Crystal Structure of Acyl-CoA Oxidase 3 from *Yarrowia lipolytica* with Specificity for Short-Chain Acyl-CoA

Sangwoo Kim¹,² and Kyung-Jin Kim¹,²*

¹School of Life Sciences, KNU Creative BioResearch Group, Kyungpook National University, Daegu 41566, Republic of Korea
²KNU Institute for Microorganisms, Kyungpook National University, Daegu 41566, Republic of Korea

**Introduction**

Fatty acid oxidation is an essential system for energy generation in cells by processing lipid constituents, carbon storage molecules, regulators of enzymes and membrane channels, ligands of nuclear receptors, hormone precursors, and substrates of α-, β-, and ω-oxidation [1, 2]. The biodegradation of fatty acids occurs mainly by β-oxidation in the mitochondria and peroxisomes. The β-oxidation of fatty acids is catalyzed by four enzymes: acyl-CoA dehydrogenase in mitochondria or acyl-CoA oxidase (ACOX) in peroxisomes, enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase, and 3-ketoacyl-CoA thiolase [3, 4]. ACOX catalyzes the first step of peroxisomal β-oxidation by converting acyl-CoA to 2-trans-enoyl-CoA. The yeast *Yarrowia lipolytica* can utilize fatty acids as a carbon source and thus has extensive biotechnological applications. The crystal structure of ACOX3 from *Y. lipolytica* (YlACOX3) was determined at a resolution of 2.5 Å. It contained two molecules per asymmetric unit, and the monomeric structure was folded into four domains; Nα, Nβ, Cα1, and Cα2 domains. The cofactor flavin adenine dinucleotide was bound in the dimer interface. The substrate-binding pocket was located near the cofactor, and formed at the interface between the Nα, Nβ, and Cα1 domains. Comparisons with other ACOX structures provided structural insights into how YlACOX has a substrate preference for short-chain acyl-CoA. In addition, the structure of YlACOX3 was compared with those of medium- and long-chain ACOXs, and the structural basis for their differences in substrate specificity was discussed.

**Keywords:** *Yarrowia lipolytica*, acyl-CoA oxidase 3, fatty acid, β-oxidation
(C4–C8). At ACX5 activity has not been biochemically characterized, and At ACX6 transcripts have not been detected experimentally. Maize ACOX exhibits activity on short- and medium-chain acyl-CoAs (C4–C8 and C10–C14) [25]. In mammals, mouse ACOX (ACO-I and ACO-II) [19, 20] exhibits activity on long- and medium-chain acyl-CoA, and ACO-II is active on C12 acyl-CoA. Human ACOXs [21] (i.e., ACOX1a or ACOX1b isoforms) exhibit specificity for long-chain acyl-CoAs (C12–18).

The yeast *Yarrowia lipolytica* is able to utilize fatty acids as a carbon source, resulting in the production of various metabolites and the secretion of numerous proteins; accordingly, the species is used for biotechnological applications, such as the generation of fatty acid-derived compounds for oil, biofuel, and green chemicals [26–28]. *Y. lipolytica* has six known isoforms of ACOX (*Yl* ACOX1–6), with different substrate specificities according to biochemical studies [29–33]. *Yl* ACOX1 prefers short-chain acyl-CoAs (C6), *Yl* ACOX2 favors long-chain acyl-CoAs (C10–C14) [34], *Yl* ACOX3 and *Yl* ACOX4 exhibit activity for short-chain acyl-CoAs (C4–C10) [35, 36], *Yl* ACOX5 accepts acyl-CoAs regardless of chain length [37], and *Yl* ACOX6 participates in the degradation of dicarboxylic acids [38].

In this study, we determined the crystal structure of *Yl* ACOX3 at 2.50 Å resolution to understand its structural features and substrate specificity. We identified the major structural features of *Yl* ACOX3, and the mechanisms underlying the use of short-chain acyl-CoA as a substrate by comparing its amino acid sequence and structure with those of other ACOXs.

**Materials and Methods**

**Cloning, Expression, and Purification**

The acyl-coenzyme A oxidase 3 gene (*Yl* ACOX3) from *Yarrowia lipolytica* was amplified with primers: sense, 5'-TATACATATGATCTCCCCCAACCTCAC-3', and antisense, 5'-TATAGCGGCCGCTCCTCGTCACGCTGC-3'. The amplified *Yl* ACOX3 gene was cloned with restriction endonucleases NdeI and NotI into the pET22b(+) vector (Merck Millipore, USA). The *Yl* ACOX3 protein was overexpressed in *Escherichia coli* BL21(DE3).

The cells were grown in an LB medium containing kanamycin at 37°C, until an absorbance of 0.7 at 600 nm. The cells were induced by 0.1 mM isopropyl-1-thio-beta-d-galactopyranoside for 20 h at 20°C. After harvest, the cell pellet was resuspended in buffer A (40 mM Tris-HCl, pH 8.0, 5 mM beta-mercaptoethanol) and disrupted by ultrasonication. The cell debris was removed by centrifugation at 12,000 × g for 20 min, and the lysate was bound to Ni-NTA agarose (Qiagen, Germany). After washing with buffer A containing 20 mM imidazole, the bound proteins were eluted with 500 mM imidazole in buffer A. *Yl* ACOX3 enzyme was purified by ion exchange (Hitrap Q-FF; GE Healthcare, USA) and (Superdex200; GE Healthcare, USA) size exclusion chromatography. The purified protein was concentrated to 40 mg/ml in 40 mM Tris-HCl, pH 8.0, containing 1 mM DTT.

**Crystallization and Data Collection**

Crystallization of the purified protein was initially performed with crystal screening kits (Hampton Research Co., USA and Emerald Biosciences Co., USA) using the hanging-drop vapor-diffusion technique at 20°C. Each experiment consisted of mixing 1.5 μl of protein solution with 1.5 μl of reservoir solution and then equilibrating it against 500 μl of the reservoir solution. The single crystal of *Yl* ACOX3 was obtained with condition of 25% pentaerythritol propoxylate, 0.2 M potassium chloride, and 0.05 M HEPES (pH 6.5). The data were collected at beamline 7A [39] of the Pohang Accelerator Laboratory (Korea) under 100K conditions, and then indexed, integrated, and scaled using the HKL2000 suite [40]. Crystals of *Yl* ACOX3 belonged to tetragonal space group I4, with unit cell parameters of a = b = 160.8 Å, c = 1.677 Å.

**Table 1. Data collection, phasing, and refinement statistics.**

<table>
<thead>
<tr>
<th></th>
<th><em>Yl</em> ACOX3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>I4</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>160.8, 160.8, 139.5</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90.00, 90.00, 90.00</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50.0–2.50 (2.54–2.50)</td>
</tr>
<tr>
<td>Rmerge or Rwork</td>
<td>7.0 (28.1)</td>
</tr>
<tr>
<td>I / σI</td>
<td>33.0 (4.9)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.8 (98.0)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>5.5 (4.3)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50.0–2.50</td>
</tr>
<tr>
<td>No. reflections</td>
<td>57,083</td>
</tr>
<tr>
<td>Rwork / Rmerge</td>
<td>18.1/22.9</td>
</tr>
<tr>
<td>No. atoms</td>
<td>11,129</td>
</tr>
<tr>
<td>Protein</td>
<td>10,855</td>
</tr>
<tr>
<td>FAD</td>
<td>106</td>
</tr>
<tr>
<td>Water</td>
<td>168</td>
</tr>
<tr>
<td>B-factors</td>
<td>45.4</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.014</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.677</td>
</tr>
</tbody>
</table>

Values in parentheses are for the highest-resolution shell.

c = 139.5 Å. Assuming two molecules of YlACOX3 per asymmetric unit, the crystal volume per unit of protein mass was approximately 2.93 Å² Da⁻¹, which corresponds to a solvent content of approximately 58.04% [41]. The crystal structure of YlACOX3 was solved by molecular replacement using MOLREP [42] from the structure of acetyl-coenzyme A oxidase 1 from Yarrowia lipolytica (PDB code 5Y9D, 66% sequence identity) as a search model. The model building was performed using the program WinCoot [43] and the refinement was performed with REFMAC5 [44]. The data statistics are summarized in Table 1. The refined model of YlACOX3 was deposited in the Protein Data Bank (PDB code 5YS9).

Results and Discussion

Overall Structure of YlACOX3

We determined the crystal structure of acyl-coenzyme A oxidase 3 from Y. lipolytica (YlACOX3) at 2.50 Å resolution (Table 1). The structure contains two molecules in an asymmetric unit, and the two molecules form a dimer (Fig. 1B). The monomer structure of YlACOX3 consists of four domains: the N-terminal α domain (NuD, Ala8–Gly155), the N-terminal β domain (NβD, Phe156–Leu294), the C-terminal α domain (Ca1D, Ala295–Leu482), and the C-terminal α2 domain (Ca2D, Gly483–Glu699) (Figs. 1A, 1C).

The Nu domain (β1–β2, α1–α7, and η1–η3) consists of one β-hairpin (β1–β2), seven α-helices, and three η-helices. The Nβ domain (β3–β12 and η4–η5) is composed of a three-stranded antiparallel β-sheet (β8–β9) connected to a six-stranded two-layered β-sandwich (β3–β7 and β10) with a single β-hairpin (β11–β12) and two η-helices. The Ca1 domain is composed of five α-helices (α8–α12) and three η-helices (η6–η8), and four α-helices (α8–α11) form a four-helix bundle. Flavin adenine dinucleotide (FAD) and acyl-CoA binding sites are primarily found in the Ca1 domain. The Ca2 domain consists of four 3η-helices (η9–η11), five α-helices (α16 and α18–α20), and four-helix bundle (α13–α15 and α17). A long arm generated by the helices α19 and α20 of molecule A is extended on the Nα and Nβ domains of molecule B.

The dimerization interface is mainly formed by the Nβ and Ca2 domains of each monomer by hydrogen bonding and salt bridge interactions. Using Protein Interfaces, Surfaces and Assemblies [45], we calculated that the dimer interface of YlACOX3 represented approximately 16.8% of the total accessible surface area, which was 5,413.0 Å² of buried accessible surface area. Furthermore, the solvation free energy gain (∆G) for the formation of the interface was -47.3 kcal/mol with a p-value of 0.146, indicating a hydrophobic interface with specific interactions that are not artefacts of crystal packing.

FAD Binding Mode of YlACOX3

The FAD molecule is located in the dimerization interface and bound to a large cleft formed by the Nβ and Ca1 domains of molecule A and the Ca1 domain of molecule B (Fig. 1B). The isoalloxazine ring of FAD is stabilized by the main chains of Phe160, Met162, Thr163, and Gly202 containing the ACOX motif (KWWIGG; residues 198–203), and the side-chain of Thr163 of the Nβ domain. The pyrophosphate moiety is bound to the main chain of Ser169 of the Nβ domain, and the side-chain of Arg331 and main chain of Gly432 of the Ca1 domain in molecule B containing the FAD-binding motif (CGGHG; residues 430–435). The ribose moiety interacts with the side-chain of Arg340 of the Ca1 domain in molecule B. In addition, the residues Gly168 and Trp200 of the Nβ domain, Val450, Trp454, Val460, and Leu463 of the Ca1 domain, and Tyr348 and Tyr435 of the Ca1 domain in molecule B assist the binding of the ring by hydrophobic interactions (Figs. 2A, 2B).

Active Site of YlACOX3

In order to elucidate the substrate-binding mode of YlACOX3, we first attempted to determine its structure in complex with the acyl-CoA substrate. However, the complex crystal with acyl-CoA as a substrate was not obtained by either co-crystallization or soaking. Instead, the active site pocket was calculated on the basis of the structure of YlACOX3 using the CAVER software tool [46]. To estimate the active site pocket, the starting point of calculation was specified as the space between the isoalloxazine ring and the catalytic residue Glu445. A minimum probe radius of 0.9 Å was specified. The putative active site pocket of YlACOX3 is located at the interface of the Nα, Nβ, and Ca1 domains near the FAD-binding site. The pocket of YlACOX3 included α5–α6 in the Nα domain, loops of β3–β4 and β5–β6 in the Nβ domain, and α3 and α11–α12 in the Ca1 domain (Fig. 2C). The pocket entrance of YlACOX3 is mainly distributed in hydrophilic residues, such as Ser169, Asn170, Thr453, and Asp457. The center of the pocket is mainly composed of hydrophobic residues, such as Val1124, Leu128, Phe160, Met162, Val171, Trp200, Ile201, Ile259, and Thr254, except for Glu455 and Arg259. The region surrounding the pocket is distributed mostly with hydrophilic residues, such as Glu101, Ser108, Gln109, Arg119, Thr126, Arg134, Thr163, Arg256, Thr308, and Asp312 (Fig. 2C).

Furthermore, we investigated the substrate-binding site of YlACOX3 by comparing the structure of the protein with...
Fig. 1. Overall structure of YlACOX3.

(A) Amino acid sequence alignment of ACOXs from *Yarrowia lipolytica* (YlACOX1–6) and *Caenorhabditis elegans* (CeACOX1–2). Secondary structure elements are marked on top of the alignment: α-helices are marked by a cylinder and β-strands by an arrow. Conserved and highly similar residues are shown in red and blue, respectively. The ACOX motif (KWWxGx), the FAD-binding motif (CGGHGY), and the catalytic residue are indicated by green, yellow, and violet boxes, respectively. (B) Dimeric structure of YlACOX3. Each molecule is shown as a cartoon model with green and magenta color, respectively. The FAD molecule is shown as a sphere model with cyan color. (C) Monomeric structure of YlACOX3. The monomer is shown as a cartoon. The Nα, Nβ, Cα1, and Cα2 domains are shown in light pink, deep green, light blue, and orange, respectively. The FAD molecule bound to the enzyme is shown as a sphere model in cyan.
Fig. 2. Summary of the active site of YlACOX3.

(A) FAD-binding mode of YlACOX3. The FAD molecule is shown in a stick model in cyan. The Fo–Fc electron density map is shown as a magenta mesh and is contoured at 3.5 σ. The residues for hydrogen bonding and van der Waals contact are presented in stick and line models in green for molecule A and orange for molecule B, respectively. The hydrogen bonds are shown in black dotted lines. (B) Hydrogen bonds are shown as green dashed lines and van der Waals contacts are denoted as red combs. This figure was produced using LIGPLOT. (C) Substrate-binding tunnel of YlACOX3. The Nα, Nβ, and Cα1 domains of YlACOX3 are shown as a ribbon model in orange, deep green, and light blue, respectively. The tunnel is shown as a mesh model in gray. The FAD molecule and the asc-ωC5-CoA model are shown in a stick model in cyan and gray. The composition of the tunnel is shown using line models, with colors indicating each domain. (D) Superposition of the active sites of YlACOX3 and CeACOX2. The asc-ωC5-CoA binding-related residues for hydrogen bond formation and van der Waals contact in CeACOX2 are presented as stick and line models in gray, and their corresponding residues in YlACOX3 are shown in green. The asc-ωC5-CoA molecule of CeACOX2 is represented as a gray-colored stick model. The FAD molecules for YlACOX3 and CeACOX2 are shown in a stick model in cyan and gray, respectively. The hydrogen bonds are shown as black dotted lines.
that of ACOX2 from *Caenorhabditis elegans* (*CeACOX2*) bound with ω-ascaroside-CoA (C5) (*CeACOX2*, PDB code 5K3J, 29% amino acid identity) [24] (Fig. 2D). In *CeACOX2*, the CoA moiety of acyl-CoA is stabilized by hydrogen bonds with the side-chain of Lys283, Arg293, Arg444, and FAD. The alkyl group is located in the hydrophobic pocket, where van der Waals interactions occur with Ala112, Ala116, Leu117, Tyr147, Phe243, Ile292, and Tyr431 (Fig. 2D). The distribution of residues in the pocket of *YlACOX3* is similar to that of *CeACOX2*. In *YlACOX3*, the residues around the CoA moiety are distributed as conserved hydrophilic residues, such as Gln296, Arg306, Glu455, and Arg467, corresponding to Lys283, Arg293, Glu432, and Arg444 in *CeACOX2*. In contrast, the residues surrounding the alkyl group differ slightly from those of *CeACOX2*. The hydrophilic residue Arg256 of *YlACOX3* corresponds to the hydrophobic residue Phe243 in *CeACOX2*. Additionally, the hydrophobic residues Gly123 and Gly127 of *YlACOX3* correspond to Ala112 and Ala116 in *CeACOX2* (Fig. 2D).

**Comparison of *YlACOX3* with Other ACOXs**

The structural comparison between *YlACOX3* and other ACOXs provided insights into the structural basis for the substrate specificity of these enzymes. To investigate why

---

**Fig. 3.** Structural comparison of ACOXs.

(A) Structural differences between the Co1 domains of *YlACOX3*, *CeACOX1*, and *CeACOX2* are shown as cartoon models in green, cyan, and light blue, respectively. The other domains are shown as a ribbon model in gray. The red arrow shows the structural differences of the three enzymes. The image on the right is rotated 90 degrees in the vertical direction. Enlarged view of differences among (B) *YlACOX3*, (C) *CeACOX1*, and (D) *CeACOX2*. The hydrogen bonds between the helices are shown as red and black dotted lines, and their residues are presented in stick models.
YlACOX3 has substrate specificity for short-chain acyl-CoA, we compared the structure of YlACOX3 with those of two ACOXs: short-chain ACOX (CeACOX2, PDB code 5K3J, 29% amino acid identity) [24] and medium- or long-chain ACOX (CeACOX1, PDB code 5K3I, 27% amino acid identity) [24]. Based on the superimposition of these three ACOXs, the overall structures of these enzymes were quite similar to each other, but local structural differences were observed in the α5 and α8–α10 regions around the active site pocket (Fig. 3A). In particular, α8, α9, and α10 of YlACOX3 differed by a maximum of ~20 degrees from the other ACOXs. YlACOX3 was stabilized by hydrogen bonding interactions between α5 and α8 (Asn109 and Thr308), α8 and α10 (Arg306 and Phe405), and α10 and η9 (Asp400 and Ser484) (Fig. 3B). Similarly, CeACOX2 formed hydrogen bonds and salt bridges between α5 and α8 (Glu93 and Lys291), and α8 and α9 (Ser294 and Thr359), but not η9 (Fig. 3D). In CeACOX1, interactions were limited to single hydrogen bonds between α8 and α9 only (Fig. 3C). The flexible regions of YlACOX3 and CeACOX2 as well as short-chain ACOX were more stable than those of CeACOX1 as well as medium- or long-chain ACOXs. Presumably, the stability of the flexible regions of ACOXs is a significant determinant of the open-closed conformation of the tunnel from the active site to the outer surface of these enzymes.

Furthermore, we investigated the outer surface size of the tunnel by comparisons with four structures. To do this, the outer surface size of the ACOX tunnels were calculated using CAVER [46]. The channel of the outer surface of YlACOX3 was formed by hydrophilic residues, such as Glu101 and Gln109 of α5, Arg119, Thr126, and Arg134 of α6, and Thr308 and Glu455 of α8 (Fig. 4A). The outer surface channel was composed of the hydrophilic residues Asp91, Glu93, Ser96, and His100 of α5, His109, Asp115, and Asn123 of α6, and Lys291 and Gln295 of α8 for CeACOX2. However, the active site had more limited access to the outer surface in CeACOX2 than in YlACOX3 owing to the hydrogen bonding and salt bridge interactions between α5, α6, and α8, such as Glu93 and Lys291 from CeACOX2 (Fig. 4C). CeACOX1 mainly consisted of the hydrophobic residues Tyr99, Tyr297, and Tyr365, except for Asn125 and His293 (Fig. 4B). Interestingly, we did not detect any hydrogen bonds or salt bridges between α5, α6, and α8 of CeACOX1. The residues of the active site channel of medium- or long-chain ACOXs, such as CeACOX1, were replaced with hydrophobic and aromatic residues.

Taken together, these results suggest that the stability of the flexible region leads to the open form of the active site in YlACOX3 and CeACOX2, but not in CeACOX1.

**Fig. 4.** The outer surface of ACOX channels. (A) The pocket channel from YlACOX3, (B) CeACOX1, and (C) CeACOX2. Enzymes and helices are shown as surface and cartoon models. The residues forming the channel are represented by the stick model. The tunnel is shown as a mesh model in gray. The asc-ωC5-CoA and FAD molecules are shown in a stick model in gray, where colors represent enzymes.
pocket, and the closed form of short-chain ACOXs was explained by the high stability compared with that of medium- or long-chain ACOXs. Additionally, the channel size of the outer surface from the active site was probably determined by the properties of the residues in α5, α6 and α8.

Acknowledgments

This work was supported by National Research Foundation of Korea (NRF) grants funded by the Korean Government (MSIP) (2014M1A2A2033626, NRF-2013R1A1A2059325, and NRF-2017R1A6A3A11031554) and was also supported by the New & Renewable Energy Core Technology Program of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) granted financial resource from the Ministry of Trade, Industry & Energy, Republic of Korea (20153030091360).

Conflict of Interest

The authors have no financial conflicts of interest to declare.

References

8. Stanway CA, Gibbs JM, Berardi E. 1995. Expression of the FOX1 gene of Saccharomyces cerevisiae is regulated by carbon source, but not by the known glucose repression genes.