Biophysical *Journal*
Article

Biophysical Society

Atomistic insight on structure and dynamics of spinach acyl carrier protein with substrate length

Marcel D. Baer,[1,*] John Shanklin,[2] and Simone Raugei[1]
[1]Pacific Northwest National Laboratory, Physical and Computational Sciences Directorate, Richland, Washington and [2]Biology Department, Brookhaven National Laboratory, Upton, New York

ABSTRACT The plant acyl-acyl carrier protein (ACP) desaturases are a family of soluble enzymes that convert saturated fatty acyl-ACPs into their *cis*-monounsaturated equivalents in an oxygen-dependent reaction. These enzymes play a key role in biosynthesis of monounsaturated fatty acids in plants. ACPs are central proteins in fatty acid biosynthesis that deliver acyl chains to desaturases. They have been reported to show a varying degree of local dynamics and structural variability depending on the acyl chain size. It has been suggested that substrate-specific changes in ACP structure and dynamics have a crucial impact on the desaturase enzymatic activity. Using molecular dynamics simulations, we investigated the intrinsic solution structure and dynamics of ACP from spinach with four different acyl chains: capric (C_{10}), myristic (C_{14}), palmitic (C_{16}), and stearic (C_{18}) acids. We found that the fatty acids can adopt two distinct structural binding motifs, which feature different binding free energies and influence the ACP dynamics in a different manner. Docking simulations of ACP to castor Δ^9-desaturase and ivy Δ^4-desaturase suggest that ACP desaturase interactions could lead to a preferential selection between the motifs.

SIGNIFICANCE Acyl carrier proteins are central proteins in fatty acid biosynthesis that deliver substrates to a variety of enzymes. They have been reported to show a varying degree of local dynamics and structural variability depending on the substrate size. Here, we use advanced molecular simulation methods to study the effect of acyl chain length on the dynamics and structure of a carrier protein in solution and its implication for the interactions with a desaturase enzyme. We show that the overall structure and dynamical behavior of acyl carrier protein is dictated by the fatty acid chain length of the bound substrate. We investigate how the chain-length-dependent structural diversity might influence the recognition process between the carrier protein and desaturase enzymes.

INTRODUCTION

Acyl carrier proteins (ACPs) are soluble proteins found in plants and bacteria that are involved in a variety of enzymatic processes, including the synthesis of fatty acids (1), polyketides, oligosaccharides, and toxins (2–5). Their role is to transport small molecules between enzymes by covalently binding them as thioester via the thiol group of a 4′-phosphathetheine. ACPs share a common fold that is variably referred to as a three- or four-helix bundle typically consisting of three major α-helices, with helix I being antiparallel to helices II and IV, plus a short 3_{10} or α-helical segment (helix III) linking helices II and IV (Fig. 1). Their structure shows a rather large conformational flexibility that has been suggested to be a fundamental requirement to accommodate, within a similar fold, a variety of diverse substrates (2,4–7).

The desaturation of fatty acids is one of the most studied processes involving ACPs. Indeed, ACPs have the ability to accommodate fatty acids and fatty acid metabolites of diverse chain length (from C_4 to C_{18}) and deliver them to polyketide and fatty acid synthases (4). It is observed that different chain lengths lead to small structural changes and that the proposed binding pocket cannot fully protect the thioester linker against hydrolysis for chain lengths longer than 10 carbons (8). It is hypothesized that these small structural changes in combination with the solvent-exposed hydrophobic tail might be important for the recognition process between the ACP and enzymatic partners. Based on available NMR and X-ray crystallography studies of acyl bound ACPs, a mechanism by which enzymes might be able to select specific acyl chains through recognition of surface features of the corresponding acyl-ACP substrates

Submitted August 11, 2020, and accepted for publication December 10, 2020.

*Correspondence: marcel.baer@pnnl.gov

Editor: Alexandr Kornev.
https://doi.org/10.1016/j.bpj.2020.12.036

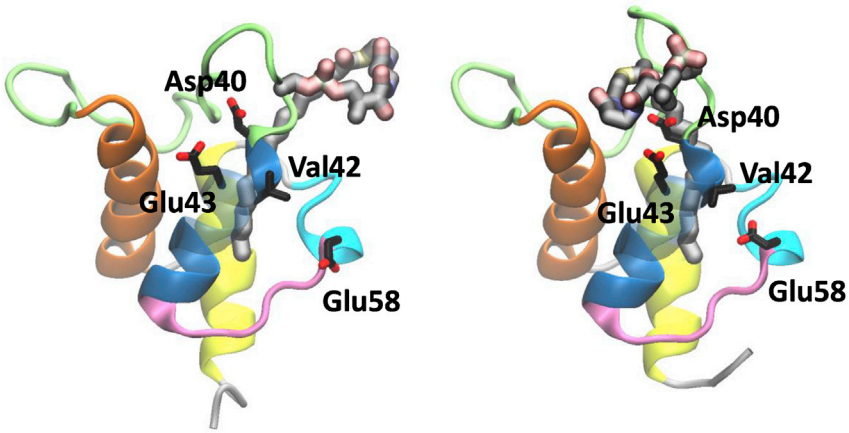

FIGURE 1 Cartoon representation of the spinach ACP NMR structure (PDB: 2FVA). The figure highlights the helices of the ACP fold, the side chains of residues in helix II that stabilize the ACP-Δ^9D complex, and the side chain of C_{18} fatty acid in the hydrophobic cavity in structural binding motif A (*left*) and motif B (*right*). The α-helices I, II, and IV are shown orange, dark blue, and yellow, and the 3_{10}-helix III is shown in cyan; the loop connecting helix I and II is shown in green, the loop connecting helix II and helix III in mauve, and the loop connecting helix III and helix IV, along with the C- and N-termini loop, in gray. To see this figure in color, go online.

was proposed (5,9). This mechanism suggests that shorter fatty acids are more strongly bound. Delivery of the substrate into the desaturase hydrophobic substrate binding channel must overcome these stabilizing interactions. It has been proposed that, for longer acyl chains, there is an increase of both solvent exposure and flexibility of the loop binding the phosphopantetheine that leads to a decrease in the barrier for the substrate transfer (9).

Sequence comparisons have revealed no common ACP binding motifs among its partner enzymes. However, there is general consensus that the acidic character of many bacterial ACPs leads to binding toward basic patches on the partner protein. In addition, in the majority of structurally characterized ACP/enzyme complexes, the conserved central helix II appears to play a dominant role in the protein-protein interaction. For this reason, it has been termed the "recognition helix" of ACP (10). Furthermore, examination of the known structures of ACP-dependent enzymes, including desaturases, has revealed a basic hydrophobic patch adjacent to the substrate channel leading to the enzyme active site, which has been implicated in an electrostatic interaction with the acidic helix II of ACP (10). Important key residues that were identified in the crystal structure of the ACP-stearoyl-ACP Δ^9-desaturase (Δ^9D) complex are highlighted in (Fig. 1). Finally, ACP exhibits relatively weak affinities with most enzymes (K_M generally in the low micromolar regime), as expected for a protein that must reversibly interact with many partners as part of its normal function (2).

The dynamical heterogeneity of the substrate-bound ACP is believed to be key for its ability to interact with diverse partners. This leads to the question of whether small changes in the ACP structure and/or its dynamic behavior can influence or control the binding to specific targets or discriminate for substrate length. Indeed, small substrate-dependent structural and dynamical changes are observed in ACP by NMR. Based on solution NMR structures of a well-studied ACP involved in the desaturation of stearic acid (C_{18}) by the $\Delta^9$18:0-ACP desaturase (EC: 1.14.99.6 Δ^9D), it was hypothesized that the partial exposure of the hydrophobic substrate is part of a specific recognition process involving a hydrophobic binding site for the fatty acid chain (9). Previous work also implicated coexpression of isoforms of ACP that partner with specific acyl-ACP desaturase variants for the production of monounsaturated fatty acids other than $\Delta^9$18:1 (11).

The Δ^9D performs a dehydrogenation reaction that results in the introduction of a *cis* double bond into a fatty acid between carbons 9 and 10. Δ^9D exhibits strong selectivity for stearoyl-ACP (18-carbon substrate, C18). A key residue involved for the regiospecificity of the dehydrogenation between carbons 9 and 10 was recently identified by X-ray crystallography. Structural comparisons between the ivy-Δ^4-desaturase (Δ^4-D) and castor-Δ^9-D revealed that the interaction of the 4′-phosphathetheine with the residue 280 of the desaturase regulates the regioselectivity of the enzyme (12). The amino acid at position 280 is an aspartate in Δ^9D and a lysine in Δ^4D. Mutation of Asp280 to Lys280 shifts the regioselectivity from Δ^9D to Δ^4D (12). These studies also identified key interactions with residues, mainly located on helix II, that stabilize the ACP desaturase complex (Fig. 1). Importantly, the crystal structures show that Δ^4-D and Δ^9-D bind ACP in two different orientations.

According to the available ACP structures, the hydrophobic acyl chain is positioned in a hydrophobic pocket formed by the four-helix bundle. During transfer of the substrate from the carrier to the active site of the enzyme, the acyl chain must overcome stabilizing hydrophobic interactions in the ACP binding pocket.

NMR experiments have established that ACP shows intrinsic structural and dynamical differences depending on the acyl chain length (8,9) and even small chemical modification to the linker (13). However, whether these observed differences are important factors governing the recognition and binding process remains an open question. Toward a better understanding of the possible functional role of these structural and dynamical changes, molecular dynamics (MD) simulations have been used to elucidate

atomistic details that are not readily accessible by X-ray or NMR investigations. One earlier MD study focused on the insertion of shorter acyl chains into *Escherichia coli* ACP binding pocket (14). It was shown that longer chains explore additional binding sites other than the proposed main pocket. These binding sites were mainly surface sites and a region between helices II and I, but there was no indication that *E. coli* has a second substrate binding channel. Another MD study focused on surface bound states for a variety of substrates and ACP structures from the mupirocin biosynthetic pathway, which confirms the binding pocket for a saturated C_{14} thioether substrate (15). Besides pure surface bound states, a new bound state, which was referred to as a tunnel state, was found. The entrance of the "tunnel" is shared with the canonical binding pocket located between helix III, the N-terminal part of helix II, and the loop N-terminal to helix II. Instead of being bound to the core of the four-helix bundle, the acyl chain is aligned along the midpart of helix I, emerging back into the solvent. Unlike all resolved ACP from *E. coli* structures for shorter substrates, the NMR structure of ACP from spinach showed two possible structural binding motifs for the substrate (herein referred to as binding motifs A and B; Fig. 1). Despite the fact that ACPs from spinach and *E. coli* have only 39% of sequence identity, their four-helix bundle topology shows a high degree of similarity (9). Kinetic studies showed an increase in the reaction rate (k_{cat}/k_M) when ACP from spinach instead of *E. coli* was used in the desaturation of C_{18} by Δ^9D (8). Differences were attributed to the existence of specific protein-protein interactions. To understand this process, it is essential that a precise picture of the substrate behavior bound to ACP is obtained. In particular, it is important to understand whether the occurrence of two distinct binding motifs plays a role in selectivity and reactivity.

In this work, we have used MD simulations to deepen our understanding of how different lengths of fatty acid chains influence the dynamics and structures under solution conditions. The recently resolved structure of ACP bound to desaturases (Δ^9D and Δ^4D) allows us to test whether the exposed hydrophobic tail should be considered a key player in the recognition process. This study provides new insights on the possible functional role of chain-length-dependent structural and dynamical changes for ACP bound with fatty acids.

RESULTS

We investigated the spinach ACP with a variety of fatty acyl chains: capric acid (C_{10}), myristic acid (C_{14}), palmitic acid (C_{16}), and stearic acid (C_{18}) bound via a phosphopantetheine group linker. We started by focusing on the dependence of the substrates' binding free energy on their length for each of the known binding motifs (motif A and motif B, Fig. 1) by performing umbrella sampling (US) simulations. US was chosen because of its computational efficiency in exploring the free energetics of processes that can be described in terms of a single-order parameter such as the one discussed in this study (16). With the free-energy landscapes for substrate binding at hand, extensive (unbiased) MD simulations were performed to refine the sampling of the bound states (free-energy minima). A number of static and dynamical properties were calculated and compared with the existing experimental quantities, such as NMR-based residue fluctuations and hydrolysis rates of the thioester bond. Consistent with NMR data, simulations reveal a rather large structural and dynamical heterogeneity. Finally, we analyzed how the different binding motifs influence the binding of ACP to its enzymatic partners.

Binding free energies

The propensity of substrates to adopt either a buried or surface binding motif was explored via US MD simulations. For each substrate, 20 US simulations were performed, starting with the substrate in either the binding motif A or motif B as inferred from available NMR structures of ACP with bound C_{10} and C_{18} substrates (see experimental section) (9). For C10, only motif A is modeled because no stable minimum was found for motif B compared with the longer chains. A total of 80 independent free-energy pathways for the movement of the substrate from an interior state to a surface state were explored (see Figs. S5–S83). Not all pathways lead to a transfer of the substrate to the ACP surface because of the chosen reaction coordinate (Fig. 2; Figs. S5–S83). Instead, some of the pathways for C_{14}, C_{16}, and C_{18} explored an unusual additional buried state, discussed below. Averaged free-energy profiles (see experimental section) are discussed in the following for each motif (Fig. 3). Indeed, the individual free-energy profiles (Figs. S5–S83) are rather sensitive to the chosen initial configuration, as indicated by the whiskers in the box plots reported in Fig. S2, and the averaged profiles provide a simpler picture of the energetics differences between the binding motifs.

All substrates prefer either the buried motif A or motif B over a surface bound state. The described tunnel state (15) is sampled for the longer (>10) acyl chains but it is not a local minimum in the calculated binding free energy. The binding motif A for all substrates is ~12–20 kJ/mol more stable than a surface state. On the other hand, the stabilization of the binding motif B with respect to a surface state has an appreciable dependence on the length of the substrate. For C_{14}, the free-energy profile follows very closely the one of motif A with an asymptotic value (surface state) of ~16 kJ/mol. Instead, C_{16} shows a larger asymptotic value (30 kJ/mol), whereas C_{18} shows a rather similar value (20 kJ/mol).

The shorter substrate (C_{10}) adopts only the binding motif A, which is ~16 kJ/mol more stable than a surface state. In fact, the length of the chain does not allow C_{10} to adopt the binding motif B. The next substrate of the series, C_{14},

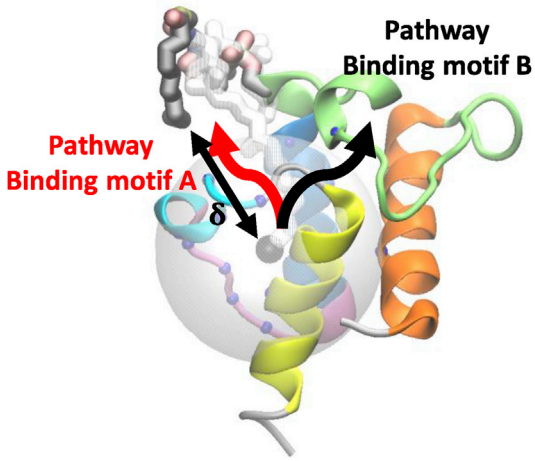

FIGURE 2 Relative binding free energies were estimated from umbrella sampling simulations using the distance (δ) between the center of mass of the last four carbon atoms of the fatty acid chain and Cα carbon atoms of the residues as order parameter ("reaction coordinate"); see text for details. Arrows indicate how the two binding motifs lead to two distinct pathways. The gray sphere highlights the boundary of the position of the tail of the fatty acid for reaction coordinate values of 1.0 nm. The specific structure of the fatty acid shown in the figure (C_{10}) has the tail located at a value of the order parameter of 1.2 nm, corresponding to the entrance of the binding pocket in binding motif A. The resolved position of the substrate in the NMR structure is shown in light-gray outline. To see this figure in color, go online.

adopts either binding motif A or motif B with (within our statistical uncertainty) equal probability. Indeed, both binding motifs are ~16–20 kJ/mol more stable than surface states. In contrast, the other two substrates, C_{16} and C_{18}, favor binding motif B. The preference is marked for C_{16}, for which the binding motif B is, on average, ~32 kJ/mol more stable than a surface state, whereas binding motif A is ~20 kJ/mol more stable than a surface state. In the case of the longer substrate (C_{18}), binding motif B is ~20 kJ/mol more stable than a surface state and is only slightly more stable than binding motif A. Interestingly, for C_{14} and C_{16}, the pathways connecting both binding motifs A and B to the surface states exhibit a local minimum at a value of the reaction coordinate $\delta \sim 0.7$ nm (Fig. 3). In this local minimum, the substrate explores an additional binding pocket located between helix I and IV, as observed in previous simulation studies for substrates with more than an eight-carbon chain length in *E. coli* ACP. The barrier between these local states decreases with chain length and is not present for all for C_{18}, resulting in a broad minimum.

Our US simulations show that, when starting from motif A, substrates C_{14}, C_{16}, and C_{18} explore additional binding states in which the hydrophobic tail of the substrate is fully buried inside the ACP. The fatty acid tail occupies the same hydrophobic pocket as seen in the NMR ensemble, but with an inverted orientation. The thioester is close to the bottom of the pocket. A representative structure for C_{18} is shown in Fig. 4. This buried state is sampled along six (C14), four (C16), and two (C18) out of the 13 independent pathways for motif A for each of the chain lengths. The overall protein structure is not affected by this inversion. As the fatty acid occupies the same ACP region, this binding motif is expected to satisfy the same NMR restraints for motif A.

If we assume that the equilibrium population is determined solely by the binding free energy as calculated here, the chain length would change the relative population of the possible binding motifs. The difference in binding energy would shift the overall population toward motif B for C_{16} and C_{18}. This could influence the overall exposed residues on the surface of the ACP, thereby influencing the recognition process. Crystal structures show differences

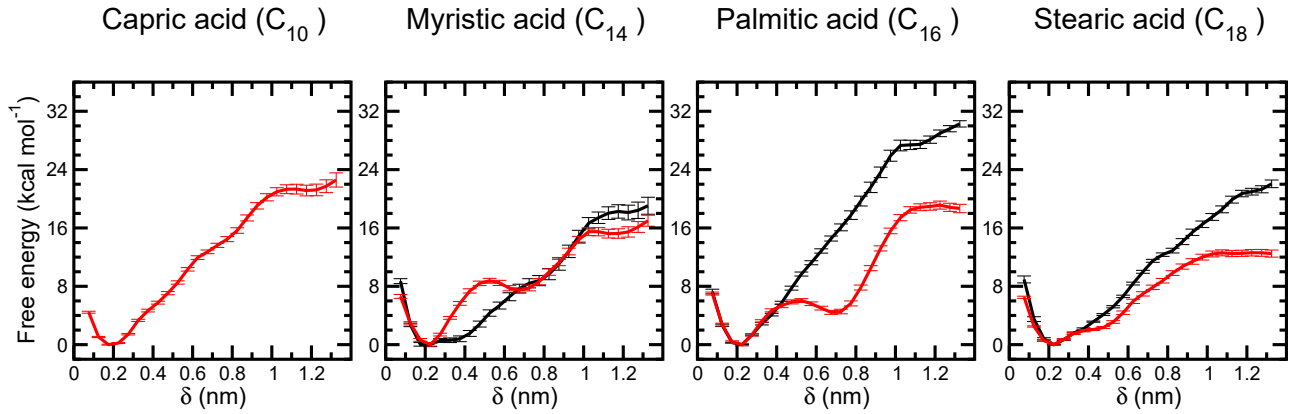

FIGURE 3 Averaged binding free energy for the fatty acid chains to the hydrophobic pocket from umbrella sampling simulations along the center of mass of the fatty acid chain and the center of mass of the binding pocket δ (see experimental procedure for details) for (*a*) capric acid (C_{10}) for binding motif A only (because no stable minimum was obtained for motif B) and (*b*) myristic acid (C_{14}), (*c*) palmitic acid (C_{16}), and (*d*) stearic acid (C_{18}) for binding motifs A and B (*red* and *black*, respectively). All free energies are relative to respective global minimum is. Mean-square deviations from the average free-energy profile are given as error bars as a measure of the dispersion of free-energy values among the different pathways. For more detailed statistics, see Fig. S2. To see this figure in color, go online.

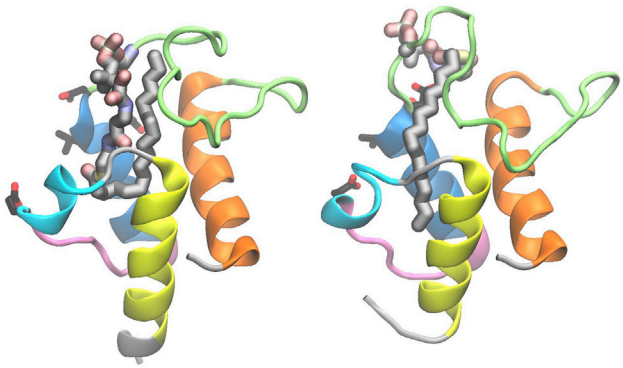

FIGURE 4 Representative structure of the newly identified buried binding motif from the biased C_{18} simulation (*left*) and representative binding of the acyl chain in the NMR structure (see Fig. 1). The thioester is close to the bottom of the binding pocket, and the hydrophobic tail occupies the same space as seen in the binding NMR structure. The α-helices I, II, and IV are shown orange, dark blue, and yellow, and the 310-helix III is shown in cyan; the loop connecting helix I and II is shown in green, the loop connecting helix II and helix III in mauve, and the loop connecting helix III and helix IV, along with the C- and N-termini loop, in gray. To see this figure in color, go online.

for the orientation of the ACP for Δ^9D and Δ^4D, which are selective for C_{18} and C_{16}, respectively. The influence of the binding motif for the transfer will be discussed below.

Protein and substrate dynamics

The free-energy calculations discussed in the previous section establish the dependence on the substrate length of the relative stability of possible binding motifs. Starting from the computed free-energy minima, we analyzed the protein and substrate dynamics using unbiased MD simulations (see experimental section for details). These simulations reproduce the average free energies (see Fig. S1) exploring states up to 8 kJ/mol from the bottom of minimum as expected for a thermal sampling at room temperature. The fluctuations of the substrate bound protein are compared with fluctuations of an apo-ACP without the linker.

The effect of the bound fatty acids on the dynamics of ACPs has been widely recognized. Extensive NMR relaxation data on both apo- and holo-ACPs have identified regions of increased fluctuation, showing a dependence for longer-timescale motions on the length of the fatty acid (8). In particular, it was found that the loop region between helices I and II as well as part of helix II has increased fluctuations on the micro- and millisecond. Specifically, increased fluctuations were observed for residues from 37 to 44 for C_{10}, 35 to 44 for C_{14}, and 27 to 44 for C_{18}. On the other hand, regions of increased fluctuations occurring on a faster (pico- and nanosecond) timescale, which include the loop region between helices I and II (residue 20–23) and helix II with surrounding residues (residues 56–65), do not show any significant substrate specificity. Consistent with these investigations, we found that some regions are influenced by the chain length more than others, as can be appreciated from the differences in the root mean-square fluctuation (RMSF) of each ACP residue reported in Fig. 5. These regions include parts of (or, depending on the substrate, the full) helix bundle enveloping the substrate. However, the part of ACP that is influenced the most is the loop that connects helix I to helix II (loop I, shown in green in Fig. 5). Notably, these simulations show that the type of binding motif (motif A versus motif B) has a rather large impact on the dynamics of ACP, which is clearly evident in Fig. 5.

A direct comparison between calculated and experimental RMSF can be made for the spinach ACP with bound C_{10} and C_{18} fatty acids, for which an ensemble of 20 NMR structures each is available. Because of the limited number of NMR structures, this comparison should be considered as purely qualitative. That said, the RMSF from the MD simulations agree well with the RMSF calculated from the 20 NMR structures. The only appreciable differences are found in the flexible (and unstructured) loop I.

Clear differences (increases and decreases) in the fluctuations of ACP with bound C_{14}, C_{16}, and C_{18} with respect to the apo-ACP are evident. For C_{10}, only an increase in the loop region (Fig. 5, green) is observed. Therefore, the binding of the substrate induces only minor changes for C_{10}. On the other hand, for C_{14} and C_{16}, larger effects due to the binding of the substrate are evident. Experimentally, it was found that an increase in substrate length leads to an increase of protein fluctuations in the micro- and millisecond timescale motions. These fluctuations are associated with motions between helix I and II. In addition, larger fluctuations of residues 20 to 25 in the loop I region are evident from NMR measurements. These changes in the protein fluctuation are consistent with the results of our simulations. Also, an increase for helix II (residues 48–56; see Fig. 5, in mauve) is seen, which is located at the bottom of the substrate channel.

Two regions, one in the loop region and the other in helix III, show marked differences for the C_{14} and C_{16}, whereas C_{18} shows similar fluctuations to those of C_{10}. The first difference is a significant increase of fluctuations in the loop region for residues 25–37 and a decrease in fluctuations for residues 37–40, highlighted in green in Fig. 5. This is in good agreement with the region identified to show substrate-specific fluctuations. For C_{14}, a decrease in helix III (residues 57–65) is observed. The changes for the medium chain length can be rationalized with the increase in chain flexibility. This allows the substrate to maximize its hydrophobic contacts with the ACP by displacing the phenylalanine ring of residue 31. This residue acts as an anchor of the loop region into the helix bundle and thereby the loop region becomes more mobile. This shift of the substrate leads to increased interactions between helix III and II, thereby reducing some of the fluctuations. For C_{18}, this change of fluctuations is not observed. Instead, the longer

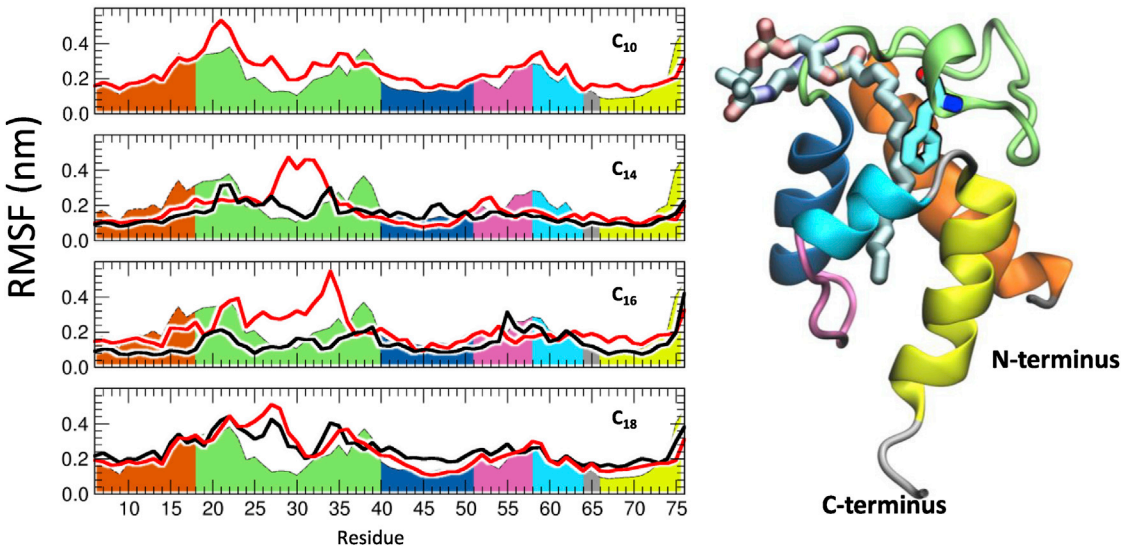

FIGURE 5 Root mean-square fluctuation (RMSF) of ACP residues with and without the substrates. For each substrate, the RMSF for binding motifs A and B are shown as red and black lines, respectively. The RMSF for the protein without the substrates is shown in each panel with colors highlighting the protein structural elements as illustrated in the cartoon representation. As a reference, in the structure to the right, the substrate C_{18} in the binding motif A is also shown, along with the side chain of Phe31. Helix I, II, III, and IV are shown orange, dark blue, cyan, and yellow; the loop connecting helix I and II is shown in green, the loop connecting helix II and helix III in mauve, and the loop connecting helix III and helix IV, along with the C- and N-termini loop, in gray. To see this figure in color, go online.

and more flexible hydrocarbon chain leads to an increased number of intramolecular contacts.

These simulations indicate that the dynamics of the substrates also depends on both the length of the fatty acid chain itself and the type of binding motif. As can be seen from RMSF of the carbon atoms of the fatty acid chain shown in Fig. 6, substrates in the binding motif B show an appreciably smaller flexibility. For both binding motifs, the flexibility of the fatty acid chain decreases up to the fifth or sixth to the last atom and then increases steadily. The lowest RMSF corresponds to atoms near Phe31, where the chain bends (see, for instance, Fig. 5). Regardless of the chain length and the binding motif, only the last five to six atoms of the chain are inserted in the core of the ACP protein, between the four-helix bundle. The rest of the fatty acid chain is folded on the ACP protein (between loop I and helix II for binding motif A or helix I and loop I for binding motif B). Depending on the chain length, the 4′-phosphathetheine moiety and the initial segment of the fatty acid fits the surface of the ACP protein in different manners and influences the dynamics of helix I, helix II, and loop I differently.

Further information on the influence of the substrate on the dynamics of ACP can be obtained from a principal component analysis (PCA), also known as "essential dynamics." The PCA uses the covariance matrix of the amino acid displacements to extract information on long-timescale, large-amplitude fluctuations of the system (17). We applied PCA to the motion of the α-carbons. The first and last five amino acids were excluded because they are unstructured and almost freely moving, thereby introducing a considerable bias in the analysis. A number of studies have shown that typically the first 20 modes are enough to capture large-scale fluctuations, thereby achieving a tremendous reduction of the dimensionality of the problem. Therefore, to compare how the essential dynamics of the protein depends on the

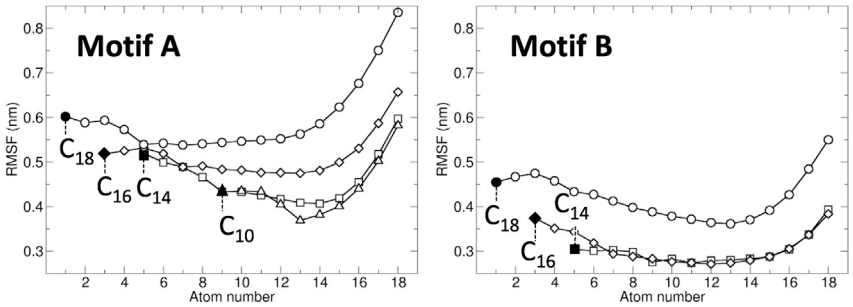

FIGURE 6 Root mean-square fluctuation (RMSF) of the fatty acid chains for the binding motifs A and B. The chains are aligned with respect to the terminal carbon atom. Atom numbering is relative to C_{18}; the first atom of the fatty acid chains (the atom linked to the thiol group 4′-phosphathetheine) is indicated with a full symbol.

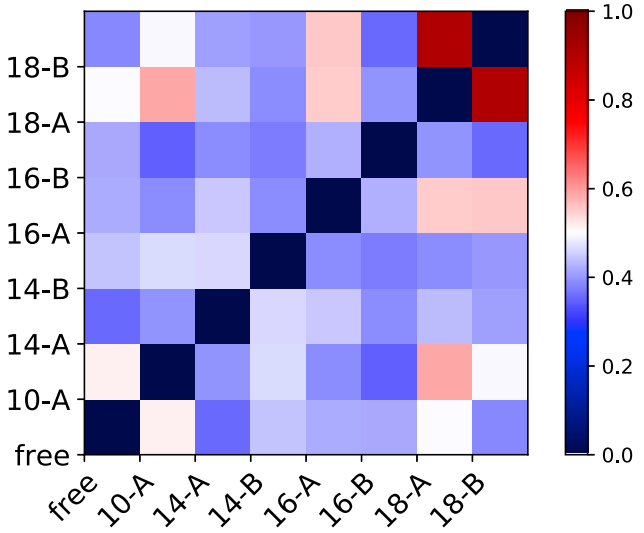

FIGURE 7 Maximal overlap between the first 20 eigenvectors obtained by PCA on the α-carbons only. Compared are all possible combinations between the apo(free) and the substrate bound ACP. PCA was performed grouping all simulations into either motif A or B. To see this figure in color, go online.

acyl chain, the overlap between the eigenvectors of only the first 20 normal modes is analyzed. A value of the overlap of 1 indicates identical dynamics, whereas values below 0.4 are indicative of very different dynamical behavior. Because the order of the eigenvectors can be different for different substrates, the maximal value of the overlap is reported for all possible combinations in Fig. 7. The largest overlap (>0.8) is observed for the simulation of C_{18} in motifs A and B. The next largest is smaller than 0.6 and is between the eigenvectors of C_{16} and C_{18} for both motifs A and B. Overall, these results, combined with the analysis of the RMSF, indicate that the length of the acyl chain and type of binding motif not only influence the amplitude of the motion of specific parts of the protein but also change the overall dynamics of the protein itself.

Solvent accessibility of the thioester bond

Experiments show that the hydrolysis of the acyl chain bound to ACP depends on the chain length (8). The longer the chain length, the faster the hydrolysis, with a noticeable increase for a chain length longer than 15 carbon atoms. It was proposed that the reaction rate correlates with the degree of solvation of the thioester (8). Long fatty acid chains cannot be fully accommodated inside the protein and therefore screen the thioester from the solvent and, as a result, have slower hydrolysis rates (8). We can test this hypothesis by analyzing the average number of water molecules close to the thioester bond along with the presence of residues nearby the bond that can act as a general base for the activation of water for hydrolysis.

The average number of water molecules is calculated using a continuous coordination number (see experimental section). The average number of water molecules can be taken as an estimate of the probability of finding water around the thioester bond and is a measure of the solvent accessibility to that site. A value of the coordination number of 1 indicates that water molecules are always present around the thioester bond. As can be seen from Fig. 8 (left panel), the solvent accessibility to the thioester increases with chain length, with a probability of finding water around the bond of ~20% for C_{10} and 50% or more (depending on the binding motif) for C_{18}. However, the trend reported (Fig. 8) does not show any steady increase for longer substrate, as would be expected from the abrupt increase in the hydrolysis rate observed for C_{16}. Conversely, the increase is rather gradual, and the distribution about the average is rather broad.

The highly negatively charged ACP is rich in aspartic and glutamic residues, which could serve as general bases assisting the hydrolysis of the thioester by water. The two binding motifs differ in the folding of the 4′-phosphathetheine moiety on the surface of ACP. In binding motif A, there are only two potential bases near (Asp40 and Glu43) the thioester bond, whereas in binding motif B, there are four potential

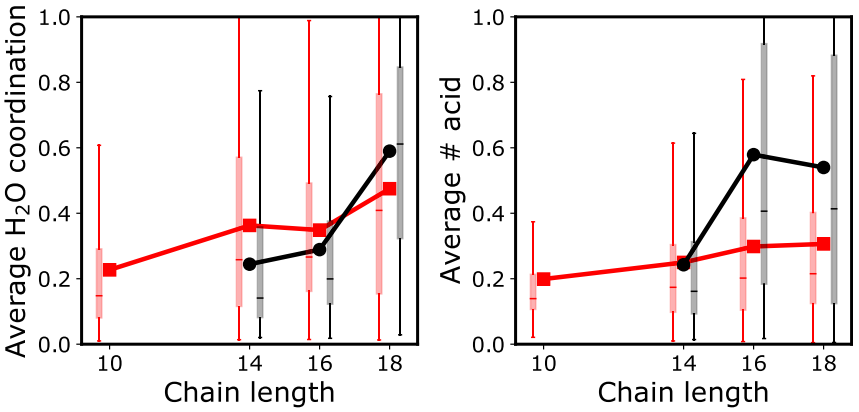

FIGURE 8 Left: chain length dependence of the average number of water molecules within 0.24 nm of the thioester group for motifs A (*red*) and B (*black*). Right: chain length dependence of the average number of carboxylic acid oxygens within 0.45 nm of the thioester group for motifs A (*red*) and B (*black*). The shadowed boxes are drawn from the first quartile to the third quartile with median as vertical line. The whiskers go from each quartile and extend by 1.5 times the interquartile range (Q3–Q1) for Q3 and (Q1–Q3) for Q1, respectively. To see this figure in color, go online.

bases (Glu30, Asp59, Asp62, and Asp63). As can be seen in Fig. 8 (right panel), the thioester bond in all substrates has on average the same probability of being close to the carboxylic group when they are in the binding motif A. Instead, substrates in binding motif B show a different behavior. Indeed, although the probability of finding a carboxylate near the thioester bond for C_{14} is similar to that observed for the binding motif A, this probability is appreciably higher for C_{16} and C_{18}. Based on these results, it could be hypothesized that the higher probability to have a general base near the thioether in C_{16} and C_{18} when they adopt a binding motif B could lead to a faster hydrolysis rate. Therefore, the experimental increase in the hydrolysis rate observed experimentally (8) could be due to the catalytic effect of nearby carboxylates when the longer substrates are in the binding motif B.

Water response and electrostatic map of the ACP with substrate bound

Based on the structural ensemble of solution structures determined by NMR, it was hypothesized that the occurrence of two different binding motifs and the exposed hydrophobic area of the substrate could be a discriminating factor determining the binding affinity. In the following, we explored this hypothesis computationally by monitoring how the solvent on the surface of the ACP protein responds to a change in the binding motif. In fact, it has been shown that changes in the solvent organization around binding sites provides valuable information on how structural perturbations influence the binding affinity between partners. The solvent response is analyzed using the method introduced by Remsing et al., which allows not only for the distinction of hydrophobic or hydrophilic regions but also for the orientation of solvating water molecules (18). Polarized water molecules can point their dipole moment toward the surface (anion-like solvation) or away from the surface (cationic-like solvation). As the two surfaces are brought closer together, polarized interfacial water may "interfere deconstructively" and effectively repel the two surfaces. Therefore, to achieve small interfacial distances, confined water must be (unfavorably) depolarized. In contrast, one might expect that bringing together surfaces with regions of opposing polarity, like the red and blue regions in Fig. 9, would instead lead to hydrogen bond networks that interfere constructively because the intersurface water is polarized in the same direction everywhere between these patches. In this case, unfavorable depolarization may not occur, minimizing this contribution to the effective repulsion between surfaces and possibly leading to an effective attraction between hydrophilic surfaces.

The binding interfaces of the ACP and desaturase should show opposing polarity for an effective recognition process. If the binding motif and the proposed exposed hydrophobic surface of the substrate play a key role, different motifs should show differences in the intrinsic orientation of the solvent. In Fig. 9, the resulting polarization of the ACP is shown for two representative C_{18} structures for binding motifs A and B. The binding motif does not significantly alter the water polarization around the recognition helix. The binding interface shows a large patch of water molecules in an anion-like solvating configuration and only small regions of hydrophobic patches, whereby water molecules adopt random orientations. Preliminary MD simulations show that the binding interface of the desaturase to which the ACP binds has water molecules in a cation-like solvating configuration. This complementarity of the water orientation between the ACP and desaturase binding regions is expected for a favorable interaction between the two proteins. However, the solvent orientation around the ACP binding site is not affected by the binding motif adopted by the substrate. Therefore, the binding motif is not expected to result in nonspecific binding differences mediated by the solvent. Instead, it might be that surface residues of the desaturase determine the preference of selectivity for different binding motifs of the ACP substrates. This hypothesis is tested in the next section by using atomistic docking simulations.

Docking

Available crystal structures of Δ^9D and Δ^4D bound to ACP show that ACP can adopt two distinct binding orientations of ACP with respect to the desaturase (12,19). It was shown that a single surface residue (Asp280) is the critical

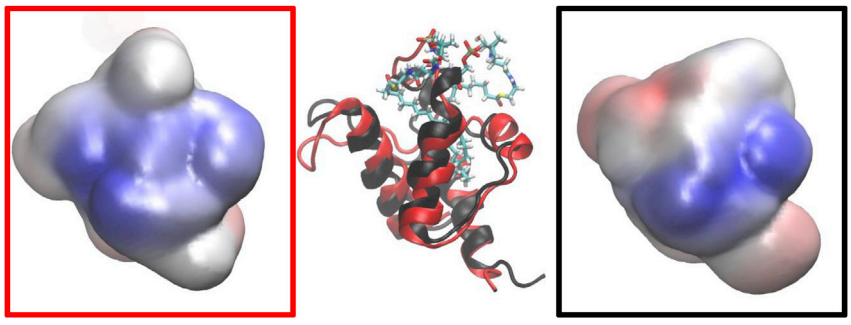

FIGURE 9 Water polarization for binding motifs A (*red*) and B (*black*) for C_{18}. The orientation is according to Fig. 1 with the recognition helix II in the front. The protein surface as described by a Willard-Chandler smoothed interface for the protein heavy atoms is shown as a surface and is colored by the value of potential. Red represents negative values of the potential, blue indicates positive values of the potential, and white corresponds to values near zero, such that slightly negative, pale red-pink indicates values of the potential observed above hydrophobic surfaces. To see this figure in color, go online.

determinant for the different binding modes of ACP and that mutation of Asp280 into Lys resulted in a major shift from Δ^9 to Δ^4 desaturation of C_{18} as substrate. The available structures of the ACP/desaturase complex represent a natural starting point to investigate how the relative orientation of ACP and the substrate binding motif can influence the overall desaturase activity. Overlaying representative structures for ACPs with either the substrate in binding motif A or B suggests that for Δ^9D direct substrate transfer is allowed only for the binding motif A. On the other hand, in the case of Δ^4D, direct substrate transfer would be allowed for binding motif B because of the different relative ACP/desaturase orientation. This is highlighted in Fig. 10 for Δ^9D. The natural substrate selection is C_{18} for Δ^9D and C_{16} for Δ^4D. Here, we explore the hypothesis that the intrinsic substrate binding motifs and their relative stability have an effect on the selectivity. Under the assumption that the calculated binding energy for the substrate to the ACP determines the relative population of possible binding motifs, C_{16} would exhibit a higher population of motif B, C_{14} would exhibit a nearly equal distribution between motifs B and A, and C_{18} would show a slightly larger population for motif B.

To test whether a preferential orientation is found not only in the crystal structure but also in encounter complexes in an aqueous environment, docking simulations for ACP to Δ^9D and Δ^4D were performed. The ACP/desaturase interaction was biased with an additional constraint to ensures residue 38 of the ACP was within the vicinity of the substrate channel to allow a transfer of the substrate; for details, see computational approach section. The trial structures are ranked based on their interfacial interaction energy as calculated within the ROSETTA docking module (20,21). The lower the interfacial interaction energy, the more probable the structure is. Of all generated structures (6000 for each desaturase), the top 5% are analyzed using a clustering algorithm (22) to identify unique ensembles of structures of ACP docked to the desaturase. Each of the resulting clusters was inspected to see whether the substrate could be transferred from the ACP (in either binding motif A or B) into the binding pocket of the desaturase. For the Δ^9D, clusters containing two-thirds of the highest ranked trial structures would allow for both motifs A and B to be transferred without obvious clashes with the ACP desaturase interface. On the other hand, clusters for the Δ^4D show a preference for motif B, indicating that the ACP-desaturase interactions could lead to a preferential selection between motif A or B.

DISCUSSION

We explored the structural and dynamical heterogeneity of the spinach ACP hosting a variety of fatty acid substrates to gain insight on how structural diversity might influence the recognition process between the ACP and desaturases. Specifically, we investigated how the length of the fatty acid chain influences substrate binding to ACP and the overall ACP structural dynamics by extensive MD simulations.

Earlier MD simulations on the *E. coli* ACP suggested that the ingress of the substrate in the hosting ACP helix bundle is a barrierless process for fatty acids with chain lengths between 4 and 18 carbon atoms. Consistent with NMR results for the spinach ACP (8,9), higher fluctuations were only observed between substrate bound ACP compared with the apo- and holo-ACP in the end of helix II into the loop. It was suggested that this might hint to differences in the accommodation of the substrate in the usual binding motif. The occurrence of the alternative binding motif B was not further discussed (14).

Here, we expanded these previous computational studies by performing US simulations to explore the binding pathways of four fatty acids (C_{10}, C_{14}, C_{16}, and C_{18}) and estimate the corresponding free energy of binding. Our simulations were started from an ensemble of NMR structure of the spinach ACP hosting C_{10} and C_{18} substrates. For each substrate, 20 independent free-energy calculations were performed. As illustrated in the Supporting material, individual simulations show a large variation in the free-energy profile, some of which include appreciable barriers between the solvent-exposed and bound state. This highlights that variations in the protein conformation influence the binding process and the sampling of a statistically relevant number of conformations (starting structures) is required for a meaningful analysis of the possible binding pathways. Despite this heterogeneity in the binding pathways, we

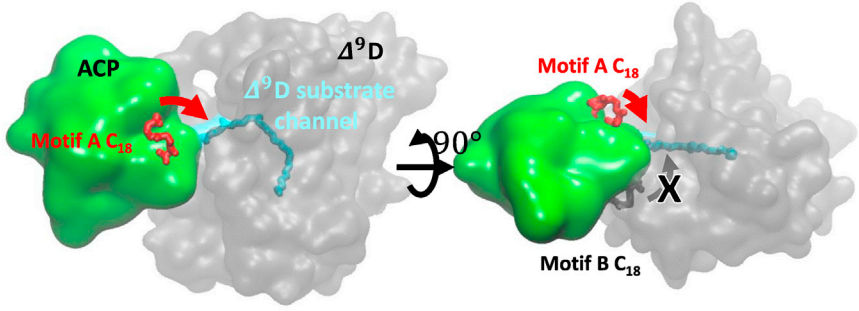

FIGURE 10 Models of the desaturase-ACP complexes. Castor desaturase-ACP complex (desaturase in *gray* and ACP in *green*), highlighting a representative configuration of the fatty acid in binding motifs A (*red*) and B (*black*). The position of the 4′-phosphathetheine moiety and fatty acid inserted into the desaturase binding channel (*cyan shaded area*) as modeled in Ref (12) is shown in cyan. Rotated view illustrates the potential clashes of the substrate in motif B with the ACP desaturase interface. To see this figure in color, go online.

found that the bound states are on average 16–20 kJ/mol more stable than states with fully solvent-exposed substrates (Fig. 3) and that binding is either barrierless or requires the crossing of a negligible barrier.

Our simulations highlighted possible binding motifs for substrate in the spinach ACP. Two of these binding motifs (motifs A and B; Fig. 1) were already identified by NMR spectroscopy and discussed in the literature (9). The new binding motif (Fig. 4) is a variant of motif A, and to the best of our knowledge, it was not reported before. Increase in chain length also allows for the alternative buried binding configuration, whereby the tail of the substrate has an inverted orientation with respect to the binding motif A. Because the fatty acid head occupies the same region within the ACP that the tail usually occupies, this structure most likely would satisfy the same NMR restrains of motif A. Within our statistical uncertainty, binding motif A shows roughly the same stability for all chain lengths investigated in this study. Instead, the stability of binding motif B, which has been uniquely described for the spinach ACP, depends appreciably on the length of the fatty acid chain. This state is not accessible for C_{10} (the chain is too short to fold according to binding motif B). In the binding motif B, C_{14} and C_{18} show binding free energy comparable to binding motif A, with binding motif B only slightly more stable for C_{14} than binding motif A and slightly more unstable for C_{18} (Fig. 3). In contrast, binding motif B is considerably more stable for C_{16}. Generally, the longer the chain, the larger the flexibility of the bound fatty acid.

The protein dynamics and fluctuations (as measured by the RMSF of the protein residues) compare well with the experimental available NMR relaxation data, even though the timescales for many of the fluctuations seen in the experiment are far beyond the simulation timescale. The regions that were experimentally identified to have increased fluctuations compared with the C_{10} are seen in the simulations as well. The large fluctuations seen for C_{14} and C_{16} in the loop region around the phenylalanine (residue 31, Fig. 5), along with the decrease in the region of helix II, observed in our simulations might be potentially relevant for the binding with the ACP partners.

The length of the fatty acid chain, and consequently differences in the binding geometry, has been proposed to impact on the solvent exposure of the thioester bond that links the fatty acids to the 4′-phosphathetheine anchor and consequently to the observed change in the hydrolysis rate of this bond. The longer the chain length, the faster the hydrolysis, with a noticeable increase for chain lengths longer than 15 carbon atoms. Based on this fact, it was concluded that this change in hydrolysis rate marks the maximal chain length that can be fully incorporated into the ACP, thereby protecting the thioester from the solvent. We do see an increase in solvent accessibility to the thioester with chain length, but within our statistical uncertainty, there is no indication that there is a significant or sudden change as a function of the chain length (or binding motif; Fig. 8). On the other hand, our simulations reveal a clear increase of the probability of having aspartic and glutamic residues nearby the thioester bond in C_{16} and C_{18} when they adopt binding motif B. These residues could facilitate hydrolysis by acting as general bases.

Taking a step further, we investigated whether the acyl chain length and binding motif alter the properties of helix II and consequently influence the recognition process. ACP is mainly negatively charged, and the substrate by itself is hydrophobic. Therefore, it can be assumed that partial exposure, especially for the longer substrates, would lead to hydrophobic patches on the ACP surface, covering negative patches. We found that the type of binding motif does not lead to a significant change in the water structure or polarization around the recognition helix. This result indicates that the binding motif does not have an intrinsic interfacial signature that would appreciably influence the binding to desaturase or other partner enzymes.

Our solved crystal structure of the castor Δ^9D (12) and modeled complex of ACP bound ivy Δ^4D (19) desaturase show that a single surface residue is the critical determinant for the different binding modes of ACP. This binding mode is proposed to ensure the proper orientation and insertion depth of the substrate into the desaturase, thereby dictating the regioselectivity. Our fit of the NMR solution structures onto the crystal structure suggest that the Δ^9D would allow for direct transfer of the substrate for motif A, whereas the Δ^4D on the other hand would allow for a direct transfer for motif B. A more accurate analysis based on docking simulations confirms that this preferential orientation is found not only in the crystal structure but also in encounter complexes in an aqueous environment. Ensembles of structures with favorable interfacial interactions between the ACP and desaturase show preferences for motif B for Δ^4D and mixed preferences (A and B) for Δ^9D. These calculated ensembles can be used in future studies to investigate the presence of on additional interface on the binding free energy of the substrate to ACP, as well as the free energy of transferring the substrate from the ACP to the desaturase.

Various acyl-ACP desaturases have arisen multiple times as variants of the archetypical 18:0-ACP desaturase (23). Such variants invariably insert double bonds closer to the acid end of the acyl chain and have preferential selectivity for alkyl chains shorter than 18 carbon atoms, in part due to the need to accommodate additional methylene groups' substrate binding cavity beyond the diiron active site (24). We previously showed that evolution of Δ^4 from Δ^9 regioselectivity required a 60-degree rotation of the ACP relative to the desaturase to allow deeper penetration of the acyl chain into the substrate binding cavity (12). Here, we identify two acyl chain binding motifs within ACP: motif A, which is better suited to binding C_{18}, i.e., stearate, and motif B, which is optimal with C_{16}, i.e., palmitate. That motif B is optimal for C_{16} and ACP rotation with respect to desaturase binding and favors regioselectivity $<\Delta^9$, which suggests

that evolution of motif B predisposed ACP to regioselectivity $<\Delta^9$ along with C_{16} chain length specificity. This provides a partial explanation for the reported difficulty in separating regioselectivity from chain length selectivity for these enzymes (25).

In summary, this study suggests that differences in the binding motif of acyl chains to ACP dictated by their length might influence desaturation. Differences in binding kinetics of ACP from *E. coli* and spinach ACP to desaturase, with the spinach ACP having a faster kinetics, support this hypothesis. To further explore this hypothesis, the binding of substrates to *E. coli* ACP must be explored computationally and/or experimentally. Using the knowledge presented herein as a basis for enzyme redesign, it is feasible that mutants of acyl-ACP desaturases, ACP, or perhaps both could be designed to create new binding modes. In combination with variations in the substrate binding cavity within the desaturase, new regioselectivities and chain length selectivity could be achieved.

Computational approach

Simulation setup

MD simulations were set up for the spinach ACP containing the phosphopantetheine group and either a capric acid (C_{10}), myristic acid (C_{14}), palmitic (C_{16}), or stearic (C_{18}) fatty acyl chain. The starting structures were obtained using available NMR structures of ACP with bound C_{10} and C_{18} substrates (PDB: 2FVA (9) for C_{18} and PDB: 2FVE (9) for C_{10}). Each NMR ensemble contains 20 independent configurations. The C_{18} shows distinct structural features (see Fig. 1). The first set of structures features the fatty acid chain in the usual binding pocket (as also found for C_{10}), which is accessed through the cleft between helix II and III (binding motif A). The second set of structures is less populated (7 out of 20 structures) and features the substrate in different binding motif (binding motif B), whereby the substrate enters the binding pocket between helix II and I. Starting structures for ACP with bound C_{14} and C_{16} were modeled using the C_{18} and shortening the fatty acid chain, obtaining a total of 20 starting structures for each substrate, with 13 in motif A and 7 in motif B. All simulations were performed using the GROMACS 4.6.2 (26) with the OPLS_AA force field (27) in conjunction with the TIP3P (28) water potential. Charges for the phosphopantetheine group were obtained using the standard RESP charge fitting approach (29). See Supporting material for topology and force field parameters for the nonstandard phosphopentetheine group and covalently bound fatty acid side chain. The protein was solvated in a cubic box of length 9 nm, and the system was neutralized by 12 sodium ions.

Binding free energy using US

The binding free energy of the fatty acid chain to the hydrophobic pocket is calculated using US. For these calculations, we chose as order parameter ("reaction coordinate") the distance between the center of mass of the last four carbon atoms of the fatty acid chain and C_α carbon atoms of the residues that were identified to be in contact with the fatty acid by NMR for the C18 structures. These are residues V9, F31, T41, I44, V45, E48, E49, I54, N55, V56, D57, K60, A61, Q62, I64, A70, V73, I74, and L77 (see Fig. 2). To facilitate a direct comparison of the simulation results, the same definition is used for all four fatty acids and binding motifs. The choice of reaction coordinate biases the system, and only relative free energies will be obtained.

The initial configurations for the US calculations were generated via steered MD simulations by applying a harmonic potential with a force constant of 1000 $kJmol^{-1}nm^{-2}$ that is pulled along the chosen order parameter at a rate of 0.0015 nm ps^{-1}. From these simulations, starting structures for the US simulations were extracted for values of the reaction coordinates ranging from 0.1 to 1.5 nm in 0.1-nm intervals. Each umbrella window was simulated for 50 ns using a force constant of 600 $kJmol^{-1}nm^{-2}$. The resulting free energy was obtained using the weighted histograms analysis as implemented within GROMACS (26,30,31). A total of 80 independent free-energy calculations are performed—20 for each of the different chain lengths. Structures for C_{10} were also modeled for motif B to test whether the short chain could adopt this configuration. However, no stable minimum was found compared with the longer chains. Therefore, C_{10} is only sampled and discussed in motif A. Uncertainties for individual free-energy profiles were estimated using a Bayesian bootstrap method with 500 trials as implemented in GROMACS (31). Uncertainties are ~4 kJ/mol. Averaged free energies for motifs A and B are obtained using the individual free energies. As the single free energies can differ significantly, the statistic over the data is shown as a box plot (Fig. S2; (32)). Shown are the minimum, first quartile (Q1), median, third quartile (Q3), and maximum as well as outliers. The box is drawn from Q1 to Q3. A vertical line goes through the box at the median. The whiskers go from each quartile and extend by 1.5 times the interquartile range (Q3–Q1) for Q3 and (Q1–Q3) for Q1, respectively. Data points beyond the whiskers are shown as outliers and indicated as circles.

For each of the 80 structures, unbiased simulations are extended using the last frame from the biased simulation around the free-energy minimum, biasing window for 0.2 nm. Each simulation is run for at least 240 ns.

Solvation structure around thioester bond

The coordination number (CN) of water molecules and carboxylic oxygens around the thioester bond are calculated as

$$\mathrm{CN} = \sum_i^N \frac{1-\left(\frac{r_i}{r_0}\right)^6}{\left(1-\left(\frac{r_i}{r_0}\right)^{12}\right)}$$

where N is the number of all water oxygen atoms or acid oxygen atoms, and r_i is the distance between a given oxygen atom to the sulfur of the thioester bond. The cutoff distance r_0 is set to 0.24 and 0.5 nm for water oxygens and acidic oxygen atoms, respectively. The coordination numbers were calculated using PLUMED (33).

Structural response of water molecules

The long-range orientational response of the solvent water around proteins can be used to predict how two proteins interact and associate. This property has been shown to be a powerful alternative to conventional analysis based on the complementarity of electrostatic potential (18). During protein binding, the water between them has to be displaced. In the initial phase of the approach, there will be an overlap between the structured solvation shells. To investigate the collective, short-frequency electrostatic response of water to the presence of the ACP surfaces and its dependence on the chain length and binding motif, we use the recently introduced method by Remsing and Weeks (18). The method uses an electrostatics-based mapping of aqueous interfaces. In addition to distinguishing between hydrophobic and hydrophilic regions of heterogeneous surfaces, this electrostatic mapping can also differentiate between hydrophilic regions that polarize nearby waters in opposing directions.

Docking of ACP to Δ^9D and Δ^4D

Following the protocol that was used to solve the crystal structure of the ACP bound to the castor Δ^9D (PDB: 2XZ1 (12)) and model the bound orientation of ACP to ivy Δ^4D (PDB: 2UW1 (19)), docking simulations are performed for the ACP without the substrate and replacing serine 38 with a glutamate to preserve the overall charge distribution. Docking simulations are performed within the docking module of ROSETTA (20,21) using the two-stage protocol. First, sampling is done in the centroid mode, and the smaller movements in the second stage are considered in full atom mode. The docking protocol assumes a fixed backbone and only allows for translation, rotation, and side-chain packing. An additional constraint was introduced to bias that the orientation of the glutamate is within the vicinity of the substrate channel. A total of 6000 trial structures are obtained using three different starting orientations for the ACP with respect to Δ^9D and Δ^4D. The first initial structure was taken from the Δ^9D, the second was from the Δ^4D crystal structure, and the third was randomly generated.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2020.12.036.

AUTHOR CONTRIBUTIONS

M.D.B. designed research, performed simulations, analyzed data, and wrote the manuscript. J.S. designed research, secured funding, and wrote the manuscript. S.R. designed research, secured funding, analyzed data, and wrote the manuscript.

ACKNOWLEDGMENTS

This work was supported by the Physical Biosciences Program, Photochemistry and Biochemistry group within the U.S. Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences. Pacific Northwest National Laboratory is operated by Battelle for DOE under Contract DE-AC05-76RL01830. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility.

REFERENCES

1. Cronan, J. E., and J. Thomas. 2009. Bacterial fatty acid synthesis and its relationships with polyketide synthetic pathways. *In* Methods in Enzymology: Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates, pp. 395–433.
2. Mercer, A. C., and M. D. Burkart. 2007. The ubiquitous carrier protein–a window to metabolite biosynthesis. *Nat. Prod. Rep.* 24:750–773.
3. Finzel, K., D. J. Lee, and M. D. Burkart. 2015. Using modern tools to probe the structure-function relationship of fatty acid synthases. *Chembiochem.* 16:528–547.
4. Crosby, J., and M. P. Crump. 2012. The structural role of the carrier protein–active controller or passive carrier. *Nat. Prod. Rep.* 29:1111–1137.
5. Roujeinikova, A., W. J. Simon, …, A. R. Slabas. 2007. Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. *J. Mol. Biol.* 365:135–145.
6. Barnwal, R. P., M. Kaur, …, G. Varani. 2020. Comparative structure, dynamics and evolution of acyl-carrier proteins from Borrelia burgdorferi, Brucella melitensis and *Rickettsia prowazekii*. *Biochem. J.* 477:491–508.
7. Lee, Y., A. Jang, …, W. C. Lee. 2020. Structural characterization of an ACP from Thermotoga maritima: insights into hyperthermal adaptation. *Int. J. Mol. Sci.* 21:2600.
8. Zornetzer, G. A., J. Tanem, …, J. L. Markley. 2010. The length of the bound fatty acid influences the dynamics of the acyl carrier protein and the stability of the thioester bond. *Biochemistry.* 49:470–477.
9. Zornetzer, G. A., B. G. Fox, and J. L. Markley. 2006. Solution structures of spinach acyl carrier protein with decanoate and stearate. *Biochemistry.* 45:5217–5227.
10. Zhang, Y. M., H. Marrakchi, …, C. O. Rock. 2003. The application of computational methods to explore the diversity and structure of bacterial fatty acid synthase. *J. Lipid Res.* 44:1–10.
11. Suh, M. C., D. J. Schultz, and J. B. Ohlrogge. 1999. Isoforms of acyl carrier protein involved in seed-specific fatty acid synthesis. *Plant J.* 17:679–688.
12. Guy, J. E., E. Whittle, …, J. Shanklin. 2011. Remote control of regioselectivity in acyl-acyl carrier protein-desaturases. *Proc. Natl. Acad. Sci. USA.* 108:16594–16599.
13. Sztain, T., A. Patel, …, M. D. Burkart. 2019. Modifying the Thioester linkage affects the structure of the acyl carrier protein. *Angew. Chem. Int.Engl.* 58:10888–10892.
14. Chan, D. I., T. Stockner, …, H. J. Vogel. 2008. Molecular dynamics simulations of the Apo-, Holo-, and acyl-forms of Escherichia coli acyl carrier protein. *J. Biol. Chem.* 283:33620–33629.

15. Farmer, R., C. M. Thomas, and P. J. Winn. 2019. Structure, function and dynamics in acyl carrier proteins. *PLoS One.* 14:e0219435.

16. Kästner, J. 2011. Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 1:932–942.

17. David, C. C., and D. J. Jacobs. 2014. Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol. Biol.* 1084:193–226.

18. Remsing, R. C., and J. D. Weeks. 2015. Hydrophobicity scaling of aqueous interfaces by an electrostatic mapping. *J. Phys. Chem. B.* 119:9268–9277.

19. Guy, J. E., E. Whittle, …, J. Shanklin. 2007. The crystal structure of the ivy Delta4-16:0-ACP desaturase reveals structural details of the oxidized active site and potential determinants of regioselectivity. *J. Biol. Chem.* 282:19863–19871.

20. Davis, I. W., and D. Baker. 2009. RosettaLigand docking with full ligand and receptor flexibility. *J. Mol. Biol.* 385:381–392.

21. Leaver-Fay, A., M. Tyka, …, P. Bradley. 2011. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol.* 487:545–574.

22. Daura, X., K. Gademann, …, A. E. Mark. 1999. Peptide folding: when simulation meets experiment. *Angew. Chem. Int. Ed.* 38:236–240.

23. Shanklin, J., J. E. Guy, …, Y. Lindqvist. 2009. Desaturases: emerging models for understanding functional diversification of diiron-containing enzymes. *J. Biol. Chem.* 284:18559–18563.

24. Shanklin, J., and E. B. Cahoon. 1998. Desaturation and related modifications of fatty Acids1. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49:611–641.

25. Cahoon, E. B., Y. Lindqvist, …, J. Shanklin. 1997. Redesign of soluble fatty acid desaturases from plants for altered substrate specificity and double bond position. *Proc. Natl. Acad. Sci. USA.* 94:4872–4877.

26. Pronk, S., S. Pall, …, R. Apostolov. 2013. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics.* 29:845–854.

27. Robertson, M. J., J. Tirado-Rives, and W. L. Jorgensen. 2015. Improved peptide and protein torsional energetics with the OPLSAA force field. *J. Chem. Theory. Comput.* 11:3499–3509.

28. Gillan, M. J., D. Alfè, and A. Michaelides. 2016. Perspective: how good is DFT for water? *J. Chem. Phys.* 144:130901.

29. Cornell, W. D., P. Cieplak, …, P. A. Kollman. 2002. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J. Am. Chem. Soc.* 115:9620–9631.

30. Hess, B., C. Kutzner, …, E. Lindahl. 2008. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4:435–447.

31. Hub, J. S., B. L. de Groot, and D. van der Spoel. 2010. g_wham—a free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J. Chem. Theory Comput.* 6:3713–3720.

32. Hunter, J. D. 2007. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* 9:90–95.

33. Tribello, G. A., M. Bonomi, …, G. Bussi. 2014. PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun.* 185:604–613.