

Starvation- and xenobiotic-related transcriptomic responses of the sulfanilic acid-degrading bacterium, *Novosphingobium resinovorum* SA1

Botond Hegedüs^{1,2} · Péter B. Kós^{1,3} · Gábor Bende^{1,2} · Naila Bounedjoum¹ · Gergely Maróti⁴ · Krisztián Laczi¹ · Márk Szuhaj¹ · Katalin Perei¹ · Gábor Rákhely^{1,2,5} 

Received: 24 July 2017 / Revised: 19 September 2017 / Accepted: 26 September 2017 / Published online: 19 October 2017
© Springer-Verlag GmbH Germany 2017

Abstract *Novosphingobium resinovorum* SA1 was the first single isolate capable of degrading sulfanilic acid, a widely used representative of sulfonated aromatic compounds. The genome of the strain was recently sequenced, and here, we present whole-cell transcriptome analyses of cells exposed to sulfanilic acid as compared to cells grown on glucose. The comparison of the transcript profiles suggested that the primary impact of sulfanilic acid on the cell transcriptome was a starvation-like effect. The genes of the peripheral, central, and common pathways of sulfanilic acid biodegradation had distinct transcript profiles. The peripheral genes located on a plasmid had very high basal expressions which were hardly upregulated by sulfanilic acid. The genomic context and the codon usage preference of these genes suggested that they were acquired by horizontal gene transfer. The genes of the central pathways were remarkably inducible by sulfanilic acid indicating the presence of a substrate-specific regulatory system in the cells. Surprisingly, the genes of the common part of

the metabolic pathway had low and sulfanilic acid-independent transcript levels. The approach applied resulted in the identification of the genes of proteins involved in auxiliary processes such as electron transfer, substrate and iron transports, sulfite oxidases, and sulfite transporters. The whole transcriptome analysis revealed that the cells exposed to xenobiotics had multiple responses including general starvation-like, substrate-specific, and substrate-related effects. From the results, we propose that the genes of the peripheral, central, and common parts of the pathway have been evolved independently.

Keywords *Novosphingobium* · Sulfonated aromatic acids · Transcriptomics · Starvation · Metabolic response

Introduction

The majority of the hazardous wastes released into the environment are incompatible with life, and their natural degradation is very slow or it does not take place at all (Dua et al. 2002). However, in the last decades, numerous microorganisms and microbial consortia capable of transforming or mineralizing highly toxic chemicals of industrial origin have been identified and characterized. The biodegradation of toxic aromatic compounds may occur either anaerobically (Heider and Fuchs 1997) or aerobically (Gibson and Parales 2000). The aerobic degradation of substituted aromatic compounds can be divided into three main steps: (1) in the peripheral pathways, the substrates are converted into a central—still aromatic—intermediate, typically into a catechol derivative; (2) in the central pathway, the ring of this intermediate is cleaved and further metabolized into smaller molecules, such as pyruvate or 3-oxoadipate (β -keto adipate) (3) which finally enter into the tricarboxylic acid (TCA) cycle

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00253-017-8553-5>) contains supplementary material, which is available to authorized users.

✉ Gábor Rákhely
rakhely.gabor@brc.mta.hu

¹ Department of Biotechnology, University of Szeged, Közép fasor 52, Szeged 6726, Hungary

² Institute of Biophysics, Biological Research Center, Temesvári krt. 62, Szeged 6726, Hungary

³ Institute of Plant Biology, Biological Research Center, Temesvári krt. 62, Szeged 6726, Hungary

⁴ Seqomics Ltd, Mórahalom, Vállalkozók útja 7, Mórahalom 6782, Hungary

⁵ Institute of Environmental and Technological Sciences, Közép fasor 52, Szeged 6726, Hungary

(Wells and Ragauskas 2012) (common pathway). Although the main principle steps of the aerobic degradation of aromatics have already been disclosed, a number of accessory/supplementary processes (substrate/product transport, electron transfer, etc.) should assist for the physiologically feasible utilization of such recalcitrant molecules.

Sulfanilic acid (SA) is a typical representative of sulfonated aromatic amines widely used and manufactured as an important intermediate in the production of azo dyes, plant protectives, and pharmaceuticals. Its natural degradation is difficult because of the sulfonate group which is a strongly charged anion. Hence, its uptake by intact bacteria is restricted and requires more or less specific transport systems. Within the cells, SA competitively inhibits folic acid synthesis (Brown 1962) and consequently arrests the cell growth.

Despite of its recalcitrant nature, several SA-degrading bacteria or bacterial consortia have been isolated from heavily polluted areas. Among these, *Hydrogenophaga intermedia* S1 with *Agrobacterium radiobacter* S2 (Feigel and Knackmuss 1988; Feigel and Knackmuss 1993; Dangmann et al. 1996; Contzen et al. 2000) and *Hydrogenophaga* sp. PBC with *Ralstonia* sp. PBA (Gan et al. 2011b) are able to perform the mineralization process only in co-cultures where the *Hydrogenophaga* strains are responsible for the degradation process. *Novosphingobium resinovorum* SA1 (formerly, *Novosphingobium subarcticum* SA1 and originally *Pseudomonas paucimobilis*) (Perei et al. 2001), *Pannonibacter* sp. W1 (Wang et al. 2009), *Agrobacterium* sp. PNS-1 (Singh et al. 2004, 2006), and *Ochrobactrum anthropi* (Zhang et al. 2012) strains can perform SA mineralization in pure cultures. Nevertheless, *N. resinovorum* SA1 is the only SA-utilizing single strain of which genome sequence has been determined (Hegedűs et al. 2017) and made publicly available in the Ensembl (assembly name: ASM174222v1), Uniprot (UPID: UP000094626), and KEGG (entry: T04490) databases allowing its deeper analysis.

Although several bacteria have been identified as SA-degrading strains, the components required for the complete bio-conversion of such a harmful chemical have only been partially characterized. The basic catabolic routes were mainly described in *Hydrogenophaga* species (Halak et al. 2006, 2007; Gan et al. 2011a), and it was shown that SA is converted via a modified β -keto adipate pathway specialized for this substrate. Genes of a similar pathway were identified in the *N. resinovorum* SA1 genome (Magony et al. 2007; Hegedűs et al. 2017) comprising four large extrachromosomal elements. The genome analysis revealed that the genes of the peripheral (*sadABX*), central (*scaABCDEF*), and oxoadipate-to-TCA (*pcaFJI*) pathways are located at distinct loci. The first two gene clusters, encoding proteins specialized for the SA degradation, are located on the same plasmid, pSA3, while the last cluster is located on the largest plasmid: pSA1. This plasmid also contains the genes of two other β -keto adipate metabolic routes responsible for the 4-hydroxybenzoate (HB) and benzoate (B) utilizations (Hegedűs et al. 2017).

To get deeper insight into the physiological responses of the cells exposed to SA, whole-cell transcriptome profiling of *N. resinovorum* SA1 cells grown on sulfanilic acid and glucose was performed. Since SA might not be an ideal substrate for the cells, its use as nutrient likely represents starvation and stress conditions for the cells. The effect of starvation was modeled with cells taken from the stationary phases where the substrates were completely consumed. Based on these data, novel elements involved in the sulfanilic acid utilization were identified and both SA-specific and starvation-like responses in the gene expression profiles were demonstrated.

Materials and methods

N. resinovorum SA1 (originally classified as *P. paucimobilis*) was isolated in our lab (deposited in NCAIM ref. no.: (P) B001265) (Perei et al. 2001). For inoculum preparation, the cells were cultivated in Luria-Bertani broth (LB) medium (Sambrook et al. 1989) supplemented with 25 $\mu\text{g}/\text{mL}$ streptomycin (Sm). The cultures were shaken at 28 °C, 180 rpm overnight and harvested at the exponential phase (optical density at 600 nm ($\text{OD}_{600\text{nm}}$) of 0.6–0.8. Before inoculation, the starter culture was centrifuged (12,000 rpm, 10 min, 4 °C) and washed twice with minimal salt media (MM) solution (Perei et al. 2001). The cells were inoculated into the specific media to a final $\text{OD}_{600\text{nm}} = 0.05$.

N. resinovorum SA1 strain was grown in 250-mL Erlenmeyer flasks in 100 mL MM (Perei et al. 2001) containing 10 mM substrates (either SA or glucose (Glc) or HB), 25 $\mu\text{g}/\text{mL}$ Sm, and 25 mM 3-(*N*-morpholino)-propanesulfonic acid (MOPS) (pH = 7.0). The addition of MOPS could stabilize the pH during the cell growth; the pH change (drop) was less than 0.2 unit in any cultivation. When glucose or HB was used as carbon source, the medium was supplemented with 10 mM NH_4Cl ; in the case of SA, the substrate was the sole nitrogen source. The cultures were shaken at 28 °C, 180 rpm, and the cell growth was monitored by measuring the $\text{OD}_{600\text{nm}}$ values. The glucose concentration was measured with HPLC (Hitachi Chromaster) equipped with refractive index detector (RI Chromaster detector 5450). The separation was performed on an ICsep ICE-COREGEL-64H (300 \times 7.8 mm) column. The temperatures of the column and detector were 50 and 41 °C, respectively. The eluent was 0.01 M H_2SO_4 (0.8 mL/min). The SA content was followed by measuring the absorbance of the cell-free medium at 248 nm. All experiments were performed in duplicates.

RNA isolation

For RNA-Seq analysis, 6-mL samples were taken from exponential- and late stationary-phase cultures grown on Glc or SA in MM. For RT-qPCR experiments, samples (6 mL)

were taken from the exponential-phase cell separately grown on Glc, SA, or HB. In each case, two independent biological replicas were used. The samples were harvested ($17,226\times g$, 5 min, 4 °C) and immediately frozen in liquid nitrogen and stored at -80 °C until RNA isolation.

The frozen samples were resuspended in 350 μ L Zymo DNA/RNA Shield (Zymo Research, Irvine, USA). The resuspended cells were mixed with 700 μ L RLT buffer (Qiagen, Venlo, Netherlands) (with β -merkaptoetanol) and were transferred into a new tube containing 0.8-g, 0.5-mm glass beads (Scientific Industries Inc., New York, USA, SI-BG05). The cells were vortexed for 5 min on maximum speed with a Vortex-Genie 2 (Scientific Industries Inc., New York, USA, SI-0236) supplemented with a TurboMix attachment (Scientific Industries Inc., New York, USA, SI-0564). Following lysis, the mixtures were centrifuged at $9600\times g$ for 10 s to remove the glass beads and cell debris. The supernatants were used for purification of total RNAs with the Qiagen RNeasy Mini Kit protocol (Qiagen, Venlo, Netherlands) with on-column DNase digestion. Then, the purified RNAs were treated again with RNase-free DNase I (Sigma, Saint Louis, USA) to remove any residual genomic DNA contamination. Finally, the RNAs obtained were repurified with the RNeasy Mini Kit Clean Up protocol (Qiagen, Venlo, Netherlands).

Total RNAs were used in RT-qPCR experiments, while for whole-transcript analysis (WTA), the ribosomal RNA (rRNA) fractions were removed using the Ribo-Zero rRNA removal kit (Gram-negative bacteria, Illumina Madison/Epicentre Biotechnologies, Madison, USA). The integrity of the RNA was checked by Agilent 2100 bioanalyzer.

RNA-Seq experiments and data processing

RNA-Seq library preparation and Illumina sequencing were carried out at SeqOmics Biotechnology Ltd. (Mórahalom, Hungary). The fragment libraries were generated with the ScriptSeq V2 RNA-Seq library preparation kit (Illumina Madison/Epicentre Biotechnologies, Madison, USA). For indexing, the ScriptSeq index PCR primers were used. The libraries were sequenced on an Illumina MiSeq platform using MiSeq reagent kit v2 (300 cycles) (Illumina, San Diego, USA). The raw sequence data have been deposited in the GEO (Gene Expression Omnibus) database (GSE102626) which has a link to the NCBI Sequence Read Archive (SRA) database (SRA study: SRP115405, bioproject: PRJNA398236).

The adapter sequences were removed from the raw reads, and low-quality sequences were filtered out. The high-quality reads were mapped onto the genome sequence of the *N. resinovorum* SA1 strain (assembly: GCA_001742225.1). The CLC Genomic Workbench 7 (Qiagen Bioinformatics, Aarhus, Denmark) was used for data pre-processing and alignment.

Reliability of the data and statistical evaluation of the replicas and the differentially expressed genes

Principal component analysis (PCA) with the count per million (CPM) data was carried out with the R Stats Package. The CPM values were calculated with the use of $\log(\text{counts} + 2)$ values and normalized library size.

The EdgeR R package (Robinson et al. 2009) was used for differential gene expression (DEG) analysis. In the calculations, only those genes which had at least one read per million in at least three samples were included. The *p* values obtained were adjusted according to the Benjamini and Hochberg approach for controlling the false discovery rate (FDR). Genes with an adjusted *p* value of ≤ 0.001 and at least log fold change ($\log\text{FC}$) ≥ 2 were considered to be differentially expressed.

Annotation, enrichment analysis, and sequence comparisons

For the bioinformatic analysis, the *N. resinovorum* SA1 genome (assembly: GCA_001742225.1) was used from the Ensembl genome database. The InterPro and Pfam annotations were obtained from the Uniprot proteome database (ProteomID UP000094626). The KEGG annotations were obtained from the KEGG database (entry: T04490). The assignments of proteins to the cluster of orthologous group (COG) categories were performed by the emapper.py python script (Huerta-Cepas et al. 2016) and were based on the eggNOG 4.5.1 database (Huerta-Cepas et al. 2015).

For the enrichment analysis, the enricher function of the clusterProfiler (Yu et al. 2012) R package was used for the COG-associated gene set overrepresentation analysis which is based on a hypergeometric statistical test. For the KEGG-related gene set enrichment analysis (GSEA) (Subramanian et al. 2005), the gseKEGG function was used with the default settings.

The protein similarities were determined with the Emboss needle pairwise sequence alignment software with default settings.

RT-qPCR analysis

Complementary DNA (cDNA) synthesis was performed using 0.5- μ g purified total RNA using M-MLV reverse transcriptase (Promega, Madison, USA). The cDNA product was diluted twofold, and qPCR was performed with Power SYBER® Green Master Mix (Life Technologies Co., Carlsbad, USA) in ABI 7500 real-time PCR system (Life Technologies Co., Carlsbad, USA). The reaction conditions were as follows: 2 min at 50 °C and 10 min denaturation at 95 °C (once) and 15 s at 95 °C and 1 min at 60 °C (40 times). Primers were used in a final concentration of 400 nM. Primer sequences are listed in the Supplementary materials, Table S1. Cycle threshold (C_t) values were determined using the ABI

sequence detection software v1.4. (Life Technologies Co., Carlsbad, USA). The C_t values of the samples were normalized (ΔC_t) with the C_t values of the *gyrB* reference gene. The expression \log_2 fold change corresponded to the $-\Delta\Delta C_t$ value which was obtained by subtracting the ΔC_t of the reference samples from the ΔC_t of the samples of interest.

Results

The *N. resinovorum* SA1 strain was the first single isolate capable of utilization of SA as sole carbon, nitrogen, and sulfur source (Perei et al. 2001). Its genome has been sequenced (Hegedűs et al. 2017), but the development of reliable and stable genetic system failed so far. It might be due to the fact that most genes of interest are located on a multicopy plasmid, pSA3. Therefore, to get deeper insight into the SA-related metabolic and physiological processes, the complete transcript profiles of cells grown on SA and glucose (as reference substrate) were analyzed and compared. For the WTA, the *N. resinovorum* SA1 strain was grown in minimal medium supplemented with either 10 mM glucose (Glc_MM) or 10 mM sulfanilic acid (SA_MM). The Glc_MM medium was supplemented with 10 mM NH_4Cl as nitrogen source, while in the case of the SA_MM, SA served as both carbon and nitrogen source. Thus, both the initial carbon and nitrogen concentrations were the same for the two substrates. No effect indicating nitrogen limitation was observed in the case of the SA (data not shown). The cell growth and the substrate concentrations were monitored in time (Fig. 1). The complete consumption of the substrates coincides with the stationary phases on the growth curves. In each case, the experiments were performed with two biological replicates.

Whole-transcriptome analysis and reliability of the raw data

Samples were collected at the exponential- and stationary-phase cultures for transcriptome profiling as indicated in Fig. 1. RNA was isolated, and WTA was performed according to the protocol described in the “Materials and methods” section. The reads obtained from the new-generation sequencing (NGS) runs were subjected to bioinformatic and statistical analyses. In the case of samples taken from the exponential phase, nearly 50% of the trimmed reads could be mapped onto the coding regions of the *N. resinovorum* SA1 reference genome (Hegedűs et al. 2017), while for the samples derived from the stationary-phase cells, this ratio was slightly lower (Table 1.). Unfortunately, presumably due to a sequencing error, the sequencing of the second replicate of the glucose-based stationary-phase sample (Glc_S2) resulted in unexpectedly low mappable read number. Although the statistical analysis revealed good correlation between the data of Glc_S1 and Glc_S2 samples (see Supplementary materials, Fig. S1),

owing to the low number of usable reads of Glc_S2, this dataset was left out from further analyses.

PCA was performed to evaluate similarities and variations in the overall gene expression datasets of the samples. Figure 2 shows that although the genome-wide variances of the biological replicates are negligible, the data derived from various sampling points remarkably differ. The overall transcript profiles of the stationary-phase cells grown on different substrates (Glc_S and SA_S) showed the highest similarity. Interestingly, the gene expression pattern of the SA_E and the Glc_S samples was in closer relationship than that of the SA_E and the SA_S. It indicates a starvation-like effect on the global gene expression in cells grown on SA.

Differential gene expression analysis

To investigate the effect of the substrate and the growth phase on the gene expression, the transcript levels of the individual genes in the various samples were compared by the EdgeR package (Robinson et al. 2009). Genes with at least fourfold changes ($\log_2\text{FC} \geq 2$) and with a FDR-corrected p value ≤ 0.001 were accepted as DEGs.

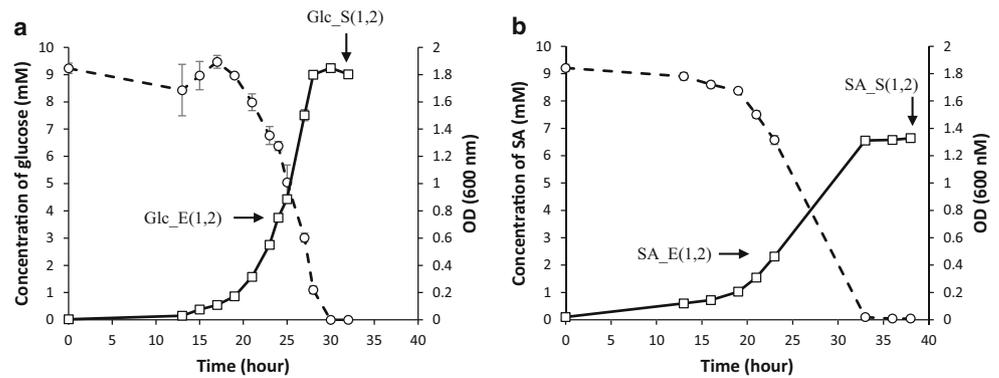
Table 2 summarizes the data of the DEGs of the three sample pairs. The shift from exponential to stationary growth phase resulted in the largest number of DEGs. In the case of cells grown on glucose and SA, 17.7 and 13.5% of the genes were differentially expressed in the two growth phases, respectively. Since in the stationary phases, no substrates were present (Fig. 1), in these cases, global starvation effect on the gene expression was expected and observed. However, SA had also substantial (specific) effect on the transcript profiles of the cells, since the comparison of the transcriptome of the exponential-phase cells (SA_E vs Glc_E) showed that a significant portion of the genes (7.3% of total genes) expressed differentially in these samples.

The global effect of SA on the gene expression pattern

For more reliable orthology assignment, the *N. resinovorum* SA1 genome was reannotated with the eggNOG-mapper (Huerta-Cepas et al. 2016) based on the eggNOG 4.5.1 database. As a result, 84.8% of the genes have been assigned into either of the COGs (Galperin et al. 2015). Based on this annotation, the DEGs described above have also been categorized into these COGs. Hypergeometric test-based gene set overrepresentation analysis was performed (Yu et al. 2012) to determine the effects of growth phase/starvation (Glc_S vs Glc_E, SA_S vs SA_E) as well as the substrate (SA_E vs Glc_E) on the relative transcript levels of the genes belonging to various COGs (Fig. 3).

It can be seen that the shift in the growth phase of the cells grown on glucose led to a global effect of starvation (Glc_S vs Glc_E). In the Glc samples, the gene expression of the most

Fig. 1 Growth of *N. resinovorum* SA1 **a** on glucose (Glc) and **b** on sulfanilic acid (SA). The error bars indicate the standard deviations of two biological replicates. The arrows mark the points where the samples were taken for transcriptome analysis



important functional groups such as the “energy production and conversion” (C), the “amino acid transport and metabolism” (E), the “nucleotide transport and metabolism” (F), the “coenzyme transport and metabolism” (H), the “translation ribosomal structure and biogenesis” (J), and the “posttranslational modification, protein turnover, chaperon” (O) COGs was substantially down-regulated in the stationary phase.

In contrast, the effect of starvation was not so pronounced in the comparison of the transcriptome of the SA_S and SA_E samples. In the SA_S cells, the transcript levels of the “C,” “F,” “H,” and “J” COGs decreased to various extents, while the genes of “secondary metabolite biosynthesis, transport, and catabolism” (Q) group were upregulated (Fig. 3).

The comparison of the SA_E vs Glc_E transcriptomes revealed that SA had a non-specific starvation-like effect on the

gene expression profile (“J,” “H,” and “E” COGs). This is especially apparent for the genes belonging to the “translation, ribosomal structure, and biogenesis” (J) functional group. The invariance in the transcript profiles of the “O” (post-translational modification, protein turnover, and chaperones) COG in the SA_E vs Glc_E comparison strongly indicates that SA has a starvation-like but not other stress-related—e.g., toxic—effect.

A similar tendency was observed after running a GSEA (Subramanian et al. 2005) with clusterProfiler R package using the KEGG pathway gene set (Kanehisa et al. 2016). Almost all significantly enriched gene sets had a decreased expression; only the transcription of the genes of the “degradation of aromatic compounds (nre01220)” pathway category had elevated expression in the presence of SA (see Supplementary materials, Fig. S2).

Table 1 Summary of Illumina RNA-Seq data

Sample ID.	Total reads	Trimmed reads in pairs	Mapped trimmed reads in pair to genes	%
Glc_E1	2 989 506	1 388 729	664 559	47.9
Glc_E2	3 151 936	1 467 686	752 965	51.3
Glc_S1	1 437 618	699 741	368 393	52.6
<i>Glc_S2</i>	<i>3 650 414</i>	<i>1 321 804</i>	<i>90 268</i>	<i>6.8</i>
SA_E1	3 319 482	1 558 226	834 220	53.5
SA_E2	2 694 052	1 319 173	736 529	55.8
SA_S1	1 822 426	784 532	268 548	34.2
SA_S2	1 736 428	802 557	292 106	36.4

The Glc_S2 sample with low usable reads is in italics

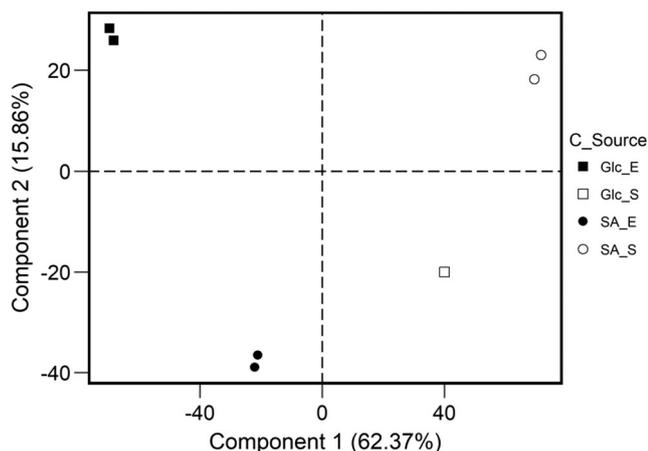


Fig. 2 Principal component analysis (PCA) of the general transcriptome characteristics. The first principal component (component 1) accounted for 62.37% and the second principal component (component 2) for 15.86% of the total variance in the dataset. The plot indicates that the transcriptome data are highly reproducible

Expression level of the genes involved in SA catabolism

From the previous sections, it was concluded that SA as a sole carbon source has specific and non-specific (starvation-like) effects on the transcriptome of the cells. It is plausible to assume that the expression of the genes encoding proteins involved in the conversion of SA to 3-oxoadipate is induced by SA. However, the genes of the peripheral part of the catabolic pathway had remarkable transcript levels in the exponential-phase cells almost independently of the substrate used (Fig. 4). The *sadA* (ring-hydroxylating dioxygenase) and *sadB* (glutamine synthase) genes were among the top 20 highly expressed genes together with genes coding for ribosomal proteins, chaperons, and translation initiation factors. The basal expression levels of the *sadAB* genes increased almost four-fold during the GlcE-to-GlcS shift, and their transcript levels were barely higher in the SA_E than in the Glc_E samples. An exponential-to-stationary-phase shift resulted in a moderate decrease in the messenger RNA (mRNA) level of these genes in the cells grown on SA.

In contrast, the genes involved in the central part of the pathway (*scaABCDEF*) had quite low basal transcription level

in the Glc_E and Glc_S cells (Fig. 4). However, in the exponential-phase cells growing on SA, a 64-fold higher expression of these genes could be observed as compared to the cells cultivated on glucose (SA_E vs Glc_E). The consumption of the substrate resulted in a dramatic drop in their expression level almost to the basal level measured in Glc_E and Glc_S cells.

The genes coding for the third part of the pathway—which is a common step in all β -ketoacid pathways—had an unexpectedly low and hardly changing transcript levels in each sample analyzed (Fig. 4).

Identification of the auxiliary genes in the SA degradation

Ferredoxin reductase/ferredoxin

The main step of the peripheral part of the SA catabolic pathway is the dihydroxylation of the aromatic ring with a concomitant oxidation of NAD(P)H and removal of the amine group. This reaction is catalyzed by a Rieske oxygenase (RO) enzyme coded by the *sadA* gene. This ring-hydroxylating enzyme belongs to the $\text{I}\alpha$ -type oxygenase group according to the Kweon classification (Kweon et al. 2008). The ring hydroxylation is linked to an electron transport chain (ETC) which transfers electrons from NAD(P)H to the oxygenase. Usually, this is a two-component system consisting of a flavoprotein reductase and a redox carrier, typically ferredoxin.

In the annotated genome of *N. resinovorum* SA1, six potential flavoprotein reductase genes were identified whose products might participate in the RO enzyme-linked electron supply: five gene products belonged to group I while one gene encoded for a group II reductase according to the classification described in Kweon et al. (2008). The comparative transcriptome analysis of the cells revealed only one flavoprotein reductase gene, the *ferB* gene (BES08_30695), which had substantial and SA-inducible expression in the exponential-phase cells (Fig. 4, SA_E vs Glc_E). Interestingly, similarly to the other peripheral elements of the SA catabolic pathway, the *ferB* had also a relatively high expression level in the exponential-phase cells

Table 2 The numbers of differentially expressed genes in various sample pairs

	Glc_S vs. Glc_E	SA_E vs. Glc_E	SA_S vs. SA_E
All DEGs	1059 (17.7%)	437 (7.3%)	808 (13.5%)
Up DEGs	534 (8.9%)	256 (4.3%)	370 (6.2%)
Down DEGs	525 (8.8%)	181 (3.0%)	438 (7.3%)

All DEG values are their sum

Up DEGs genes being upregulated, Down DEGs genes being downregulated

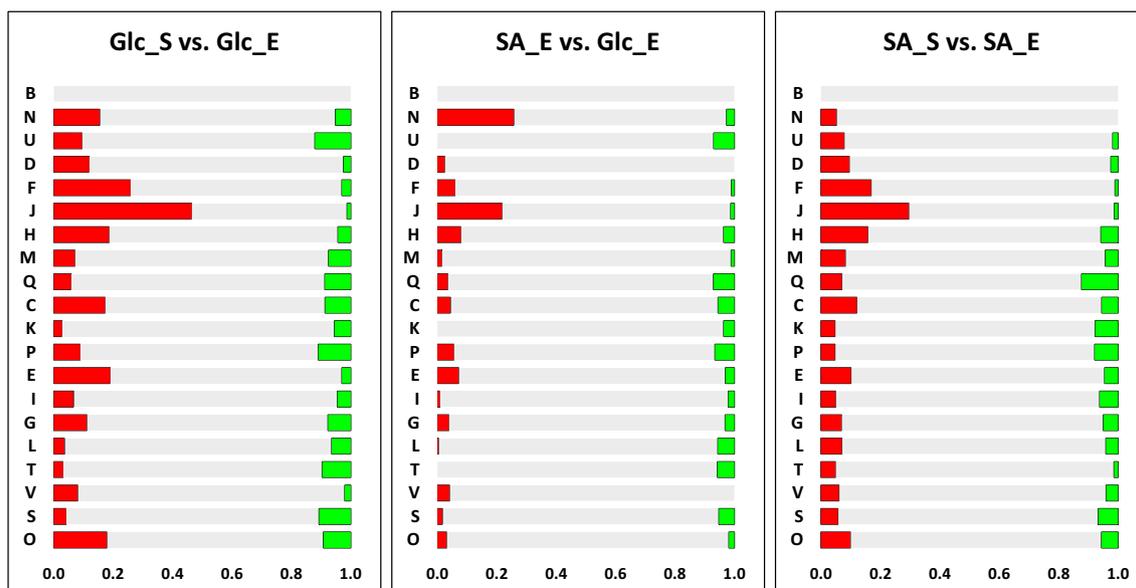


Fig. 3 COG gene set overrepresentation analysis of differentially expressed genes in *N. resinovorum* SA1. The *x*-axis represents the normalized COG category size. The left side bars (dotted) indicates the proportion of the genes with reduced expression while the right side bars (slope lined) indicates the proportion of the genes with increased expression. Genes classified in more than one category were counted more than once. The letters mean the following COGs: N, cell motility; D, cell cycle control, cell division, and chromosome partitioning; J,

translation, ribosomal structure, and biogenesis, M, cell wall/membrane/envelope biogenesis; C, energy production and conversion; P, inorganic ion transport and metabolism; I, lipid transport and metabolism; L, replication, recombination, and repair; O, post-translational modification, protein turnover, and chaperones. Significance levels are indicated with asterisks: * $p < 0.05$; ** $p < 0.001$ by hypergeometric test (Color figure online)

grown on Glc which further increased upon the exponential-to-stationary growth transition (G_S vs. G_E). The other flavoprotein reductase genes had relatively low and invariant expression levels in all samples examined. The *ferB* gene is located on the pSA3 plasmid, around 3 kbp downstream of the *sadAB* genes. It is clustered together with a ferredoxin coding gene, *ferA* (BES08_30690), having a Fer2 (PF00111) pfam domain. A similar gene-encoding ferredoxin (*sadD*) (Gan et al. 2012) was identified in the *sad* gene cluster of the *Hydrogenophaga* PBC strain; however, the ferredoxin reductase component was missing from that locus. The similarity of the FerA and SadD proteins of the two strains was 57%.

Transporters

Since SA is a polar, zwitterionic molecule, its cellular uptake is difficult (Hwang et al. 1989) and it requires active transport system. According to the KEGG BRITE functional classification of the *N. resinovorum* SA1 genome, 360 genes encode proteins that apparently have transporting function (“Transporters,” nre02000). Among these, only 13 genes were upregulated by SA; 12 of them were related to ion transport especially to iron transport (Table 3). One of the iron transporter genes, the BES08_23675 gene, had prominently sensitive expression for SA. This might be expected since the aromatic catabolic pathways comprise many iron-containing proteins. However, one

gene product (BES08_30350) belongs to the organic acid transporters, the major facilitator superfamily (MFS) (Pao et al. 1998), and particularly to the aromatic acid/H⁺ symporter (AAHS) family. Many transporters of this family have already been identified which had a role in the admission of aromatic acid compounds through the membrane. The members of the AAHS family contain a characteristic MFS_1 (PF07690) pfam motif. Using this motif, 67 transporter genes were identified in the genome and only the BES08_30350 gene had SA-dependent expression. The expression of this gene was high in the presence of glucose (Glc_E) and did not change with the growth phase. However, it was substantially elevated in the SA_E samples (Table 3). After consumption of SA, its transcript level significantly decreased (data not shown).

The BES08_30350 is located on the pSA3 plasmid clustered together with the genes presumably encoding the benzoate 1,2-dioxygenase large (BES08_30344) (*benA*) and small (BES08_30345) (*benB*) subunits which further corroborates its proposed role in the aromatic acid transport.

The fate of sulfite

The assimilation of SA as nutrient source is inevitably associated with excess sulfite formation which is toxic for the cells. Therefore, the sulfite must be converted to another sulfur compound (reduction to sulfur or sulfide or oxidation to sulfate). A

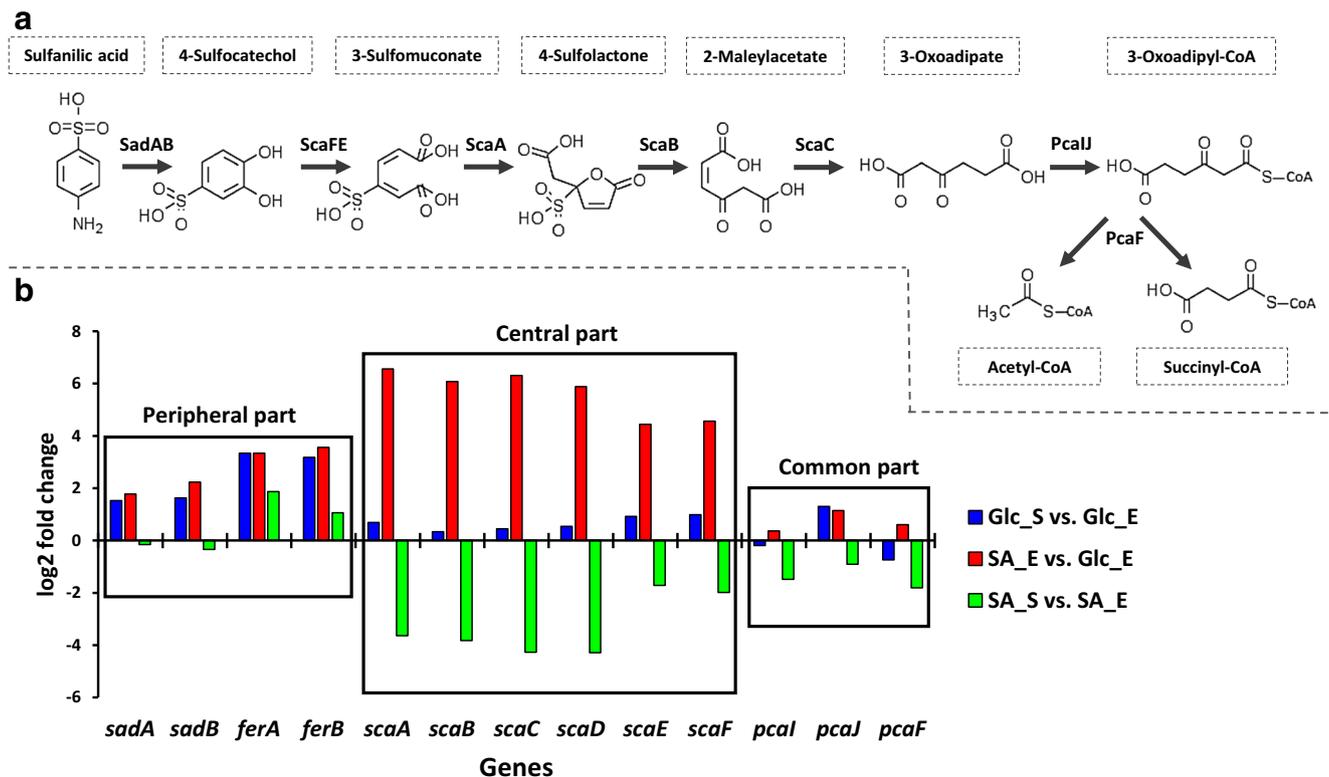


Fig. 4 Proposed enzymatic reaction set (**a**) and the relative changes in the transcript level of the genes involved in the peripheral, central, and common parts of SA conversion (**b**). The genes coding for the following enzymes are displayed: *sadA*, 4-aminobenzenesulfonate 3,4-dioxygenase; *sadB*, glutamine synthetase; *ferA*, ferredoxin; *ferB*, ferredoxin reductase; *scaA*, 3-sulfo-cis-cis-muconate cycloisomerase; *scaB*,

sulfolactone hydrolase; *scaC*, maleylacetate reductase; *scaD*, putative oxidoreductase; *scaE*, sulfocatechol 3,4-dioxygenase beta subunit; *scaF*, sulfocatechol 3,4-dioxygenase alpha subunit; *pcaI*, 3-oxoadipate CoA transferase subunit A; *pcaJ*, 3-oxoacid CoA-transferase subunit B; *pcaF*, 3-oxoadipyl-CoA thiolase

small part of sulfur derived from SA is utilized in the cellular materials, but *N. resinovorum* SA1 strain releases substantial amount of sulfate from SA to the medium (Magony et al. 2007). Thus, in this bacterium, sulfite oxidation and sulfate

export are the dominant detoxification processes which are common in bacteria (Kappler 2011). Sulfite oxidation can take place through the cytoplasmic formation of adenylylphosphosulfate (APS) (indirect pathway), or it can be oxidized in the periplasm

Table 3 Transporter genes having SA-inducible transcription

Locus_tag	Category	Function	Log2FC
BES08_09020	Prokaryotic type ABC transporters	Iron complex transport system ATP-binding protein [EC:3.6.3.34]	2.14
BES08_30350	Major Facilitator Superfamily (MFS)	MFS transporter, AAHS family, 4-hydroxybenzoate transporter	2.43
BES08_08830	Other Transporters	Iron complex outermembrane receptor protein	2.55
BES08_14745	Other Transporters	Outer membrane immunogenic protein	2.29
BES08_18055	Other Transporters	Iron complex outermembrane receptor protein	2.92
BES08_18430	Other Transporters	Iron complex outermembrane receptor protein	2.43
BES08_18580	Other Transporters	Iron complex outermembrane receptor protein	2.33
BES08_22625	Other Transporters	Iron complex outermembrane receptor protein	2.85
BES08_23675	Other Transporters	Iron complex outermembrane receptor protein	6.71
BES08_24205	Other Transporters	Outer membrane immunogenic protein	2.20
BES08_26825	Other Transporters	Iron complex outermembrane receptor protein	2.11
BES08_28345	Other Transporters	Magnesium transporter	2.46
BES08_28525	Other Transporters	Na ⁺ :H ⁺ antiporter, NhaA family	6.23

The SA (aromatic acid) transporter is in italics

by the members of the sulfite oxidase (SO) family (direct route). The members of the SO enzyme family contain a characteristic Oxidored_molyb (PF00174) pfam domain. The direct route is more widespread in bacteria; therefore, we searched for SA-inducible sulfite oxidase genes in the genome and transcriptome of *N. resinovorum* SA1.

Two genes were found (BES08_29795, BES08_29830) likely coding for SO enzymes. The comparative transcriptome analysis revealed that the expression of both genes could be induced by SA (Fig. 5). The mRNA level of these genes decreased in the stationary phase of the growth curve. Both genes are located on the pSA3 plasmid relatively close to each other. According to in silico analysis, both proteins have Mo-co_dimer (PF03404) pfam domain and therefore belong to the group 2 especially to the “2C subgroup” of SO enzymes (Kappler 2011). The deduced sequence of BES08_29830 is more similar to the SorA (Kappler et al. 2000) protein (around 60% similarity), which is a well-characterized member of the 2C subgroup. The SorA has a twin arginine-type (TAT) leader sequence (IPR006311) which allows the TAT-dependent export of the enzyme to the periplasm. Moreover, the gene is clustered together with another *orf* encoding a cytochrome c-like protein (BES08_29825 with IPR009056 domain) which is characteristic to the heme-containing subunit (SorB) of SOs. Concerning the other gene product (BES08_29795), its further classification was difficult due to the low sequence similarities. Thus, it is named as SorX.

Since SO enzymes are located in the periplasm, sulfite should be exported from the cytoplasm by proper transporters. Sulfite exporters can be divided into two groups according to

their pfam motifs. The members of the first group where the TauZ/CuyZ/SuyZ transporters belong possess a Cons_hypoth698 (PF03601) pfam domain (Brüggemann et al. 2004; Rein et al. 2005; Denger et al. 2006; Denger et al. 2009) while the members of the second group possess a TauE (PF01925) pfam domain (Weinitschke et al. 2007).

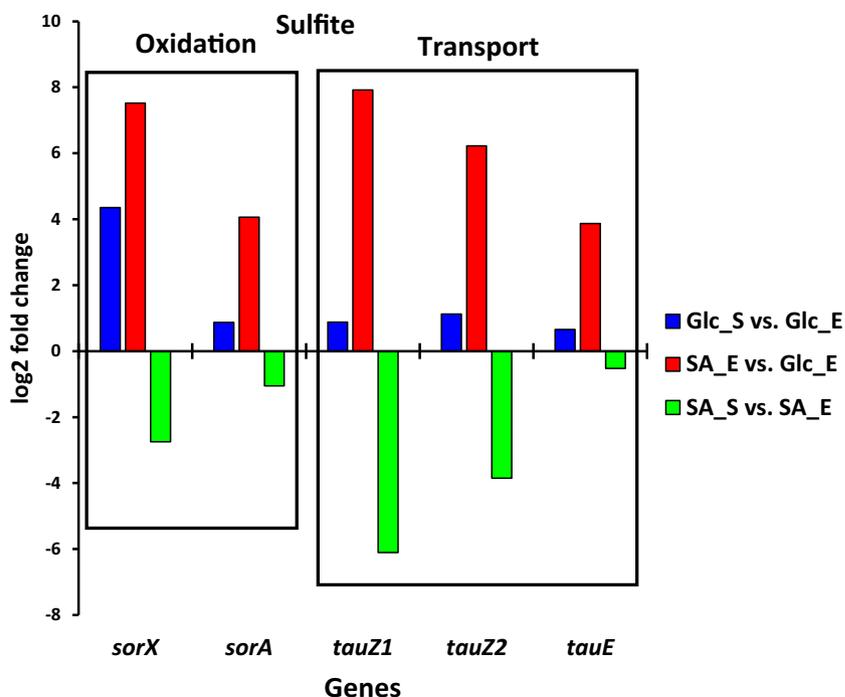
In the *N. resinovorum* SA1 strain, 13 genes coding for potential sulfite transporter were found. Among these, four and nine genes code for proteins containing the PF03601 or PF01925 domains, respectively. Three transporter genes (BES08_29790, BES08_30485, and BES08_30490) had substantially higher expression in the SA_E than the Glc_E samples (Fig. 5). In each case, the transcript levels diminished in the stationary phase. The deduced proteins of BES08_29790 and BES08_30485 have the PF03601 domains and resemble the TauZ proteins (46.4 and 79.5% similarities). The third transporter comprising a PF01925 domain (BES08_30490) apparently is a TauE-type protein.

All these SA-inducible transporter genes are located on the pSA3 plasmid harboring the SO genes, as well. Furthermore, the BES08_29790 (TauZ1) (PF03601 group) occurs in the vicinity of the BES08_29795 (SorX) SO gene while the other two putative transporter genes, the BES08_30485 (TauZ2) and BES08_30490 (TauE), are located immediately upstream of the *sea* operon of the SA catabolism.

Genes of 3-oxoadipate conversion

The third part of SA degradation is a common biochemical reaction set converting 3-oxoadipate to its CoA derivative

Fig. 5 Relative changes in the transcript level of the genes likely involved in sulfite oxidation and transport



then to succinyl-CoA and acetyl-CoA. The PcaIJ and PcaF enzymes are involved in these reactions. The corresponding genes were found in the *N. resinovorum* SA1 genome; they were localized on the pSA1 plasmid. Surprisingly, the transcript analyses revealed that the basal expressions of these genes were unexpectedly low and apparently did not change in the presence of SA.

RT-qPCR analysis of selected genes

In order to confirm the RNA-Seq results described above, RT-qPCR analyses were carried out using exponential-phase cells grown on glucose and sulfanilic acid (Fig. 6, SA_E vs. Glc_E comparisons). Hydroxybenzoate was additionally included in the experiments, since it is also an aromatic xenobiotic which is converted by another independent enzyme set to 3-oxoadipate (Fig. 6, HB_E vs Glc_E comparison). Therefore, the expression levels of genes selected from various parts of the various metabolisms could be monitored and compared. The following target genes were chosen: *sadA*, which codes for the oxygenase of the peripheral SA pathway; *scaE* and *pcaH* genes of the ring-cleaving dioxygenase subunit of the SA and HB pathways, respectively; *pcaJ*, which encodes a subunit of the enzyme-converting 3-oxoadipate to 3-oxoadipate-CoA; *ferB*, ferredoxin reductase gene (BES08_30695); *pcaK*, SA-inducible transporter gene

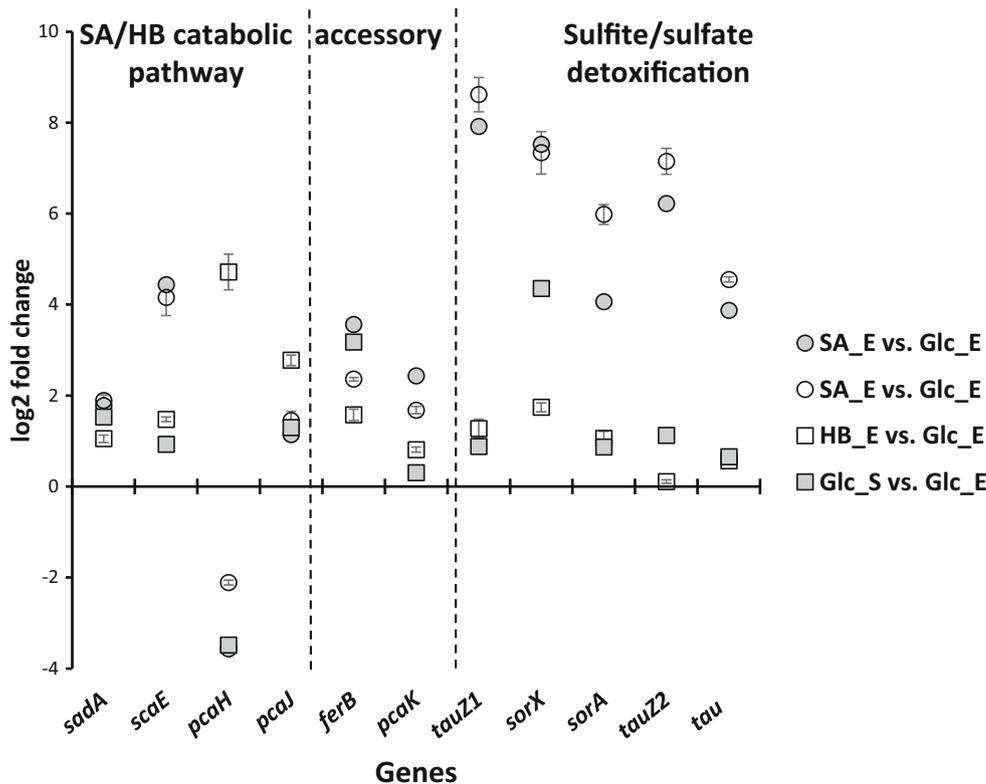
(BES08_30350); and the genes coding for the sulfite transporters and oxidase described above.

The comparison of the RNA-Seq and RT-qPCR analyses is illustrated in Fig. 6. It can be seen that the RT-qPCR experiments confirmed the conclusions of the WTA. There are some deviations in the fold change values, but these do not modify the conclusions. There is a strong substrate dependency of the expression of the genes of the SA and HB central pathways (*scaE*, *pcaH*). The sulfite/sulfate detoxification and the transporter genes had apparently higher mRNA levels in cells grown on SA as compared to the other samples. The substrate-specific induction of the *sadA* and *ferB* genes is less obvious; the elevated expression might be a combined effect of the substrate and starvation. The expression level of the *pcaK* transporter was slightly upregulated by SA but not by HB.

For most of the genes selected, similar tendencies could be observed for the HB_E vs Glc_E and Glc_S vs Glc_E samples. Since the latter comparison represents the starvation effect, it might be concluded that using HB as substrate might also trigger a starvation-like response.

The low expression level of the *pcaJ* gene coding for an enzyme involved in the common part of both SA and HB metabolisms was confirmed by the RT-qPCR experiments. However, slightly elevated—but still low—mRNA level could be measured in the cells cultivated on HB.

Fig. 6 RT-qPCR analysis of selected genes in cells grown on SA, Glc, and HB. Changes in the transcript level of the genes involved in the catabolism (left), substrate and electron transport (middle), and sulfite removal (right). The filled and open shapes represent the results obtained from RNA-Seq and RT-qPCR, respectively. Circles: validation of the RNA-Seq (filled) experiments with RT-qPCR (open) for the SA_E vs Glc_E sample pair. Open squares: novel dataset for HB_E relative to Glc_E (RT-qPCR). Filled squares illustrate (remind) the starvation effect on the expression of these genes (RNA-Seq)



Discussion

Sulfanilic acid is a widely used compound in numerous industrial processes. It cannot be easily taken up by microbes, but within the cell, it is strongly toxic by inhibiting the folic acid biosynthesis. In spite of the incidence of this compound, there are only few microbial strains or co-cultures capable of its complete biodegradation (Feigel and Knackmuss 1988, 1993; Dangmann et al. 1996; Contzen et al. 2000; Perei et al. 2001; Singh et al. 2004, 2006; Wang et al. 2009; Gan et al. 2011b; Zhang et al. 2012). The main components of the SA bioconversion have already been disclosed in microbial mono- and co-cultures.

N. resinovorum SA1 was the first single isolate which could utilize SA as sole carbon, nitrogen, and sulfur source. Its genomic sequence has recently been published; thus, the strain is an ideal candidate to study the global and specific transcriptomic responses of the cells exposed to SA. Since SA (or other substituted aromatic xenobiotics) is far from being an optimal carbon source for the cells, thus, a starvation-like global effect might also be expected in cells growing on SA. It has been estimated that the ATP gained from a monoaromatic compound is around 50% of that from glucose (Haußmann et al. 2009).

In the sections above, the transcriptional responses of *N. resinovorum* SA1 cells were presented, where the reference culture was grown on Glc. Two growth phases were chosen: the exponential and late stationary phases; the latter represented the starvation effect. The transcript profiles were compared at the global as well as the single-gene levels.

According to the previous studies (Hegedűs et al. 2017), the enzymes directly involved in the specific metabolic conversion of SA are the following: SadAB and ScaABCDEF. Moreover, there must be accessory proteins assisting in related processes, such as electron transport, substrate and byproduct transport, and supplementary redox reactions among others. However, the comparison of the SA_E vs Glc_E samples revealed 437 differentially expressed genes. A statistical comparison of the expression profiles of Glc_S vs Glc_E, SA_E vs Glc_E, and SA_E vs SA_S samples revealed that the SA_E cells had higher global similarities to the Glc_S than to the Glc_E samples, i.e., SA as substrate provoked a global starvation-like, substrate-unspecific effect. A similar physiological effect was described for *Corynebacterium glutamicum* grown on protocatechuate (Haußmann and Poetsch 2012).

On the other hand, there are genes and gene products which are more or less specifically required, consequently, specially evolved, for the metabolic conversion and assimilation of SA. The peripheral and central pathways should comprise the most specific enzymes, while the steps of the 3-oxoadipate route are common for various substituted aromatic compounds. In *N. resinovorum* SA1, the genes coding for the peripheral (*sadAB*) and central (*scaA-F*) enzymes are located on the

pSA3 plasmid in distinct loci (Fig. 7). Similar degradative plasmids are not rare in the *Sphingomonadaceae* family (Stolz 2013). The fragmentation of these gene clusters might be due to the frequent genetic change of the plasmid which is indicated by the large number of mobile genomic element around the operons of interest. To our surprise, the basal transcript level of the *sadAB* genes is high and their expression is hardly upregulated by SA. Moreover, this effect is hardly specific for SA; therefore, it might be assumed that this region has been acquired by horizontal gene transfer and the regulatory elements have been destroyed during the genetic transfer/rearrangement. A wide-scale comparison of the codon usages of the genes on the pSA3 plasmid and other genetic elements revealed that the codon usage of the peripheral genes substantially differs from the other genes located on either the chromosome or the plasmid (data not shown). Thus, the peripheral genes seem to be younger than the genes of the central and common pathways. In contrast to the old genes, the “young” ones are not clustered and regulated (Cases and de Lorenzo 2001). Thus, the fact that the peripheral genes coding for proteins responsible for the conversion of SA to a catecholic intermediate (4-sulfocatechol) seem to be constitutively expressed (Fig. 4) is not astonishing for such “new” enzymes located on a plasmid. A similar phenomenon was observed for the *atzABC* genes in the *Pseudomonas* sp. ADP strain (de Souza et al. 1998). The high expression level of the peripheral genes should be evolutionarily advantageous during the adaptation to a novel substrate (Szappanos et al. 2016). It is also to note that the efficacy of the SA-sulfocatechol conversion must be high since any free SA in the cytoplasm would be toxic for the cell. Therefore, the SadAB proteins must convert all SA entering into the cells which require their sufficient expression levels. Since the comparative global expression profiling did not indicate a stress (toxic) response in the SA_E vs Glc_E

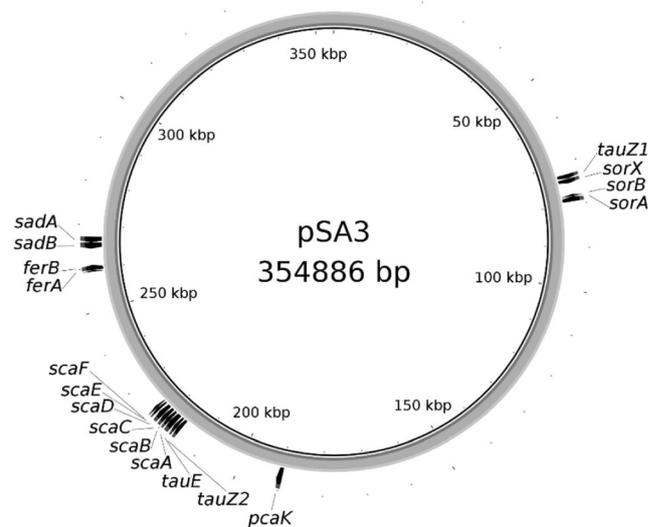


Fig. 7 Genes related to SA metabolism on the pSA3 plasmid

comparison, it was concluded that the SA conversion was efficient, likely complete, in the case of *N. resinovorum* SA1.

Similar expression pattern—as well as codon usage—could be observed for the electron transfer (ET) elements which are important for the function of the RO enzymes. The RO-specific flavoprotein reductases can be divided into three groups according to their domain composition (Kweon et al. 2008). The group I reductases consist of glutathione reductase (GR) with Pyr_redux_2 (PF07992) and Reductase_C (PF14759) pfam domains. The group II and III reductases include the ferredoxin-NADPH+ (FNR)-type reductases with FAD_binding_6 (PF00970), NAD_binding_1 (PF00175), and Fer2 (PF00111) pfam domains. Although the group II and III reductases share the same domain composition, they differ in those arrangements (Kweon et al. 2008). In *N. resinovorum* SA1, the ET components consist of a GR-type flavoprotein reductase and a Fer2-type ferredoxin component. This type of ETCs is mostly characteristic for the type 4 oxygenases (Kweon et al. 2008), but the SadA is a type I oxygenase. In contrast, in *Hydrogenophaga* sp. PBC, an FNRc reductase (SadC) was assumed which is a typical flavoprotein reductase of the type 1 oxygenases (Kweon et al. 2008). However, this reductase contains a plant-type ferredoxin domain; therefore, an independent ferredoxin component is not essential for its proper function. Nevertheless, a ferredoxin-coding gene, *sadD*, is part of the peripheral pathway encoding the *sadABD* operon in the *Hydrogenophaga* sp. PBC strain. Interestingly, the recombinant SadABD and SadAB + SadC protein combinations in *Escherichia coli* had similar activities (Gan et al. 2012). The SadD had significant identity (57%) with the ferredoxin component of the ETC identified in this work. As a sum, the electron transfer components of *N. resinovorum* SA1 and *Hydrogenophaga* species are apparently different. While in the *Hydrogenophaga* strains, a single FNRc reductase (with a ferredoxin domain) can support the activity of the SadA oxygenase, in *N. resinovorum* SA1, two components, a GR-type flavoprotein reductase and an independent ferredoxin, are required for the SadA-coupled redox reaction.

Contrary to the genes of peripheral elements, the expression of the *sca* genes (central pathway) was specifically inducible by SA, indicating the presence of a substrate-specific regulatory mechanism in the strain. For *Hydrogenophaga* sp. PBC, it was shown that the transcription of both the *sadABD* and the *sca* gene clusters was upregulated by SA or sulfocatechol (Gan et al. 2012). The gene of a GntR family transcriptional factor is located upstream of the *sadABD* operon in opposite orientation. This gene product was assumed to play a role in the SA-specific regulation (Gan et al. 2012). However, no transcription factor-coding gene was found in a reasonable

proximity of the *sadAB* genes in our strain. Thus, the transcriptional regulatory mechanism of the SA-dependent control is still to be disclosed.

The expression of the substrate transporter genes should also be substrate-inducible. A number of aromatic acid transporters have been identified for various compounds, such as the PcaK for 4-hydroxybenzoate and protocatechuate (Harwood et al. 1994; Nichols and Harwood 1995, 1997; Ditty and Harwood 1999, 2002; Pernstich et al. 2014), BenK for benzoate (Collier et al. 1997), OphD for phthalate (Chang and Zylstra 1999), MhbT for 3-hydroxybenzoate (Xu et al. 2012), and TfdK for 2,4-dichlorophenoxyacetate (Leveau and Zehnder 1998). The transporter, likely responsible for the SA uptake, belongs to the aromatic AAHS family. The transcription of the gene was apparently upregulated by SA, but it is clustered together with the genes of RO α and β subunits of the benzoate pathway (*benAB*). Therefore, this transcriptional inducibility might not be fully SA-specific.

There are other genes which have virtually SA-sensitive gene expression, such as the genes of transporters and sulfite oxidases. Intracellular sulfite release during the SA degradation generates serious burden for the cells. The direct oxidation of the reactive sulfite to inert sulfate is the most common detoxification approach in bacteria. The oxidation process basically takes place in the periplasm by sulfite-oxidizing enzymes (SOEs) where the intracellular-formed sulfite reaches through sulfite exporters. Three possible sulfite exporters and two SOEs have been identified which likely play a role in the sulfite detoxification. Two of the sulfite exporters showed similarity to the TauZ/SuyZ/CuyZ sulfite exporters and one to TauE sulfite exporter. One SOE is similar to SorA while the typing of other SOE (SorX) is still uncertain. Although these genes were induced by SA, the signal of the control was rather sulfite- than SA-specific. Nevertheless, the tight genetic linkage of the *tauE3* and *tauE2* with the *sca* genes reflects a functional relationship. In *Hydrogenophaga* sp. PBC, the sulfite oxidation gene set was similar, but only one sulfite transporter (TauE) could be identified by BLAST search.

As a summary, the transcriptional response of *N. resinovorum* SA1 cells exposed to SA could be related to various regulatory mechanisms: (a) a global profile change was linked to a nonspecific starvation-like effect; (b) a specific SA-dependent regulation could be observed for the genes of the central pathway; (c) aromatic-specific control of the substrate transport genes was noticed which was apparently adapted to the novel substrate, SA; (d) transcript level changes induced by the by-products of SA metabolism could also be outlined, such as sulfite-induced genes of sulfite oxidation and export; and (e) one might expect the aromatic acid-dependent inducibility of the genes of the 3-oxoadipate conversion (*pcaJIF*), but this could not be observed in our case.

Acknowledgements The authors thank Balázs Bálint and Balázs Horváth for their kind help in the data processing of the first preliminary RNA-Seq experiments.

Funding information This study was funded by the European Union and co-financed by the European Social Fund (grant agreement nos. TÁMOP-4.2.4.A/2-11/1-2012-0001 “National Excellence Program,” (Krisztián Laczi), TÁMOP-4.1.1.C-12/1/KONV-2012-0012, and EFOP-3.6.2-16-2017-00010) and by the Norway Grant (grant agreement no. HU09-0044-A1-2013).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent The research presented did not involve any human participants.

References

- Brown G (1962) The biosynthesis of folic acid: II. Inhibition by sulfonamide. *J Biol Chem* 237:536–540
- Brüggemann C, Denger K, Cook AM, Ruff J (2004) Enzymes and genes of taurine and isethionate dissimilation in *Paracoccus denitrificans*. *Microbiology* 150:805–816. <https://doi.org/10.1099/mic.0.26795-0>
- Cases I, de Lorenzo V (2001) The black cat/white cat principle of signal integration in bacterial promoters. *EMBO J* 20:1–11
- Chang HK, Zylstra GJ (1999) Characterization of the phthalate permease OphD from *Burkholderia cepacia* ATCC 17616. *J Bacteriol* 181:6197–6199
- Collier LS, Nichols NN, Neidle EL (1997) *benK* encodes a hydrophobic permease-like protein involved in benzoate degradation by *Acinetobacter* sp. strain ADP1. *J Bacteriol* 179:5943–5946
- Contzen M, Moore ERB, Blumel S, Stolz A, Kämpfer P (2000) *Hydrogenophaga intermedia* sp. nov., a 4-aminobenzenesulfonate degrading organism. *Syst Appl Microbiol* 23:487–493. [https://doi.org/10.1016/S0723-2020\(00\)80022-3](https://doi.org/10.1016/S0723-2020(00)80022-3)
- Dangmann E, Stolz A, Kuhm AE, Hammer A, Feigel B, Noisommit-Rizzi N, Rizzi M, Reuss M, Knackmuss H-J (1996) Degradation of 4-aminobenzenesulfonate by a two-species bacterial coculture. *Biodegradation* 7:223–229. <https://doi.org/10.1007/BF00058181>
- de Souza ML, Wackett LP, Sadowsky MJ (1998) The *atzABC* genes encoding atrazine catabolism are located on a self-transmissible plasmid in *Pseudomonas* sp. strain ADP. *Appl Environ Microbiol* 64:2323–2326
- Denger K, Mayer J, Buhmann M, Weinitschke S, Smits THM, Cook AM (2009) Bifurcated degradative pathway of 3-sulfolactate in *Roseovarius nubinhibens* ISM via sulfoacetaldehyde acetyltransferase and (S)-cysteate sulfolylase. *J Bacteriol* 191:5648–5656. <https://doi.org/10.1128/JB.00569-09>
- Denger K, Smits THM, Cook AM (2006) L-Cysteate sulphonylase, a widespread pyridoxal 5'-phosphate-coupled desulphonative enzyme purified from *Silicibacter pomeroyi* DSS-3(T). *Biochem J* 394:657–664. <https://doi.org/10.1042/BJ20051311>
- Ditty JL, Harwood CS (1999) Conserved cytoplasmic loops are important for both the transport and chemotaxis functions of PcaK, a protein from *Pseudomonas putida* with 12 membrane-spanning regions. *J Bacteriol* 181:5068–5074
- Ditty JL, Harwood CS (2002) Charged amino acids conserved in the aromatic acid/H⁺ symporter family of permeases are required for 4-hydroxybenzoate transport by PcaK from *Pseudomonas putida*. *J Bacteriol* 184:1444–1448. <https://doi.org/10.1128/jb.184.5.1444-1448.2002>
- Dua M, Singh A, Sethunathan N, Johri A (2002) Biotechnology and bioremediation: successes and limitations. *Appl Microbiol Biotechnol* 59:143–152. <https://doi.org/10.1007/s00253-002-1024-6>
- Feigel BJ, Knackmuss H-J (1988) Bacterial catabolism of sulfanilic acid via catechol-4-sulfonic acid. *FEMS Microbiol Lett* 55:113–117. <https://doi.org/10.1111/j.1574-6968.1988.tb02807.x>
- Feigel BJ, Knackmuss HJ (1993) Syntrophic interactions during degradation of 4-aminobenzenesulfonic acid by a 2 species bacterial culture. *Arch Microbiol* 159:124–130
- Galperin MY, Makarova KS, Wolf YI, Koonin EV (2015) Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* 43:D261–D269. <https://doi.org/10.1093/nar/gku1223>
- Gan HM, Ibrahim Z, Shahir S, Yahya A (2011a) Identification of genes involved in the 4-aminobenzenesulfonate degradation pathway of *Hydrogenophaga* sp. PBC via transposon mutagenesis. *FEMS Microbiol Lett* 318:108–114. <https://doi.org/10.1111/j.1574-6968.2011.02245.x>
- Gan HM, Shahir S, Ibrahim Z, Yahya A (2011b) Biodegradation of 4-aminobenzenesulfonate by *Ralstonia* sp. PBA and *Hydrogenophaga* sp. PBC isolated from textile wastewater treatment plant. *Chemosphere* 82:507–513. <https://doi.org/10.1016/j.chemosphere.2010.10.094>
- Gan HM, Shahir S, Yahya A (2012) Cloning and functional analysis of the genes coding for 4-aminobenzenesulfonate 3,4-dioxygenase from *Hydrogenophaga* sp. PBC Microbiol (United Kingdom) 158:1933–1941. <https://doi.org/10.1099/mic.0.059550-0>
- Gibson DT, Parales RE (2000) Aromatic hydrocarbon dioxygenases in environmental biotechnology. *Curr Opin Biotechnol* 11:236–243
- Halak S, Basta T, Burger S, Contzen M, Wray V, Pieper DH, Stolz A (2007) 4-Sulfomuconolactone hydrolases from *Hydrogenophaga intermedia* S1 and *Agrobacterium radiobacter* S2. *J Bacteriol* 189:6998–7006. <https://doi.org/10.1128/JB.00611-07>
- Halak S, Basta T, Bürger S, Contzen M, Stolz A (2006) Characterization of the genes encoding the 3-carboxy-cis,cis-muconate-lactonizing enzymes from the 4-sulfocatechol degradative pathways of *Hydrogenophaga intermedia* S1 and *Agrobacterium radiobacter* S2. *Microbiology* 152:3207–3216. <https://doi.org/10.1099/mic.0.29136-0>
- Harwood CS, Nichols NN, Kim M, Ditty JL, Parales RE (1994) Identification of the *pcaRKF* gene cluster from *Pseudomonas*: involvement in chemotaxis, biodegradation, and transport of 4-hydroxybenzoate. *BioTechniques* 17:6479–6488
- Haußmann U, Poetsch A (2012) Global proteome survey of protocatechuate- and glucose-grown *Corynebacterium glutamicum* reveals multiple physiological differences. *J Proteome* 75:2649–2659. <https://doi.org/10.1016/j.jprot.2012.03.005>
- Haußmann U, Qi SW, Wolters D, Röner M, Liu SJ, Poetsch A (2009) Physiological adaptation of *Corynebacterium glutamicum* to benzoate as alternative carbon source—a membrane proteome-centric view. *Proteomics* 9:3635–3651. <https://doi.org/10.1002/pmic.200900025>
- Hegedűs B, Kós PB, Bálint B, Maróti G, Gan HM, Perei K, Rákhely G (2017) Complete genome sequence of *Novosphingobium resinovorum* SA1, a versatile xenobiotic-degrading bacterium capable of utilizing sulfanilic acid. *J Biotechnol* 241:76–80. <https://doi.org/10.1016/j.jbiotec.2016.11.013>
- Heider J, Fuchs G (1997) Anaerobic metabolism of aromatic compounds. *Eur J Biochem* 243:577–596. <https://doi.org/10.1196/annals.1419.010>
- Huerta-Cepas J, Forslund K, Szklarczyk D, Jensen LJ, von Mering C, Bork P (2016) Fast genome-wide functional annotation through

- orthology assignment by eggNOG-mapper. bioRxiv. <https://doi.org/10.1101/076331>
- Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ, von Mering C, Bork P (2015) eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* 44:286–293. <https://doi.org/10.1093/nar/gkv1248>
- Hwang SY, Berges DA, Taggart JJ, Gilvarg C (1989) Portage transport of sulfanilamide and sulfanilic acid. *J Med Chem* 32:694–698
- Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 44:D457–D462. <https://doi.org/10.1093/nar/gkv1070>
- Kappler U (2011) Bacterial sulfite-oxidizing enzymes. *Biochim Biophys Acta Bioenerg* 1807:1–10. <https://doi.org/10.1016/j.bbabi.2010.09.004>
- Kappler U, Bennett B, Deutzmann R, Mcewan AG, Dahl C (2000) Sulfite:cytochrome c oxidoreductase from *Thiobacillus novellus*. *J Biol Chem* 275:13202–13212
- Kweon O, Kim S-J, Baek S, Chae J-C, Adjei MD, Baek D-H, Kim Y-C, Cerniglia CE (2008) A new classification system for bacterial Rieske non-heme iron aromatic ring-hydroxylating oxygenases. *BMC Biochem* 9:11. <https://doi.org/10.1186/1471-2091-9-11>
- Leveau JHJ, Zehnder AJB (1998) The *tfdK* gene product facilitates uptake of 2,4-dichlorophenoxyacetate by *Ralstonia eutropha* JMP134 (pJP4). 180:2237–2243
- Magony M, Kákonyi I, Gara A, Rapali P, Perei K, Kovács KL, Rákhely G (2007) Overlaps between the various biodegradation pathways in *Sphingomonas subarctica* SA1. *Acta Biol Hung* 58(Suppl):37–49. <https://doi.org/10.1556/ABiol.58.2007.Suppl.4>
- Nichols NN, Harwood CS (1995) Repression of 4-hydroxybenzoate transport and degradation by benzoate: a new layer of regulatory control in the *Pseudomonas putida* β -ketoadipate pathway. *J Bacteriol* 177:7033–7040
- Nichols NN, Harwood CS (1997) PcaK, a high-affinity permease for the aromatic compounds 4-hydroxybenzoate and protocatechuate from *Pseudomonas putida*. *J Bacteriol* 179:5056–5061
- Pao SS, Paulsen IANT, Saier MH (1998) Major facilitator superfamily. *Microbiol Mol Biol Rev* 62:1–34
- Perei K, Rákhely G, Kiss I, Polyák B, Kovács KL (2001) Biodegradation of sulfanilic acid by *Pseudomonas paucimobilis*. *Appl Microbiol Biotechnol* 55:101–107
- Pernstich C, Senior L, Macinnes KA, Forsaith M, Curnow P (2014) Expression, purification and reconstitution of the 4-hydroxybenzoate transporter PcaK from *Acinetobacter* sp. ADP1. *Protein Expr Purif* 101:68–75. <https://doi.org/10.1016/j.pep.2014.05.011>
- Rein U, Gueta R, Denger K, Ruff J, Hollemeyer K, Cook AM (2005) Dissimilation of cysteate via 3-sulfolactate sulfo-lyase and a sulfate exporter in *Paracoccus pantotrophus* NKNCYSA. *Microbiology* 151:737–747. <https://doi.org/10.1099/mic.0.27548-0>
- Robinson MD, McCarthy DJ, Smyth GK (2009) edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140. <https://doi.org/10.1093/bioinformatics/btp616>
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, second edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Singh P, Birkeland NK, Iyengar L, Gurunath R (2006) Mineralization of 4-aminobenzenesulfonate (4-ABS) by *Agrobacterium* sp. strain PNS-1. *Biodegradation* 17:495–502. <https://doi.org/10.1007/s10532-005-9017-7>
- Singh P, Mishra LC, Iyengar L (2004) Biodegradation of 4-aminobenzenesulphonate by a newly isolated bacterial strain PNS-1. *World J Microbiol Biotechnol* 20:845–849. <https://doi.org/10.1007/s11274-004-9010-z>
- Stolz A (2013) Degradative plasmids from sphingomonads. *FEMS Microbiol Lett* 350:1–10. <https://doi.org/10.1111/1574-6968.12283>
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102:15545–15550. <https://doi.org/10.1073/pnas.0506580102>
- Szappanos B, Fritzemeier J, Csörgő B, Lázár V, Lu X, Fekete G, Bálint B, Herczeg R, Nagy I, Notebaart RA, Lercher MJ, Pál C, Papp B (2016) Adaptive evolution of complex innovations through stepwise metabolic niche expansion. *Nat Commun* 7:11607. <https://doi.org/10.1038/ncomms11607>
- Wang YQ, Zhang JS, Zhou JT, Zhang ZP (2009) Biodegradation of 4-aminobenzenesulfonate by a novel *Pannonibacter* sp. W1 isolated from activated sludge. *J Hazard Mater* 169:1163–1167. <https://doi.org/10.1016/j.jhazmat.2009.04.002>
- Weinitschke S, Denger K, Cook AM, Smits THM (2007) The DUF81 protein TauE in *Cupriavidus necator* H16, a sulfite exporter in the metabolism of C2 sulfonates. *Microbiology* 153:3055–3060. <https://doi.org/10.1099/mic.0.2007/009845-0>
- Wells T, Ragauskas AJ (2012) Biotechnological opportunities with the β -ketoadipate pathway. *Trends Biotechnol* 30:627–637. <https://doi.org/10.1016/j.tibtech.2012.09.008>
- Xu Y, Gao X, Wang SH, Liu H, Williams PA, Zhou NY (2012) MhbT is a specific transporter for 3-hydroxybenzoate uptake by Gram-negative bacteria. *Appl Environ Microbiol* 78:6113–6120. <https://doi.org/10.1128/AEM.01511-12>
- Yu G, Wang L-G, Han Y, He Q-Y (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *Omi A J Integr Biol* 16:284–287. <https://doi.org/10.1089/omi.2011.0118>
- Zhang RENH, Jing YW, Cheng ZX (2012) Isolation and characterization of a sulfanilic acid degrading bacterial strain. *Appl Mech Mater* 149:46–49. <https://doi.org/10.4028/www.scientific.net/AMM.148-149.46>