PROTEINS: Structure, Function, and Bioinformatics 64:817–822 (2006)

STRUCTURE NOTE

Crystal Structure of Fatty Acid-CoA Racemase from *Mycobacterium tuberculosis* H37Rv

Ki Seog Lee,[1] Seon Mi Park,[1] Kyung-Hee Rhee,[2] Won-Gi Bang,[1] Kwang Yeon Hwang,[1] and Young-Min Chi[1]*

[1]*Division of Biotechnology and Genetic Engineering, Life and Environmental Sciences, Korea University, Seoul, South Korea*
[2]*Biomedical Research Center, Division of Life Science, Korea Institute of Science and Technology, Seoul, South Korea*

Introduction. The classic catabolic pathway by which fatty acids are degraded is β-oxidation, and a mitochondrial as well as a peroxisomal β-oxidation pathway is known. Very long-chain fatty acids, 2-methyl-branched fatty acids, the side-chains of bile acid intermediates, and eicosanoids are mainly handled by the peroxisomal pathway, whereas short- and medium-chain fatty acids are oxidized mainly in mitochondria.[1,2] Phytanic acid and other 3-methyl-branched fatty acids cannot undergo β-oxidation, because the 3-methyl group prevents the formation of a 3-keto substituent in the dehydrogenation step. Therefore, 3-methyl-branched fatty acids first undergo α-oxidation.[3] In the case of phytanic acid, this results in the generation of 2-methyl-branched pristanic acid (2,6,10,14-tetramethyl pentadecanoic acid), which is then shortened to 4,8-dimethyl nonanoic acid via peroxisomal β-oxidation. The dimethyl fatty acid is then degraded further via mitochondrial β-oxidation.[4] The α-oxidation of fatty acids is not a stereoselective process,[5,6] so that after α-oxidation of branched-chain fatty acids, both (*R,R,R*)- and (*S,R,R*)-isomers are formed. However, the β-oxidation system is stereoselective because only the (2*S*)-isomer is accepted as substrate by branched-chain acyl–coenzyme A (CoA) oxidase, the first enzyme of the β-oxidation system. Thus, the (*R,R,R*)-isomers of the metabolites of α-oxidation need to be converted to their (*S,R,R*)-isomers by enzymatic racemization prior to further degradation.[6,7]

The fatty acid-CoA racemase (FAR) encoded by the *far* gene of *Mycobacterium tuberculosis* H37Rv converts several (2*R*)-branched-chain fatty acid-CoA's to their (2*S*)-stereoisomers in the metabolism of fatty acids. FAR is thought to be involved in this racemization prior to the stereospecific β-oxidation, although this has not yet been demonstrated experimentally. Bhaumik et al.[8] found three genes—*mcr*, *far*, and *Rv3727*—in the *M. tuberculosis* genome that encode proteins that are homologous to mammalian α-methylacyl-CoA racemase (AMACR), which is a mitochondrial and peroxisomal enzyme that is essential in the β-oxidation of bile acid intermediates and branched-chain fatty acids.[7,9] Among them, only one crystal structure of α-methylacyl-CoA racemase (MCR) from *M. tuberculosis* relative to mammalian AMACR has been reported to date.[10] Thus, the FAR from *M. tuberculosis* was taken as a model protein to provide the structural information of another protein corresponding to mammalian AMACR. Moreover, regarding the strong sequence similarity of AMACR with CoA transferases of family III,[11] the question arises whether this enzyme also catalyzes CoA trans-

TABLE I. Summary of Data Collection and Refinement Statistics

Data collection[a]	
Resolution (Å)	50.0–2.70 (2.82–2.70)
Unique reflections	52589
Redundancy	3.7 (4.0)
Data completeness (%)	95.8 (86.2)
R_{sym}[b]	0.080 (0.306)
$I/\sigma(I)$	9.9 (2.0)
Space group	$P3_2$
Lattice constant (Å)	$a = b = 109.56$, $c = 147.97$
No. of molecules per asymmetric unit	6
Refinement statistics	
R_{factor} (%)[c]	24.2
R_{free} (%)[c]	29.0
No. of protein residues	2118
No. of water molecules	472
Mean of overall B factor (Å^2)	55.9
RMSD bond length (Å)	0.008
RMSD bond angle (°)	1.218
Ramachandran plot	
Most allowed region (%)	83.5
Additional allowed region (%)	14.6
Generously allowed region (%)	1.9
Disallowed region (%)	0.0

[a]Values in parentheses are for the highest resolution shell.
[b]$R_{sym} = \Sigma |I - \langle I \rangle| / \Sigma \langle I \rangle$, where I is the observed intensity and $\langle I \rangle$ is the average intensity.
[c]$R_{factor} = \Sigma ||F_{obs}| - |F_{calc}|| / \Sigma |F_{obs}|$, where F_{obs} are the observed structure factors and F_{calc} are the calculated structure factors. The crystallographic R_{factor} is based on 95% of the data used in refinement, and the R_{free} is based on 5% of the data withheld for cross-validation test.

Grant sponsor: Ministry of Educational and Human Resources Development in Korea (2005); Grant sponsor: Functional Proteomics Center 21C Frontier Program of the Korean Ministry of Science and Technology.

*Correspondence to: Young-Min Chi, Division of Biotechnology and Genetic Engineering, College of Life and Environmental Sciences, Korea University, 1,5-ka, Anam-dong, Sungbuk-gu, Seoul 136-701, South Korea. E-mail: ezeg@korea.ac.kr

Received 13 February 2006; Revised 17 March 2006; Accepted Accepted 17 March 2006

Published online 5 June 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21029

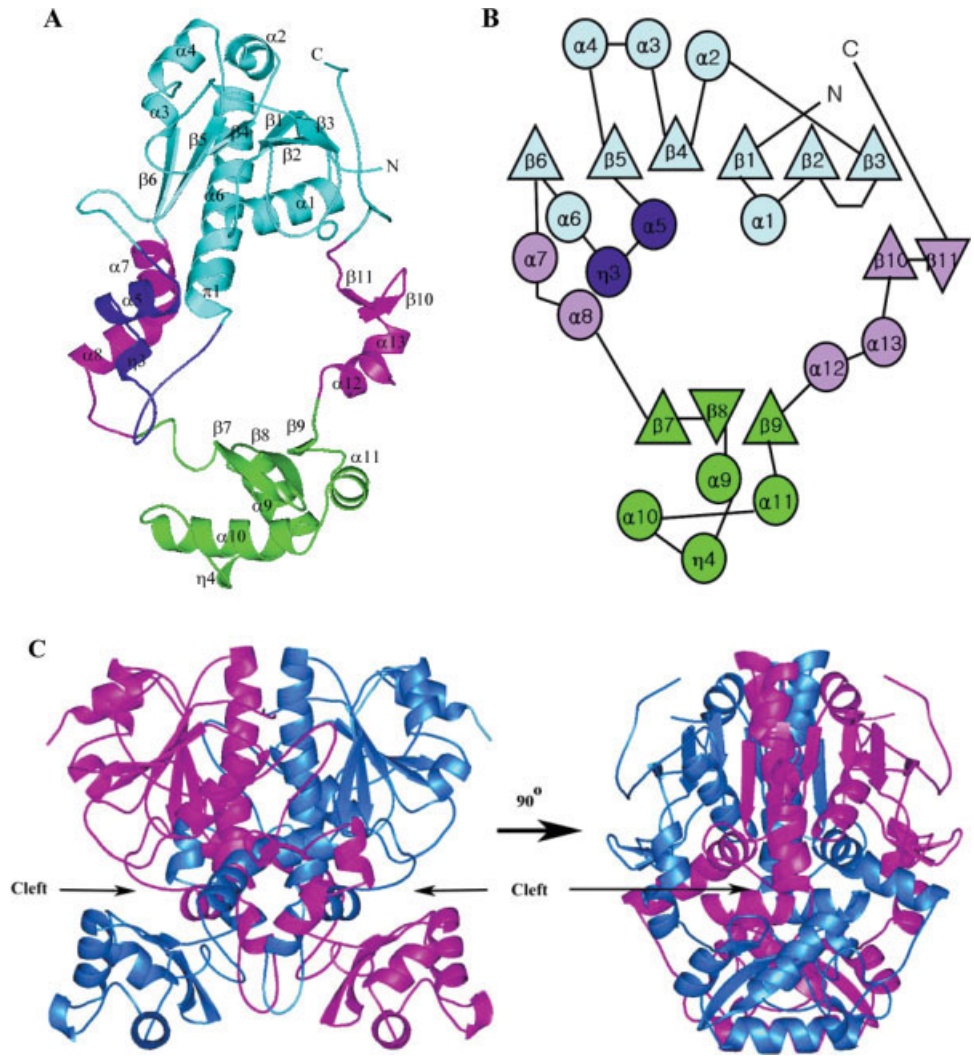

Fig. 1. (**A**) Ribbon diagram of the monomeric structure of FAR. (**B**) A topology diagram of the FAR monomer. The large domain is cyan, and the small domain is green; the chain protruding from the large domain to small domain is blue, and the two linkers between the two domains are purple. (**C**) Ribbon diagram for homodimeric arrangement of FAR. Each monomer is colored blue or purple. η-Helices denote 3_{10}-helices.

fer reactions. However, CoA transferase activity has not been detected in the enzyme assays used so far.[7,12,13] Therefore, a possible new role is suggested for AMACR as a functional enzyme that racemizes its substrates.[11]

In this work, we determined the crystal structure of FAR from *M. tuberculosis* H37Rv, revealing an interlocked dimer with an unusual protein fold. Our results indicate that the FAR has amino acid sequences that represent CoA transferase family III, and is structurally similar to this family of proteins.

Materials and Methods. *Protein production, crystallization, and X-ray diffraction data collection.* Protein expression, purification, and crystallization have been described elsewhere.[14] The crystals of the apo-protein were obtained by hanging-drop vapor-diffusion method at 295 K in 2 days. The reservoir solution used 20% (w/v) polyethylene glycol (PEG) 4,000, 0.05 *M* Tris-HCl, pH 8.5, 0.2 *M* ammonium chloride, and 0.1 *M* calcium chloride. The crystal was transferred into cryo-protection solution containing 30% (w/v) PEG 4,000 as precipitant, 0.05 *M* Tris-HCl, pH 8.5, 0.2 *M* ammonium chloride, 0.01 *M* calcium chloride, and 30% (v/v) glycerol, then it was scooped up in a cryoloop and frozen in liquid nitrogen. It was then mounted on the goniometer in a stream of cold nitrogen at 100 K. X-ray diffraction data were collected from the cooled crystal using a Bruker Proteum 300 charge-coupled device (CCD) at beamline 6B at Pohang Light Source (PLS; South Korea). Diffraction data were integrated and scaled with DENZO and SCALEPACK crystallographic data-reduction package.[15]

Structure determination. The structure of FAR was solved by a molecular replacement method using the Crystallography & NMR System (CNS) package[16] with the monomer of the *M. tuberculosis* MCR [Protein Data Bank (PDB) ID: 1X74][13] as a search model. The model shares 57% sequence identity with *M. tuberculosis* FAR equally distrib-

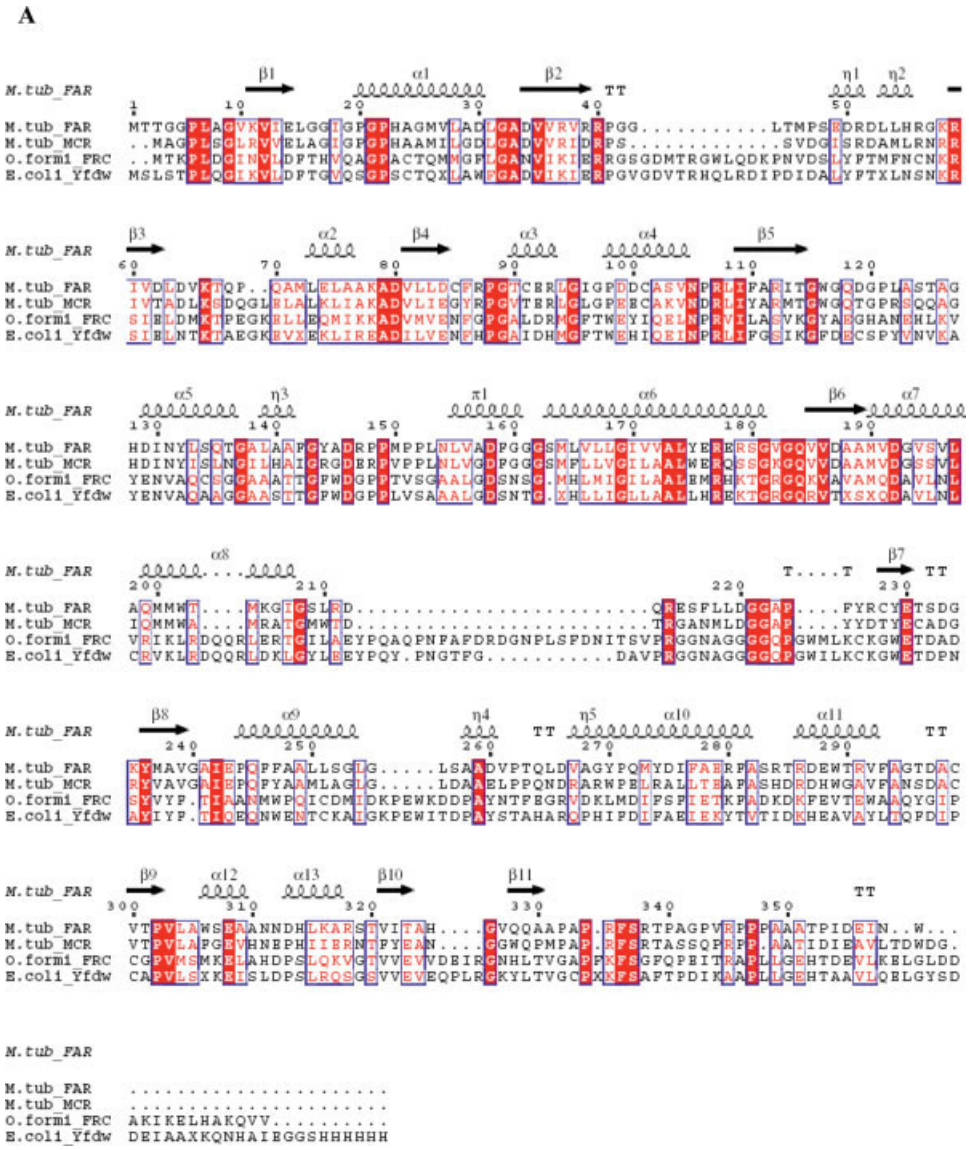

Fig. 2. (**A**) Structure-based sequential alignment of *M. tuberculosis* H37Rv FAR and MCR, *O. formigenes* FRC, and *E. coli* YfdW with available crystal structure. The secondary structure assignment for FAR is indicated by helices (α- and 3_{10}-helices) and arrow (β-strands). Conserved residues are boxed. (**B**) The two clefts of dimeric FAR. Each monomer is cyan or purple. Conserved residues (Arg93, His128, Asp158, Glu243) are shown by a stick model. (**C**) Superposition of three crystallographically independent dimeric structures [FAR (blue), FRC (purple), and YfdW (green)] drawn by ribbon models of Cα trace. Although slight structural changes are found among these molecules (FAR-FRC, RMSD is 2.96 Å; FAR-YfdW, RMSD is 2.87 Å), overall fold shows almost identical structures except for some loop regions.

uted over the sequence.[14] The CNS was run using the data with a resolution range of 50.0–2.7 Å in the space group $P3_2$. Reflection data chosen randomly from 5% of the observed data were used for cross-validation with the R_{free} value. We initially refined the rigid body model with the program CNS, then performed model refinement using COOT.[17] Water molecules were automatically picked up by the CNS package, and they were then confirmed based on peak height and distance criteria in the *Fo–Fc* and 2*Fo–Fc* maps. We analyzed the quality of the final and intermediate models using PROCHECK.[18] The coordinates have been deposited in the PDB with the accession number 2G04.

Structural comparison. We carried out structure analysis using the following computer programs: Swiss-PdbViewer[19] and COOT[17] for superposition of molecules and secondary structure-based alignment; SSAP Server[20] for calculation of root-mean-square deviations (RMSDs); ESPript[21] for preparation of the secondary structure alignment figure; PyMol[22] for the depiction of structure; and GRASP[23] for calculation and depiction of electrostatic potential at the molecular surface. Other orthologs used in the structural comparison were *M. tuberculosis* MCR (PDB code: 1X74),[13] *Escherichia coli* YfdW (PDB code: 1PT7),[24] and *Oxalobacter formigenes* FRC (PDB code: 1P5H).[25]

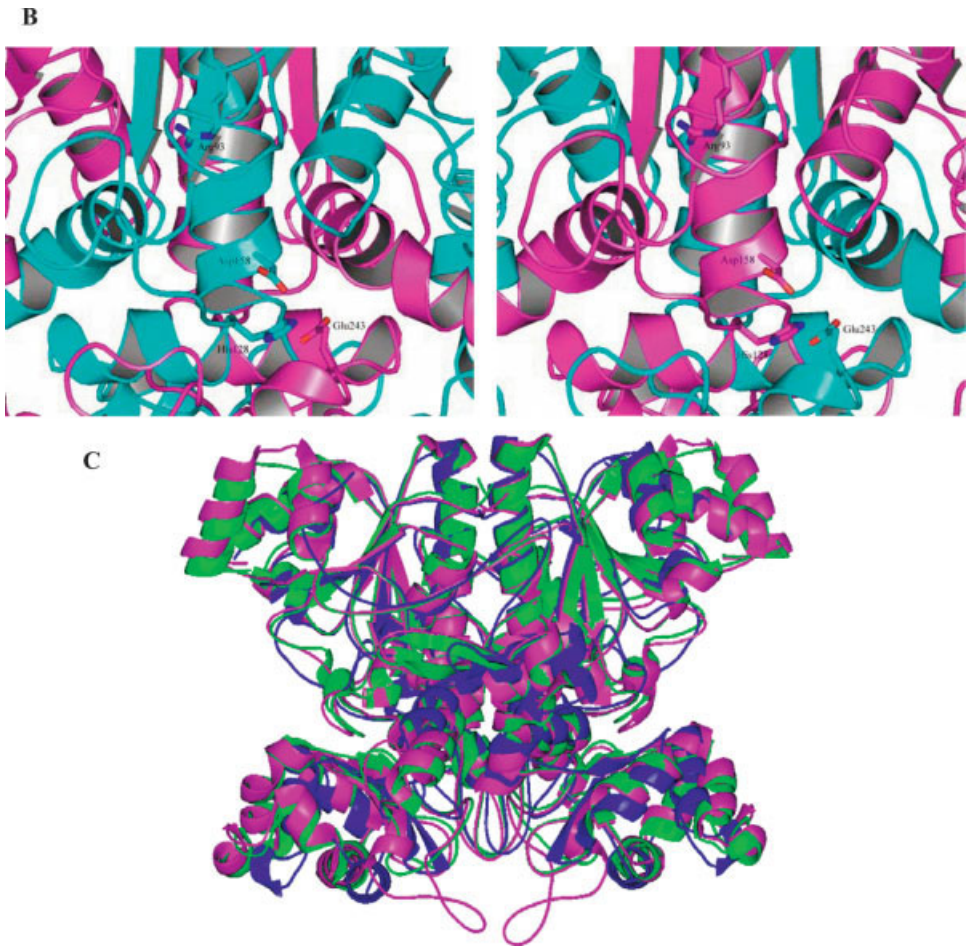

Figure 2. (Continued.)

Results and Discussion. *Crystallographic data and refinement statistics.* The crystals of FAR were found to belong to the trigonal space group $P3_2$, with unit cell parameters $a = b = 109.56$ Å and $c = 147.97$ Å. Six monomers were contained per asymmetric unit ($V_M = 2.15$ Å^{3}Da^{-1}, ca. 42.8% solvent).[14] The crystal structure of FAR was determined at 2.7 Å resolution by molecular replacement. After the phases were improved, the resulting electron density map was readily interpretable and allowed model building. Several rounds of model adjustment, simulated annealing, positional refinement, noncrystallographic symmetry (NCS) restraints refinement, and individual thermal factor refinement were performed. The refined model of FAR gave R_{factor} and R_{free} values[26] of 24.2 and 29.0%, respectively. The final apo-protein model consists of 2118 residues (353 residues/monomer of the 359 in sequence), except for residues 1–3 and 66–68 in the main-chain of each subunit. In the Ramachandran plot,[27] analysis of the final model with PROCHECK reveals that 83.5%, 14.6%, 1.9%, and 0.0% of residues were in the most favored, additional allowed, generously allowed, and disallowed regions, respectively. The data collection and refinement statistics are summarized in Table I.

Overall structure and dimer configuration. The overall structure of the FAR monomer consists essentially of two domains, the large domain and the small domain, and two linker regions (α7–8 and α12–13, β10–11) that connect the two domains. The large domain comprises both N-terminal (α1–4, α6, β1–6) and C-terminal parts of the protein chain (residues 335–359), while the small domain contains α9–11, η4, and β7-9 [Fig. 1(A)]. The core of the large domain has an open sheet structure with a Rossmann-like fold consisting of a central, six-stranded, parallel β-sheet (β1–6) with helices packed to both sides. The loop after β2 is disordered. The connection between strains β5 and β6 is left-handed and is connected by three helices (α5, η3, and α6): After β5, the polypeptide chain, including two helices (α5 and η3), makes a protrusion toward the small domain. A long loop leads back to the large domain where α6 and β6 complete the α/β structure. The first linker α7–8, a long, very bent helix, connects the large domain with the small domain. The core of the small domain is a three-stranded, antiparallel β-sheet that starts and ends with strands β7 and β9, respectively. The first two strands (β7 and β8) are joined by a short loop, while four helices (α9–11 and η4) are inserted between the second and the third strand.

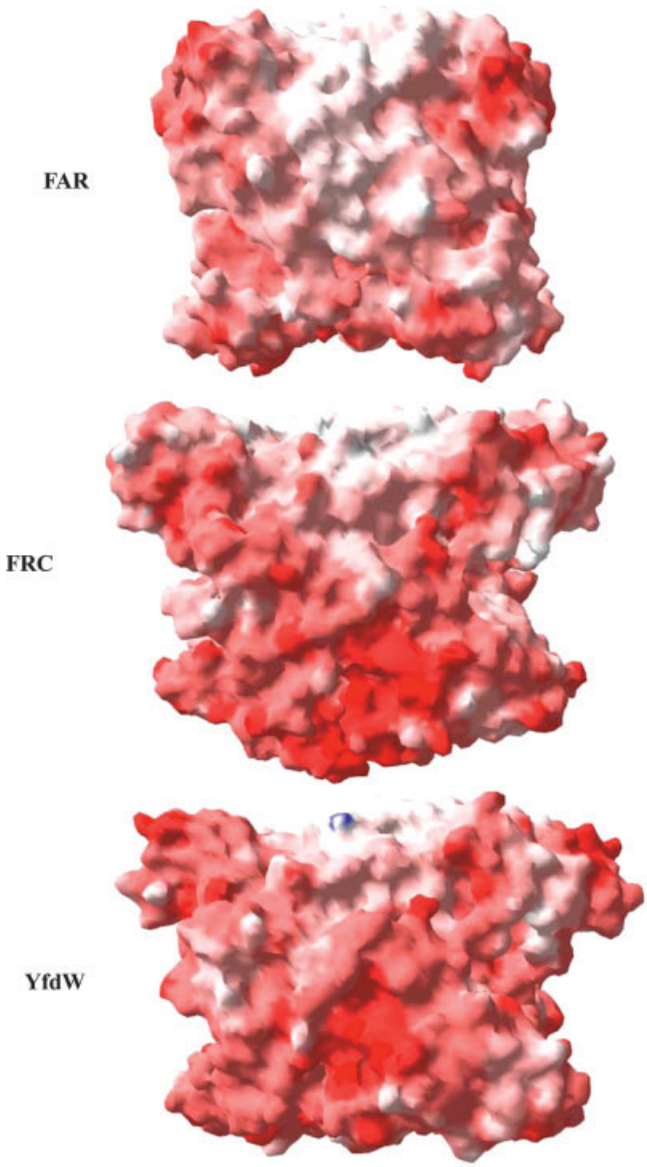

Fig. 3. Electrostatic potential diagrams of *M. tuberculosis* H37Rv FAR, *O. formigenes* FRC, and *E. coli* YfdW. Positive and negative potentials are represented by blue and red with thresholds of ±10 kT/e, respectively.

After strand β9, the second linker (α12–13, β10–11) connects the small domain back to the large domain, such that a large hole is created by the folded polypeptide chain of one subunit [Fig. 1(B)].

An unexpected feature of the subunit is the hole formed between the domains. The connecting, kinked helix α7–8 of the second subunit, as well as the protruding helices α5 and η3 of the second subunit, passes through this hole, making a very tight, interlocked dimer, in which the two subunits are related by a local twofold axis. The twofold related subunits in the dimer embrace each other with their C-terminal parts, giving rise to an interlocked dimer. The linker helix α7–8 of the first monomer passes through the hole present in the second monomer and vice versa. This puts two domains of one subunit on top of and below the hole of the other subunit. The hole is filled further by the stretch α5 and η3, but this does not give rise to a knot, because this is only a protrusion from the large domain. The FAR dimer is formed by two elongated rings that interpenetrate as the first two rings of a chain. The dimeric assembly of two subunits results in an interlocked mode of interaction between the two polypeptide chains, because the C-terminal tail of one subunit folds back on its own N-terminal large domain, after a long excursion in which the small domain is formed [Fig. 1(C)]. Consequently, the crystal structure of FAR exists as a tightly interlocked dimer, and the crystal packing contains three dimers in a crystallographic asymmetry unit. The accessible surface areas (ASAs) of the monomer and the dimer were 18637.2 $Å^2$ and 24592.5 $Å^2$, respectively. Dimer formation decreases the ASA by 6340.9 $Å^2$ per monomer, which amounts to 34.0% of the total surface of the monomer. Among the buried surface residues, 62.0% are nonpolar, 24.1% are polar, and 13.9% are charged, indicating that the majority of the dimeric interface is formed by hydrophobic interactions. A long α-helix (α6) from each monomer is located at the dimeric interface contributing hydrophobic interactions from the following residues; Val156, Ala157, Leu165, Val166, Leu167, Leu168, Ile170, Val171, Val172, Ala173, Leu174. From the structural aspects, dimerization appears to be important for the racemase function and stability of *M. tuberculosis* H37Rv FAR. Most of the dimeric interface is nonpolar, so the formation of a dimer would be energetically favorable compared to the monomer in solution.

The two clefts. Dimeric FAR presents two surface clefts, located in the interface between the large domain of one subunit and the small domain of the other subunit [Fig. 1(C)]. Some conserved residues—Arg93, His128, Asp158, and Glu243—are all located in the two clefts [Fig. 2(B)]. The arrangement of residues in the cleft of FAR is quite similar to that of *M. tuberculosis* MCR[13] and other family III CoA transferase members, suggesting that the clefts of FAR would be suitable for substrate binding and catalytic reaction, as observed in the structure of the FRC–CoA complex.[25,28]

Functional and structural similarities to other proteins. The multiple alignment of amino acid sequence identities between FAR and FRC or YfdW, as the known structure of the family III CoA transferases, were 29.8% and 32.7%, respectively [Fig 2(A)]. The pairwise structure comparison using the SSAP Server[20] revealed that the FAR possesses the structural resemblance to family III of CoA transferases. In the superposition of three crystallographically independent dimeric structures with FAR, FRC (PDB ID: 1P5H), and YfdW (PDB ID: 1PT7), the RMSDs for the Cα atoms were 2.96 and 2.87 Å, respectively. For the superposition of the FAR-dimer on the FRC-dimer and YfdW-dimer as shown in Figure 2(C), the investigation of structural similarities found that the six β-strands of the large domain and the three β-strands of the small domain, as well as the α-helices α1–4, and α6 of the large domain, of both subunits were used. Inspection of Figure 2(C) shows

that the quaternary structure of the interlocked dimer is very well preserved between FAR and FRC or YfdW. Despite the low sequence identity between FAR and FRC or YfdW, both of these CoA transferases are folded as interlocked dimers. It is very likely that most other members of the family III of CoA transferases have this intriguing fold. It is unlikely that monomeric proteins of this fold exist. In addition, the surface electrostatic potential of dimeric FAR is negative and is similar to those of dimeric FRC and YfdW as the known structures of the family III CoA transferases [Fig. 3].

These results indicated that the FAR from *M. tuberculosis* H37Rv belongs to member of the recently identified family III of CoA transferases. It is shown that the FAR has unique structural features which not seen in the identities of the sequence related FRC or YfdW. Although additional experimental evidence must be obtained, we suggest that the family III CoA transferases might be subdivided in at least two classes, being racemases and CoA transferases.

Acknowledgments. We thank Dr. H. S. Lee and his staff for assistance during data collection at beamline BL-6B of Pohang Light Source, South Korea.

REFERENCES

1. Mannaerts GP, Van Veldhoven PP, Casteels M. Peroxisomal lipid degradation via α- and β-oxidation in mammals. Cell Biochem Biophysics 2000;32:73–87.
2. Ackman RG, Hansen RP. The occurrence of diastereo-isomers of phytanic and pristanic acids and their determination by gas–liquid chromatography. Lipids 1967;2:357–362.
3. Wanders RJA, Jansen GA, Lloyd MD. Phytanic acid alpha-oxidation, new insights into an old problem: a review. Biochim Biophy Acta 2003;1631:119–135.
4. Casteels M, Foulon V, Mannaert GP, Van Veldhoven PP. Alpha-oxidation of 3-methyl-substituted fatty acids and its thiamine dependence. Eur J Biochem 2003;270:1619–1627.
5. Tsai SC, Steinberg D, Avigan J, Fales HM. Studies on the stereospeficicity of mitochondrial oxidation of phytanic acid and of α-hydroxyphytanic acid. J Biol Chem 1973;246:1091–1097.
6. Croes K, Casteels M, Dieuaide-Noubhani M, Mannaerts GP, Van Veldhoven PP. Stereochemistry of the α-oxidation of 3-methyl-branched fatty acids in rat liver. J Lipid Res 1999; 40:601–609.
7. Schmitz W, Conzelmann E. Stereohemistry of peroxisomal and mitochondrial β-oxidation of α-methylacyl-CoAs. Eur J Biochem 1997;244:434–440.
8. Bhaumik P, Kursula P, Ratas V, Conzelmann E, Hiltunen JK, Schmitz W, Wierenga RK. Crystallization and preliminary X-ray diffraction studies of an α-methylacyl-CoA racemase from *Mycobacterium tuberculosis*. Acta Crystallogr D Biol Crystallogr 2003;59: 353–355.
9. Ferdinandusses S, Denis S, Clayton PT, Graham A, Rees JE, Allen JT, Mclean BN, Brown AY, Vreken P, Waterham HR, Wanders RJA. Mutation in the gene encoding peroxisomal alpha-methylacyl-CoA racemase cause adult-onset sensory motor neuropathy. Nat Genet 2000;24:188–191.
10. Savolainen K, Bhaumik P, Schmitz W, Kotti TJ, Conzelmann E, Wierenga RK, Hiltunen J K. α-Methylacyl-CoA racemase from *Mycobacterium tuberculosis*—mutational and structural characterization of the active site and the fold. J Biol Chem 2005;280:12611–12620.
11. Heider J. A new family of CoA-transferases. FEBS Lett 2001;509: 345–349.
12. Shieh W-R, Chen C-S. Purification and characterization of novel "2-arylpropionyl-CoA epimerases" from rat liver cytosol and mitochondria. J Biol Chem 1993;268:3487–3493.
13. Reichel C, Brugger R, Bang H, Geisslinger G, Brune K. Molecular cloning and expression of a 2-arylpropionyl-coenzyme A epimerase: a key enzyme in the inversion metabolism of ibuprofen. Mol Pharmacol 1997;51:576–582.
14. Rhee K-H, Lee KS, Priyadarshi A, Kim EE, Hwang KY. Crystallization and preliminary X-ray crystallographic studies of fatty acid-CoA racemase from *Mycobacterium tuberculosis* H37Rv. Acta Crystallogr F 2005;61:1017–1019.
15. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
16. Brünger AT, Adams PD, Clore GM, Delano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges N, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR System: a new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54:905–921.
17. Paul E, Kevin C. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004;60:2126–2132.
18. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. Procheck—a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1996;26:283–291.
19. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 1997;18:2714–2723.
20. Orengo CA, Michie AD, Jones S, Jones DT, Swindells MB, Thornton JM. CATH—a hierarchic classification of protein domain structures. Structure 1997;5:1093–1108.
21. Gouet P, Courcelle E, Stuart DI, Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 1999; 14:305–308.
22. DeLano WL. The PYMOL Molecular Graphics System (http://pymol.sourceforge.net). San Carlos, CA: DeLano Scientic LLC; 2002.
23. Nicholls A, Sharp KA, Honig B. Protein folding and association: insight from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991;11:281–296.
24. Gruez A, Roig-Zamboni V, Valencia C, Campanacci V, Cambillau C. The crystal structure of the *Escherichia coli* YfdW gene product reveals a new fold of two interlaced rings identifying a wide family of CoA transferases. J Biol Chem 2003;278:34582–34586.
25. Ricagno S, Jonsson S, Richards N, Lindqvist Y. Formyl-CoA transferase encloses the CoA binding site at the interface of an interlocked dimer. EMBO J 2003;22:3210–3219.
26. Brünger AT. The free R-value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 1992;355:472–474.
27. Ramachandran GN, Sasisekharan V. Conformation of polypeptides and proteins. Adv Protein Chem 1968;23:283–438.
28. Jonsson S, Ricagno S, Lindqvist Y, Richards NGJ. Kinetics and mechanistic characterization of the formyl-CoA transferase from *Oxalobacter formigenes*. J Biol Chem 2004;279:36003–36012.