Supplemental Data

Relation of the enzymes and regulator to known proteins

The gene clusters are compared in Fig. S2, showing that the core conserved module consists of the genes for the coenzyme B12-dependent mutase (*iaaGH*), the hydantoinase-like enzyme (*iaaCE*) and the acyl-CoA dehydrogenase (*iaaF*). The mutases from these four genomes even seem to define a new functional subclass of these enzymes, judging from their close clustering in a common subbranch (Fig. S3) that is distinct from the subbranches containing the biochemically defined methylmalonyl-CoA-, isobutyryl-CoA- and ethylmalonyl-CoA mutases (2) as well as from another subbranch containing a putative alkylsuccinyl-CoA mutase from the recently sequenced *Desulfatibacillum alkenivorans* that is possibly involved in anaerobic alkane metabolism (1, 3, 8). There is no obvious gene for an epimerase, which is usually associated with methylmalonyl-CoA mutase, but this function may be taken over by another epimerase (2). The acyl-CoA dehydrogenases (IaaF) encoded in the four gene clusters form a similarly close subbranch in the phylogenetic tree obtained from BLAST searches of these enzymes (data not shown), suggesting that the enzymes are specifically adapted to the reaction catalysed, which may include the proposed simultaneous decarboxylation function (Fig. 5).

The sequences of the hydantoinase-like enzymes (IaaCE) are only close between "*A. aromaticum"* and *R. palustris*, whereas the orthologs of the other two strains are further remote on the phylogenetic tree. Therefore, the evolutionary origins of the *iaaCE* genes may diverge, whereas the *iaaF* and *iaaGH* genes are most probably derived from the same origin in all four species. Only the gene cluster of "*A. aromaticum"* harbours the small *iaaD* gene between *iaaC* and *iaaE*, suggesting that the partial gene duplication involved in its generation took place very recently and that the two large subunits are sufficient for the function of the

enzyme. The significance of IaaD in "*A. aromaticum"* can only be assessed by further biochemical investigation.

The gene clusters of the three bacterial strains contain genes for putative periplasmic IAA-binding proteins of ABC-type transporters (*iaaM*) and CoA ligases (*iaaB*) potentially activating 2-(2'-aminophenyl)succinate (or 2-aminobenzylmalonate) for further conversion reactions. In both cases, the gene products of "*A. aromaticum"* and *R. palustris* are closely related, whereas those of strain NaphS2 are further remote (Tab. S1). Moreover, further genes for potentially matching ABC transport systems are present in the gene clusters of *R. palustris* and strain NaphS2, but not in "*A. aromaticum"* (Fig. S2).

The *iaaIJK* genes coding for the three subunits of a molybdoenzyme of the xanthine dehydrogenase family are also conserved in the gene clusters of the three bacteria, but not the archaeal species. Their sequences form a separate common subbranch within the enzyme family, which is close to the subbranch of 4-hydroxybenzoyl-CoA reductases that are involved in anaerobic phenol and 4-hydroxybenzoate metabolism (5, 6). The two small subunits of the enzyme in strain NaphS2 appear to be fused to a single protein (Fig. S2). The function of this molybdenum enzyme may be taken over by a putative tungsten-containing enzyme of the aldehyde:ferredoxin oxidoredutase family (AOR) in *F. placidus*, which is encoded in the gene cluster together with its putative electron-accepting ferredoxin (Fig. S2).

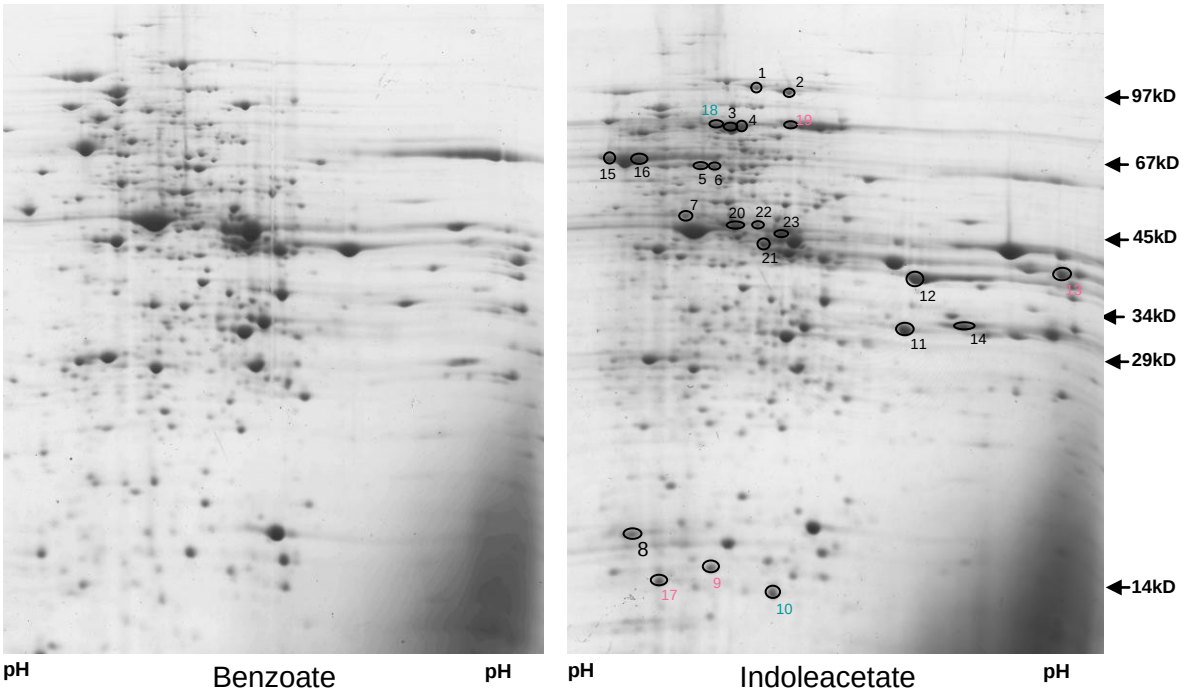
A closely related CoA-transferase of family III (4) is encoded in the operons of "*A. aromaticum"* and *R. palustris* (*iaaL*), and the same is observed for the enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase fusion protein (*iaaP*) and the thiolase (*iaaA*). These latter genes are missing in the gene clusters of strain NaphS2 and *F. placidus*, but *F. placidus* contains alternative genes for a two-subunit thiolase (*iaaA1A2*) similar to a thiolase involved in anaerobic toluene metabolism (7) and a 3-hydroxyacyl-CoA dehydrogenase (*adh*) in the gene cluster (Fig. S2).

Finally, all gene clusters of the bacterial strains contain a gene for a regulator of the GntR family (*iaaR*), which is most likely involved in IAA-specific induction of gene expression. The significance of further potential regulators encoded in the gene clusters, e.g. the TetR-family protein IaaQ in "*A. aromaticum"* or the Rpe_3221 gene product of *R. palustris*, is not clear. Note that the gene cluster of *F. placidus* also contains a gene for a potential regulator (Ferp_2332) of a still unnamed archaeal family (DUF128).

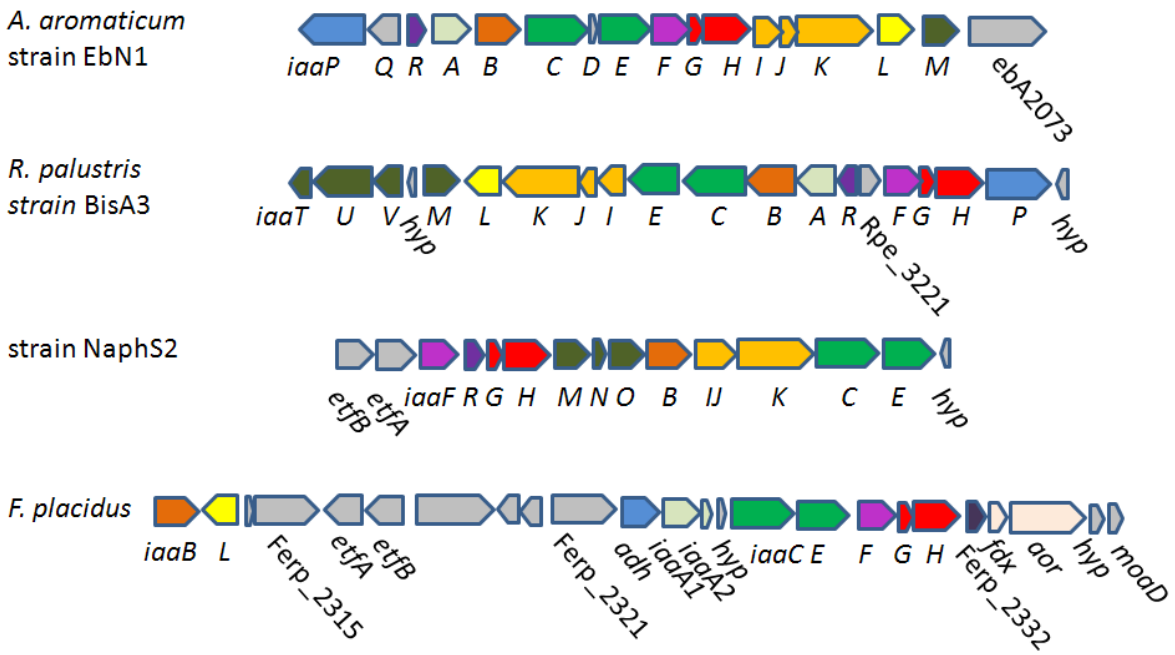
The only "missing" genes in strain NaphS2 and *F. placidus* are those coding for CoA-thioester-forming enzymes and beta-oxidation functions, which can easily be supplied by redundant genes located somewhere else in the genomes. Actually, genes coding for a CoA ligase and a family III CoA-transferase (Ferp_2312 and 2313) are located quite close to the *iaa*-related gene cluster in the *F. placidus* genome (Fig. S2). Moreover, the presence of these genes in only a few isolated and widely separated bacterial strains (e.g. only one of over 10 sequenced strains of *R. palustris*) indicates their distribution via lateral gene transfer. It seems that a few key enzymes, such as the molybdoenzyme (possibly replaceable by a tungsten enzyme), the coenzyme B12-containing mutase and the IaaF-type acyl-CoA dehydrogenase constitute the core of the pathway, whereas the other functions may be taken over by proteins from different origin during evolution.

1. **Callaghan, A. V., B. E. Morris, I. A. Pereira, M. J. McInerney, R. N. Austin, J. T. Groves, J. J. Kukor, J. M. Suflita, L. Y. Young, G. J. Zylstra, and B. Wawrik.** 2012. The genome sequence of *Desulfatibacillum alkenivorans* AK-01: a blueprint for anaerobic alkane oxidation. Environ Microbiol. **14:**101-113.
2. **Erb, T. J., J. Rétey, G. Fuchs, and B. E. Alber.** 2008. Ethylmalonyl-CoA mutase from *Rhodobacter sphaeroides* defines a new subclade of coenzyme B12-dependent acyl-CoA mutases. J. Biol. Chem. .**283:**32283-32293.

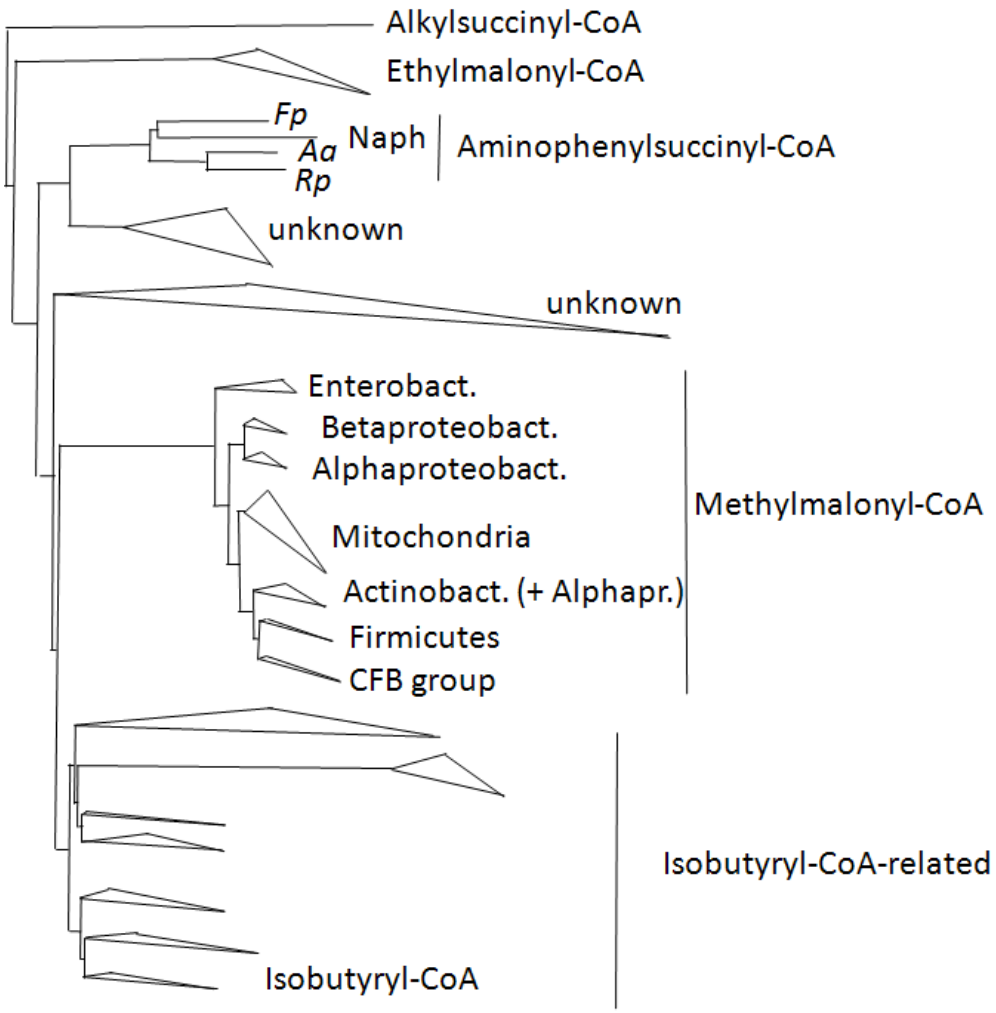
3. **Grundmann, O., A. Behrends, R. Rabus, J. Amann, T. Halder, J. Heider, and F. Widdel.** 2008. Genes encoding the candidate enzyme for anaerobic activation of *n*-alkanes in the denitrifying bacterium, strain HxN1. Environ Microbiol. **10:**376-385.
4. **Heider, J.** 2001. A new family of CoA-transferases. FEBS Lett. **509:**345-349
5. **Heider, J., M. Boll, K. Breese, S. Breinig, C. Ebenau-Jehle, U. Feil, N. Gad'on, D. Laempe, B. Leuthner, M. Mohamed, S. Schneider, G. Burchhardt, and G. Fuchs.** 1998. Differential induction of enzymes involved in anaerobic metabolism of aromatic compounds in the denitrifying bacterium *Thauera aromatica.* Arch. Microbiol. **170:**120-131.
6. **Heider, J., and G. Fuchs.** 1997. Anaerobic metabolism of aromatic compounds. Eur. J. Biochem. **243:**577-596.
7. **Leuthner, B., and J. Heider.** 2000. Anaerobic toluene catabolism of *Thauera aromatica*: the bbs operon codes for enzymes of beta oxidation of the intermediate benzylsuccinate. J. Bacteriol. **182:**272-277.
8. **Wilkes, H., R. Rabus, T. Fischer, A. Armstroff, A. Behrends, and F. Widdel.** 2002. Anaerobic degradation of n-hexane in a denitrifying bacterium: further degradation of the initial intermediate (1-methylpentyl)succinate via C-skeleton rearrangement. Arch. Microbiol. **177:**235-243.



SI Fig. 1. Comparative two-dimensional gel electrophoresis of soluble proteins. Left panel, cells grown on benzoate and nitrate. Right panel, cells grown on IAA and nitrate. 230 µg of protein each was applied. Isoelectric focusing pH 3-10, SDS-PAGE 11 %. The circled spots indicate the main IAA induced proteins. Spots 9, 13, 17, and 19 that were also present in cells grown on 2-aminobenzoate (not shown here), but were missing in benzoate grown cells, were not considered.



SI Fig. 2. Comparison of conserved clusters of *iaa*-related genes in "*A. aromaticum*" strain EbN1, *R. palustris* strain BisA3, deltaproteobacterium NaphS2, and *F. placidus*. The conserved genes are named after the orthologs in the "*A. aromaticum*" gene cluster. The same colors indicate conserved functions of the gene products. Some other genes are indicated that may be functionally related to IAA metabolism and are absent in the "*A. aromaticum*" gene cluster, but present in those of other organisms: *iaaTUV* and *iaaNO* code for the membrane and ATPase subunits of ABC transporters with IaaM encoding the putative periplasmic binding protein, *etfAB* code for the subunits of ETF, *iaaA1*/*A2* code for the subunits of an alternative thiolase, *fdx* and *aor* code for a tungsten-containing aldehyde:ferredoxin oxidoreductase and its electron acceptor, and the *moaD* gene product is needed for tungsten cofactor synthesis. *hyp*: hypothetical genes.



SI Fig. 3. Phylogenetic tree of the large subunits of coenzyme B12-dependent mutases. The tree was generated from 500 sequences retrieved from a BLAST analysis of IaaH, using the tree building function provided by the program. The mutase families are indicated by their respective substrates. A simplified version of the main branches is shown, in which positions of enzymes with biochemically defined and hypothetic functions are indicated. The subbranches of the methylmalonyl-CoA mutases contain predominantly sequences from organisms affiliated to the indicated phylogenetic groups. Such phylogenetic preferences are not visible in any other subbranch of the tree. Note that isobutyryl-CoA mutase activity has only been demonstrated with enzymes belonging to the small subgroup indicated, whereas the rest of this cluster was designated as related enzymes. Additionally, the functions of alkylsuccinyl-CoA and aminophenylsuccinyl-CoA mutases are not yet confirmed biochemically. *Fp: Ferroglobus placidus*, Naph: strain NaphS2, *Aa: "Aromatoleum aromaticum", Rp: Rhodopseudomonas palustris.*

SI Table S1. Presence of orthologous genes in IAA-related gene clusters of various microorganisms. + indicates presence of an orthologous gene, (+) the presence of a gene coding for a similar function. Further differences are directly indicated. For the orthologs, the similarity values of the gene products from the BLAST analysis are indicated in % (identical/similar amino acids).

Gene	Predicted function	*"Aromatoleum aromaticum"* EbN1	*Rhodopseudo-monas palustris* BisA3	Deltaproteobacterium Strain NaphS2	*Ferroglobus placidus*
iaaA	Thiolase	+	+ 72/84 %		(+)
iaaB	CoA-Ligase	+	+ 54/68 %	+ (39/56 %)	
iaaC	Hydantoinase A	+	+ 75/83 %	+ 37/54 %	+ 40/58 %
iaaD	hypothetic	+			
iaaE	Hydantoinase B	+	+ 84/93 %	+ 29/47 %	+29/48 %
iaaF	Acyl-CoA DH	+	+ 72/84 %	+ 45/64 %	+ 49/67 %
iaaG	Mutase B12-binding	+	+ 59/79 %	+ 57/74 %	+ 50/71 %
iaaH	Mutase large subunit	+	+ 69/80 %	+ 53/73 %	+ 58/77 %
iaaI	Mo-enz. FAD-bdg	+	+ 57/74 %	+ (fused w/ J)	
iaaJ	Mo-enz. FeS-bdg	+	+ 68/80 %	+ (50/65 %) (fused w/ I)	
iaaK	Mo-enz. large	+	+ 74/84 %	+ (41/58 %)	(W-enzyme)
iaaL	CoA-transferase	+	+ 68/78 %		
iaaM	binding protein (*livK*)	+	+ 64/74 %	+	
iaaP	ech_adh fusion	+	+ 65/77 %		(only *adh*)
iaaQ	Regulator TetR	+			
iaaR	Regulator GntR	+	+ 50/66 %	+	
etfA	Electron-transfering flavoprotein A			+	+
etfB	Electron-transfering flavoprotein B			+	+
	Transport ATPase LivF family		+	+	
	Transport ATPase LivG family		+ fused w/ *livM*		
	LivM type transport membrane protein		+ fused w/ *livG*	+	
	LivH type transport membrane protein		+		
	Sorbosone DH	+			