Structural insights into catalytic mechanism and product delivery of cyanobacterial acyl-acyl carrier protein reductase

Yu Gao1,5, Hongmei Zhang1,5, Minrui Fan1,3, Chenjun Jia1,4, Lifang Shi1, Xiaowei Pan1, Peng Cao1, Xuelin Zhao1, Wenrui Chang1,2 & Mei Li1

Long-chain alk(a/e)nones represent the major constituents of conventional transportation fuels. Biosynthesis of alkanes is ubiquitous in many kinds of organisms. Cyanobacteria possess two enzymes, acyl-acyl carrier protein (acyl-ACP) reductase (AAR) and aldehyde-deformylating oxygenase (ADO), which function in a two-step alkane biosynthesis pathway. These two enzymes act in series and possibly form a complex that efficiently converts long chain fatty acyl-ACP/fatty acyl-CoA into hydrocarbon. While the structure of ADO has been previously described, structures of both AAR and AAR-ADO complex have not been solved, preventing deeper understanding of this pathway. Here, we report a ligand-free AAR structure, and three AAR-ADO complex structures in which AARs bind various ligands. Our results reveal the binding pattern of AAR with its substrate/cofactor, and suggest a potential aldehyde-transferring channel from AAR to ADO. Based on our structural and biochemical data, we proposed a model for the complete catalytic cycle of AAR.
The demands for petrol-based fuels has remained high, calling for alternative fuel sources. Biofuels are of particular interest as a renewable supplement for gradually diminishing fossil fuels. Long-chain alk(a/e)nes are the major constituents of conventional transportation fuels. Biosynthesis of alkanes represents a promising way to produce substitutes for petroleum-based fuels. Long-chain alk(a/e)nes are commonly produced from fatty acid metabolites in various species. The earlier discovery of the pathway responsible for alkane biosynthesis in cyanobacteria allowed the conversion from solar energy to alkanes through photosynthesis, without providing external carbon sources. This two-step pathway utilizes two water-soluble enzymes, namely acyl-acyl carrier protein (acyl-ACP) reductase (AAR) and aldehyde-deformylating oxygenase (ADO). AAR catalyzes the first step of this pathway, reducing the substrate fatty acyl-ACP/fatty acyl-coenzyme A (acyl-CoA) into the corresponding aldehyde in a NADPH-dependent manner. Subsequently, ADO catalyzes the conversion of fatty aldehyde into alk(a/e)ne with one carbon shorter (Cn-1 hydrocarbon) and the for-mate as a Cn-derived form.

However, this pathway yields little alkane products, primarily because of the slow turnover of the two enzymes. Thus, structure-based engineering of the two enzymes is a practical way to optimize the efficiency of alkane synthesis. Several crystal structures of ADO from different species of cyanobacteria were previously solved, revealing that ADO adopts a ferritin-like eight-helix (henceforth named helix 1–8) architecture, with a di-iron center buried inside the molecule. An occluded substrate binding cavity accessing the di-iron center was identified in these structures. In addition, a hydrophobic tunnel featuring an opening between helix 7 and helix 8 was suggested to be the substrate entering tunnel based on the structures of ADO from Prochlorococcus marinus (PmADO). Together, the substrate cavity and the entering tunnel form a "T-shaped" channel inside ADO. ADO is considered to be the rate-limiting enzyme in this pathway. Based on the available structural information, a number of ADO mutants were generated and shown to possess higher activity compared with the wild type.

The catalytic reaction of AAR was shown to occur in two steps. Firstly, the substrate (fatty acyl-ACP or fatty acyl-CoA) was found to be associated with AAR, with the fatty acyl moiety covalently bound with the residue C294 of AAR (ref. [18]). As a result, an acyl-enzyme intermediate is formed, and the ACP or CoA moieties are released. Subsequently, a corresponding aldehyde is produced from the acyl-enzyme intermediate in the presence of NADPH. AAR was previously shown to prefer stearoyl-CoA among acyl-CoA esters of different length and saturation. It was shown that the highest activity toward stearoyl-ACP is obtained at 25°C, and subsequently it was shown to exhibit activity toward stearoyl-ACP.

Previously reported results suggested that AAR utilizes a "Ping-pong" mechanism, binding fatty acyl substrate prior to binding the cofactor NADPH during the catalysis. To date, no structure of either AAR alone or AAR in complex with substrate or NADPH has been solved. In addition, due to the poor solubility of long chain aldehydes, it had remained unclear how are the aldehydes transferred between two water-soluble enzymes in the hydrophilic environment. A recent report suggested that together, AAR and ADO forms a tight complex that, by facilitating the aldehyde transferring from AAR to ADO either in a direct or an indirect way, increases the efficiency of the alkane biosynthesis. However, due to the current lack of the complex structure, the precise details of the pathway of aldehyde transferring from AAR to ADO have remained undetermined.

We report the crystal structure of AAR from Synechococcus elongatus PCC 7942 (SeAAR) in its apo form (AARapo), together with three structures of AAR in complex with ADO from the same species of cyanobacteria (SeADO), in which AAR binds either substrate or substrate/cofactor. Our results provide details about the binding mode of AAR with ligands, the interaction between AAR and ADO, as well as the potential aldehyde-transferring pathway between the two enzymes. AAR in the four structures likely represents different state during its catalytic process, and our results shed light on the catalytic mechanism of AAR.

Results

The AARapo structure represents the resting state of AAR. In this work, we solved the AARapo structure at 2.8 Å resolution (Table 1), and modeled the complete AAR protein containing 341 amino acid residues. As shown in Fig. 1a, AAR is composed of three domains: the N-terminal domain (NTD, residues 1–130), the middle domain (mid-domain, residues 131–264), and the C-terminal domain (CTD, residues 265–341). A conserved loop, in the mid-domain was designated as the dinucleotide recognition loop. In addition, the classical sequence motif GX(1–3)GX(1–2)G for binding NAD(P)⁺ or NAD(P)H (Supplementary Fig. 2). Residue C294, which is strictly conserved in AAR homologues and has been identified as the key residue for catalysis, is located approximately at the center of AAR molecule. Above C294, a wide cleft that includes the dinucleotide recognition loop is visible (Fig. 1a). This structural observation suggests that the cleft constitutes the binding site for NADPH and the entrance for substrate access to the catalytic residue C294.

In the AARapo structure, we built five molecules, which show highly similar overall architectures, in an asymmetric unit. However, three fragments, namely Fr-I (residues R19–Q29) and Fr-II (S113–T120) from the NTD, and Fr-III (V221–D235) from the mid-domain, show higher structural diversity (Supplementary Fig. 3a), indicating that these regions possess an intrinsic flexibility. In fact, we found that the mobility of these fragments is pivotal for the function of AAR, as they are involved in either the interaction with ADO (Fr-I and Fr-II), or NADPH binding (Fr-III), as discussed later.

The AAR–ADO complex and the intermolecular interactions. We determined three ternary complex structures of AAR–ligand bound with ADO, which are similar in their overall architecture (Supplementary Fig. 3b). We named these structures based on which ligands associated with AAR, i.e., AARligand–ADO (Table 1). In these complex structures, ADO exhibits a nearly identical conformation to that observed in the previously reported structures of SeADO alone. ADO adopts an eight-helix folding, and contains two iron atoms and a hydrocarbon chain bound at the substrate cavity (Supplementary Fig. 3b). While helix 7 of ADO interacts with the NTD of AAR (Fig. 1b, c), the acidic residues of this helix form strong electrostatic interactions with the basic residues from the long helical region (R73–H91), as well as R118 from Fr-II of AAR (Fig. 1c, Supplementary Fig. 3c). These residues are highly conserved in homologues of AAR (Supplementary Fig. 2) and ADO, highlighting their crucial roles in forming the complex.

To further verify the AAR–ADO interface observed in the complex structures, we mutated each of the five acidic residues from helix 7 of ADO, and further generated a double mutant E200A/D201A and triple mutant E196A/E200A/D201A (Supplementary Fig. 4). We then measured the binding affinities of AAR with the wild type and the seven mutant forms of ADO using isothermal titration calorimetry (ITC; Table 2, Supplementary Fig. 5). The binding assays revealed that the wild-type forms of AAR and ADO formed a complex with a measured K_D value of...
Table 1 Summary of data collection and structure refinement.

| PDB code | AAR-His (Hg-derivative) | AARapo | AARhydro−ADO | AARNADPH−ADO | AARapothy−CoA−ADO |
|----------|------------------------|--------|--------------|--------------|-------------------|
|          | 6JZQ                   | 6JZU   | 6JZY         | 6JZZ         |                   |
| Data collection |                    |        |              |              |                   |
| Space group | C2                  | C2     | P2, 2, 1     | P2, 2, 1     | P2, 2, 1          |
| Cell dimensions | a, b, c (Å)  | 56.7, 180.9, 90.3  | 92.2, 201.4, 124.6 | 66.4, 71.3, 137.1 | 65.6, 71.3, 135.7 |
| Resolution (Å) | 50.0 – 2.30 (2.38 – 2.30) | 0.08 (0.49) | 0.086 (1.08) | 48.59 – 2.18 (2.26 – 2.18) | 48.26 – 2.10 (2.17 – 2.10) |
| Completeness (%) | 99.9 (99.9) | 99.5 (99.2) | 99.6 (99.3) | 99.1 (99.0) | 99.9 (100) |
| Redundancy | 6.8 (6.9) | 7.6 (7.7) | 13.1 (13.1) | 12.7 (12.8) | 14.0 (14.4) |
| Rmerge | 0.08 (0.49) | 0.086 (1.08) | 0.08 (0.99) | 0.08 (0.99) | 0.089 (>1) |
| I/σI | 22.6 (4.2) | 23.6 (2.8) | 30 (2.8) | 32 (2.4) | 27.4 (2.5) |
| Resolution (Å) | 50.0 – 2.30 (2.38 – 2.30) | 53,258 (5159) | 34,321 (3189) | 37,568 (3434) | 38,668 (3121) |
| No. reflections | 46.11 – 2.80 (2.90 – 2.80) | 49.43 – 2.18 (2.26 – 2.18) | 49.13 – 2.10 (2.17 – 2.10) | 38.66 – 3.01 (3.12 – 3.01) | 12,814 (721) |
| Rwork/Rfree | 0.22/0.25 | 0.21/0.24 | 0.21/0.27 | 0.22/0.25 | 0.21/0.27 |
| No. atoms | Protein | 12948 | 38 | 49 | 74.23 |
|           | Ligand/ion | 4338 | 36.1 | 163 | 36.41 |
|           | Water | 38 | 57.93 | 55.02 | 57.93 |
| B-factors | Protein | 74.23 | 36.41 | 57.93 | 50.30 |
|           | Ligand/ion | 39.39 | 55.02 | 47.74 | 47.74 |
|           | Water | 75 | 36.99 | 36.99 | 36.99 |
| R.m.s. deviations | Bond lengths (Å) | 0.012 | 0.008 | 0.012 | 0.012 |
|           | Bond angles (°) | 1.45 | 0.89 | 1.36 | 1.36 |

Values in parentheses are for the highest-resolution shell.

**Fig. 1** Structures of AAR and AAR-ADO complex. **a** Cartoon (left) and surface (right) representation of AARapo structure. NTD, mid-domain, and CTD of AAR are shown in slate, yellow and magenta, respectively. The dinucleotide recognition loop in the mid-domain is highlighted in green. The key residue C294 is shown in stick-ball mode and labeled. The cleft above the residue C294 was indicated by an arrow in the right panel. **b** Cartoon representation of AAR-ADO complex structure (shown by AARNADPH−ADO structure). NTD, mid-domain, and CTD of AAR are shown in slate, yellow and magenta, respectively. ADO is shown in cyan and the eight helices are labeled as 1–8. The two iron atoms in ADO are shown in orange spheres. The residues involved in the intermolecular interaction are shown in sticks. The salt bridges and hydrogen bonds are shown by black dashed lines. The bound ligands are omitted for clarity. **c** Details of the intermolecular interaction between AAR and ADO. For clarity, only fragments of AAR (R73–H91 and L114–L122) and ADO (helix 7) are shown. The residues involved in the intermolecular interactions are shown in sticks and labeled.
2.0–2.2 μM, which is comparable to that of the similar AAR–ADO complex from *Nostoc punctiforme* (Np)\(^{18}\). The mutants E196A and E200A exhibit significantly reduced binding affinity with AAR, whereas the E200A/D201A and E196A/E200A/D201A mutations completely abolished the capability of ADO to bind with AAR (Table 2, Supplementary Fig. 5). We further mutated the basic residues of AAR involved in the interactions with E196 and E200, creating three AAR single mutants, R73A, R79A, and K80A (Supplementary Fig. 4). The binding assay between AAR (wild type and mutants) and ADO confirmed the essential role of the R73(AAR)–E196(ADO), R79(AAR)–E196(ADO), and K80(AAR)–E200(ADO) pairs for the AAR–ADO complex formation (Table 2, Supplementary Fig. 6). Together, our results demonstrate that the AAR–ADO complex observed in our structure is representative of the complex that exists in physiological conditions.

### The AARthioester–ADO structure
During catalysis, AAR first binds its substrate acyl-CoA or acyl-ACP to form an acyl-enzyme intermediate with a covalent thioester bond between the acyl moiety and C294 of AAR. In our AARthioester–ADO structure solved at 2.18 Å resolution, we observed a long tube-shaped electron density, which can be well fitted with a stearic chain, within an interior tunnel of AAR (Fig. 2a, b, Supplementary Fig. 7a). No electron density corresponding to the CoA moiety was observed, despite the incubation of stearoyl-CoA with AAR prior to crystallization. This finding indicates that the acyl-enzyme intermediate was spontaneously formed, and the CoA moiety had already been released. The structure showed that the C1 atom of the stearoyl chain (stearoyl-C1) had formed a covalent thioester bond with the hydrosulfide group of C294 (C294-SH; Fig. 2b, Supplementary Fig. 7a). Mass spectrometric analysis of the stearoyl-CoA-incubated AAR sample that was used for crystallization confirmed the thioester bond formation (Supplementary Fig. 7b), thus the AARthioester–ADO structure represents the acyl-enzyme intermediate state of AAR. In the AARthioester–ADO structure, the stearic chain is buried inside a hydrophobic tunnel (acyl-tunnel) formed mainly by non-polar residues from the NTD and CTD of AAR (Fig. 2c). One end of the acyl-tunnel connects with the wide cleft above C294, forming an L-shaped, long channel (Fig. 2b). At the other end, the acyl-tunnel can be seen to extend toward the AAR–ADO interface with its opening

### Table 2 Binding affinity between AAR and ADO measured through ITC method.

| Binding with AAR | ADO wild type | 2.0 ± 0.50 | N.D. | 4.9 ± 1.40 | 2.8 ± 0.85 | N.D. | 14.8 ± 3.70 | 4.9 ± 1.40 | 2.8 ± 0.85 |
|------------------|---------------|------------|-----|------------|-----------|-----|--------------|------------|-----------|
|                  | ADO E196A     | N.D.       |     | N.D.       | N.D.      |     | N.D.         | N.D.       | N.D.      |
|                  | ADO E200A     | N.D.       |     | N.D.       | N.D.      |     | N.D.         | N.D.       | N.D.      |
|                  | ADO D201A     | 4.9 ± 1.40 |     | N.D.       | N.D.      |     | N.D.         | N.D.       | N.D.      |
|                  | ADO E208A     | 14.8 ± 3.70|     | N.D.       | N.D.      |     | N.D.         | N.D.       | N.D.      |
|                  | ADO E211A     | 2.8 ± 0.85 |     | N.D.       | N.D.      |     | N.D.         | N.D.       | N.D.      |
|                  | ADO E200A/D201A| N.D.       |     | N.D.       | N.D.      |     | N.D.         | N.D.       | N.D.      |
|                  | ADO E196A/E200A/D201A | N.D. |     | N.D.       | N.D.      |     | N.D.         | N.D.       | N.D.      |

| Binding with ADO | AAR wild type | 2.2 ± 0.30 | N.D. | N.D. | N.D. | N.D. |
|------------------|---------------|------------|-----|-----|-----|-----|
|                  | AAR Y247F    | 2.2 ± 0.40 | N.D. | N.D. | N.D. | N.D. |
|                  | AAR R73A     | N.D.       | N.D. | N.D. | N.D. | N.D. |
|                  | AAR R79A     | N.D.       | N.D. | N.D. | N.D. | N.D. |
|                  | AAR K80A     | N.D.       | N.D. | N.D. | N.D. | N.D. |
|                  | AAR Y247A    | N.D.       | N.D. | N.D. | N.D. | N.D. |

*N.D. not detected.*

---

**Fig. 2 AARthioester–ADO structure.**

- **a** Chemical structure of the substrate stearoyl-CoA. The red dashed line indicates the thioester bond between stearic chain and CoA moiety.
- **b** Surface representation of AARthioester–ADO structure. The L-shaped channel within AAR is indicated. The stearic chain (shown as magenta stick), covalently bound with C294 (shown in stick-ball mode), is located in the hydrophobic tunnel, which is a part of the L-shaped channel. Residues Y23 and Y26 from AAR, and M221 from ADO, forming the terminal end of the tunnel, are shown in stick and labeled. The color code of AAR and ADO is the same as that in Fig. 1b. c The hydrophobic residues, mainly from the NTD and the CTD, are involved in the substrate tunnel formation and shown as sticks. The color code of AAR is the same as that in Fig. 1b, and ADO is shown in green.
surrounded by residues Y23 and Y26 from the NTD of AAR and sealed by M221 of ADO (Fig. 2b). In the AARapo structure, this acyl-tunnel is broken by the side chains of M279 and M21 (Supplementary Fig. 7c). Both M279 and M21 undergo significant conformational changes in the AARSkoester–ADO structure, where they participate in forming the acyl-tunnel. Together, these findings indicate that substrate binding induces formation of the acyl-tunnel in AAR, as a result of that the stearic chain is well accommodated in the tunnel.

The AARNADPH–ADO structure. The molecule NADPH is an essential cofactor for catalysis, as it provides a hydride to the thioester bond. The stearic chain is located in the acyl-tunnel of AAR, as was observed in the AARSkoester–ADO structure. However, the weaker electron density of the stearic chain in the AARNADPH–ADO structure suggests that it only partially occupies the acyl-tunnel. Previous studies reported that AAR requires divalent metal ions (Mg2+) for catalysis, and that its activity is stimulated 10–20-fold by monovalent metal ions, including K+ ions. In the present work, ions such as Mg2+ and K+ were always provided in the purification and crystallization buffers of AAR. The introduction of NADPH should enable AAR to proceed with catalysis in the presence of cation ions. This process may have occurred slowly within the crystals, as a result of that aldehyde was produced and released in a proportion of the AAR molecules, while the thioester bond was preserved in other AAR molecules in the crystals. Together, these results suggest that the AARNADPH–ADO structure represents the catalytic state of AAR.

Table 3 Binding affinity and enzymatic activity of the wild type and mutants of AAR toward stearoyl-CoA.

| Protein         | Binding assay $K_d$ value (µM) | Activity assay Octadecanal yield (µg ml⁻¹) |
|-----------------|--------------------------------|--------------------------------------------|
| AAR-wt (−K⁺)    | 2.3 ± 0.38                     | 6.0 ± 0.76                                 |
| AAR-wt (+K⁺)    | 0.1 ± 0.05                     | 110.9 ± 41.25                              |
| AAR-Y247F (+K⁺) | 0.5 ± 0.15                     | 61.1 ± 6.76                                |
| AAR-Y247A (+K⁺) | N.D.                           | -                                          |

$^a$ The binding affinity of wild type (wt) or mutants of AAR with the substrate stearoyl-CoA measured by ITC.

$^b$ $K^+$ and (+K⁺) mean that the assays were measured in the absence and presence of potassium ion, respectively.

$^c$ N.D. not detected.

Source data are provided as a Source Data file.

In the AARNADPH–ADO structure, the nicotinamide moiety of NADPH is located close to C294 of AAR, with its C4 atom being located only 3.14 Å away from the stearoyl-C1 atom (Fig. 3b). The short distance supports the possibility that C4 of NADPH provides a hydride to break the thioester bond. The stearic chain is located in the acyl-tunnel of AAR, as was observed in the AARSkoester–ADO structure. However, the weaker electron density of the stearic chain in the AARNADPH–ADO structure suggests that it only partially occupies the acyl-tunnel. Previous studies reported that AAR requires divalent metal ions (Mg2+) for catalysis, and that its activity is stimulated 10–20-fold by monovalent metal ions, including K+ ions. In the present work, ions such as Mg2+ and K+ were always provided in the purification and crystallization buffers of AAR. The introduction of NADPH should enable AAR to proceed with catalysis in the presence of cation ions. This process may have occurred slowly within the crystals, as a result of that aldehyde was produced and released in a proportion of the AAR molecules, while the thioester bond was preserved in other AAR molecules in the crystals. Together, these results suggest that the AARNADPH–ADO structure represents the catalytic state of AAR.
same ADP moiety, with the ribose phosphate group located at a slightly different position (Supplementary Fig. 10c). Our structures showed that their ADP parts are well superposed. The 3′-ribose phosphate group of CoA almost overlapped on the 2′-ribose phosphate of NADPH (Supplementary Fig. 10c), therefore is able to interact with the three positively charged residues of AAR. These results suggested that NADPH and CoA occupy the same site within AAR. In agreement with our structural observation, our binding assay showed that the wild-type AAR binds NADPH with a $K_D$ value of $\sim 7.99 \, \mu M$, whereas the C294S mutant is unable to bind NADPH after incubated with stearoyl-CoA (Supplementary Fig. 11), suggesting the unreleased stearoyl-CoA occupies the NADPH binding site. Our results are in line with previously suggested "ping-pong" catalytic mechanism employed by AAR\textsuperscript{15,18}.

It is noteworthy that stearoyl-CoA adopts an extended conformation, resulting in a shift of the aliphatic chain of $\sim 6 \, \AA$ toward the AAR–ADO interface compared with that in other two AAR–ADO complex structures described above. The main cause of this shift is that the C294S mutant is not able to form a thioester bond, and thus fails to fix the stearic chain at the active site of AAR (Fig. 4a). This extended conformation of stearoyl-CoA induces movement of the Fr-I region of AAR (Supplementary Fig. 12a). Thus, Y23 and Y26, which form the terminal end of the acyl-tunnel in the AAR\textsubscript{stearoyl-CoA}–ADO structure, are distant from the stearic moiety in the AAR\textsubscript{stearoyl-CoA}–ADO structure. These changes result in a wider opening at the surface of the AAR molecule (Supplementary Fig. 12b). Furthermore, M21, Y23, and Y26 of AAR form additional hydrogen bonds with R218 and M221 from helix 8 of ADO (Supplementary Fig. 12c), and the conformation of M221 of ADO alters so that it no longer seals the acyl-tunnel of AAR as it does in the AAR\textsubscript{stearoyl-CoA}–ADO structure (Supplementary Fig. 12b). Together, these structural changes facilitate the release of produced aldehydes from AAR, suggesting that the AAR\textsubscript{stearoyl-CoA}–ADO structure mimics the aldehyde-releasing state of AAR, in which the free aldehyde is released, ready to be transferred to ADO.

A potential aldehyde-transferring channel from AAR to ADO. Previous studies\textsuperscript{18}, together with our enzymatic assay, demonstrated that both AAR and ADO alone are functional (Table 3). In addition, we found that when both incubated with AAR, the wild-type ADO yields two times more product (heptadecane) than the triple mutant (E196A/E200A/D201A) of ADO (Supplementary Fig. 13), which is unable to form the AAR–ADO complex (Table 2, Supplementary Fig. 5). These results together indicate that the formation of AAR–ADO complex is to facilitate the direct transfer of the aldehyde product from AAR to ADO, which increases the efficiency of alkane production in this pathway. Direct transfer would require a contiguous channel from AAR to ADO. In our AAR\textsubscript{ligand}–ADO structures, the acyl-tunnel of AAR has an opening at the surface of AAR and toward AAR–ADO interface (Supplementary Fig. 12b), suggesting the release direction of the aldehyde. However, in these structures, ADO exhibits an occluded conformation, therefore, we were not able to identify the entry point for the aldehyde to access the active site of ADO based on our AAR–ADO complex structures. Previously reported crystal structures of PmADO showed that a T-shaped channel is present inside ADO, with an opening between helices 7 and 8 at the ADO surface, which was suggested to represent the entry point for aldehyde\textsuperscript{20}. The helices 7 and 8 of
ADO are at the AAR–ADO interface and helix 7 is essential for the complex formation. On the basis of these results, we superposed the PmADO structure (PDB code: 4PGI) with the ADO part of our AARstearoyl-CoA–ADO structure, and found that the two openings of both SeAAR and PmADO face each other (Fig. 4b), thus the acyl-tunnel of AAR and the T-shaped channel inside ADO are linked, forming a long contiguous channel in the superposed model. The opening at ADO surface leads to the entering tunnel, which corresponds to one half of the horizontal arm of the T-shaped channel. The other half of the horizontal arm, together with the vertical arm, forms the substrate cavity, which was observed in several ADO structures solved in both the present study, as well as in earlier reports19,21,22 (Fig. 4c). Based on these structural observations, we propose a model for the process of transferring aldehyde from AAR to ADO. The alkyl group enters the T-shaped channel of ADO, as a result of which the entire aldehyde molecule passes through the entering tunnel, before it reaches the distal end of the whole horizontal arm of the T-shaped channel. The aldehyde group proceeds toward the inside the horizontal arm after the transfer is completed, and upon the bending of the alkyl chain in the middle, accesses the di-iron center of ADO (Fig. 5).

R191 switches its rotamer and as a result, opens the entry of the entering tunnel (PDB code: 4PGI; Supplementary Fig. 14a). When we superposed the structure of PmADO (PDB code: 4PGI) on SeADO in our complex structure, we found that W178 in SeADO and R191 in PmADO are located at the same positions, with their main chain atoms and Cβ atoms well superposed (Supplementary Fig. 14b). Our structural comparison indicated that W178 could easily change its rotamer to swing its indole ring away, thus opening the entry without inducing large movement of the main chain atoms. Together, these results suggest that W178 (or R191 of PmADO) functions as a gate residue; it may shield the hydrophobic entering tunnel of ADO from the aqueous environment, whereas the approaching aldehyde produced by AAR may result in its sidechain swinging away, leading to the opening of the entry at the ADO surface. These conformational changes result in the formation of a long hydrophobic channel through which the aldehyde could be transferred from AAR to ADO.

Our structures provide strong evidence that the interactions between the NTD of AAR and the C-terminal region of ADO are essential for the complex formation and the aldehyde transfer. This is consistent with the previous analysis showing that introducing the ADO–AAR fusion protein (C-terminus of ADO linked with N-terminus of AAR) into Escherichia coli increased its alkane yield by 4.8-fold compared with the strain expressing both AAR and ADO, but separately8. The fused protein may
Structural variations of AAR in active and resting states. Superposition of the AARapo structure onto the AAR part of the AAR–ADO complex structures showed that the Fr-II region, which is flexible in the AARapo structure (Supplementary Fig. 3a), is stabilized by the hydrogen bond between R118 in Fr-II and D201 of ADO (Fig. 1c). In addition, three other regions of AAR, namely Fr-I, Fr-III, and Fr-IV (Q277–M284), exhibit dramatic conformational changes between the apo- and ligand-bound states (Supplementary Fig. 15a). The N-terminal segment (H7–A37) containing Fr-I is more loosely packed, potentially hindering AAR–ADO complex formation in the AARapo structure. Nevertheless, it adopts a conformation that facilitates intermolecular interactions in the AAR–ADO complex structures (Supplementary Fig. 15a). The Fr-III region (W221–T235) in the mid-domain, which is involved in NADPH binding, swings toward the central cleft in the AAR–ADO complex structures, resulting in a better fit of NADPH in the cleft (Supplementary Fig. 15b). Moreover, a short fragment containing M279 in the CTD (Fr-IV), which potentially clashes with the stearic chain in the AARapo structure (Supplementary Fig. 15c), changes its conformation to shape the acyl-tunnel in the AAR–ADO complex structures (Supplementary Figs. 7c and 15c). The three fragments tend to move toward the active center, creating a more compact architecture of AAR within the AAR–ADO complex structures (Supplementary Fig. 15a). The compact folding allows AAR to better accommodate the substrate/NADPH, therefore AAR in the AAR–ADO complex structures is likely to adopt an active conformation, which allows the catalytic reaction to proceed, whereas the AARapo structure represents the resting state. Furthermore, our structures suggest that the active form (the ligand-bound form) is more favorable for AAR to bind ADO, compared with the loosely packed apo form of AAR.

In addition, our structural analysis suggested a pivotal role of Y247 for maintaining AAR in a compact conformation. It forms a hydrogen bond with H7 from the NTD and hydrophobic interactions, with aromatic residues from both CTD and NTD of AAR (Supplementary Fig. 15d), thus drawing the mid-domain, CTD, and NTD together. In addition, Y247 may also be involved in the catalytic process, as it participates in the formation of the acyl-tunnel via its benzene ring and stabilizes NADPH via its carbonyl group (Fig. 3b). In agreement with these structural observations, the Y247F mutant of AAR exhibited reduced binding affinity with stearoyl-CoA, and maintained only ~50% of the catalytic activity of the wild type (Table 3). Furthermore, Y247A completely failed to bind stearoyl-CoA (Table 3, Supplementary Fig. 9). Interestingly, we found that Y247A mutant also loses the capability to bind ADO, while the binding affinity of Y247F mutant with ADO is almost the same as that of AAR wild type (Table 2, Supplementary Fig. 6). This observation suggests that mutation of the Tyr residue to Ala induces the structural or conformational change of the whole AAR protein, resulting in the complete loss of its function.

Discussion
We solved the crystal structures of AAR and the AAR–ADO complex bound with several, structurally different ligands, thus greatly advancing the understanding of the detailed mechanisms of the catalytic cycle of AAR(–ADO) responsible for alkane production. Interestingly, we noticed that neither the AAR protein with ligands nor the AAR–ADO complex without ligands yielded crystals, implying that both types of binary complexes are flexible, and only the ternary complex (AAR–ligand–ADO) is sufficiently stable to yield crystals. These results suggest that AAR shows a synergistic effect of substrate binding and ADO association. In agreement with this observation, our structural data suggested that the compact folding of AAR in the complex is favorable for substrate/NADPH binding and ADO association, and suggest that the formation of AAR–ADO complex facilitates the catalysis of AAR by relieving its substrate/product inhibition.
the resting state exhibiting a loose conformation (S0), detach from ADO, and prepare for the next cycle. The allele-hydro in ADO moves further to occupy the substrate cavity with its head group close to the di-iron center of ADO (Fig. 5). Our suggested mechanism explains the increased alkane yield when providing ADO, especially in higher molar ratio to AAR (ref. 25), and highlights the functional importance of the reversible association and dissociation of AAR and ADO in the alkane biosynthesis pathway.

Methods

Protein expression and purification. The AAR gene (Genbank ID:CP000100.1) was synthesized (GeneScript, China) and cloned (Supplementary Tables 1 and 2) into the modified version of vector pET24a (Novagen), that encodes a tobacco etch virus (TEV) protease cleavage site with the sequence of ENLYQQL, followed by a downstream 21-residue peptide, including a His(6)-tag with the sequence of SQVSLALEHHHHHH. Upon verification of the correct sequence, the plasmid containing AAR gene was transformed into E. coli strain BL21(DE3) (TransGen Biotech) along with the plasmid pG-KJE8 (TaKaRa), which encodes five molecular chaperones of dnaK, dnaJ, grpE, groES, and groEL. The transformed cells were cultured in Luria–Bertani medium containing 50 µg mL^{-1} kanamycin, 1 µg mL^{-1} streptomycin, 2 µg mL^{-1} L-arnobinol, and 10 mM (NH4)2SO4. The cultures were grown overnight at 37 °C. Next day, 1% (v/v) of the culture was used for induction by 1 mM IPTG (isopropyl-β-D-1-thiogalactopyranoside) and continued shaking at ~200 rpm for overnight at 37 °C. The cells were harvested and disrupted by ultra-sonication in lysis buffer containing 50 mM Tris–HCl pH 8.0, 50 mM NaCl, 1 mM TCEP, G2-enzyme (2-carboxyethylphosphothioate)-modified chymotrypsinogen A, 0.1% Triton X-100, 15% glycerol, 1 mM PMSF (phenylmethylsulfonyl fluoride), 5 mM KCl and 2 mM MgCl2 and 10 mM imidazole. Cell debris was discarded after centrifugation at 38,000 × g for 30 min at 4 °C. The supernatant containing the expressed AAR protein with His-tag (AAR-His) was loaded onto a Ni affinity column (GE Healthcare) pre-equilibrated with lysis buffer. The AAR-His protein was then purified through size-exclusion chromatography with elution buffer of 10 mM HEPES (pH 7.5), 50 mM NaCl, 5 mM KCl, 2 mM MgSO4, and 1 mM TCEP. The protein was concentrated to ~30 mg mL^{-1} through measuring its absorption at 280 nm using a calculated extinction coefficient of 43.430 M^{-1} cm^{-1} (http://cda.expasy.org). SeADO gene (Genbank ID:CP000100.1) was constructed (Supplementary Tables 1 and 2) and expressed similarly as that of SeAAR gene, except that the chaperone-encoded plasmid pG-KJE8 (TaKaRa) was not used. The expression of ADO protein was equilibrated with 1 mM IPTG and 100 µM FeSO4, and IPTG was added to the 50% of the total concentration. The ADO protein was treated with TEV protease overnight at 4 °C, and then washed and eluted with the lysis buffer containing 50 mM and 300 mM imidazole, respectively. To obtain the His-tag free AAR protein, AAR-His eluted from the Ni affinity column was treated with TEV protease for 12 h at 4 °C, and then flown through the Ni affinity column. This treatment removes the C-terminal 22-residue peptide, including the His-tag, encoded by the vector, thus results in the AAR protein alone (AAR). Both AAR-His and AAR proteins were further purified through size-exclusion chromatography with elution of 10 mM HEPES (pH 7.5), 50 mM NaCl, 5 mM KCl, 2 mM MgSO4, and 1 mM TCEP. The protein was concentrated to ~30 mg mL^{-1} through measuring its absorption at 280 nm using a calculated extinction coefficient of 43.430 M^{-1} cm^{-1} (http://cda.expasy.org). The over-expressed AAR and ADO mutant proteins were purified using the same protocol as the wide-type AAR and ADO proteins, respectively. The purity of all the wild type and mutant forms of AAR and ADO are verified through SDS–PAGE.

ITC measurement. The binding affinity of AAR with ADO or stearyl-CoA was measured by isothermal titration calorimetry (ITC) using a MicroCal Auto-ITC200 microcalorimeter (GE Healthcare Bio-sciences, Pittsburgh, PA). To measure the affinity between AAR and ADO, all the protein samples were prepared in the same buffer containing 25 mM Tris–HCl, pH 7.5, and 1 mM TCEP. Wild type or mutants of ADO (or AAR) were placed into the sample cell (200 µL working volume) at a concentration of 1 µM and 4 µM, respectively, and centrifuged at 15,000 rpm for 10 min. The dilution buffer contains 20 mM Tris–HCl pH 7.5, 0.2 M KCl, 1 mM TCEP, and 0.5 mM MgCl2. A two-fold dilution series was prepared for ligand NADPH with dilution buffer. The experiment about ligand NADPH was accompanied by a step-by-step gradient dilution process. Subsequently, 10 µL of AAR protein with wild type or C294S mutant was mixed with ligand NADPH with different concentration. Samples were filled into hydrophilic capillaries (Monolith NT. labelfree capillary). The label free MST assay was performed with 30% excitation power and the high MST power controlled by the MO.Control Software Version 1.6.1. Kd, fit function of Mo. Affinity Analysis v3.2 was used to fit curve and calculate the value of dissociation constant (Kd).

Enzymatic assay. For measuring the enzymatic activity of AAR, the AAR protein at the concentration of 20 µM was mixed with 2 mM NADPH and 0.5 mM stearyl-CoA in a 100 µL reaction buffer containing 20 mM HEPES pH 7.5, 200 mM KCl, 0.5 mM MgCl2, and 0.5 mM TCEP. For measuring the enzymatic activity of AAR–ADO complex, additional 40 µM ADO, 150 µM ferredoxin, and 50 µM Fd-NADP+ oxidoreductase (FNR) were added to the AAR reaction buffer, with increased NADPH concentration of 4 mM. The ferredoxin and FNR are recombinant proteins from cyanobacterial strains Synechocystis sp. PCC 6803 and Synechocystis sp. PCC 7942, respectively, and were added into the reaction buffer for supplying electrons for the reduction of ADO. The mixture was incubated at 30 °C for 5 h. The produced octadecanal and heptadecane were extracted by addition of an equal volume of ethyl acetate (Sigma Aldrich, HPLC grade) and centrifuged at 13,000 × g for 15 min at room temperature. The extraction step was repeated for three times. The extracted supernatant was pooled together, dried with nitrogen gas, and then resolved in 20 µL ethyl acetate for the GC–MS measurement.

Preparation of the AAR–ADO complex. To get the AAR in complex with stearyl-CoA and ADO, AAR was first incubated with its substrate stearyl-CoA at the molar ratio of 1:3 for 30 min at 18 °C, and was further incubated with ADO protein at 1:1 molar ratio for additional 4 h. The unbound stearyl-CoA was removed using an AAR–ADO complex through gel filtration chromatography on Sephadex G200 10/500GL column, with elution buffer containing 10 mM HEPES (pH 7.2), 50 mM NaCl, 5 mM KCl, 2 mM MgSO4, and 1 mM TCEP. The eluted protein complex was characterized through SDS–PAGE and concentrated to 40 mg mL^{-1} for crystallization. The molecular weight of both the AAR protein alone and AAR protein covalently bound with stearyl-CoA were determined by mass spectrometry (HPLC-Q-TOF-MS).

Crystallization, structure determination, and refinement. Crystallization screens were carried out at 18 °C by mixing equal volume of protein and reservoir solution, using the sitting-drop vapor diffusion method. Crystals of AAR–His were grown in condition of 0.1 M Tris–HCl, pH 8.5, and 1.2 M NaAc. For phasing, the AAR–His crystals were soaked in mother liquor (0.1 M Tris–HCl, pH8.5, 200 mM NaAc, and 25% PEG3350) containing 10 mM ethyl mercury phosphate for 10 s. Crystals of AAR were harvested in reservoir solution of 0.1 M citric acid, pH6.0, and 1 M (NH4)2SO4. The crystals of AAR–ADO in complex with substrate (AARsubstrate–ADO and AARsubstrate–CoA–ADO) were harvested in reservoir solution containing 0.1 M HEPES, pH 7.5, and 20% PEG8000. To obtain the crystal of the AAR–ADO, the crystal of AARsubstrate–ADO was soaked with 1 M NADPH in reservoir solution overnight. Crystals were flash-cooled in liquid nitrogen at 100 K for data collection using 2.5 µm NaAc in mother liquor as cryo-protectant for AAR–His crystals, 20% glycerol in mother liquor as cryo-protectant for AARsubstrate crystals, and the reservoir solution as cryo-protectant for AARsubstrate–ADO crystals. Crystallographic data were processed and scaled with HKL–2000 package24. The initial phase of AAR–His

Microscale thermophoresis measurement. The binding affinity of AAR with NADPH was measured by microscale thermophoresis (MST) experiment, which was performed on a Monolith NT.labelfree system (NanoTemper Technologies, Germany). Proteins of wild type and C294S mutant of AAR were diluted at a stock concentration of 1 µM and 4 µM, respectively, and centrifuged at 15,000 rpm for 10 min. The dilution buffer contains 20 mM Tris–HCl pH 7.5, 0.2 M KCl, 1 mM TCEP, and 0.5 mM MgCl2. A two-fold dilution series was prepared for ligand NADPH with dilution buffer. The experiment about ligand NADPH was accompanied by a step-by-step gradient dilution process. Subsequently, 10 µL of AAR protein with wild type or C294S mutant was mixed with ligand NADPH with different concentration. Samples were filled into hydrophilic capillaries (Monolith NT. labelfree capillary). The label free MST assay was performed with 30% excitation power and the high MST power controlled by the MO.Control Software Version 1.6.1. Kd, fit function of Mo. Affinity Analysis v3.2 was used to fit curve and calculate the value of dissociation constant (Kd).
was obtained from the Hg-resistant crystals by single isomorphous replacement anomalous scattering method using AutoBuild in PHENIX package. In the AAR-His structure, the C-terminal peptide encoded by the vector forms an α-helix, which is participating the molecular packing in the crystals, leading to an artifact of α-carbon monoxide, as the stoichiometric coproduct in conversion of fatty aldehydes to alkanes by a cyanobacterial aldehyde decarbonylase. Proc. Natl Acad. Sci. USA 101, 514 – 518 (2004).

15. Lin, F., Das, D., Lin, X. N. & Marsh, E. N. Aldehyde-forming fatty acyl-CoA reductase from cyanobacteria, expression, purification and characterization of the recombinant enzyme. J. Biol. Chem. 280, 4773 – 4781 (2005).

16. Warui, D. M. et al. Detection of formate, rather than carbon monoxide, as the stoichiometric reaction of diiron enzymes. J. Am. Chem. Soc. 133, 3316 – 3319 (2011).

17. Das, D., Eser, B. E., Han, J., Sciore, A. & Marsh, E. N. Oxygen-independent decarbonylation of aldehydes by cyanobacterial aldehyde decarbonylase: a new reaction of diiron enzymes. Angew. Chem. Int. Ed. Engl. 50, 7148 – 7152 (2011).

18. Warui, D. M. et al. Efficient delivery of long-chain fatty aldehydes from the Nostoc punctiforme acyl-acetyl carrier protein reductase to its cognate aldehyde-deformylating oxygenase. Biochemistry 54, 1006 – 1015 (2015).

19. Khara, B. et al. Production of propane and other short-chain alkanes by structure-based engineering of ligand specificity in aldehyde-deformylating oxygenases. ChemBioChem 14, 1204 – 1208 (2013).

20. Buer, B. C., Paul, B., Das, D., Stuckey, J. A. & Marsh, E. N. Insights into substrate and metal binding from the crystal structure of cyanobacterial aldehyde deoxyformylating oxygenase with substrate bound. ACS Chem. Biol. 9, 2584 – 2593 (2014).

21. Demmer, U. et al. Structural insights into the catalytic mechanism of aldehyde-deformylating oxygenases. Protein Cell 6, 55 – 67 (2015).

22. Park, A. K. et al. Crystal structures of aldehyde deoxyformylating oxygenase from Limnothrix sp. KNUA012 and Oscillatoria sp. KNUA011. Biochem. Biophys. Res. Commun. 477, 395 – 400 (2016).

23. Wang, Q. et al. Identification of residues important for the activity of aldehyde-deformylating oxygenase through investigation into the structure-activity relationship. BMC Biotechnol. 17, 31 (2017).

24. Li, N. et al. Evidence for only oxygen-free cleavage of aldehydes to alkane/ene and formate by cyanobacterial aldehyde deoxyformylases. Biochemistry 51, 7908 – 7916 (2012).

25. Demmer, U. et al. Structural basis for a bispecific NADP+ and CoA binding site in an archaeal malonyl-coenzyme A reductase. J. Biol. Chem. 288, 6363 – 6370 (2013).

26. Constantinides, P. P. & Stein, J. M. Solubility of palmityl-coenzyme A in acyltransferase assay buffers containing magnesium ions. Arch. Biochem. Biophys. 250, 267 – 270 (1986).

27. Rahman, Z. et al. Enhanced production of n-alkanes in Escherichia coli by spatial organization of biosynthetic pathway enzymes. J. Biotechnol. 192 Pt A, 187 – 191 (2014).

28. Wang, Q.-S. et al. Upgrade of macromolecular crystallography beamline BL17U1 at SSRF. Nucl. Sci. Tech. 29, 68 (2018).

29. Okrzesiński, Z. M. W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307 – 326 (1997).

30. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D. Biol. Crystallogr. 66, 213 – 221 (2010).

31. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D. Biol. Crystallogr. 60, 2126 – 2132 (2004).

Acknowledgments

We are grateful to the staffs at the Shanghai Synchrotron Radiation Facility (Shanghai, China) for technical support during diffraction data collection. We thank Y. Chen, Z. Yang, and B. Zhou from IBP, CAS for MST experiment. The authors thank Dr. Torsten Juelich (Peking University, China) for linguistic assistance during the preparation of this manuscript. The project was funded by the Strategic Priority Research Program of CAS (XDB27020106 and XDB08020302), the National Key R&D Program of China (2017YFA0503702 and 2011CB909000), the National Natural Science Foundation of China (31707077 and 31900646), and the Key Research Program of Frontier Sciences of CAS (QYZDB-SSW-SMC05).

Author contributions

M.L. and W.C. conceived the project; Y.G. and C.J. did the plasmid construction; Y.G. expressed, purified AAR and ADO, prepared AAR-ADO complex and did the crystalization; L.S. and X.Z. assisted the protein expression and purification; H.Z., Y.G., P.C., and K.P. collected diffraction data; M.F. determined the AAR-His structure; H.Z. solved, built, and refined the AAR<sub>re</sub> and AAR-ADO complex structures; M.L. and H.Z. analyzed the structures, and wrote the manuscript; all authors discussed and commented on the results and the manuscript.

Competing interests

The authors declare no competing interests.
