was sealed with transparent tape and was incubated at 22°C. It could take days or months to visualize any crystals. In case that crystals were observed in any condition, a second screening would be performed in a 24 wells plate with the conditions in which the crystals appeared, varying the concentration of precipitant and the pH of the buffer solution. Then, the crystals should be send to the synchrotron Alba situated in Barcelona to diffract and by using diverse informatic software analyze and determine the 3D structure of our protein.

3.9. Bioinformatic Software

The principal bioinformatic software used to design, visualize, and analyze oligonucleotides and sequence alignments was Vector NTI. To determine homology sequences of DNA or protein, Blast online server was used. Furthermore, the program used to translate DNA code into amino acids was ExPASy SIB Bioinformatics resource. In addition, the ProtParam tool from ExPASy was utilized to analyze protein parameters like the extinction coefficient, molecular weight, isoelectric point or number of amino acids. Moreover, the structural modeling software used to visualize and make predictions was Phyre2 server. This software was also used to identify structural homologs and to generate 3D structural models of our proteins. Similarly, Swiss model server was also used to determine protein structures, as well as, other homologous from the protein data bank. The main software utilized to visualize and create videos and quality photos of protein structures was Pymol (www.pymol.org).

4. Results

4.1. PfaA Constructions

PfaA is formed by ten individual domains that shows five different structural folds. PfaA is the most complex protein of all the pfa cluster due to its length. In the N-terminal domain of PfaA, it was observed a thiolase-like folding that appears to be a keto synthase domain (aa 29-487). This domain is adjacent to an acyltransferase domain (aa 601-923), like in the case of fatty acid synthesis previously described. Then, there is five tandem acyl carrier protein domains (aa 1255-1768). The ACP domains are followed by two keto reductase domains. The final domain is predicted to be a dehydratase domain. The domain prediction for PfaA is show in figure R-1 that contains a scheme of its locations and structural predictions.

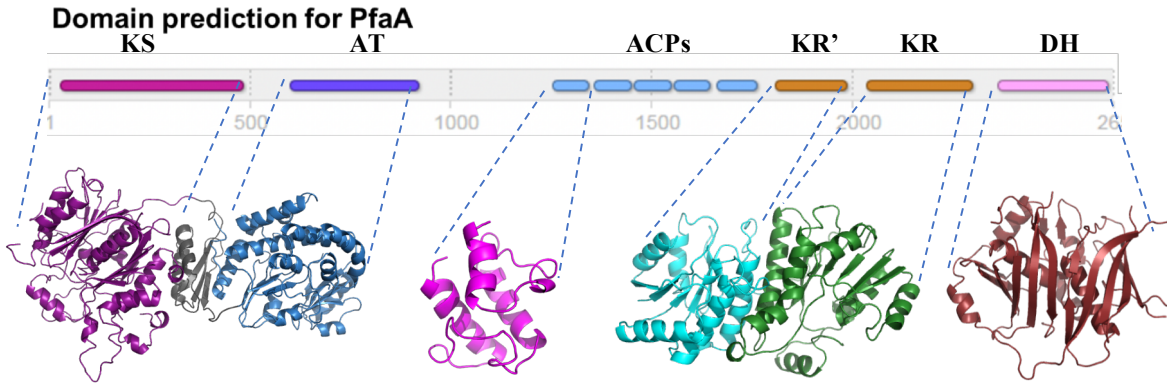


Figure R-1: Domain prediction for PfaA. Each domain predicted is represented in a different color with its correspondent predicted structures. The templates for the model predictions of all the structures of this figure were obtained from the mammalian FAS.

The constructions used in this work to perform the crystallization and biochemical experiments were previously designed and provided by Dr. Omar Santín Martínez. Of all the constructions made in his doctoral thesis, for this project we decided to choose the constructions summarized in table R-1. This work was centered in the cluster of PfaA of M. marina. Our aim was to study the biochemical structure of the KSAT and ACP domains. The constructions were tested for stability and solubility by a small-scale expression protein protocol. The proteins were

soluble and were able to be expressed. This step was highly important as it is necessary to know if the proteins would be capable of being purified.

Table R-1: Summery of the constructions designed in this work.

Number	Name	Amino Acid Range
1	KSAT	**M1-1150**
2	1ACP	**A1258-V1348**
3	**5ACP**	**A1258-A1748**

The following figure (figure R-3) shows a model prediction of the KS-AT domain as a single di-domain. The figure shows the surface electrostatic potential of KSAT. We can also observe in the figure the active site of the AT where the malonyl-CoA should bind to the protein to initiate the synthesis of polyunsaturated fatty acids.

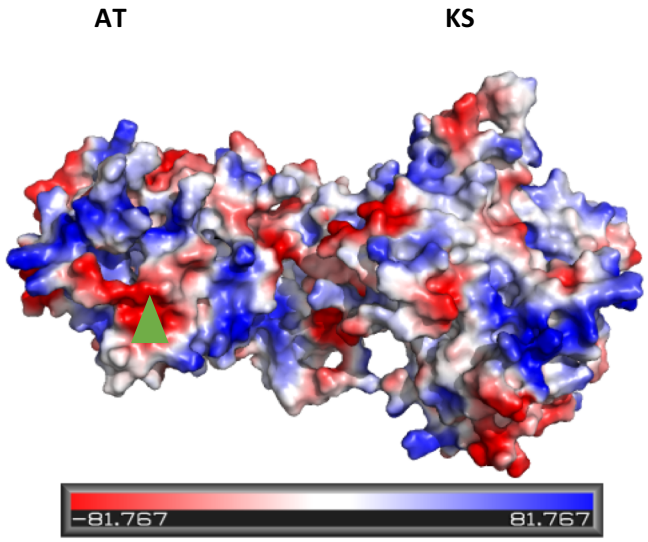


Figure R-2: The figure shows an image of the protein contact potential of KSAT domain. It is a visualization option of the PyMol software to represent with the red/blue colors the charge-smoothed surface on the protein. The structural template used was from the mammalian FAS (2VZ8).

In this project, our objectives were to study the function of the KSAT domain in the PfaA cluster. Due to the number of examples in the literature (Khosla et al. 2007; Maier, Leibundgut, and Ban 2008b), to perform the experiments, the KS and AT domains were treated as a single di-

domain. To make a deeper understanding of the biochemical interactions between KSAT and ACPs, a one-point mutant of the KSAT (ATS703A) was chosen to observe the effects on substrate binding of the tandem acyl carrier proteins.

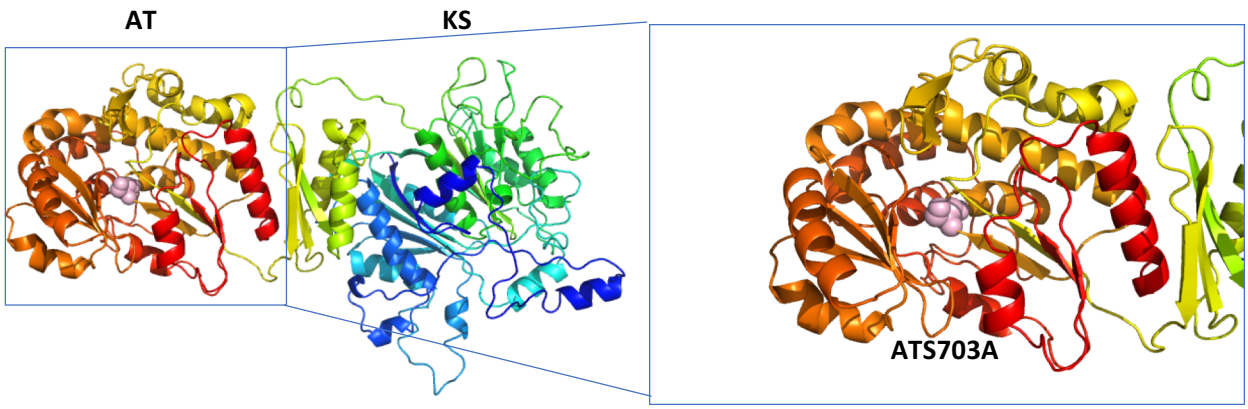


Figure R-3: Structural model for the ATS703A mutation of KSAT of PfaA. The cartoon KSAT domain representation was obtained using PyMol. The pink sphere in the AT domain is the one-point mutation ATS703A. The mammalian FAS (2VZ8) was used as the structural template.

The structural model of the ATS703A mutant is showed in figure R-3. The mammalian FAS (identification number: 2VZ8) was used as the structural template to create this structural prediction of the mutation. In the figure R-3, the ATS703A (pink sphere) is located in the AT domain, more precisely in the active site of the protein. The 703 serine of the active site was modified to an alanine. The location of the ATS703A was the reason why this exact residue was chosen. The AT domain supposed to bind with the malonyl units at this active site. The objective of the following experiments was to analyze the biochemical alterations due to the one-point mutation ATS703A on the substrate binding specificity.

4.2. Purification of KSAT and ACPs by Affinity Chromatography

Firstly, to be able to perform the biochemical and crystallography experiments, we performed a series of purifications using affinity chromatography of the proteins previously described in table R-1. To perform the experiments of biochemical interaction and alterations of substrate specificity, we also purified the ATS703A mutant. KSAT, 1ACP, 5ACP, and AS703A were purified with a nickel column as described in the Materials and Methods section. All the proteins were soluble, stable, and overexpressed correctly. To identify and isolate the protein molecules of our samples, we performed an electrophoresis with gels of 12% SDS-PAGE. The SDS-PAGE gels denatured the protein and allows to separate the protein by its molecular mass.

In all the purifications performed, we added a well with the marker (M) and we collected the pellet (P) and supernatant (S) to verify if our protein was attached to the column or was lost in the process. The marker (M) used to measure the molecular masses of the proteins was a "PageRuler Prestained Protein Ladder Plus" (ThermoFisher) that measures specific molecular masses between 10 and 250 kDa. The pellet (P) was collected after the process of ultracentrifugation during the purification and it should contain the insoluble proteins and the broken cells. The flow through (F) was collected after the target protein was placed into the nickel column and it should contain the soluble proteins that could not attach to the column.

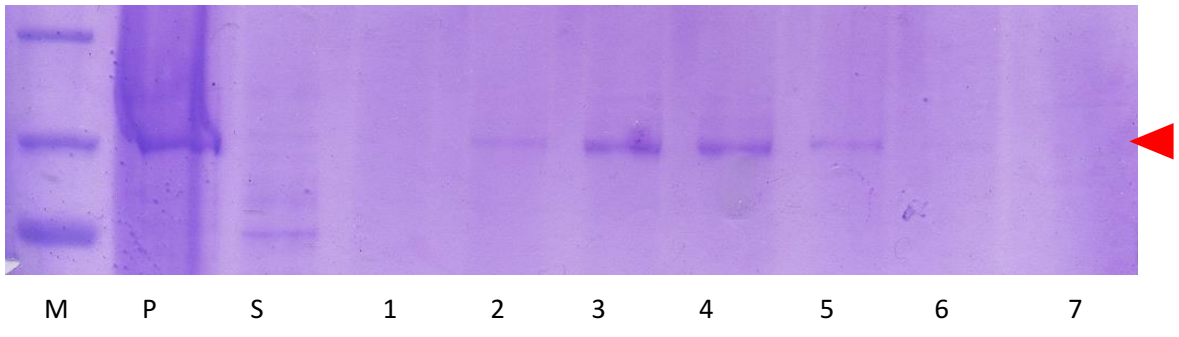


Figure R-4: Purification of the KSAT domain of PfaA. KSAT domain has a molecular mass of 130 kDa that can be observed in the figure in the fractions 2, 3, 4, and 5. The first three wells of the gel were filled with the protein marker (M), the pellet (P), and the flow through (F).

First, we performed a purification of the KS-AT domains (Figure R-4), which had a molecular mass of 130 kDa. The final concentration of protein obtained from the fractions was 8.15 mg/ml. The KSAT di-domain was constructed with a histidine tail at the end of its sequence that permitted its attached to the nickel column. As observed in figure R-4, in this purification we did not obtained many residues attached to our target protein. In the following experiment, we wanted to purify 1ACP to study the biochemical interaction and the overall structure of just one ACP.

The purification of 1ACP is showed in figure R-5. 1ACP has a molecular mass of 28 kDa and it wasobtained with a concentration of 6,8mg/ml. The purpose to purify a single ACP was to perform biochemical studies to observe the interaction of a single ACP with different substrates in different assays. In the figure R-5, it was also observed another band of residues with a molecular mass of 10 kDa that we supposed is a residue of the protein PfaE that was co-expressed with 1ACP.

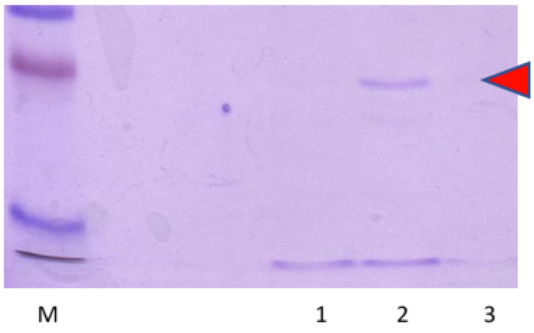


Figure R-5: Purification of 1ACP of the ACP domain. 1ACP has a molecular mass of approximately 28 kDa that can be observed in the fraction 2. In the gel can also be observed another band that we identify as PfaE with a molecular mass of 10 kDa. The first three wells of the gel were filled with the protein marker (M), the pellet (P), and the flow through (F).

Next, we performed a purification of 5ACP to study how the entire ACP domain interact with the substrates in further experiments. 5ACP was obtained with a final concentration of 5,75mg/ml. The amount of protein obtained was minimal but enough to perform the biochemical experiments. In the figure R-6, we can observe a small amount of protein with a molecular mass of 55 kDa.

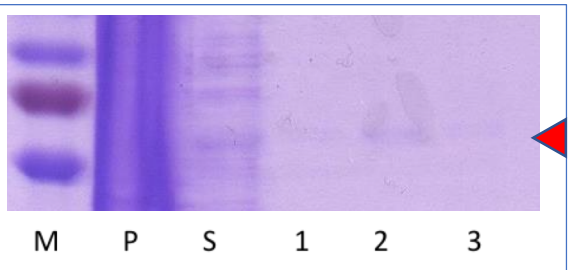


Figure R-6: Purification of the 5ACPs of the ACP domain. 5ACP has a molecular mass of 55 kDa that can be observed in the fractions 1, 2, and 3 in the gel. The first three wells of the gel were filled with the protein marker (M), the pellet (P), and the flow through (F).

In the pellet, it is observed a big quantity of the protein, indicating that a huge amount of protein was lost in the process of ultracentrifugation.

Later, we performed the purification of the mutant ATS703A. The mutant ATS703A was constructed with the domains KSAT as a single di-domain. We obtined a fair amount of protein during the purifications. In figure R-7, we observed ten fractions containnig ATS703A with a molecular mass of 130 kDa. The final concentration obtained was 11,714mg/ml. it was also observed in the gel that some residues at the top of our protein.

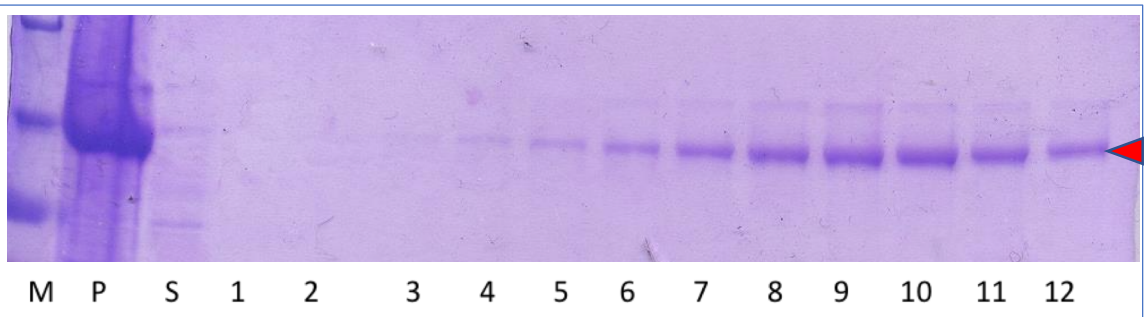


Figure R-7: Purification of ATS703A mutant of the AT domain. ATS703A has a molecular mass of 130 kDa that is observed in the fractions 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12. The first three wells of the gel were filled with the protein marker (M), the pellet (P), and the flow through (F).

After the purifications of KSAT, 1ACP, 5ACP and ATS703A, we concentrated the proteins for crystallography and biochemical experiments, such as native-PAGE electrophoresis assays and radioactivity assays. The buffer was also changed in order to eliminate all the imidazole and salts that could interfere with the following experiments. The process of purification, buffer exchange, and concentration procedure is described in detail in the *Materials and Methods* section of this work.

4.3. Biochemical assays of KSAT and ACPs

4.3.1. Native gel electrophoresis of ACPs

The native gel electrophoresis of ACPs was used to determine the biochemical difference between the Apo- and Holo-form of the ACP domain. The inactive form of the ACP domains is

the Apo-form and the active form of the ACPs is the Holo-form. The objective of these experiments was to observe a molecular mass different between the different species of ACP by using electrophoresis in native conditions. Also, another aim of these experiments was to determine when the malonyl-CoA binds with the ACP. The Holo-form of ACP was co-expressed with the PfaE gene coding for a M. marina phosphopantetheinyl transferase (PPTase).

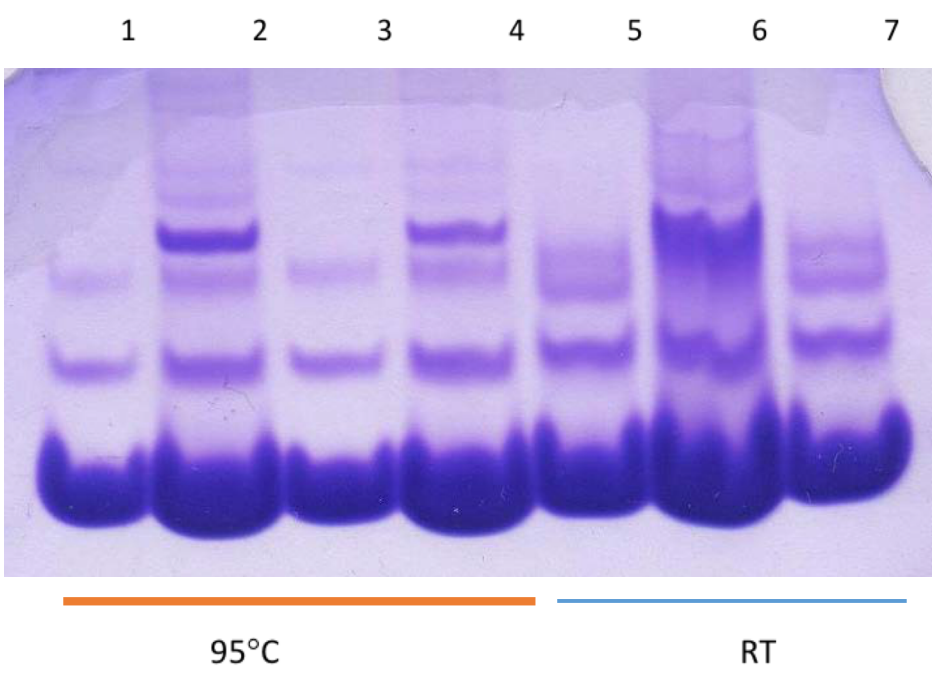


Figure R-8: A 17,5% native urea-PAGE gel electrophoresis of Apo and Holo-forms of ACP. In this gel, the samples had different concentrations of DTT (2mM and 50mM respectively). The wells 1, 2,5, and 6 were the samples with the 2mM DTT concentration and the wells 3,4,7, and 8 were the samples with the 50mM DTT concentration. The samples were boiled at 95°C or tested at room temperature.

In our next experiment, we changed the percentage of the polyacrylamide gel to 20% in order to try to visualize a relevant difference between the Apo- and Holo-form of ACP. The objective of this experiment was to classify the different species of ACP by testing the migration of the different forms of ACP, Apo and Holo.

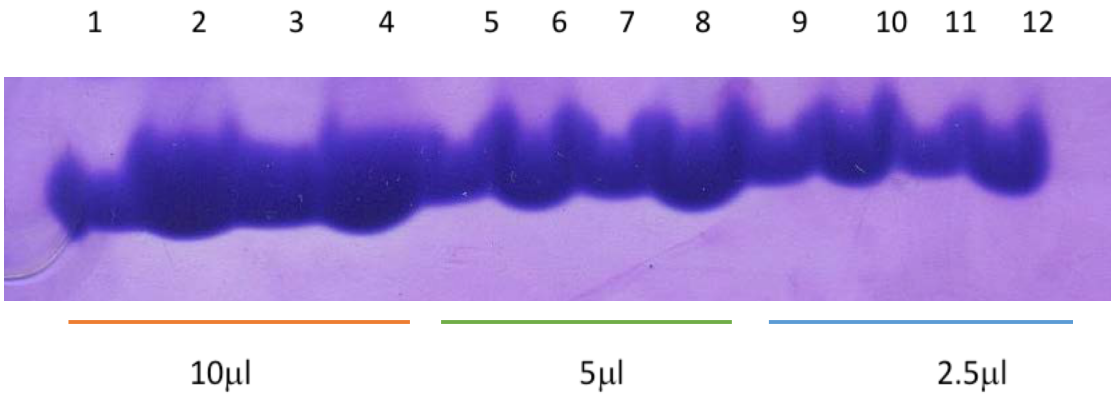


Figure R-9: A 20% native polyacrylamide gel with Apo and Holo-forms of ACP. The wells 1, 2, 5, 6, 9, and 10 contained 2mM of DTT. The wells 3, 4, 7, 8, 11, and 12 contained 50mM of DTT. Different amounts of protein (10µl, 5µl, and 2.5µl) were tested in order to observe any different.

In this experiment, we varied the amount of protein added and the DDT concentration. However, the results obtained in figure R-9 concluded that there was no different between the Apo- and Holo-form of the ACP domain. Even with a different acrylamide gel percentage the results were the same as in figure R-8.

In our next experiment, we tried to analyze the interaction between the Apo- and Holo-form with malonyl-CoA.

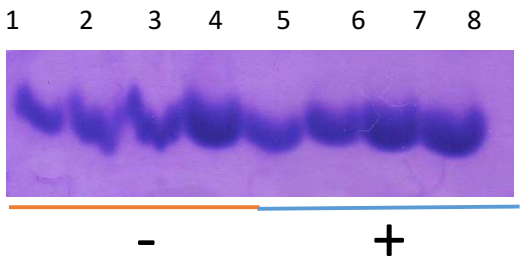


Figure R-10: 20% native-PAGE gel of the Apo and Holo-form of ACPs with different malonyl-CoA concentrations. The first four wells just contained the Apo and Holo ACPs without the malonyl-CoA. The 5 and 6 wells contained 1mM malonyl-CoA concentration. The wells 7 and 8 contained 10mM malonyl-CoA. The wells containing Apo-form are impairs and the par wells contain the Holo-form.

The results obtained from figure R-10 didn't show any different between the Apo- and Holo-form of the ACP when malonyl-CoA is added. Even though, we increased the concentration of malonyl-CoA to 10mM, we still didn't observe any difference between the Apo- and Holo-form. To observe any difference between the Apo- and Holo-form of the ACP with malonyl-CoA other assays must be performed. In the next section, we performed radioactive assays to analyze the interaction of the different domains with malonyl-CoA and Acetyl-CoA.

4.3.2. Radioactivity Assays

The radioactivity assays protocol is described in detail in the *Material and Methods* section of this work. In these experiments, we hoped to observe the difference between the ACP species Apo and Holo. Also, we expected to observe which if KSAT will promote the binding of malonyl-CoA to the ACPs. Similarly, we expected to observe a change in substrate binding specificity when the serine active site of the AT is mutated.

To observe an accurate difference between the Apo- and Holo-forms of 5ACP, we carried out a radioactivity assay where the malonyl-CoA was radiolabeled. The results of the radioactivity assay (figure R-11), determined that there was a considerable difference between the Apo- and Holo-forms of 5ACP. The Apo-form could not bind with the malonyl-CoA as it does not have the PPT moiety where the malonyl-CoA binds to the protein. Therefore, the radioactive substrate migrates much faster than the protein (indicated by a red arrow). The Holo-form of 5ACP bound with the malonyl-CoA and was observed in the band much higher up than the one from Apo.

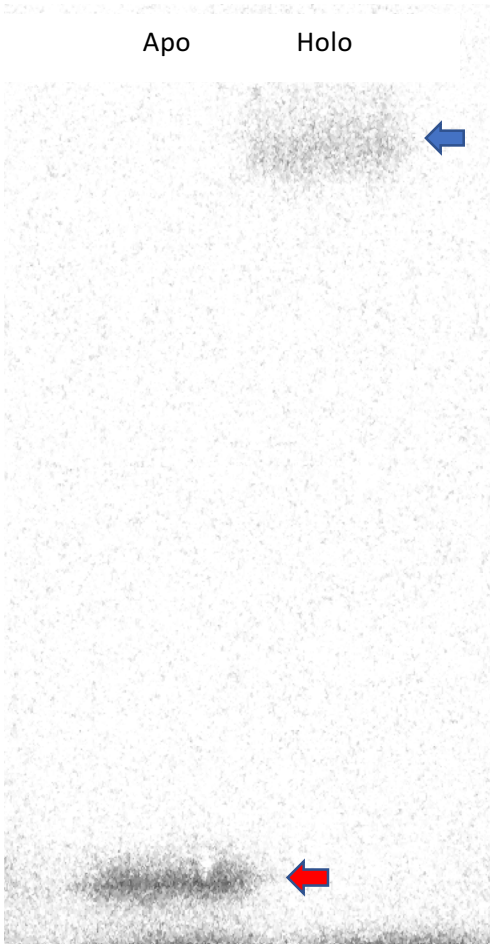


Figure R-11: Radioactivity assay of Apo- and Holo-forms of ACP. The first band represents the Apo form (red arrow) and the second band is the Holo-form (blue arrow) of the 5ACP. Both Apo and Holo were incubated with malonyl-CoA in order to determine the ability of the two forms to bind to the substrate.

Next, we wanted to analyze how the KSAT and the five tandem ACPs coordinated their enzymatic reactions to extend and condensate the carbon chain when malonyl-CoA or acetyl-CoA is added. The following radioactive assay was performed to study the substrate specificity of KSAT and 5ACP.

The results of the radioactivity assay of KSAT and 5ACP (figure R-12) showed that malonyl-CoA binds to KSAT and 5ACP while acetyl-CoA does not bind to the protein. In conclusion, the specific substrate for KSAT and 5ACPs is malonyl-CoA. KSAT is capable of promoting the binding of substrate to the ACP when that substrate is malonyl-CoA but not with acetyl-CoA.

To study why the KSAT and 5ACP domains have a malonyl-CoA specificity, we performed several radioactivity assays using the ATS703A mutant. The following experiments was performed by adding radiolabeled malonyl- and acetyl-CoA to KSAT, 5ACP, and ATS703A.

In this experiment (figure R-13), we observed that the KSAT domain is able to selectively bind the malonyl-CoA units but not the acetyl-CoA (figure R-9). In figure R-13, we could also observe in lanes 1 and 2 where 5ACP + KSAT (wild type) were incubated with radiolabeled malonyl-CoA and acetyl-CoA, respectively the same as in figure R-9. On the other hand, as shown in the lanes 3 and 4 of the gel, the KSAT mutant ATS703A when incubated with malonyl-CoA, a band is not observed indicating that the malonyl- CoA was not able to bind to the ACPs and that the AT domain did not promote the binding of the substrate to the ACPs domains. However, in the lanes 5 and 6 of the SDS-PAGE gel when 5ACP and ATS703A were incubated with radiolabeled acetyl-CoA, it was observed a change in substrate specificity as the mutant promoted the binding of the acetyl-CoA to the ACPs. Conclusively, we could observe a change in substrate binding specificity when the serine of the active center of the AT domain is mutated. Further experiments must be performed in order to study the specific binding site of the acetyl-CoA in the mutant ATS703A.

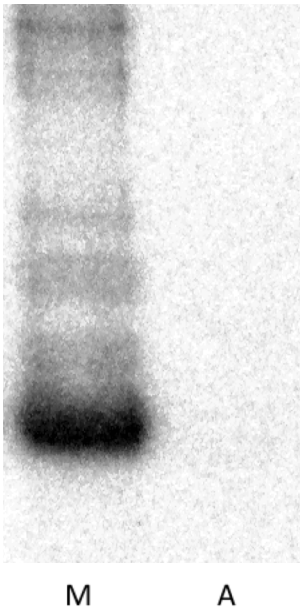


Figure R-12: Radioactivity assay of KSAT and 5ACP. KSAT and 5 ACPs were incubated with radiolabeled malonyl-CoA (M) and acetyl-CoA (A) to determine which substrate were to bind.

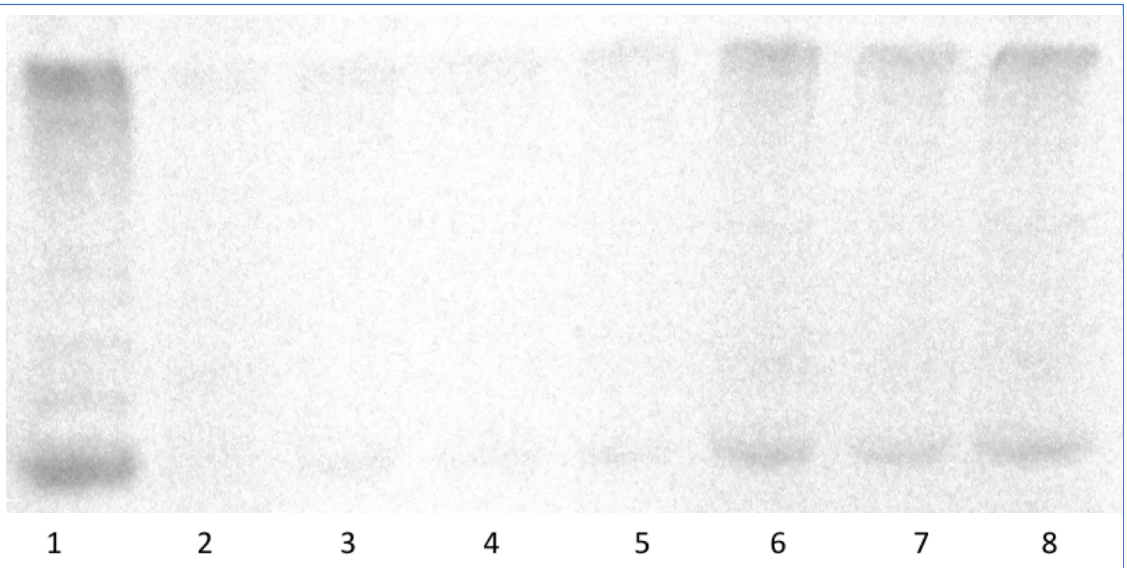


Figure R-13: Radioactivity assay of KSAT, 5ACP, and ATS703A. Lane 1: 5ACP + KSAT + Malonyl- CoA; lane 2: 5ACP + KSAT + Acetyl-CoA; lane 3 and 4: 5ACP + ATS707A + Malonyl-CoA at concentrations (1µl and 2µl, respectively); lane 5 and 6: 5ACP + ATS707A + Acetyl-CoA at 1µl and 2µl, respectively; lane 7 and 8: 5ACP + ATS707A + Malonyl-CoA + Acetyl-CoA.

4.4. Site-Directed Mutagenesis

One-point site-directed mutagenesis is a technique used to insert a mutation in an amino acid. To create the oligonucleotides to perform this experiment we used a series of software's specialized in the creation of primers like Vector NTI Express Designer, PyMol, and QuikChange Primer Design Program. Also, the QuikChange Primer Design Program was used to design the one-point mutation primers. The protocol specified that the primers must be between 25 and 45 base pairs long, in order to avoid second structures or a bad folding of the protein. Also, the anneal temperature must be between 50 °C and 75 °C, which were calculated using the NEB Tm calculator. The figure R-14 shows the visualization of the KSAT domain of PfaA with the one-point mutagenesis ATS607A and ATS863A. The next constructions were designed to analyze the changes in substrate binding specificity of the ACP domain observed in the previous section of the Results.

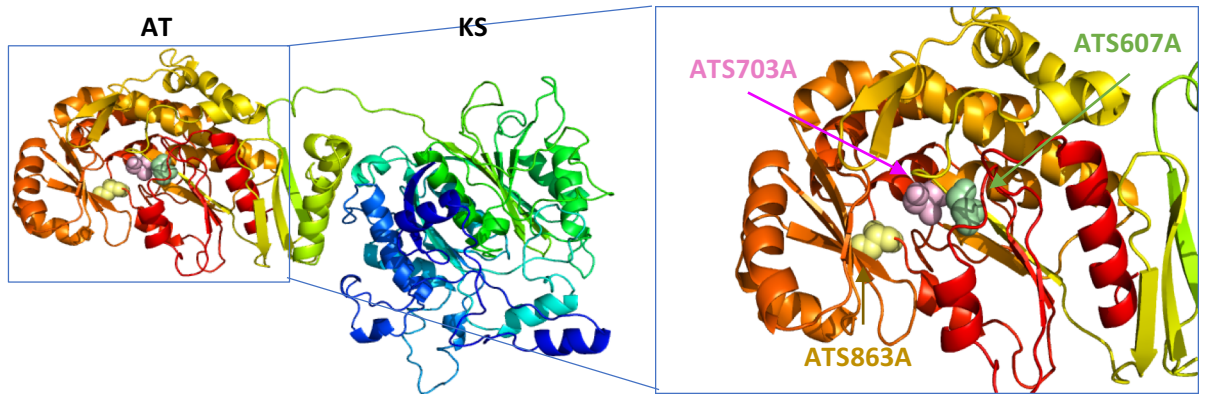


Figure R-14: The figures show two the structural model of ATS703A, ATS607A, and ATS863A mutations on the AT domain of the di-domain KSAT. The mutation ATS703A is represented in the figures by a pink sphere in the center of the active center of the protein. The green and yellow spheres represent the ATS603A and ATS863A one-point mutations respectively. The mammalian FAS (identification number: 2VZ8) was used as the structural template.

The results obtained with the QuikChange II side-directed kit were not successful because as we did not obtain any mutant clone. Multiple protocols were performed in order to obtain the one-point mutation but neither of them was successful. The protocols must be tuned up in order to pursue with further experiments.

4.5. X-Ray Crystallography

To obtain diffraction quality crystals, it is necessary to have a highly purity and concentrated protein. The crystallization of any protein requires that the protein is at least 90% pure. The crystallization conditions were first tested with various commercially available screening kits "crystal screen 1 and 2" (Hampton research). The kits contained multiple reagents with a wide range of precipitants, pH and other organic compounds. The crystallization process is explained in the Materials and Methods section of this work. By performing x-ray crystallography, we wanted to obtain crystals of ATS703A to solve its structure.

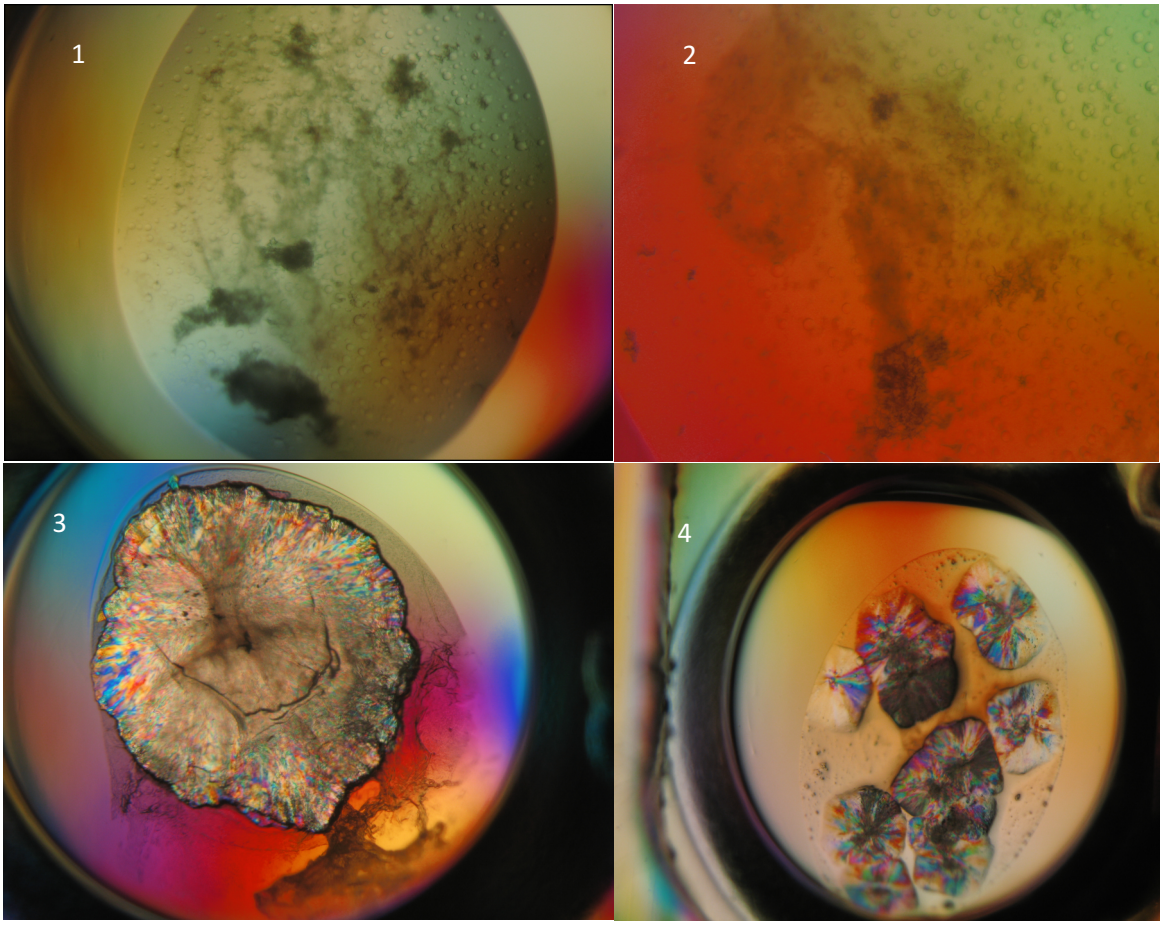


Figure R-15: The figures above show microscopic images of droplets of the 96 well plates of the mutated protein ATS703A. In image number 1 and 2, it can be observed what is known as the phase separation and precipitate. In the 3 and 4 image, it can be observed a cluster of precipitates.

We tested some parameters such as the concentration of the protein, the proportion of protein and crystallization solution, or the total volume of the drop. Unfortunately, after many attempts, no viable crystal of the protein was obtained. The figure R-15 shows various events occurred during the crystallization process of the protein. All of the images correspond to precipitation events or phase separation of the buffer. Therefore, further experiments with different reagents must be performed in order to crystalize the ATS703A.

5. Discussion

In the process of PUFA synthesis, the ACP domains covalently attach the fatty acids chain to transfer it in a series of iterative reactions from one module to the next. There are multiple condensation cycles where the acyl chain, which is covalently attached to the ACPs, eleongates two-carbon units per round. The fatty acid is attached to the sulfhydryl group of the ppt prostetic group of the ACP, which at the same time is attached to the serine residue of the active center of the ACP.

The functionality of Apo- and Holo-form of the ACP domains was the main study in this work. A series of experiments were performed to identify the different species of ACP. After achieving a good purification of both Apo- and Holo-forms of a single ACP and the 5ACPs in tandem, a tive electrophoresis experiment was performed. I showed that there was no visible difference in molecular weight between both active and inactive forms of the single ACP. Later, we performed a series of radioactivity assays, in which we observed that it is necessary that the ACP carries the PPT moiety in order to load the substrates of the pathway. Thus, in accordace with other authors the ACP protein can exist in two forms, the Apo-ACP, that is not activated so there is no presence of the PPT arm, and the Holo- ACP, totally active and able to be malonylated (Evans et al. 2008; Morris et al. 1993).

Next, we studied the initiation of the synthesis of LC-PUFAs. The elongation starts when the AT domain promotes the binding of malonyl-CoA onto the ACPs. It was determined that malonyl-CoA was able to bind to the 5ACPs in presence of the wild type KS-AT di-domain. Also, we observed that the ACPs can self-acylate with malonyl-CoA but not acetyl-CoA. To study the substrate specificity of the AT domain, we chosen to study the AT domain, where the serine from the active site was mutated to an alanine (ATS703A). By analyzing the ability of the mutated AT domain to bind radioactive substrates, we observed a change in the substrate binding specificity of the AT, which transferred an acetyl residue to the ACP but not malonyl.

To further understand this alteration in the substrate specificity when a one-point mutation (ATS703A) is introduced in the AT domain of the single di-domain construction of KSAT, we designed two additional mutations of other serine residues (ATS607A and ATS863A) that are close to the ATS703A residue to specify the binding center of the substrates. Unfortunately, due to difficulties with the site-directed mutagenesis we could not obtain the mutants.

Finally, various attempts were made to crystallize the ATS703A mutant but all the effort were unsuccessful. We tried to change some crystallization parameters like protein concentration and reagents amounts but no crystals grew in the plates. Further attempts are in progress, in order to test more parameters.

6. Conclusions & Future Research

- It was determined that ACPs self-acylate with malonyl-CoA buy not acetyl-CoA. We have proved that KSAT promotes the binding of malonyl-CoA to the ACPs but not acetyl-CoA.
- When the serine active site of the AT domain is mutated (ATS703A), the AT domain transfers an acetyl-CoA residue to the ACPs instead of malonyl-CoA.
- This change in substrate binding specificity presents a new path to further understand the mechanism by which KSAT promotes the binding of acetyl-CoA to the ACPs when its active site is mutated.

7. Bibliography

Alberts, A W, and P R Vagelos, (1968). Acetyl CoA Carboxylase. I. Requirement for Two Protein Fractions. Proceedings of the National Academy of Sciences of the United States of America 59(2): 561–568.

Chakravarty, Bornali, Ziwei Gu, Subrahmanyam S. Chirala, Salih J. Wakil, and Florante A. Quiocho, (2004). Human Fatty Acid Synthase: Structure and Substrate Selectivity of the Thioesterase Domain. Proceedings of the National Academy of Sciences of the United States of America 101(44): 15567–15572.

Chan, David I., and Hans J. Vogel, (2010). Current Understanding of Fatty Acid Biosynthesis and the Acyl Carrier Protein. The Biochemical Journal 430(1): 1–19.

Chowdhury, R., Stevens, S., Gorman, D., Pan, A., Warnakula, S., Chowdhury, S., ... & Franco, O. H. (2012). Association between fish consumption, long chain omega 3 fatty acids, and risk of cerebrovascular disease: systematic review and meta-analysis. *Bmj*, *345*, e6698.

Cole, Greg M., Giselle P. Lim, Fusheng Yang, et al., (2005). Prevention of Alzheimer's Disease: Omega-3 Fatty Acid and Phenolic Anti-Oxidant Interventions. Neurobiology of Aging 26(1, Supplement). Supplement: Aging, Diabetes, Obesity, Mood and Cognition, "SPARK" Workshop: 133–136.

Covington, M. B. (2004). Omega-3 fatty acids. *Atlantic*, *1*(2.0).

Dyall, S. C. (2015). Long-chain omega-3 fatty acids and the brain: a review of the independent and shared effects of EPA, DPA and DHA. *Frontiers in aging neuroscience*, *7*, 52.

Evans, Simon E., Christopher Williams, Christopher J. Arthur, et al., (2008). An ACP Structural Switch: Conformational Differences between the Apo and Holo Forms of the Actinorhodin Polyketide Synthase Acyl Carrier Protein. ChemBioChem 9(15): 2424– 2432.

Flugel, Roger S., Yon Hwangbo, Ralph H. Lambalot, John E. Cronan, and Christopher T. Walsh, (2000). Holo-(Acyl Carrier Protein) Synthase and Phosphopantetheinyl Transfer in Escherichia Coli. Journal of Biological Chemistry 275(2): 959–968.

Ghayourmanesh, Soraya, and Soma Kumar, (1981). Synthesis of Acetoacetyl-CoA by Bovine Mammary Fatty Acid Synthase. FEBS Letters 132(2): 231–234.

Heath, Richard J., and Charles O. Rock, (1995). Enoyl-Acyl Carrier Protein Reductase (fabI) Plays a Determinant Role in Completing Cycles of Fatty Acid Elongation in Escherichia Coli. Journal of Biological Chemistry 270(44): 26538–26542.

Holub, B. J. (2002). Clinical nutrition: 4. Omega-3 fatty acids in cardiovascular care. *Canadian Medical Association Journal*, *166*(5), 608-615.

Jiang, Hui, Scott R. Rajski, and Ben Shen, (2009). Tandem Acyl Carrier Protein Domains in Polyunsaturated Fatty Acid Synthases. Methods in Enzymology 459: 79–96.

Khosla, Chaitan, Yinyan Tang, Alice Y. Chen, Nathan A. Schnarr, and David E. Cane, (2007). Structure and Mechanism of the 6-Deoxyerythronolide B Synthase. Annual Review of Biochemistry 76: 195–221.

Kris-Etherton, P. M., Harris, W. S., & Appel, L. J. (2002). Fish consumption, fish oil, omega-3 fatty acids, and cardiovascular disease. *circulation*, *106*(21), 2747-2757.

Maier, Timm, Marc Leibundgut, and Nenad Ban, (2008). The Crystal Structure of a Mammalian Fatty Acid Synthase. Science 321(5894): 1315–1322.

Massengo-Tiassé, R. Prisca, and John E. Cronan, (2009). Diversity in Enoyl-Acyl Carrier Protein Reductases. Cellular and Molecular Life Sciences: CMLS 66(9): 1507.

Metz, James G., Paul Roessler, Daniel Facciotti, et al., (2001). Production of Polyunsaturated Fatty Acids by Polyketide Synthases in Both Prokaryotes and Eukaryotes. Science 293(5528): 290–293.

Morris, S. A., W. P. Revill, J. Staunton, and P. F. Leadlay, (1993). Purification and Separation of Holo- and Apo-Forms of Saccharopolyspora Erythraea Acyl-Carrier Protein Released from Recombinant Escherichia Coli by Freezing and Thawing. Biochemical Journal 294(2): 521–527.

Nichols, David, John Bowman, Kevin Sanderson, et al., (1999). Developments with Antarctic Microorganisms: Culture Collections, Bioactivity Screening, Taxonomy, PUFA Production and Cold-Adapted Enzymes. Current Opinion in Biotechnology 10(3): 240– 246.

Nishida, Takanori, Yoshitake Orikasa, Yukiya Ito, et al., (2006). Escherichia Coli Engineered to Produce Eicosapentaenoic Acid Becomes Resistant against Oxidative Damages. FEBS Letters 580(11): 2731–2735.

Olsen, Johan Gotthardt, Anders Kadziola, Penny von Wettstein-Knowles, Mads Siggaard-Andersen, and Sine Larsen, (2001). Structures of β-Ketoacyl-Acyl Carrier Protein Synthase I Complexed with Fatty Acids Elucidate Its Catalytic Machinery. Structure 9(3): 233–243.

Orikasa, Y., A. Yamada, R. Yu, et al., (2004). Characterization of the Eicosapentaenoic Acid Biosynthesis Gene Cluster from Shewanella Sp. Strain SCRC-2738. Cellular and Molecular Biology (Noisy-Le-Grand, France) 50(5): 625–630.

Orikasa, Yoshitake, Takanori Nishida, Akiko Yamada, et al., (2006). Recombinant Production of Docosahexaenoic Acid in a Polyketide Biosynthesis Mode in Escherichia Coli. Biotechnology Letters 28(22): 1841–1847.

Orikasa, Yoshitake, Mika Tanaka, Shinji Sugihara, et al., (2009). pfaB Products Determine the Molecular Species Produced in Bacterial Polyunsaturated Fatty Acid Biosynthesis. FEMS Microbiology Letters 295(2): 170–176.

Pereira, Suzette L, Amanda E Leonard, and Pradip Mukerji, (2003). Recent Advances in the Study of Fatty Acid Desaturases from Animals and Lower Eukaryotes. Prostaglandins, Leukotrienes and Essential Fatty Acids 68(2): 97–106.

Rose, David P., and Jeanne M. Connolly, (1999). Omega-3 Fatty Acids as Cancer Chemopreventive Agents. Pharmacology & Therapeutics 83(3): 217–244.

Rock, Charles O., and John E. Cronan, (1996). Escherichia Coli as a Model for the Regulation of Dissociable (Type II) Fatty Acid Biosynthesis. Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism 1302(1): 1–16.

Shulse, Christine N., and Eric E. Allen, (2011). Widespread Occurrence of Secondary Lipid Biosynthesis Potential in Microbial Lineages. PloS One 6(5): e20146.

Sugihara, Shinji, Yoshitake Orikasa, and Hidetoshi Okuyama, (2010). The Escherichia Coli Highly Expressed entD Gene Complements the pfaE Deficiency in a Pfa Gene Clone Responsible for the Biosynthesis of Long-Chain N-3 Polyunsaturated Fatty Acids. FEMS Microbiology Letters 307(2): 207–211.

Swanson, D., Block, R., & Mousa, S. A. (2012). Omega-3 fatty acids EPA and DHA: health benefits throughout life. *Advances in nutrition*, *3*(1), 1-7.

Tapiero, H, G Nguyen Ba, P Couvreur, and K. D Tew, (2002). Polyunsaturated Fatty Acids (PUFA) and Eicosanoids in Human Health and Pathologies. Biomedicine & Pharmacotherapy 56(5): 215–222.

Wall, Rebecca, R. Paul Ross, Gerald F. Fitzgerald, and Catherine Stanton, (2010). Fatty Acids from Fish: The Anti-Inflammatory Potential of Long-Chain Omega-3 Fatty Acids. Nutrition Reviews 68(5): 280–289.

Wakil, S. J., J. K. Stoops, and V. C. Joshi, (1983). Fatty Acid Synthesis and Its Regulation. Annual Review of Biochemistry 52: 537–579.

Witkowski, Andrzej, Anil K. Joshi, and Stuart Smith, (1997). Characterization of the Interthiol Acyltransferase Reaction Catalyzed by the β-Ketoacyl Synthase Domain of the Animal Fatty Acid Synthase. Biochemistry 36(51): 16338–16344.

Yazawa, K., (1996). Production of Eicosapentaenoic Acid from Marine Bacteria. Lipids 31 Suppl: S297-300.

Yoshida, Kiyohito, Mikako Hashimoto, Ryuji Hori, et al., (2016). Bacterial Long-Chain Polyunsaturated Fatty Acids: Their Biosynthetic Genes, Functions, and Practical Use. Marine Drugs 14(5).

PyMOL: www.pymol.org

Vector, N. T. I. Express Designer Software.

QuikChange II Site-Directed Mutagenesis Kit