On the enigma of glutathione dependent styrene degradation in *Gordonia rubripertincta* CWB2

Thomas Heine\(^1,#\), Juliane Zimmerling\(^1\), Anne Ballmann\(^1\), Sebastian Bruno Kleeberg\(^1\), Christian Rückert\(^2\), Tobias Busche\(^2\), Anika Winkler\(^2\), Jörn Kalinowski\(^2\), Ansgar Poetsch\(^3,4\), Anika Scholtissek\(^1\), Michel Oelschlägel\(^1\), Gert Schmidt\(^5\) and Dirk Tischler\(^1,6,#\)

\(^1\)Institute of Biosciences, TU Bergakademie Freiberg, Leipziger Str. 29, Freiberg, Germany.
Technologieplattform Genomik, Centrum für Biotechnologie (CeBiTec), Universität Bielefeld, Universitätsstr. 27, Bielefeld, Germany.

Plant Biochemistry, Ruhr University Bochum, 44801 Bochum, Germany

School of Biomedical and Healthcare Sciences, Plymouth University, Plymouth PL4 8AA, UK

Institut für Keramik, Glas- und Baustofftechnik, TU Bergakademie Freiberg, Agricolastr. 17, Freiberg, Germany.

Microbial Biotechnology, Ruhr University Bochum, 44801 Bochum, Germany

Running title: Styrene degradation via glutathione-S-transferase
Among bacteria, only a single styrene specific degradation pathway has been reported so far. It comprises the activity of styrene monooxygenase, styrene oxide isomerase and phenylacetaldehyde dehydrogenase yielding phenylacetic acid as central metabolite. The alternative route comprises ring-hydroxylating enzymes and yields vinyl catechol as central metabolite, which undergoes meta-cleavage. This was reported to be unspecific and also allows the degradation of benzene derivatives. However, some bacteria had been described to degrade styrene but do not employ one of those routes or only parts of them. Herein we describe a novel “hybrid” degradation pathway for styrene located on a plasmid of foreign origin. As putatively also unspecific, it allows metabolizing chemically analogous compounds (e.g. halogenated and/or alkylated styrene derivatives). *Gordonia rubripertincta* CWB2 was isolated with styrene as sole source of carbon and energy. It employs an assembled route of the styrene side chain degradation and isoprene degradation pathways that also funnels into phenylacetic acid as central metabolite. Metabolites, enzyme activity, genome, transcriptome and proteome data reinforce the observation and allow to understand this biotechnologically relevant pathway which can be used for the production of ibuprofen.

Degradation of xenobiotics by bacteria of high interest for bioremediation, but also as involved enzymes are potential catalysts in biotechnological applications. This study reveals a novel degradation pathway for the hazardous organic compound styrene in *Gordonia rubripertincta* CWB2. It is an impressive illustration of horizontal gene transfer, which enables novel metabolic capabilities. This study presents glutathione-dependent styrene metabolization in an (actino-)bacterium. Further,
genomic background of the ability of strain CWB2 to produce ibuprofen is demonstrated.

INTRODUCTION

Styrene is a monoaromatic compound that naturally occurs as a component of tar, volatile and oily substances from plants and food, but can also be produced by microorganisms. Styrene is of high relevance in industry and produced in million tonne scale causing substantial anthropogenic release. This is problematic as it is hazardous for living organisms (1–3). Due to the disposability, it is corollary that organisms evolved strategies to detoxify and/or use styrene as a source of energy and carbon (3–5).

Styrene can be channelled through different unspecific degradation pathways due to relaxed substrate specificity of the respective enzymes (see supplemental material and 3 for details). However, only one styrene specific degradation pathway is known and seems to be favoured by microorganisms under aerobic conditions (3, 5, 6). This upper degradation pathway is initiated by oxidation of the vinyl side chain. A styrene monooxygenase (SMO) produces (S)-styrene oxide, which is converted by a membrane bound styrene oxide isomerase (SOI) to phenylacetaldehyde. A phenylacetaldehyde dehydrogenase (PAD) oxidizes the aldehyde to phenylacetic acid (PAA) (7). PAA is a central catabolite and metabolized in the so-called lower degradation pathway, which is present in about 16% of all genome-sequenced microorganisms (8, 9). That route has been described for several proteobacteria (Pseudomonas, Xanthobacter, Sphingopyxis), actinobacteria (Rhodococcus, Corynebacterium) and fungi (Exophiala) (reviewed by 3). Rhodococcus sp. ST-10 has an incomplete degradation cluster lacking the SOI,
while still being able to use styrene as sole source of carbon and energy (10–13). This gene cluster comprises the SMO and a putative (partial) open reading frame (ORF), designated as “ORF3”. It was hypothesized that the SOI can be bypassed by chemical conversion of styrene oxide to phenylacetaldehyde or enzymatically (12–14). However, chemical conversion is unlikely and no probable enzymes were identified (3), thus the degradation pathway for strain ST-10 remains unclear.

The genus *Gordonia* is known to be a versatile degrader of aromatic compounds (15, 16) and *Gordonia rubripertincta* CWB2 in particular is able to metabolize styrene and related compounds. As previously described, strain CWB2 was obtained from a soil sample and separated via styrene-enrichment culture (17–19). Moreover, it was shown that it is able to produce ibuprofen from 4-isobutyl-α-methylstyrene in a co-metabolic process (17, 18). Oelschlägel *et al.* 2015 reported that other styrene degraders are not capable to catalyze this reaction and therefore proposed substantial differences in the respective enzymatic cascades. Interestingly, *Gordonia rubripertincta* CWB2 has a cluster that is homolog to the partial one of strain ST-10.

In this study, we identified the complete gene cluster, which enables styrene degradation in strain CWB2. Therefore, we studied the transcriptome and the proteome under styrene exposure. Further, we measured the activity of key enzymes to clarify the root of the metabolic potential of *Gordonia rubripertincta* CWB2. Moreover, this gene cluster seems to be alien as it is located on a plasmid and assigned as genomic island if compared to the rest of the genome. It embodies a hybrid of several homolog epoxide and aromatic compound degradation clusters from different actinobacteria.
RESULTS

Identification and annotation of gene clusters associated with styrene degradation. Genes that might be involved in styrene degradation in G. rubripertincta CWB2 were identified and annotated by homology search using the BLASTP algorithm (20) on the non-redundant protein database or the UniProtKB database (NCBI). The annotation of the putative styrene degradation cluster with respect to the closest (characterized) homolog is listed in Table 2 and Dataset S1. A 32424-bp cluster with 36 putative open reading frames (orf) was identified on the plasmid pGCWB2 (~ 100 kbp), which is framed by a styrene monoxygenase (GCWB2_24100) and phenylacetaldehyde dehydrogenase (GCWB2_23925). Interestingly, the average GC content of this cluster is 62.11% and thus 5% lower as for the whole genome. The GC content of the whole plasmid is 3% lower compared to the chromosome. Besides a high amount of hypothetical proteins, pGCWB2 contains 4 transposase-family proteins, 2 integrase-like proteins, one relaxase-like protein and one type IV secretory system as an inventory for gene mobility. Genomic island analysis on the whole genome illustrates that at least parts of this styrene degradation cluster have foreign origin (Fig. 1).

The cluster can be separated into 4 subclusters comprised as follows: cluster S1 contains a styrene monoxygenase, which is known to initiate the styrene degradation at the vinyl side chain. The closest characterized homolog of this protein was found in Rhodococcus sp. ST-10 (StyA – 86% identity at amino acid level; StyB – 82%). A phylogenetic analysis of the amino acid sequence classifies it as an E1-type SMO (Fig. S2). The closest match within Gordonia species is a putative monoxygenase from Gordonia polyisoprenivorans NBRC 16320 (GAB22407 – 45%; GAB22406 – 39%). However, also the genetic organisation refers to a close relation...
to the *Rhodococcus* cluster as the partial sequence of ORF_3 from strain ST-10 shows 67% identity to GCWB2_24090. Transmembrane domain prediction (TMHMM) of GCWB2_24090 identified 4 transmembrane helices what classifies this protein as hypothetical membrane associated. Beyond that, no characterized homologs and no known domains are present in the database for this ORF. It has to be mentioned that no styrene oxide isomerase (*styC*) gene was found on the genome of strain CWB2.

The second cluster S2 embeds 7 hypothetical proteins (GCWB2_24085 - GCWB2_24055). Two of them are presumably soluble and the others are annotated as membrane proteins, while each has one transmembrane domain. Members of these clusters appear to be rare within the database and are predominantly present in rhodococci. *Rhodococcus opacus* 1CP owns a homologous cluster downstream of its StyABCD cluster (WP_065493732 - WP_045063326; 56 - 68%). Further, *Gordonia* sp. i37 has a homologous gene cluster in the neighbourhood of a recently recorded isoprene degradation cluster (contig257: WP_079929940 - WP_079929944, contig258: OPX14963 - OPX14964; 56 - 74%) (21).

Cluster S3 encodes for proteins that might be involved in glutathione and isoprene metabolism. GCWB2_24050 and GCWB2_24045 show highest identity to a glutamate-cysteine ligase GshA (P9WPK7 – 33%) and a glutathione synthetase GshB (P45480 – 50%) followed by a putative coenzyme A-disulfide reductase (CoA-DSR). The other genes of this cluster (GCWB2_24050 - GCWB2_24010) encode for a putative MarR-like transcriptional regulator, a coenzyme A-transferase, a dehydrogenase, a glutathione S-transferase, a disulfidebond oxidoreductase and an aldehyde dehydrogenase. The closest characterized homolog of the latter gene product is the phenylacetaldehyde dehydrogenase from *Pseudomonas fluorescens* ST (O06837 – 36%). Cluster S3 can also be found in *Aeromicrobium* sp. Root495.
wherein a styrene monooxygenase is located between the CoA-DSR and the transcriptional regulator. The same is true for *Nocardioides* sp. Root240 except for a 13983-bp insertion right after the styrene monooxygenase gene. Interestingly, the closest characterized homologs of GCWB2_24025 and GCWB2_24020 can be found in *Rhodococcus* sp. AD45 (WP_045063294 – 59%; WP_045063292 – 49%) and are known to be a functional part of an isoprene degradation cluster which is located on a megaplasmid (300 kbp) (22). In addition, homologs of other genes from cluster S3 can be found on this plasmid, even in a similar arrangement of parts from this cluster. However, strain AD45 does not encode for a styrene monooxygenase on its genome and on the other hand strain CWB2 is lacking the isoprene monooxygenase. Homologous proteins from cluster S3 were recently found in *Gordonia* sp. i37 next to a homologous to cluster S2 (21). The 13983-bp insertion of *Nocardioides* sp. Root240 comprises a putative mce operon, a cluster of membrane proteins whose specific function is unclear.

The putative styrene degradation cluster is completed by a fourth gene set S4 (GCWB2_24005 - GCWB2_23925) that encodes for proteins which are required for the lower styrene degradation pathway (phenylacetic acid catabolism). The proposed pathway of strain CWB2 is displayed in Figure 3b. Besides some regulatory elements and a partial gene (GCWB2_24000), this cluster is homolog to that of *Rhodococcus opacus* 1CP and can also be found in *Gordonia soli* NBRC 108243 as well as *Gordonia* sp. i37 (contig69). The AraC-like transcriptional regulator (GCWB2_24005) shows homology to regulators of strain 1CP, that are located bevor the SMO and behind the PAA degradation cluster (ANS32446 – 46%; WP_061046101 – 41%). However, in contrast to strain CWB2 the upper and lower styrene degradation pathways are not associated in this strain. Here the genes for the conversion of styrene to phenylacetic acid are located on a plasmid pR1CP1 (NZ_CP009112),
whereas the subsequent metabolism is encoded on the chromosome (NZ_CP009111). The second transcriptional regulator (GCWB2_23980) belongs to the TetR-family. Cluster S4 is terminated by a phenylacetaldehyde dehydrogenase (GCWB2_23925) that is highly similar to StyD from *Rhodococcus* sp. ST-5 (BAL04135 – 76%). So far, a comparable genetic environment of styrene monooxygenases can only be found in *Aeromicrobium* sp. Root495 and in five strains of *Nocardioides* sp. (Root79, Root190, Root240, Root614, Root682). All of them were isolated from *Arabidopsis* root microbiota (23).

As already mentioned, the whole cluster embeds 3 HTH-type regulators, which are known to respond to aromatic compounds (24). They are not similar to the regulation machinery as described for pseudomonads (3). Besides this, the genome of strain CWB2 was examined for other genes and clusters that might enable degradation of styrene or metabolites. As no styrene oxide isomerase gene (*styC*) is located on the genome of strain CWB2, it might be possible that this degradation step is bypassed by a styrene oxide reductase (SOR) and a phenylacetaldehyde reductase (PAR). So far, there is no enzyme characterized that has SOR activity and thus no comparison with strain CWB2 on DNA level is possible. However, two putative ORFs (GCWB2_12345 – 70%; GCWB2_18410 – 35%) show similarity to the PAR of *Rhodococcus* sp. ST-10 (BAD51480). Interestingly, the latter one is part of a cluster with 8 ORFs (GCWB2_18380 - GCWB2_18415) that is homolog to one in *Gordonia* sp. TY-5 (BAD03956 - BAD03963; 87 - 96%) (25). It comprises a chaperonin, a putative alcohol dehydrogenase, two hypothetical proteins and a putative propane monooxygenase. The monooxygenase has also resemblance with a propene monooxygenase from *Mycobacterium* sp. M156 (28 - 38%) (26) but also a methane monooxygenase from *Methylococcus capsulatus* Bath (29 - 34%).
These binuclear iron monooxygenases are able to epoxidize styrene. Homologous clusters can be found in several actinobacteria (28–30), for instance in *R. opacus* 1CP and *Gordonia* sp. i37 (21). The other putative PAR is not part of a cluster. In addition, several cytochrome P450 monooxygenases can be found on the CWB2 genome, which might also be able to perform the epoxidation of styrene.

**Transcriptome and proteome analysis of the styrene degrader CWB2.** The transcriptome and proteome of strain CWB2 was analysed to reveal the global profile of genes and proteins that are involved in styrene metabolism. Therefore, fructose-grown cultures served as reference condition. The transcriptome output is summarized in Table S4.

If assuming a threshold of ≥ 1.5 (M-value), then 2.5% of the genes of strain CWB2 are overexpressed under styrene exposure and 30% of these are located on the plasmid (Fig. S5). It is known, that the transcriptome as the total amount of mRNA does not necessarily reflect the total amount of proteins abundance in the cell. However, we were able to identify 3691 proteins in the proteome of strain CWB2. If assuming a threshold of 1.5 (log2 ratio), then 7% of the proteins where highly abundant, when strain CWB2 was grown on styrene (Fig. S6). The gene cluster, which is framed by the styrene monooxygenase and the phenylacetaldehyde dehydrogenase, is highly upregulated on transcriptome (increased on average 7-fold) as well as proteome level (increased on average 6.6-fold) (Table 2 and Dataset S1).

**Validation of enzymatic activity of selected members of the styrene degradation pathway.** After analysis of the genome, transcriptome and proteome of strain CWB2, we screened for enzyme activities that enable styrene degradation on different pathways. For that, crude extract from styrene grown biomass was prepared, proteins were separated and enriched by different chromatography
The activities were measured directly or indirectly on a spectrophotometer or by quantification of the products on the reverse phase HPLC (RP-HPLC), respectively (Table S5). It was possible to detect SMO activity in the crude extract and to enrich the enzyme 36-fold to an activity of 6.82 mU mg\(^{-1}\). Due to the missing SOI, it was proposed that the conversion of styrene oxide is bypassed by the activity of a SOR, which produces 2-phenylethanol. This is supposed to be converted to phenylacetaldehyde by a PAR. Only minor activity of a PAR with 2-phenylethanol was detected in crude extract. Higher activities were determined in crude extracts, when styrene oxide was applied as substrate. However, this might be due to activity of a GST, while residual glutathione (GSH) is present in the crude extract of strain CWB2. To further prove this assumption, crude extract was assayed for GST activity after supply of additional GSH. Thereby, a GST activity of 44.23 U mg\(^{-1}\) was reached for the conversion of styrene oxide (Fig. 2).

Further, crude extracts were assayed for vinylcatechol-2,3-dioxygenase and cis,cis-muconate cycloisomerase (MCI) to exclude other degradation pathways. However, there was no detectable activity for one of these enzymes.

Two putative SMOs of strain CWB2 were cloned and expressed for initial characterization. Of the two putative SMOs, only one was expressed and synthesised in an active form. It is part of the styrene degradation cluster (S1; GCWB2_24100) and produces (S)-styrene oxide with a specific activity of 0.42 ± 0.02 U mg\(^{-1}\). The SMO can be classified as E1-type SMO (Fig. S2). PADs are aldehyde dehydrogenases that catalyze the formation of the central intermediate phenylacetic acid. Two aldehyde dehydrogenases of strain CWB2 were recombinantly expressed in *E. coli*. Aldh1 which originates from the isoprene degradation cluster (S3) catalyzes the conversion of phenylacetaldehyde with an activity of 0.29 ± 0.01 U mg\(^{-1}\). StyD,
which is encoded in cluster S4 is 10-times slower and has an activity of $0.026 \pm 0.001$ U mg$^{-1}$.

**DISCUSSION**

Adaption of *G. rubripertincta* CWB2 to styrene exposure. Only few reports exist for *Gordonia* considering the metabolization of styrene (18, 31). Further, the limited amount of SMOs that are encoded on genomes of this genus indicate that styrene degradation is no common feature. *Gordonia rubripertincta* CWB2 is able to withstand and degrade high amounts of styrene ($520$ g m$^{-3}$ in $21$ h), even compared to other efficient styrene degraders (32, 33). Some bacteria produce surfactants, when they are exposed to hydrophobic substrates, to increase their accessibility (34–36). However, we found no indication that strain CWB2 exports biosurfactants into the media but it seems to have a hydrophobic cell surface, which improves substrate uptake. This is supported by a tendency to form agglomerates during growth in liquid media.

There is no complete prokaryotic transcriptome under styrene exposure available yet. So far, studies focused on the transcriptional regulation of styrene degradation and a small number of target genes, solely with respect to *Pseudomonas* strains (37–44). A proteome of *R. jostii* RHA1, which employs an unspecific styrene degradation route, is available (45). So far, only one system level proteome analysis for styrene degradation in *P. putida* CA-3 exists (46). It was the first time where all of the respective enzymes of the upper and lower degradation pathway were detected, when a strain was grown on styrene (46).

Omic analysis of strain CWB2 in this study outlines the biological background for its adaption to styrene as source of carbon and energy. This was found to be totally
different to so far characterized styrene degraders. Initially styrene has to be
imported into the cell. The only specific styrene transporter StyE was found in
pseudomonads (47). However, the *styE* gene is not encoded in most other styrene
degraders and thus, other transport mechanisms as well as diffusion have to be
considered (3, 46). Cluster S2 of strain CWB2 contains several membrane proteins
that are highly upregulated and the same cluster is also present in the styrene
degraders *R. opacus* 1CP (Fig. 1). Thus, it is likely that these proteins might also be
involved in substrate transport or cell membrane adaption. Interestingly, *Gordonia* sp.
i37 owns a similar cluster in proximity of an isoprene degradation cluster (21). However, there are no characterized homologs available in the database and thus
the specific function of these proteins remains unclear.

**Strain CWB2 merged clusters to form a hybrid that enables styrene
degradation.** The genetic organisation of the putative styrene degradation cluster
compared to other clusters with homolog proteins can be found in Figure 1. The
“classical” styrene degradation cluster of *Pseudomonas* sp. Y2 differs to
*Rhodococcus* clusters as well as the recently reported cluster of *S. fribergensis* Kp5.2
(17, 38, 48). Thus, it is obvious that the arrangement and regulation is variable
among different organisms. The styrene degradation cluster of strain CWB2 is highly
upregulated on mRNA and protein level, when strain CWB2 grows on styrene (Table
2 and Dataset S1). In *Pseudomonas* putida CA-3 the SMO and PAD were the most
abundant proteins (46). In strain CWB2 they are also highly upregulated but in a
comparable range to the rest of the genes and proteins of this cluster. It can be seen
that the transcriptional regulators (GCWB2_24035; GCWB2_23980) are less
expressed and synthesized. It is also obvious that regulation of gene expression
differs in strain CWB2 as no StyR/StyS homolog is associated to this cluster.
However, further studies with different inducers are necessary to clarify the regulation.

There is no evidence that strain CWB2 performs direct ring cleavage of styrene or activation by an epoxide hydrolase as the respective parts of these pathways are not present on the genome or upregulated in the transcriptome or proteome, when cultivated on styrene (supplemental material). In contrast, initial epoxidation of styrene was found to be catalyzed by a SMO. Enzyme activity was detected in crude extracts of styrene-grown cells and the SMO was successfully enriched by ion-exchange chromatography and hydrophobic interaction chromatography (Table S5). The SMO (GCWB2_24100; ASR05591) was recombinantly expressed and purified. The specific epoxidation activity is about 0.4 U mg$^{-1}$ and thus higher than for most other characterized SMOs (48–50). However, epoxidation of styrene is usually the rate-limiting step due to the relative low activity of the SMOs (51).

The SMO is part of cluster S1, which is highly similar to the partial styrene degradation cluster of *Rhodococcus* sp. ST-10. Toda and co-workers proposed chemical conversion of styrene oxide or the cooperation of a styrene oxide reductase (SOR) and phenylacetaldehyde reductase (PAR) (12, 13) as no SOI is present in this strain. However, previous as well as this study indicate that this is rather unlikely, as we detected only minor SOR and PAR activity in the crude extract of strain CWB2. Both assumptions would not explain fast degradation of styrene as found in these strains (3, 52, 53). However, strain ST-10 accumulated the epoxide when incubated with styrene and thus it remains to be shown if the rest of the genes are also homolog to the styrene degradation cluster in strain CWB2.

To circumvent this missing link of enzymatic styrene oxide isomerization, strain CWB2 seems to have incorporated a cluster (S3), which is very similar to ones from
Aeromicrobium sp. Root495 and Nocarioides sp. Root240. Interestingly, both were isolated at the same site (23) and both clusters are as well closely located to a styrene monooxygenase in these strains. The genes of cluster S3 may originate from an isoprene degradation cluster as found on a megaplasmid in Rhodococcus sp. AD45 (AJ249207) but also in Gordonia sp. i37 (21, 54). Actinobacteria from the genera Mycobacterium, Rhodococcus and Gordonia were constantly detected in different environments as isoprene degraders (21, 54, 55). Rhodococcus sp. AD45 initially epoxidizes isoprene by the activity of an isoprene monooxygenase. Then it uses a glutathione S-transferase to convert the epoxide to a glutathione-alcohol adduct, which is further metabolized by a dehydrogenase to form an aldehyde and subsequently an acid (22, 54, 56–59). Remarkably, strain AD45 is also able to metabolize styrene but has no SMO (56). Derived from these observations, it might be possible that styrene is also channelled through the isoprene degradation pathway in G. rubripertincta CWB2. Further, strain CWB2 owns genes that are necessary for glutathione synthesis and reduction in cluster S3 (60). Interestingly, strain CWB2 does not possess an isoprene monooxygenase and has no ability to catabolize isoprene (Table 1).

Styrene oxide is channelled into a novel glutathione dependent degradation pathway. To proof, whether glutathione dependent metabolization occurs, we assayed crude extract from styrene grown cells for GST activity with (S)-styrene oxide as substrate. We found that the epoxide was degraded fast with an activity of 44 U mg\text{crude extract}^{-1} (Fig. 2). Only minor activity was detected when no reduced glutathione was added to the reaction.

Therefore, we propose a novel degradation pathway for styrene via initial epoxidation by a SMO to (S)-styrene oxide and addition of glutathione by the GST
StyI (Fig. 3a). The resulting (S)-(1-Phenyl-2-hydroxyethyl) glutathione (CAS: 64186-97-6) will be further converted by the dehydrogenase StyH to (S)-(1-Phenyl-2-acetaldehyde) glutathione and (S)-(1-Phenyl-2-acetic acid) glutathione. It might be possible that the phenylacetaldehyde dehydrogenase (PAD) and/or the aldehyde dehydrogenase (Adh1) are also involved in this step as both showed activity with phenylacetaldehyde (61). As the glutathione adduct is not easily accessible it has to be verified if the Adh1 and the PAD can also catalyze this reaction. It was shown that the aldehyde dehydrogenase of cluster S3 are induced in strain AD45 but no specific role had been ascribed (22). Subsequently the glutathione is removed from the adduct what might occur by the activity of StyJ and StyG (58, 62, 63). The product of this process will be phenylacetic acid or phenylacetyl-CoA, which will be degraded via several enzymes from cluster S4 to yield acetyl-CoA and succinyl-CoA (Fig. 3b).

We suppose that the 2-phenylethanol and phenylacetaldehyde, that can be detected during growth, results from side-product formation of this novel pathway due to instability of the glutathione adducts or enzymatic removal of glutathione in an earlier step (Fig. 3a).

It should be mentioned that it is unusual for actinobacteria to produce glutathione, as mycothiol is the dominant thiol in these organisms (64–66). However, it was reported that strain AD45 additionally produces substantial amounts of glutathione and it was suggested that this ability was gained by horizontal gene transfer of isoprene degradation genes (54, 56, 58, 65). It is likely that the same is true for strain CWB2 due to the plasmid uptake. This is supported by the finding that the GC content of the plasmid and the styrene degradation cluster is much lower compared to the whole genome. The GC content of that cluster S3 is close to that of strain AD45 (61.7%; (22). Recently, genes of cluster S3 were found in *Gordonia* sp. i37 (21). Further, strain CWB2 as well as strain 1CP encode for several mobile elements.
in direct neighbourhood of the styrene degradation cluster on their plasmid what suggests horizontal gene transfer. In addition, the cluster S4 is highly similar that that of *R. opacus* 1CP and strain RHA1 and degradation of PAA likely takes place in the same way (67, 68) (Fig. 3b).

**CONCLUSION**

Omic analyses imply that strain CWB2 incorporated a plasmid which contains an assembly of different gene clusters and forms a “hybrid” that enables to metabolize styrene and analogous compounds. Our study illustrates the possibilities of horizontal gene transfer for Gram-positive bacteria and an ongoing adaptation to glutathione as cofactor in actinobacteria. This adaption is coupled with a high biotechnological potential of this organism, as *G. rubripertincta* CWB2 can produce ibuprofen, which is not possible through the classical styrene degradation pathway (17, 18). The involved SMO shows higher activities than reported for other SMOs so far. This might be interesting as these enzymes are known to catalyze a variety of valuable reactions (3). Further, bacterial GST are known to be involved in degradation of (halogenated) xenobiotics and other chemical transformations and therefore the GST of strain CWB2 might open a new field of possible biochemical reactions to this class of enzymes (63, 69, 70).
MATERIALS AND METHODS

Isolation and cultivation of styrene-degrading strains. Styrene degrading bacteria were isolated from (contaminated) soil. A small amount of the soil was transferred into a 1-l Erlenmeyer flask and suspended in 100 ml water. Portions of 10 to 40 µl of styrene were supplied via an evaporation adaptor as sole source of carbon and energy. The growth media was dosed with 0.02 mg ml⁻¹ nalidixin acid and 0.075 mg ml⁻¹ cycloheximide to prevent growth of Gram-negative bacteria or fungi, respectively. 10 ml of the culture was plated on solid mineral medium (MM) (71) without carbon source and incubated at room temperature in a 5-l desiccator under styrene containing atmosphere. The grown colonies were repeatedly transferred on fresh solid mineral media and incubated for 2 - 3 days in the desiccator. The isolates were stored at - 80°C in 40% (v/v) glycerol.

Liquid cultures were kept in Erlenmeyer flasks containing mineral media (71). The respective carbon source was added either directly into the media or in case of volatile compounds via an evaporation adapter.

Characterization of G. rubripertincta CWB2. Growth of strain CWB2 in liquid MM was assayed on various substrates (Table S1). Production of surfactants with fructose, hexadecane or styrene as carbon source was examined as published earlier (36, 72). Siderophage production was determined by the CAS-agar plate test (73). Analysis of the mycol- and fatty acid composition was done by the Deutsche Sammlung für Mikroorganismen und Zellkulturen (DSMZ) (Table S2). 16S rRNA analysis and Scanning electron microscopy (SEM) was done as described earlier (18) (Fig. S1 and S3). The in silico DNA-DNA-hybridization was performed by the Genome-to-Genome Distance Calculator 2.1 (DSMZ) (Fig. S7). G. rubripertincta
CWB2 was assayed for antibiotic resistance on chloramphenicol, ampicillin, tetracycline, nalidixic acid, gentamycin, streptomycin and kanamycin.

**DNA extraction and genome sequencing, annotation and bioinformatic analysis.** A 50-ml culture of *G. rubripertincta* CWB2 was grown on fructose in mineral media. Cells were harvested at an OD$_{600}$ of 0.6 by centrifugation (5 000 x g; 15 min), washed once with 100 mM phosphate buffer (pH 7.5) and centrifuged again. The cell Pellet was resuspended in 2.7 ml buffer (10 Tris-HCl, 10% sucrose, 30 mg lysozyme; pH 7.8) and incubated for 1.5 h at 37°C. After centrifugation the supernatant was discarded and the pellet was resuspended in 2.8 ml TE-buffer 10.1 (10 mM Tris, 1 mM EDTA; pH 8) with 100 µg ml$^{-1}$ proteinase K and 150 µl SDS (10%) and incubated for 2 h at 37°C. Chromosomal DNA was extracted successively with 3 to 5 ml phenol (equilibrated; 2 times), phenol:chloroform:isoamyl alcohol (25:24:1) and chloroform:isoamyl alcohol (24:1; 2 times). Each extraction was followed by a centrifugation (20 000 x g; 10 min) and separation of the aqueous phase into a new tube. Afterwards, RNA was digested by adding 100 µg ml$^{-1}$ RNase A at 37°C for 10 min. The DNA was precipitated on ice by the addition of 1/10 volume of 3 M sodium acetate (pH 5.2) and 3 volumes of 100% ethanol. The precipitate was centrifuged at 4°C, resuspended in 70% ice-cold ethanol and centrifuged again. The supernatant was discarded and the pellet was dried for 10 min in a SpeedVac vacuum concentrator. The pellet was dissolved in TE-buffer 10.1 for storage at 4°C and purity was controlled by agarose-gel electrophoresis and on the Nanodrop.

To obtain the complete genome sequence, two sequencing libraries were prepared, a TruSeq PCR-free whole genome shotgun library and a 8k Nextera MatePair library (Illumina Inc, Netherlands). Both libraries were sequenced on an Illumina MiSeq desktop sequencer with 2x 300 bp. The obtained reads were
assembled using the Newbler (v2.8) de novo assembler (Roche). The initial assembly consisted of just one scaffold of 90 contigs, with 120 contigs larger than 500 bp in total. Manual inspection and assembly was performed using CONSED (74, 75), which revealed a misassembly: 3 contigs representing a 105 kbp plasmid, henceforth called pGCWB2, were wrongly "attached" to the scaffold representing the chromosome. The sequence of both replicons could be completely established. The sequences for both replicons were annotated using PROKKA (76), see Table S3 for details. The annotated replicons were submitted to GenBank, the accession numbers are CP022580 (chromosome) and CP022581 (pGCWB2).

Additional functional annotation of proteins was done by using the BLASTP algorithm (20) on the non-redundant protein database or the UniProtKB database (NCBI) (date of search: 01.09.2017). Membrane association of proteins was verified by prediction on the TMHMM server (77). Island viewer 4 (78) was applied to detect genomic islands and foreign genes on the genome of strain CWB2 (Fig. S4).

**RNA extraction and transcriptome sequencing.** A pre-culture of *G. rubripertincta* CWB2 was grown at 30°C in minimal media with fructose or styrene as sole source of carbon, respectively. After 5 days of cultivation the culture was diluted 1/10 in fresh media and incubated for 24 h with the respective substrate in a set of four Erlenmeyer flasks. Prior harvesting the cells, 10% of an ice-cold STOP-solution (10% buffered phenol in ethanol) was added to the culture followed by centrifugation at 11 000 x g for 5 min at 4°C. The supernatant was discarded and the pellet was stored until RNA isolation at - 80°C. To break up the cells, 150 µl of a 5 mg ml⁻¹ lysozyme solution were added to the pellet, mixed and incubated at room temperature for 5 min. 450 µl of buffer RLT (Qiagen) and 50 mg of (0.1 mm) glass beads were added to resuspend and break the cells by repeated vortexing at 4°C.
The suspension was applied to QIAshredder column for homogenization and to remove particles from the sample. Extraction of total RNA Extraction was done by applying the RNeasy Mini Kit including on-column DNA digestion (Qiagen). Isolated RNA was stored at -80°C and quality was controlled on the 2100 Bioanalyzer using the RNA 6000 Nano Kit (Agilent).

RNA quality and quantity was again checked by an Agilent 2100 Bioanalyzer run (Agilent Technologies, Böblingen, Germany) and Trinean Xpose sytem (Gentbrugge, Belgium) prior and after rRNA depletion by Ribo-Zero rRNA Removal Kit (Bacteria) (Illumina, San Diego, CA, USA). TruSeq Stranded mRNA Library Prep Kit from Illumina, (San Diego, CA, USA) was used to prepare the cDNA libraries to analyze the whole transcriptome. The resulting cDNAs were then sequenced paired end on an Illumina MiSeq and HiSeq 1500 system (San Diego, CA, USA) using 2 x 75 nt read length. The raw sequencing read files are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number: E-MTAB-6012.

Reads were mapped on the reference G. rubripertincta CWB2 (CP022580, CP022581) with Bowtie2 (79) using standard settings. ReadXplorer 2.2.0 (80) was used for visualization of short read alignments and data analysis. Differential gene expression analysis was performed based on normalized read count using TPM values (Transcripts Per Million) of CDS calculated by ReadXplorer 2.2.0. The signal intensity value (a-value) was calculated by 0.5*(log2 TPM condition A + log2 TPM condition B) of each CDS and the signal intensity ratio (m-value) by the difference of (log2) TPM. CDS with m-values of higher/equal than +1.5 or lower/equal than −1.5 were considered to be differentially transcribed.

Preparation of protein samples and identification by LC-ESI-MS/MS mass spectrometry. G. rubripertincta CWB2 was grown the same way as for RNA
extraction in a set of four samples per carbon source. After cultivation, two samples were pooled and centrifuged at 5 000 x g for 30 min at 4°C and resuspended in 2.5 ml 50 mM PP, pH 7.26. The cells were disrupted by sonication on ice (10 cycles, 1.5 min; power: 70%; BANDELIN Sonoplus Homogenisator HD2070) after adding 40 U DNAseI and 1 mg ml⁻¹ lysozyme. The suspension was centrifugated at 50 000 x g for 1 h at 4°C to separate soluble from insoluble matter. The proteins were separated by size using discontinuous sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) with 50 μg of each sample per lane. Lanes were cut into 8 slices and de-stained, alkylated and digested with trypsin as previously described (81, 82). Peptides were extracted from the gel pieces with acetonitrile, loaded onto STAGE tips for storage, and eluted from the tips shortly before MS analysis (83).

By using an EASY-nLC 1000 (Thermo Scientific) LC system, peptides were separated at a flow rate of 400 nl/min on a 18 cm self-packed column (75 μm ID, 1.9 μm Reprosil-Pur 120 C-18AQ beads, Dr Maisch Germany) housed in a custom-built column oven (84) at 45°C. Peptides were separated using gradient of buffers A (0.1% formic acid) and B (80% acetonitrile, 0.1% formic acid) from 1% to 60% B. The column was interfaced with a Nanospray Flex Ion Source (Thermo Scientific) to a Q-Exactive HF mass spectrometer (Thermo Scientific). MS instrument settings were: 1.5 kV spray voltage, Full MS at 60K resolution, AGC target 3e6, range of 300 - 1750 m/z, max injection time 20 ms; Top 15 MS/MS at 15K resolution, AGC target 1e5, max injection time 25 ms, isolation width 2.2 m/z, charge exclusion +1 and unassigned, peptide match preferred, exclude isotope on, dynamic exclusion for 20s.

Mass spectra were recorded with Xcalibur software 3.1.66.10 (Thermo Scientific). Using a custom database containing 4831 predicted protein sequences, proteins were identified with Andromeda and quantified with the LFQ algorithm embedded in
MaxQuant version 1.5.3.17 (85). The following parameters were used: main search max. peptide mass error of 4.5 ppm, tryptic peptides of min. 6 amino acid length with max. two missed cleavages, variable oxidation of methionine, protein N-terminal acetylation, fixed cysteine carbamidomethylation, LFQ min. ratio count of 2, matching between runs enabled, PSM and (Razor) protein FDR of 0.01, advanced ratio estimation and second peptides enabled. Proteins with a log2 ratio of higher/equal than +1.5 or lower/equal than −1.5 were considered to be differentially synthesized.

Cloning, expression and purification of recombinant styAs, styD and aldh1.

The *styA* (GCWB2_24100, GCWB2_21620), *styD* (GCWB2_23925) and *aldh1* (GCWB2_24010) genes were purchased from Eurofins MWG (Ebersberg) in a pEX-K2 vector system allowing for kanamycin resistance selection. The DNA sequences were optimized for the codon usage and GC content of *Acinetobacter baylyi* ADP1 with the OPTIMIZER tool (48, 86). 5′-NdeI and 3′-NotI restriction sites were added and used for subcloning into pET16bP to obtain the expression constructs pSGrA1_P01, pSGrA2_P01 and pSGrD1_P01, pSGrD2_P01 from which recombinant proteins can be obtained as His10-tagged proteins. *Escherichia coli* strain DH5α and strain BL21 (DE3) pLysS were cultivated for cloning and expression purposes as described elsewhere (87). Plasmids are listed in Table 3.

Expression of StyA’s took place in a 3-l biofermenter. *E. coli* BL21 strains with the respective plasmids were cultivated in LB media (100 µg ml⁻¹ ampicillin and 50 µg ml⁻¹ chloramphenicol) at 30°C until an OD₆₀₀ of 0.4 was reached. The batch was subsequently cooled to 20°C. Expression was induced at an OD₆₀₀ of 0.6 by adding 0.1 mM of IPTG (isopropyl-β-D-thiogalactopyranoside) to the culture and grown for 20 h at 20°C (120 rpm). Cells were harvested by centrifugation (5 000 x g, 30 min, 4°C), resuspended in 10 mM Tris-HCl buffer (pH 7.5) and stored at - 80°C. Formation of
the blue dye indigo is observable if active SMOs are produced during expression in LB media (88). As this was not the case for expression of *GCWB2_21620* we assumed that the protein is not synthesized or active.

For purification of StyA, crude extracts were prepared from freshly thawed biomass by disruption in a precooled French Pressure cell, followed by centrifugation to remove cell debris (50 000 x g, 2 h, 4°C). The supernatants were applied to a 1-ml HisTrap FF column. The column was washed with 10 CV of binding buffer (10 mM Tris-HCl, 0.5 M NaCl, 25 mM imidazole, pH 7.5) to remove nonspecific bound proteins. Enzymes were eluted with a linear imidazole gradient up to 500 mM over 30 CV. Fractions with respective enzyme activity (see 2.6.) were pooled and concentrated using Sartorius Vivaspin 20 filters (5 000 MWCO) at 4°C. The concentrates were passed through a 10-ml Econo-Pac 10DG desalting gravity-flow column (Bio-Rad) to remove remaining imidazole and sodium chloride. Protein obtained was kept in storage buffer (10 mM Tris-HCl, 50% [v/v] glycerol, pH 7.5) at -20°C. Expression of *GCWB2_21620* did not yield active protein as already mentioned. Preparation of StyD and Aldh1 was done according to Zimmerling *et al.* 2017 (61).

**Purification of wild-type proteins.** All following purification steps were performed on an ÄKTA fast-performance liquid chromatographer (GE Healthcare). Selected wild-type enzymes were enriched from crude extract by ion-exchange chromatography. Therefore, strain CWB2 was cultivated on styrene and soluble crude extract was prepared as described above. The supernatant was loaded with buffer A (20 mM Tris-HCl; pH 7.5) on a MonoQ HR 5/5 column (GE Healthcare) at a flow rate of 1 ml min⁻¹. Nonspecific bound proteins were removed by washing with 5 column volumes (CV) of buffer A. Elution of proteins was done over 25 CV with a
linear gradient of buffer B (20 mM Tris-HCl, 1 M NaCl; pH 7.5). Fractions of 1 ml were collected and tested on the respective enzyme activity. A second purification step was applied for some enzymes by using hydrophobic interaction chromatography. Therefore, the fractions that showed the respective enzyme activity were pooled and (NH$_4$)$_2$SO$_4$ was added to a final concentration of 460 mM. The sample was loaded with buffer C (20 mM Tris-HCl, 0.8 M (NH$_4$)$_2$SO$_4$; pH 7.5) on a 1-ml Phenyl HP HiTrap column (GE Healthcare) at a flow rate of 1 ml min$^{-1}$. Nonspecific bound proteins were removed by washing with 5 column volumes (CV) of buffer A. Elution of proteins was done over 25 CV with a linear gradient of buffer A (20 mM Tris-HCl; pH 7.5). Fractions of 1 ml were tested on enzyme activity.

For VC12DO gel filtration was done after hydrophobic interaction chromatography. Therefore, the fraction containing VC12DO activity were pooled and applied with buffer D (25 mM Tris-HCl, 0.5 M NaCl; pH 7.5) to a Superdex 200 HR 10/30 column at a flow rate of 0.4 ml min$^{-1}$. Fractions of 1 ml were tested for VC12DO activity.

Recombinant and wild-type proteins were subjected to discontinuous sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (87) in order to determine purity and subunit molecular size.

**Enzyme assays.** Crude extracts, enriched or purified protein preparations from *G. rubripertincta* CWB2 were assayed for enzyme activities that are representative for known degradation pathways of styrene.

Wild-type catechol 1,2-dioxygenase, catechol 2,3-dioxygenase and cis,cis-muconate cycloisomerase activity was measured spectrophotometrically (Cary 50, Varian) by following the product formation or substrate depletion according to Warhurst *et al.* 1994 (89) using catechol, protocatechuate and cis,cis-muconate as
Styrene monooxygenase (SMO) activity of wild-type and recombinant enzyme preparations with styrene were measured by quantification of the reaction product styrene oxide on a RP-HPLC system as described previously (90).

Styrene oxide reductase (SOR) and phenylacetaldehyde reductase (PAR) wild-type activity with styrene oxide and 2-phenylethanol was determined by quantification of the reaction products phenylacetaldehyde or 2-phenylethanol on a RP-HPLC system following the protocol as described previously for the styrene oxide isomerase (SOI) (17, 18).

Wild-type phenylacetaldehyde dehydrogenase (PAD) and wild-type PAR activity was assayed indirectly on a spectrophotometer (Cary 50, Varian) by following the reduction of NAD⁺ to NADH at 340 nm (ε340 nm = 6.22 mM⁻¹ cm⁻¹) (91). The 1 ml assay mixture contained 0.5 mM phenylacetaldehyde or 2-phenylethanol in 10 mM Tris-HCl (pH 7.5), 1 mM NAD⁺ and 50 µl protein containing sample, respectively. Recombinant PAD activity was assayed according to (61).

GST wild-type activity was assayed in soluble crude extract by following the (S)-styrene oxide consumption over time. Therefore, G. rubripertincta CWB2 was grown on MM with styrene as sole source of carbon. A 100 ml pre-culture was prepared and used to inoculate the main culture 1:50 in 500 ml fresh MM. The main culture was incubated at 30°C for 5 days by adding 20 to 80 µl portions styrene via gas-phase. Cells were harvested by centrifugation at 5 000 x g for 20 min at 4°C. The supernatant was discarded and the pellet was resuspended and washed 2 times in 10 ml 20 mM PP (pH 8). The cells were disrupted by sonication on ice (10 cycles, 1 min; power: 70%; BANDELIN Sonoplus Homogenisator HD2070) after adding 40 U DNAseI and 1 mg ml⁻¹ lysozyme. Soluble crude extracts were obtained by
centrifugation at 50 000 x g at 4°C for 1 h and separation from the insoluble matter. The reaction mix (600 µl) contained 20 mM PP (pH 8), 4 mM (S)-styrene oxide, 5 mM GSH and an appropriate amount of soluble crude extract. Blank measurements were carried out by omitting either GSH or enzyme preparation. Samples were tempered for 10 min at 30°C and the reaction was initiated by the addition of the substrate (S)-styrene oxide. 25 µl samples were quenched at certain time points in 50 µl ice cold acetonitrile:methanol (1:1) and centrifuged at 16 000 x g for 10 min at 4°C to remove precipitates. Supernatants were applied to RP-HPLC by injection of 10 µl samples. All measurements were done in triplicates. Enzyme activities are given in 1 U mg⁻¹ representing the conversion µmol substrate per min per mg protein.

All RP-HPLC measurements were done with a Eurospher C_{18} column (125 mm length by 4 mm i.d., 5 µm particle size, 100 Å pore size; Knauer, Germany). The protein content was determined by means of the Bradford method (92), using BradfordUltra reagent (Expedeon) and bovine serum albumin (Sigma) as reference protein.

**Accession numbers.** Genome and assembly of *Gordonia rubripertincta* CWB2 is deposited at NCBI (BioProject Accession: PRJNA394617; URL: https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA394617) with the chromosome (CP022580) and plasmid (CP022581) sequences.

The raw sequencing read files are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number: E-MTAB-6012.


ACKNOWLEDGEMENTS

We appreciate the funding of this project by the Deutsche Bundesstiftung Umwelt (DBU), the European Social Fund and the Saxonian Government (GETGEOWEB: 100101363).

We hereby declare no conflicting interests among all of us the co-authors.

SUPPLEMENTAL MATERIAL

Supplemental material is available at AEM's website.
REFERENCES


799  styrene and benzene catabolism in *Rhodococcus jostii* RHA1. J Bacteriol
800  190:37–47.
802  Analysis of the *Pseudomonas putida* CA-3 proteome during growth on styrene
803  under nitrogen-limiting and non-limiting conditions. Microbiology (Reading, Engl.)
804  155:3348–3361.
806  characterization of the styE gene, involved in styrene transport in *Pseudomonas
808  1309.2006.
810  Tischler D. 2015. Catalytic and hydrodynamic properties of styrene
811  monooxygenases from *Rhodococcus opacus* 1CP are modulated by cofactor
812  binding. AMB Express 5:112.
814  characterization of StyAB from *Pseudomonas* sp. strain VLB120 as a two-
817  styrene monooxygenases of *Rhodococcus* sp. ST-5 and ST-10 for synthesizing
820  production by over-expression of the styrene monooxygenase in *Pseudomonas
823  of (S)-styrene oxide using styrene oxide isomerase negative mutant of
826  Styrene oxide isomerase of *Rhodococcus opacus* 1CP, a highly stable and
828 54. El Khawand M, Crombie AT, Johnston A, Vavlline DV, McAuliffe JC, Latone JA,
830  isoprene degrading bacteria from soils, development of *isoA* gene probes and
831  identification of the active isoprene-degrading soil community using DNA-stable
834  Characterization of marine isoprene-degrading communities. Environ Microbiol
835  11:3280–3291.
837  glutathione S-transferase with activity towards cis-1,2-dichloroepoxyethane is
838  involved in isoprene utilization by *Rhodococcus* sp. strain AD45. Appl Environ
839  Microbiol 64:2800–2805.
841  glutathione S-transferase and a glutathione conjugate-specific dehydrogenase
842  involved in isoprene metabolism in *Rhodococcus* sp. strain AD45. J Bacteriol
843  181:2094–2101.


**FIGURE LEGENDS**

**Fig. 1.** Comparison of the styrene degradation cluster of *Gordonia rubripertincta CWB2* with homologous clusters as found in the strains *Rhodococcus opacus* PD630 (Accession: NZ_CP003949) (22), *Rhodococcus* sp. AD45 (NZ_CM003191) (58), *Nocardioides* sp. Root240 (NZ_LMIT01000013), *Aeromicrobium* sp. Root495 (NZ_LMFJ01000002), *Rhodococcus* sp. ST-10 (AB594506) (13), *Rhodococcus opacus* 1CP (NZ_CP009112, NZ_CP009111) (48), *Pseudomonas* sp. Y2 (AJ000330) (38, 93) and *Sphingopyxis fribergensis* sp. Kp5.2 (CP009122) (17). Subclusters of strain CWB2 are indicated (S1-S4) and gene products are given in the legend coloured by their (predicted) function. Relevant homologous genes and clusters are emphasized by interspaced conjunctions. Clusteres of marked strains are reported to be involved in isoprene (●) or styrene (♯) degradation.

**Fig. 2.** Degradation of 4 mM (S)-styrene oxide with crude extract of styrene grown biomass of *Gordonia rubripertincta* CWB2 and 5 mM reduced glutathione (●). Only minor consumption was detected when excluding either reduced glutathione (○) or crude extract (X) from the reaction mix.

**Fig. 3a.** Proposed novel degradation pathway of styrene in *Gordonia rubripertincta CWB2* (see text for details). **3b.** Proposed phenylacetic acid degradation pathway of *Gordonia rubripertincta* CWB2. The genes of the involved enzymes are present on the genome (cluster S4) and upregulated on transcriptome and the proteins on proteome level, respectively (see Table 2). Starting from the product of the upper degradation pathway, phenylacetic acid, strain CWB2 is able to metabolize styrene to acetyl-CoA or succinyl-CoA (adapted to 68).
TABLE LEGENDS

Table 1. Substrate spectra with focus on ones that might be related to styrene degradation in *G. rubripertincta* CWB2.

Table 2. Functional categorization of proteins from *G. rubripertincta* CWB2 that are supposed to be involved in styrene degradation and regulation on RNA and protein level. Enzymes and proteins with reported activity or function are underlined. For further details, see Dataset S1 in supporting material.

Table 3. Plasmids used in this study.
### 3a. Upper degradation pathway

- **Styrene** → **(S)-styrene oxide**
  - StyA/StyB: NADH + H⁺ + O₂ → NAD⁺ + H₂O

- **(S)-1-Phenyl-2-hydroxyethyl glutathione**
  - GSH

- **(S)-1-Phenyl-2-acetaldehyde glutathione**
  - Aldh1, StyD, StyH: NAD⁺ + H⁺

- **(S)-1-Phenyl-2-acetic acid glutathione**
  - GSSG → GSH

### 3b. Lower degradation pathway

1. Phenylacetic acid → phenylacetyl-CoA → PaaABCDE → 1,2-epoxyphenylacetyl-CoA → oxepin-CoA → PaaG → 3-oxo-5,6-dehydrosuberyl-CoA → PaaJ → succinyl-CoA → 3-oxoadipyl-CoA → PaaJ → 2-acetyl-CoA

- **3-hydroxyadipyl-CoA**
- **2,3-dehydroadipyl-CoA**
- **3-oxo-5,6-dehydrosuberyl-CoA**
Table 1. Substrate spectra with focus on ones that might be related to styrene degradation in *G. rubripertincta* CWB2.

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Styrene</td>
<td>++</td>
</tr>
<tr>
<td>Styrene oxide</td>
<td>++</td>
</tr>
<tr>
<td>2-Phenylethanol</td>
<td>++</td>
</tr>
<tr>
<td>Phenylacetaldehyde</td>
<td>++</td>
</tr>
<tr>
<td>Phenylacetic acid</td>
<td>+++</td>
</tr>
<tr>
<td>Succinate</td>
<td>++</td>
</tr>
<tr>
<td>Citrate</td>
<td>++</td>
</tr>
<tr>
<td>Isoprene</td>
<td>-</td>
</tr>
<tr>
<td>Mandelic acid</td>
<td>-</td>
</tr>
<tr>
<td>D-Fructose</td>
<td>+++</td>
</tr>
</tbody>
</table>

(++) vigorous growth; (+++) good growth; (+) growth; (-) no growth
Table 2. Functional categorization of proteins from *G. rubripertincta* CWB2 that are supposed to be involved in styrene degradation and regulation on RNA and protein level. Enzymes and proteins with reported activity or function are underlined. For further details see supporting information.

<table>
<thead>
<tr>
<th>ORF</th>
<th>Gene</th>
<th>A</th>
<th>M</th>
<th>Cyt</th>
<th>Mem</th>
<th>Name*</th>
<th>Function</th>
<th>Accession</th>
<th>% Id</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>23925</td>
<td>styD</td>
<td>7.7</td>
<td>8.6</td>
<td>7.1</td>
<td>5.7</td>
<td><strong>styD</strong></td>
<td>Phenylacetaldehyde dehydrogenase</td>
<td>BAL04135</td>
<td>76</td>
<td>(13, 50)</td>
</tr>
<tr>
<td>23930</td>
<td>paaK</td>
<td>8.1</td>
<td>6.0</td>
<td>8.1</td>
<td>5.9</td>
<td><strong>paaK</strong></td>
<td>Phenylacetate-coenzyme A ligase</td>
<td>Q9L9C1</td>
<td>68</td>
<td>(94)</td>
</tr>
<tr>
<td>23935</td>
<td>paaE</td>
<td>8.0</td>
<td>6.2</td>
<td>11.3</td>
<td>8.2</td>
<td><strong>paaE</strong></td>
<td>1,2-phenylacetly-CoA epoxidase, subunit E</td>
<td>P76081</td>
<td>43</td>
<td>(8)</td>
</tr>
<tr>
<td>23940</td>
<td>paaD</td>
<td>8.1</td>
<td>6.7</td>
<td>NaNf</td>
<td>NaNf</td>
<td><strong>paaD</strong></td>
<td>1,2-phenylacetly-CoA epoxidase, subunit D</td>
<td>P76080</td>
<td>42</td>
<td>(8)</td>
</tr>
<tr>
<td>23945</td>
<td>paaC</td>
<td>8.3</td>
<td>6.7</td>
<td>10.5</td>
<td>7.4</td>
<td><strong>paaC</strong></td>
<td>1,2-phenylacetly-CoA epoxidase, subunit C</td>
<td>P76079</td>
<td>42</td>
<td>(8)</td>
</tr>
<tr>
<td>23950</td>
<td>paaB</td>
<td>8.0</td>
<td>6.8</td>
<td>8.8</td>
<td>9.1</td>
<td><strong>paaB</strong></td>
<td>1,2-phenylacetly-CoA epoxidase, subunit B</td>
<td>P76078</td>
<td>67</td>
<td>(8)</td>
</tr>
<tr>
<td>23955</td>
<td>paaA</td>
<td>8.5</td>
<td>7.1</td>
<td>11.4</td>
<td>9.0</td>
<td><strong>paaA</strong></td>
<td>1,2-phenylacetly-CoA epoxidase, subunit A</td>
<td>P76077</td>
<td>66</td>
<td>(8)</td>
</tr>
<tr>
<td>23960</td>
<td>paaG</td>
<td>7.7</td>
<td>7.4</td>
<td>6.9</td>
<td>6.8</td>
<td><strong>paaG</strong></td>
<td>1,2-epoxyphenylacetly-CoA isomerase</td>
<td>P77467</td>
<td>37</td>
<td>(8)</td>
</tr>
<tr>
<td>23965</td>
<td>paaH</td>
<td>7.6</td>
<td>7.5</td>
<td>7.1</td>
<td>6.2</td>
<td><strong>paaH</strong></td>
<td>3-hydroxyadipyl-CoA dehydrogenase</td>
<td>P76083</td>
<td>36</td>
<td>(8)</td>
</tr>
<tr>
<td>23970</td>
<td>tetR</td>
<td>7.3</td>
<td>7.4</td>
<td>7.1</td>
<td>5.8</td>
<td><strong>tetR2</strong></td>
<td>HTH-type transcriptional repressor KstR2</td>
<td>A0R4Z6</td>
<td>25</td>
<td>(95)</td>
</tr>
<tr>
<td>23980</td>
<td>paaZ</td>
<td>7.9</td>
<td>6.4</td>
<td>6.8</td>
<td>6.0</td>
<td><strong>paaZ</strong></td>
<td>bifunctional aldehyde dehydrogenase</td>
<td>P77455</td>
<td>51</td>
<td>(8)</td>
</tr>
<tr>
<td>23985</td>
<td>paaI</td>
<td>6.6</td>
<td>8.2</td>
<td>7.4</td>
<td>6.6</td>
<td><strong>paaJ</strong></td>
<td>3-oxoadipyl-CoA/3-oxo-5,6-dehydrosuberyl-CoA thiolase</td>
<td>P0C7L2</td>
<td>55</td>
<td>(8)</td>
</tr>
<tr>
<td>23990</td>
<td>ethD</td>
<td>8.4</td>
<td>6.3</td>
<td>5.1</td>
<td>5.7</td>
<td><strong>ethD</strong></td>
<td>Uncharacterized 11.0 kDa protein</td>
<td>P43491</td>
<td>48</td>
<td>(96)</td>
</tr>
<tr>
<td>24000</td>
<td>partial</td>
<td>7.3</td>
<td>6.4</td>
<td>NaN</td>
<td>NaN</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>24005</td>
<td>araC</td>
<td>8.0</td>
<td>6.0</td>
<td>4.1</td>
<td>8.4</td>
<td><strong>phR</strong></td>
<td>Transcriptional activator NphR</td>
<td>B1Q2A8</td>
<td>31</td>
<td>(97)</td>
</tr>
</tbody>
</table>

**Transcriptome** | **Proteome** | **Best hit on the Uniprot Database at amino acid level**

<table>
<thead>
<tr>
<th>ORF</th>
<th>Gene</th>
<th>A</th>
<th>M</th>
<th>Cyt</th>
<th>Mem</th>
<th>Name*</th>
<th>Function</th>
<th>Accession</th>
<th>% Id</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>24010</td>
<td><strong>aldh1</strong></td>
<td>9.6</td>
<td>6.4</td>
<td>6.0</td>
<td>5.5</td>
<td><strong>styD</strong></td>
<td>Phenylacetaldehyde dehydrogenase</td>
<td>O06837</td>
<td>36</td>
<td>(98)</td>
</tr>
<tr>
<td>24015</td>
<td><strong>styG</strong></td>
<td>10.1</td>
<td>6.1</td>
<td>7.7</td>
<td>5.9</td>
<td><strong>yfcG</strong></td>
<td>Disulfide-bond oxidoreductase YfcG</td>
<td>P77526</td>
<td>47</td>
<td>(99)</td>
</tr>
<tr>
<td>24020</td>
<td><strong>isoI</strong></td>
<td>10.4</td>
<td>6.0</td>
<td>5.0</td>
<td>5.3</td>
<td><strong>isoI</strong></td>
<td>Glutathione-S-transferase</td>
<td>WP_045063292</td>
<td>49</td>
<td>(57)</td>
</tr>
<tr>
<td>24025</td>
<td><strong>isoH</strong></td>
<td>9.8</td>
<td>6.3</td>
<td>5.0</td>
<td>4.8</td>
<td><strong>isoH</strong></td>
<td>1-hydroxy-2-glutathionyl-2-methyl-3-butene DH</td>
<td>WP_045063294</td>
<td>59</td>
<td>(57)</td>
</tr>
<tr>
<td>24030</td>
<td><strong>vdE</strong></td>
<td>10.0</td>
<td>6.3</td>
<td>6.8</td>
<td>5.5</td>
<td><strong>vdE</strong></td>
<td>Acetyl-CoA:oxalate CoA-transferase</td>
<td>P76518</td>
<td>33</td>
<td>(100)</td>
</tr>
<tr>
<td>24035</td>
<td><strong>marA</strong></td>
<td>10.0</td>
<td>3.5</td>
<td>2.1</td>
<td>2.7</td>
<td><strong>marR</strong></td>
<td>regulatory protein</td>
<td>CAA52427</td>
<td>31</td>
<td>(101)</td>
</tr>
<tr>
<td>24040</td>
<td><strong>dsr</strong></td>
<td>9.1</td>
<td>4.7</td>
<td>8.5</td>
<td>4.8</td>
<td><strong>dsr</strong></td>
<td>Coenzyme A disulfide reductase</td>
<td>O58308</td>
<td>34</td>
<td>(102)</td>
</tr>
<tr>
<td>24045</td>
<td><strong>gshB</strong></td>
<td>8.9</td>
<td>4.9</td>
<td>5.7</td>
<td>3.8</td>
<td><strong>gshB</strong></td>
<td>Glutathione synthetase</td>
<td>P45480</td>
<td>50</td>
<td>(103)</td>
</tr>
<tr>
<td>24050</td>
<td><strong>gshA</strong></td>
<td>9.0</td>
<td>3.4</td>
<td>5.5</td>
<td>2.2</td>
<td><strong>gshA</strong></td>
<td>Glutamate-cysteine ligase EgtA</td>
<td>P9WPK7</td>
<td>33</td>
<td>(104)</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Description</td>
<td>Chromosome Position</td>
<td>Protein Accession</td>
<td>Protein Name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>---------------------</td>
<td>-------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24055</td>
<td>Hyp</td>
<td>7.8 8.5  8.8</td>
<td>S2</td>
<td>StyB-SMO flavin oxidoreductase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24060</td>
<td>Hyp</td>
<td>7.0 9.3  7.5 10.3</td>
<td>S2</td>
<td>StyA-SMO styrene monoxygenase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24065</td>
<td>Hyp</td>
<td>6.6 9.3  8.8 10.3</td>
<td>S2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24070</td>
<td>Hyp</td>
<td>7.1 8.4  8.2 8.4</td>
<td>S2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24075</td>
<td>Hyp</td>
<td>6.8 9.5  7.9 8.2</td>
<td>S2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24080</td>
<td>Hyp</td>
<td>7.8 8.0  8.0 8.2</td>
<td>S2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24085</td>
<td>Hyp</td>
<td>9.2 10.7 5.8 6.2</td>
<td>S2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24090</td>
<td>Hyp</td>
<td>7.4 10.4 9.2 9.0</td>
<td>S2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24095</td>
<td>styR</td>
<td>8.9 9.4  11.2 8.5</td>
<td>BAL04130</td>
<td>(13, 50)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24100</td>
<td>styA</td>
<td>9.0 9.1  8.1 6.0</td>
<td>BAL04129</td>
<td>(13, 50)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18380</td>
<td>pmrA</td>
<td>5.3 3.5  3.4 3.2</td>
<td>BAD03956</td>
<td>(25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18385</td>
<td>pmrB</td>
<td>3.1 4.8  4.9 2.0</td>
<td>BAD03957</td>
<td>(25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18390</td>
<td>pmrC</td>
<td>4.1 4.0  3.7 2.4</td>
<td>BAD03958</td>
<td>(25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18395</td>
<td>pmrD</td>
<td>4.1 4.3  4.7 4.4</td>
<td>BAD03959</td>
<td>(25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18400</td>
<td>adh</td>
<td>3.7 3.7  4.9 2.3</td>
<td>BAD03962</td>
<td>(25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18405</td>
<td>groL</td>
<td>3.3 3.8  3.8 1.7</td>
<td>P28598</td>
<td>(105)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14195</td>
<td>ahpD</td>
<td>7.1 3.2  4.7 4.5</td>
<td>Q50441</td>
<td>(106)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14190</td>
<td>ahpC</td>
<td>7.7 4.9  4.1 2.5</td>
<td>A0R1V9</td>
<td>(107)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14185</td>
<td>oxyR</td>
<td>2.7 5.1  1.2 2.6</td>
<td>Q9X5P2</td>
<td>(108)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Transcript abundance (A) and fold-change (M) as well as protein fold-change in cytosol (Cyt) and membrane fraction (Mem) is given as log2 ratio. NaNf – not detected in the fructose proteome; NaN – not detected.
### Table 3. Plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevant characteristic(s)</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET16bP</td>
<td>pET16b with additional multicloning site; allows synthesis of recombinant proteins with an N-terminal His&lt;sub&gt;10&lt;/sub&gt;-tag</td>
<td>U. Wehmeyer*</td>
</tr>
<tr>
<td>pEX-K2-pSGrA1</td>
<td>GCWB2_24100 (1284-bp NdeI/NotI fragment) cloned into pEX-K2 vector</td>
<td>MWG Eurofins</td>
</tr>
<tr>
<td>pEX-K2-pSGrA2</td>
<td>GCWB2_21620 (1359-bp NdeI/NotI fragment) cloned into pEX-K2 vector</td>
<td>MWG Eurofins</td>
</tr>
<tr>
<td>pEX-K2-pSGrD1</td>
<td>GCWB2_23925 (1488-bp NdeI/NotI fragment) cloned into pEX-K2 vector</td>
<td>MWG Eurofins</td>
</tr>
<tr>
<td>pEX-K2-pSGrD2</td>
<td>GCWB2_24010 (1437-bp NdeI/NotI fragment) cloned into pEX-K2 vector</td>
<td>MWG Eurofins</td>
</tr>
<tr>
<td>pSGrA1_P01</td>
<td>GCWB2_24100 (1284-bp NdeI/NotI fragment) cloned into pET16bP</td>
<td>This study</td>
</tr>
<tr>
<td>pSGrA2_P01</td>
<td>GCWB2_21620 (1359-bp NdeI/NotI fragment) cloned into pET16bP</td>
<td>This study</td>
</tr>
<tr>
<td>pSGrD1_P01</td>
<td>GCWB2_23925 (1488-bp NdeI/NotI fragment) cloned into pET16bP</td>
<td>This study</td>
</tr>
<tr>
<td>pSGrD2_P01</td>
<td>GCWB2_24010 (1437-bp NdeI/NotI fragment) cloned into pET16bP</td>
<td>This study</td>
</tr>
</tbody>
</table>

* personal communication