

Supplementary Materials

to "*Phylogenomic reconstruction of archaeal fatty acid metabolism*"

by Daria V. Dibrova, Michael Y. Galperin, and Armen Y. Mulkidjanian

Figure S1. Reactions of β -oxidation of fatty acids (modified from the charts in KEGG and MetaCyc databases).

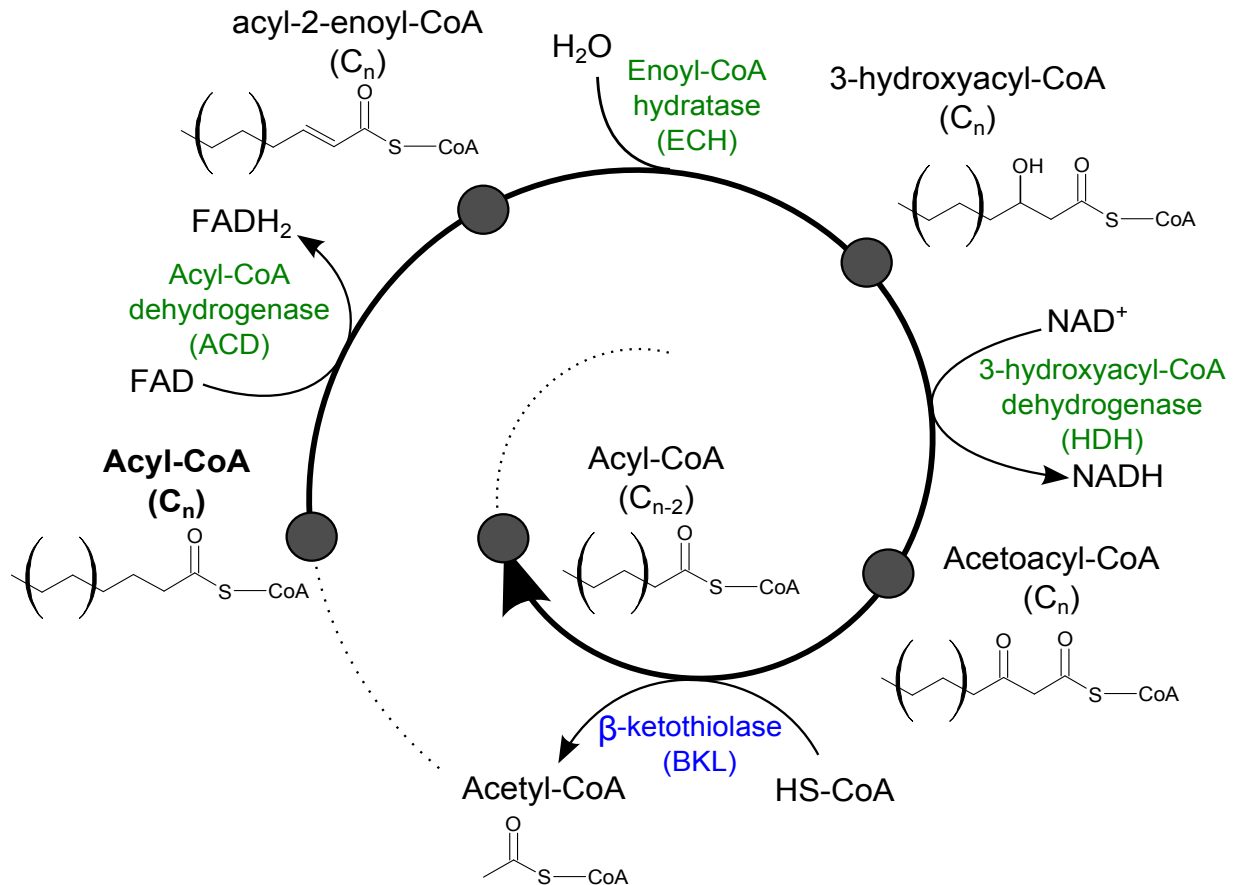


Figure S2. Phylogenetic tree of β -ketothiolase and its homologs.

The tree was constructed based on conserved blocks (total 227 positions) from 511 sequences. Each protein sequence on the tree is listed under its gene identifier (gi) in the NCBI's RefSeq database and the name of the source organism. Eukaryotic proteins and clades are marked by bold orange branches. The names of bacterial and archaeal proteins are colored according to their taxonomy, as follows:

Archaea:

Crenarchaeota
Euryarchaeota
Korarchaeota
Nanoarchaeota
Thaumarchaeota

Well-sampled bacterial phyla*:

Actinobacteria
Aquificae
Bacteroidetes/Chlorobi group
Chlamydiae/Verrucomicrobia group
Chloroflexi
Cyanobacteria
Deinococcus-Thermus
Firmicutes
Proteobacteria
Thermotogae

Poorly sampled bacterial phyla*:

Chrysiogenetes
Deferribacteres
Dictyoglomi
Elusimicrobia
Fibrobacteres/Acidobacteria group
Fusobacteria
Gemmatimonadetes
Nitrospirae
Planctomycetes
Spirochaetes
Synergistetes
Thermodesulfobacteria
unclassified Bacteria

* The terms "poorly sampled" and "well-sampled" refer to the number of full genome sequences available for the respective phyla.

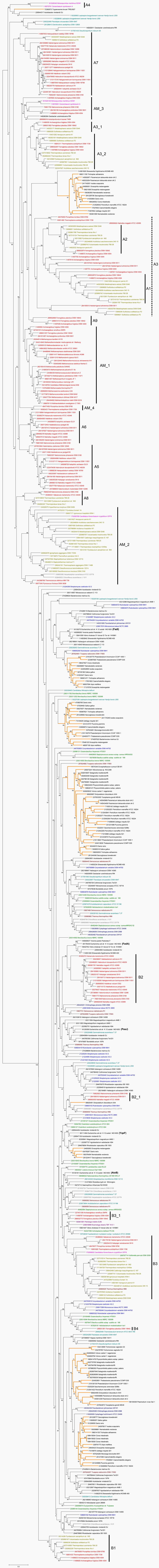


Figure S3. Reactions of fatty acids biosynthesis in bacteria (modified from the charts in KEGG and MetaCyc databases). The gene-based enzyme names in *Escherichia coli* and/or *Bacillus subtilis* are in red; the same names are used in the **Table S2**. Abbreviations: ACP, acyl-carrier protein, CoA, coenzyme A.

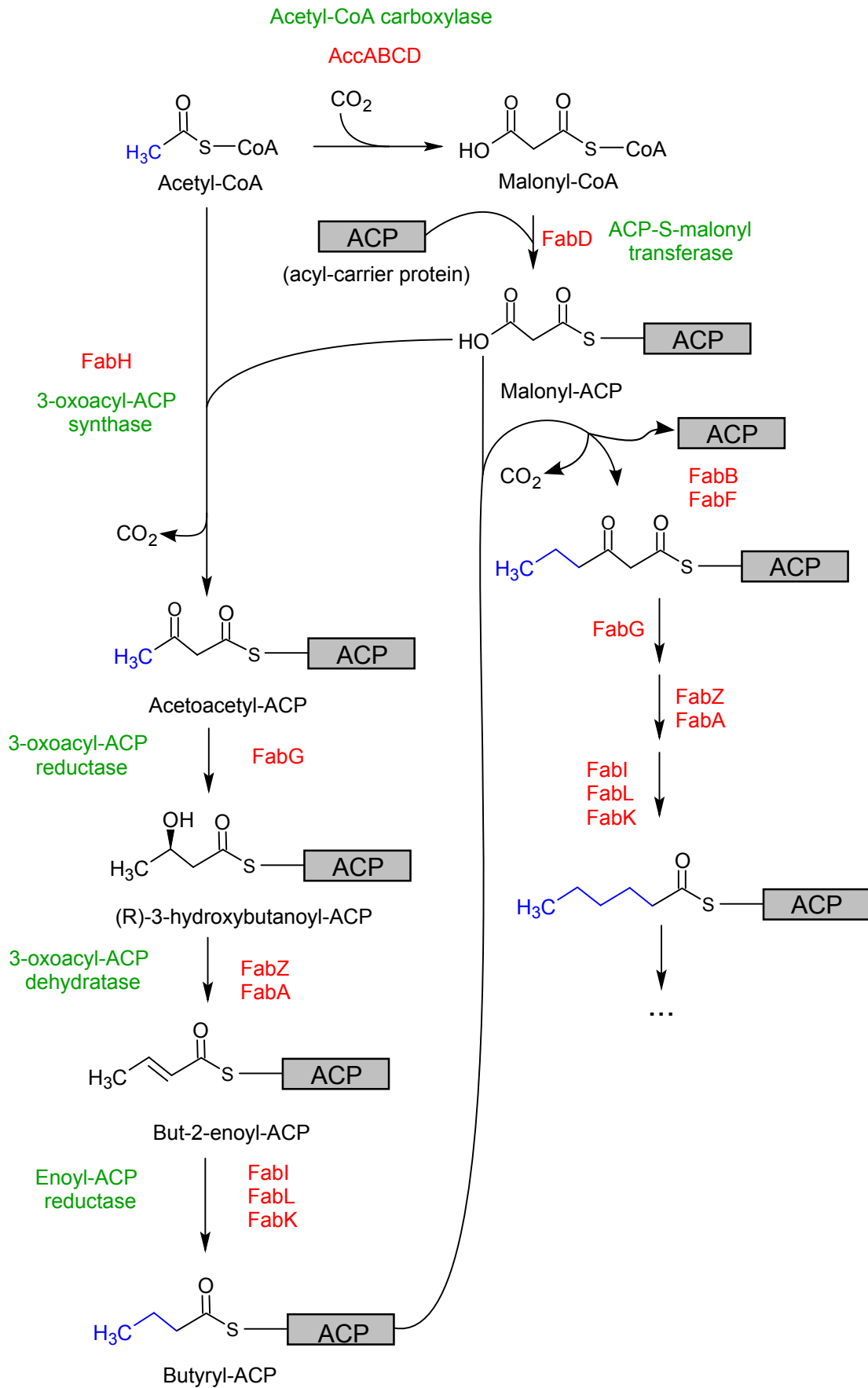


Figure S4. Phylogenetic tree of the 4-hydroxybutyryl-CoA dehydratase and its homologs (COG2368) from bacteria and archaea.

The tree was constructed based on conserved blocks (total 227 positions) from 511 sequences. Each protein sequence on the tree is listed under its gene identifier (gi) in the NCBI's RefSeq database and the name of the source organism. The names of bacterial and archaeal proteins are colored as in Fig. S2.

