Characterization of Anaerobic Catabolism of 
P-coumarate in Rhodopseudomonas palustris 
by Integrating Transcriptomics and 
Quantitative Proteomics*§

Chongle Pan‡§, Yasuhiro Oda¶, Patricia K. Lankford||, Bing Zhang$**, 
Nagiza F. Samatova§‡‡, Dale A. Pelletier||§§, Caroline S. Harwood¶¶¶, 
and Robert L. Hettich‡||

In this study, the pathway for anaerobic catabolism of 
P-coumarate by a model bacterium, Rhodopseudomonas 
palustris, was characterized by comparing the gene expression 
profiles of cultures grown in the presence of 
P-coumarate, benzoate, or succinate as the sole 
carbon sources. Gene expression was quantified at the mRNA 
level with transcriptomics and at the protein level with 
quantitative proteomics using 15N metabolic labeling. Protein relative abundances, along with their confidence intervals for statistical significance evaluation, were estimated with the software ProRata. Both -omics measurements were used as the transcriptomics provided near-full genome coverage of gene expression profiles and the quantitative proteomics ascertained abundance changes of over 1600 proteins. The integrated gene expression data are consistent with the hypothesis that P-coumarate is converted to benzoyl-CoA, which is then degraded via a known aromatic ring reduction pathway. For the metabolism of P-coumarate to benzoyl-CoA, two alternative routes, a β-oxidation route and a non-β-oxidation route, are possible. The integrated gene expression data provided strong support for the non-β-oxidation route in R. palustris. A putative gene was proposed for every step in the non-β-oxidation route. Molecular & Cellular Proteomics 7:938–948, 2008.

Lignin constitutes almost one-third of all plant dry mass, making it the second most abundant organic compound on earth after cellulose. Biodegradation of lignin during the decay of plant residues in natural environments is a massive biological process within the global carbon cycle (1). Lignin biodegradation is also of great practical significance because of its potential application to biological treatment and the reuse of agricultural wastes. Lignin is a polymer of phenylpropanoid units, and its biodegradation involves depolymerization and subsequent catabolism of the derived aromatic monomers (2). p-Coumarate, or 4-hydroxybenzoinic acid, is one of the main aromatic monomers (3). The biodegradation of p-coumarate can take place not only under oxygen-rich environments but also under anoxic environments as in aquifers, aquatic sediments, and submerged soils.

The purple, non-sulfur phototrophic bacterium Rhodopseudomonas palustris has served as a model organism for studies of anaerobic aromatic compound degradation (4, 5). R. palustris can grow aerobically by respiration or anaerobically using photophosphorylation. Under anaerobic conditions R. palustris generates energy from light and uses organic compounds, including aliphatic and aromatic compounds, as its source of carbon for growth. Structurally diverse single aromatic ring compounds, such as p-coumarate and p-hydroxybenzoate, are proposed to be degraded to a central intermediate, benzoyl-CoA, via various peripheral pathways (6, 7). Then benzoyl-CoA is degraded to acetyl-CoA through a central pathway in the steps of ring reduction, ring cleavage, and β-oxidation (see Fig. 1) (6). Most enzymes in the benzoyl-CoA pathway have been identified and characterized. However, the peripheral pathway used by R. palustris to anaerobically convert p-coumarate to benzoyl-CoA has not been elucidated.

Different metabolic routes have been discovered for catalyzing ferulate, (4-hydroxy-3-methoxycinnamic acid), which differs from P-coumarate by an additional methoxyl group on the aromatic ring. The two major routes are 1) a β-oxidation route found in Pseudomonas acidovorans (8) and Rhodotorula rubra (9) and 2) a non-β-oxidation route found in Pseudomonas fluorescens (10, 11), Deftia acidovorans (12), and Pseudomonas sp. strain HR199 (13). Fig. 1 illustrates the two probable routes for p-coumarate catabolism in R. palustris. Both routes convert the intermediate p-coumaroyl-CoA to 4-hydroxybenzoyl-CoA, yielding an acetyl-CoA. However, the β-oxidation route uses a β-oxidation sequence of reactions for the conversion, whereas the non-β-oxidation route directly
clenches an acetyl-CoA from p-coumaroyl-CoA to generate 4-hydroxybenzaldehyde, which is then oxidized to 4-hydroxybenzoate in a CoA-independent manner. Both routes could easily function under anaerobic conditions. Furthermore the genome sequence of *R. palustris* indicates the genetic potential for both routes. There are many β-oxidation genes in the *R. palustris* genome. Similarly there are several candidate genes for the enoyl-CoA hydratase/lyase and aldehyde dehydrogenase needed for the non-β-oxidation route. In this study, we aimed to determine which of the two routes is likely used for anaerobic p-coumarate catabolism and furthermore to identify the probable genes for every step of this catabolic pathway. To this end, the global gene expression profile of *R. palustris* grown with p-coumarate as the sole organic carbon source was compared with those of *R. palustris* grown with succinate or benzoate (Figs. 1 and 2).

Gene expression activity can be measured at the mRNA level with transcriptomics and at the protein level with proteomics. In many of the past studies, proteomics measurements have largely been qualitative with the aim of detecting the presence of as many proteins as possible from a proteome (14). However, it is not straightforward to correlate the presence of proteins detected by qualitative proteomics with the relative abundances of mRNAs quantified by transcriptomics (15). With the development of quantitative proteomics using stable isotope labeling, the relative protein abundances can now be quantitatively measured for thousands of genes (16). In this study, transcriptomics and quantitative proteomics were integrated for global gene expression profiling. The comparison of the p-coumarate growth condition with the succinate growth condition yielded the relative expression level for 1680 genes at both the mRNA level and the protein level. 1324 genes had no significant expression change for 1680 genes at both the mRNA level and the protein level. 58 genes were up-regulated significantly at both levels. 41 genes were both levels. Interestingly 41 genes were both levels. Interestingly 41 genes were both levels. Interestingly 41 genes were both levels. Interestingly 41 genes were both levels. Interestingly 41 genes were both levels. Interestingly 41 genes were both levels. Interestingly 41 genes were both levels.
Anaerobic Catabolism of p-Coumarate in *R. palustris*

(LC Packings, a division of Dionex, San Francisco, CA). The two-dimensional LC separation was performed with 12 salt pulses, each of which was followed by a 2-h reverse-phase gradient elution. The LC eluent was directly electrosprayed into a linear trapping quadrupole linear ion trap mass spectrometer (ThermoFinnigan, San Jose, CA). Each full scan (400–1700 m/z) was followed by three data-dependent MS/MS scans at 35% normalized collision energy with dynamic exclusion enabled. The full scans were averaged from five microscans, and the MS/MS scans were averaged from two microscans.

Quantitative Proteomics Data Analysis—A protein database from the annotated *R. palustris* genome containing a total of 4836 protein entries (Version 3; www.genome.ornl.gov) was used for protein identification with database searching. No redundant protein entry was present in the protein database. An in-house data preprocessing pipeline based on Xcalibur Development Kit was used to generate DTA files in the same way as the ThermoFinnigan Bioworks software. All DTA files were searched in two iterations against an *R. palustris* protein sequence database (18) using the SEQUEST program Version 27 (26) (enzyme type, trypsin; parent mass tolerance, 3.0; fragment ion tolerance, 0.5; up to four missed cleavages allowed; fully tryptic peptides only). In the first iteration, unlabeled amino acids were used. In the second iteration, 15N-labeled amino acids were used (0.98 Da added to glycine, alanine, serine, proline, valine, threonine, cysteine, leucine, isoleucine, aspartic acid, glutamic acid, methionine, phenylalanine, and tyrosine; 1.96 Da added to asparagine, glutamine, lysine, and tryptophan; 2.94 Da added to histidine; and 3.92 Da added to arginine). The peptide identifications (OUT files) from the two iterations were combined. The DTASelect 1.9 program (27) was used to filter peptide identifications and assemble peptides into proteins using the following parameters: retaining the duplicate MS/MS scans for each peptide sequence (DTASelect option: -t 0); fully tryptic peptides only; a 3Cn threshold of at least 0.08; cross-correlation score (Xcorr) thresholds of at least 1.8 (+1), 2.5 (+2), and 3.5 (+3); and a minimum of two identified peptides for a protein. These widely accepted SEQUEST-DTASelect filtering criteria have been shown to provide a high identification confidence with a maximum false-positive rate of 1–2% (27–29). Identification results from the DTASelect program are provided in supplemental Table 1. Selected ion chromatogram extraction, peptide abundance ratio estimation, and protein abundance ratio estimation were completed with the program ProRata 1.1 using default parameters as described previously (30, 31). Briefly for each biological replicate of a direct comparison, identified peptides were quantified and filtered with a minimum profile signal-to-noise ratio cutoff of 2. Then proteins consisting of at least two quantified peptides were evaluated for quantification. 90% confidence intervals were calculated for every protein as an error bar of their abundance ratio estimate. Two biological replicates of a direct comparison were combined, and proteins were further filtered with a maximum confidence interval width cutoff of 3 (supplemental Tables 2 and 3). The two direct comparisons were combined for the succinate-versus-benzoate indirect comparison, and proteins were also filtered with a maximum confidence interval width cutoff of 3 (supplemental Table 4). The protein abundance of a gene was considered significantly up- or down-regulated if the log ratio (log2 abundance ratio) of the protein is greater than 1 (up-regulated) or less than −1 (down-regulated) and the 90% log ratio confidence interval excludes zero. The MS raw data and the data analysis configuration files and results are available upon request.

RESULTS AND DISCUSSION

Experimental Design and Results Overview—Both transcriptomics and quantitative proteomics measure gene expression changes between a reference condition and a treatment condition. The treatment condition in this study was anaerobic photosynthetic cell growth with p-coumarate as the sole organic carbon source. To identify the genes activated for p-coumarate catabolism, two reference conditions were selected in which succinate or benzoate replaced p-coumarate as the sole organic carbon source. As succinate is a simple dicarboxylic acid, the succinate growth condition provides a base-line gene expression profile of *R. palustris* without aromatic degradation activity; thus, comparison of the p-coumarate condition with the succinate condition should help elucidate the entire pathway for p-coumarate catabolism (Fig. 1). The comparison between the benzoate condition and

---

Fig. 1. Proposed *R. palustris* p-coumarate catabolic pathway. The degradation of p-coumarate is proposed to proceed through either a β-oxidation route or a non-β-oxidation route. Steps in the two hypothetical routes are shown with dashed arrows. Both routes reach the common intermediate 4-hydroxybenzoyl-CoA, which is degraded through the known benzyol-CoA pathway outlined with solid arrows. The degradation pathways of benzoate and succinate are also shown for comparison.
The experimental scheme is shown in Fig. 2. At the cell culture step, *R. palustris* was grown with $^{15}$N metabolic labeling for the two reference conditions. Two biological replicates were prepared for each condition and analyzed independently to capture biological variability. Each biological replicate was divided for microarray and quantitative proteomics analysis. We found that the correlation between transcriptomics results and quantitative proteomics results was improved by dividing the same sample for the two measurements as compared with using separate samples for the two measurements (data not shown).

Microarray analysis was performed with a custom-designed GeneChip, which contained probes for all 4836 predicted genes and 3190 intergenic regions in the *R. palustris* genome. Each biological replicate was analyzed with a GeneChip microarray. The reproducibility of raw signal intensities between two biological replicates (measured in Pearson correlation coefficient) was above 0.97 for all three growth conditions. The duplicate results for each comparison were combined with the Cyber-T program (Fig. 2). Note that RNA molecules from the succinate and benzoate conditions contained $^{15}$N-enriched nitrogen as a result of $^{15}$N metabolic labeling.

The unlabeled $p$-coumarate culture was mixed with the $^{15}$N-labeled succinate or benzoate cultures for quantitative proteomics analysis (Fig. 2). Direct comparisons between the $p$-coumarate condition and the two reference conditions were obtained with biological replication at the cell growth stage and technical replication at the LC-MS/MS measurement stage. The ProRata program was used to merge technical replicates for each biological replicate, then to combine the two biological replicates for each direct comparison, and finally to integrate the two direct comparisons for the indirect comparison. The numbers of identified proteins and quantified proteins, before and after combination, are shown in Table 1. For the combination of biological replicates, proteins identified in only one replicate were filtered out, and proteins quantified with inconsistent log ratio confidence intervals between the two replicates were also filtered out. The quantification reproducibility between the two biological replicates of a direct comparison was measured with the Pearson correlation coefficients ($r$) between the replicate protein abundance ratios in log$_2$ scale (protein log ratio) (Fig. 3). The level of reproducibility was comparable between transcriptomics and quantitative proteomics.

**Integration of mRNA Abundance Profiles and Protein Abundance Profiles**—The transcriptomics results and the quantitative proteomics results were cross-matched by gene locus (supplemental Tables 2–4). Due to the incomplete genome coverage by proteomics, "not available" (N/A) values were assigned to undetermined protein log ratios. Correlations between the mRNA log ratios and the protein log ratios of cross-matched genes are shown in the Fig. 4, left column; all have positive Pearson correlation coefficients. The majority of genes were distributed around the center within the square area of mRNA log ratio range ($-1, 1$) and protein log ratio range ($-1, 1$). This shows that there is little correlation between minor mRNA abundance changes and minor protein abundance changes. The relatively lower correlation coefficient of the $p$-coumarate-versus-benzoate comparison is a result of the lower number of genes with large expression changes. To quantify the discrepancy between mRNA log ratio and protein log ratio of individual genes, the differences between the mRNA log ratio and protein log ratio ($\Delta$log ratio) of every gene were calculated. Histograms of the log ratio differences are shown for the three comparisons in Fig. 4, right column.

The log ratios measured by transcriptomics and quantitative proteomics were statistically assessed to categorize genes by their regulation directions, namely, up-, down-, or null-regulation at the mRNA level or the protein level. Fig. 5 shows the number of genes in each category. The categories were color-coded to illustrate the complementarity between transcriptomics and proteomics. The categories in *yellow* highlight the advantage of transcriptomics, near full coverage of the *R. palustris* genome, as opposed to ~30% genome coverage by proteomics. Membrane proteins and low abundance proteins are particularly difficult for proteomics to measure. The categories in *red* contain genes with concordant results from transcriptomics and proteomics. The agreement between the two independent measurements enhances...
the confidence in the expression change of these genes and alleviates the need for validating the expression change of these genes individually with other biochemical experiments, such as RT-PCR and Western blotting. The categories in green contain genes with inconsistent regulation directions between the mRNA level and the protein level. Only a few genes were up-regulated at the mRNA level and down-regulated at the protein level or vice versa. However, many genes showed significant abundance change at one level but insignificant abundance change at the other level. This indicates that one must be cautious in concluding that the presence or absence of a significant mRNA abundance change of a gene detected by transcriptomics necessarily corresponds to the presence or absence of a significant protein abundance change detected by proteomics. This is especially true when considering relatively small changes in mRNA abundance. The discrepancy between mRNA log ratio and protein log ratio of a gene can stem from measurement errors, sustained protein presence from transient transcriptional induction, post-transcriptional regulation, or any combinations of these causes (32).

Measurement of Up-regulated Expression of 4-Hydroxybenzoyl-CoA Degradation Enzymes during p-Coumarate Catabolism—Previous studies have suggested that p-coumarate is converted to 4-hydroxybenzoyl-CoA, which is then degraded through the benzoyl-CoA pathway (33). The known steps for the 4-hydroxybenzoyl-CoA degradation are shown in Fig. 1 with solid arrows. Expression of the genes involved in each step was compared among the three R. palustris growth conditions: p-coumarate, benzoate, and succinate (Table II). Both mRNA abundances and protein abundances of all these genes were greatly up-regulated in the p-coumarate condition compared with the succinate condition, showing the induction of 4-hydroxybenzoyl-CoA degradation activity during p-coumarate catabolism. This supports the part of the proposed p-coumarate pathway downstream of 4-hydroxybenzoyl-CoA as shown in Fig. 1.

The comparison between the p-coumarate condition and the benzoate condition is also interesting. 4-Hydroxybenzoyl-CoA reductase (RPA0670–0672) for dehydroxylation of 4-hydroxybenzoyl-CoA to benzoyl-CoA was more abundant in the p-coumarate condition than in the benzoate condition; this agrees with published data showing that this enzyme is not required for benzoate degradation (34) (Fig. 1).
tion of both \( p \)-coumarate and benzoate is known to proceed through the benzoyl-CoA pathway. The comparison shows that this pathway was insignificantly or moderately up-regulated according to transcriptomics, whereas it was significantly down-regulated according to quantitative proteomics (Table II). Because the discrepancy between mRNA abundance change and protein abundance change was found for nine genes across multiple operons, it is probably of biological significance.

Genes encoding the two forms of ribulose-bisphosphate carboxylase (RPA1559, RPA1560, and RPA4641), along with their associated genes for carbon dioxide fixation (RPA1561 and RPA4642–4645), were induced in both \( p \)-coumarate and benzoate conditions (supplemental Tables 1 and 3). 6 reducing eq of \([\text{H}]\) are generated when a molecule of benzoyl-CoA is degraded into acetyl-CoA. Carbon compounds like \( p \)-coumarate and benzoate, which are electron-rich relative to cell material, cannot be fully assimilated into biomass unless an external electron acceptor like carbon dioxide is available. Thus, it makes sense that the main carbon dioxide-assimilat-

---

**Fig. 4. Comparison of mRNA log ratios and protein log ratios.** The protein log ratios and mRNA log ratios of quantified genes are shown as scatter plots for the three comparisons. The histograms of the differences between the protein log ratio and the mRNA log ratio (\( \Delta \log \text{ratio} \)) were also constructed for the three comparisons.
identified enzymes of the Calvin cycle were expressed at higher abundances to serve as a reducing equivalent sink during growth on aromatic compounds.

The global gene expression profiling also revealed other induced cellular activities that are indirectly related to aromatic degradation in _R. palustris_. Several methyl-accepting chemotaxis proteins (RPA0139, RPA0142, RPA1678, RPA3185, RPA4302, and RPA4639) were induced in both aromatic degradation conditions (supplemental Tables 1 and 3). _R. palustris_ is a motile bacterium, and these proteins could serve as chemoreceptors that enable cells to sense and swim toward plant-derived aromatic compounds (35). A multitude of different types of membrane transporters were also induced in both aromatic degradation conditions (Table III and supplemental Tables 1 and 3). Some of them may facilitate the uptake of aromatic compounds. In addition, _R. palustris_ often transiently excretes partially degraded aromatic intermediates into the growth media (36). Active transport may be required for the reuptake of these compounds into cells.

**Fig. 5.** Summary of integrated transcriptomics and quantitative proteomics results. Genes were categorized into up-, down-, and null-regulated at the protein level and the mRNA level. Genes not quantified by proteomics have an “N/A” protein abundance change. The 12 categories were color-coded to show the complementarity between transcriptomics and proteomics.

The non–β-oxidation route would consist of 1) an enoyl-CoA hydratase/lyase (active for both hydration of the enoyl-CoA C–C double bond and ligation of the resultant C–C single bond to release an acetyl-CoA (11)), 2) an aldehyde dehydrogenase, and 3) the known 4-hydroxybenzoate-CoA ligase (RPA0669). Whichever route is used, the genes encoding the enzymes in that route were expected to be up-regulated in the p-coumarate condition in comparison with both the benzoate condition and the succinate condition. Table III lists all the genes measured by quantitative proteomics to be significantly up-regulated at the protein level in the p-coumarate condition in comparison with the other two reference conditions.

Based on the _R. palustris_ genome sequence, the two most likely gene sets for the β-oxidation route, if this route is used, are RPA1703–1706 and RPA0674–0676. Both gene sets could encode all β-oxidation enzymes. More importantly, the gene set RPA1703–1706 lies immediately next to a possible p-coumarate-CoA ligase, RPA1707. It seems reasonable to arrange these genes into an operon. The gene set RPA0674–0676 is in a region of the _R. palustris_ genome where many known aromatic degradation genes are located (see the loci of the aromatic degradation genes listed in Table II). However, neither gene set was found to be significantly up-regulated in the p-coumarate condition (supplemental Tables 1 and 2).

Furthermore out of the up-regulated genes in the p-coumarate condition, we failed to find a β-oxidation gene set that is likely to be used for p-coumarate degradation.

On the other hand, every enzyme in the non-β-oxidation route has a candidate gene identified from Table III with enhanced protein abundance in the p-coumarate condition (Fig. 6). Due to the similarity between p-coumarate and ferulate, their CoA ligases have a high sequence similarity, and three loci are annotated as ferulate- or p-coumarate-CoA ligases in the _R. palustris_ genome, including RPA1707, RPA1787, and RPA4421. Among the three loci, RPA1787 is the only one up-regulated in the p-coumarate condition, making it the probable gene responsible for the p-coumarate-CoA ligation.

Only 9 base pairs upstream of RPA1787 in the _R. palustris_ genome is similarly up-regulated RPA1786. The short intergenic distance between the two loci suggests their co-transcription and close functional link. RPA1786 belongs to a group of enzymes, the enoyl-CoA hydratases/isomerasers (pfam00378), whose members have diverse activities. Thus, it is difficult to know with certainty what the specific function of RPA1786 might be based on its deduced amino acid sequence. We suggest that the enoyl-CoA hydratase/lyase in the non-β-oxidation route could be encoded by RPA1786 (Fig. 6).

It is known that _R. palustris_ readily oxidizes 4-hydroxybenzaldehyde to 4-hydroxybenzoate (33). RPA1206 is the only annotated aldehyde dehydrogenase with significantly up-regulated protein abundance in the p-coumarate condition. In-
| Metabolic step* | Locus          | mRNA p value | Protein p value | mRNA CI | Protein CI | Gene annotation                                      |
|---------------|----------------|--------------|-----------------|--------|------------|-----------------------------------------------------|
| Dehydroxylation | RPA0670       | 6.3          | 4 e-08          | [2.6,3.7] | [0.7,1.3] | 4-Hydroxybenzoyl-CoA reductase, first subunits      |
|               | RPA0671       | 5.6          | 2 e-10          | [3.4,4.2] | [1.3,1.7] | 4-Hydroxybenzoyl-CoA reductase, second subunits     |
|               | RPA0672       | 4.5          | 2 e-08          | [4.3,5.0] | [2.2,3.4] | 4-Hydroxybenzoyl-CoA reductase, third subunits      |
| CoA ligation  | RPA0661       | 3.0          | 5 e-08          | [1.9,2.4] | [-0.8,-0.5] | Benzoate-CoA ligase                                 |
| Ring reduction| RPA0657       | 3.2          | 6 e-10          | [4.0,4.7] | [-1.3,-1.1] | Benzoate-CoA reductase subunit badD                |
|               | RPA0658       | 4.2          | 1 e-12          | [3.6,4.2] | [-1.5,-1.1] | Benzoate-CoA reductase subunit badE                 |
|               | RPA0659       | 4.8          | 5 e-13          | [4.2,4.6] | [-1.9,-1.3] | Benzoate-CoA reductase subunit badF                 |
|               | RPA0660       | 6.4          | 9 e-09          | [4.2,4.7] | [-1.8,-1.5] | Benzoate-CoA reductase subunit badG                 |
|               | RPA0662       | 3.4          | 3 e-08          | [2.8,3.6] | [-1.2,-0.5] | Ferredoxin                                          |
| Ring cleavage | RPA0650       | 3.9          | 2 e-07          | N/A     | [-1.5,-1.2] | Cyclohex-1-ene-1-carboxyl-CoA hydratase            |
|               | RPA0653       | 4.8          | 5 e-10          | [3.7,4.0] | [-1.3,-1.1] | 2-Ketocyclohexanecarboxyl-CoA hydrolase            |
|               | RPA0654       | 4.6          | 1 e-08          | [3.5,4.0] | [-1.5,-1.3] | 2-Hydroxycyclohexanecarboxyl-CoA dehydrogenase     |
| β-Oxidation   | RPA3713       | 1.4          | 2 e-04          | [1.5,1.8] | [0.0,0.2] | Pimeloy-CoA dehydrogenase small subunit            |
|               | RPA3714       | 1.2          | 7 e-04          | [1.6,1.9] | [0.0,0.2] | Pimeloy-CoA dehydrogenase large subunit            |
|               | RPA3715       | 1.3          | 2 e-04          | [1.4,1.7] | [-0.2,0.1] | Acetyl-CoA acetyltransferase                        |
|               | RPA3717       | 1.8          | 5 e-05          | [1.3,1.6] | [-0.1,0.2] | Enoyl-CoA hydratase                                |

* See Fig. 1 for reactants and products of each metabolic step.

b Log ratio: log2 abundance ratio.

CI, confidence interval.
### Table III

| Locus | mRNA | Protein | mRNA | Protein | mRNA | Protein | Gene annotation |
|-------|------|---------|------|---------|------|---------|-----------------|
|       | Log ratio | p value | Log ratio | CI | Log ratio | p value | Log ratio | CI | |
| RPA0665 | 3.0 | 2e-06 | 2.9 | [2.7, 3.2] | 3.1 | 2e-06 | 1.6 | [1.3, 1.9] | -0.1 | 7e-01 | 1.4 | [1.0, 1.7] | Putative ABC\(^c\) transporter subunit, ATP-binding component |
| RPA0668 | 2.4 | 2e-07 | 2.6 | [2.4, 2.7] | 2.0 | 9e-07 | 1.8 | [1.7, 1.9] | 0.5 | 5e-02 | 0.8 | [0.6, 0.9] | Putative ABC transporter subunit, substrate-binding component |
| RPA0669 | 5.6 | 8e-10 | 4.1 | [3.8, 4.4] | 2.2 | 6e-08 | 2.0 | [1.9, 2.1] | 3.5 | 3e-09 | 2.1 | [1.8, 2.4] | 4-Hydroxybenzoyl-CoA ligase |
| RPA0670 | 6.3 | 4e-08 | 3.1 | [2.6, 3.7] | 2.5 | 7e-07 | 1.0 | [0.7, 1.3] | 3.9 | 5e-09 | 2.2 | [1.6, 2.7] | 4-Hydroxybenzoyl-CoA reductase, first subunits |
| RPA0671 | 5.6 | 2e-10 | 3.8 | [3.4, 4.2] | 2.5 | 4e-09 | 1.5 | [1.3, 1.7] | 3.1 | 4e-09 | 2.3 | [1.9, 2.7] | 4-Hydroxybenzoyl-CoA reductase, second subunits |
| RPA0672 | 4.5 | 2e-08 | 4.6 | [4.3, 5.0] | 2.6 | 1e-07 | 2.7 | [2.2, 3.4] | 1.9 | 5e-09 | 1.9 | [1.2, 2.6] | 4-Hydroxybenzoyl-CoA reductase, third subunits |
| RPA1009 | 3.3 | 3e-11 | 3.2 | [3.1, 3.4] | 2.0 | 9e-10 | 3.3 | [3.2, 3.4] | 1.2 | 8e-05 | -0.1 | [-0.2, 0.1] | Possible cytochrome P450 |
| RPA1206 | -0.5 | 3e-01 | 1.4 | [1.2, 1.5] | -1.0 | 5e-02 | 1.7 | [1.5, 1.8] | 0.5 | 3e-01 | -0.3 | [-0.5, -0.1] | Aldehyde dehydrogenase |
| RPA1414 | -1.0 | 1e-04 | 1.1 | [0.9, 1.4] | -0.7 | 2e-03 | 1.2 | [0.9, 1.4] | -0.3 | 2e-01 | 0.0 | [-0.3, 0.3] | Maco-like dehydratase |
| RPA1782 | 1.2 | 4e-04 | 1.7 | [1.5, 1.9] | 0.3 | 2e-01 | 1.4 | [1.2, 1.6] | 0.9 | 6e-05 | 0.3 | [0.0, 0.5] | C\(_5\)-dicarboxylate periplasmic binding protein, dctP subunit, |
| RPA1786 | 4.3 | 7e-06 | 4.1 | [3.7, 4.4] | 1.5 | 6e-04 | 2.2 | [2.1, 2.4] | 2.8 | 1e-08 | 1.8 | [1.5, 2.2] | Putative 3-hydroxybutyl-CoA dehydratase |
| RPA1787 | 4.1 | 2e-05 | 4.3 | [4.1, 4.5] | 1.7 | 6e-04 | 2.5 | [2.4, 2.6] | 2.4 | 7e-07 | 1.8 | [1.6, 2.0] | Putative feruloyl-CoA synthetase |
| RPA1788 | 2.4 | 1e-03 | 4.1 | [3.8, 4.4] | 2.7 | 9e-04 | 1.7 | [1.4, 2.0] | -0.3 | 2e-01 | 2.4 | [2.0, 2.8] | Possible 4-hydroxybenzoyl-CoA thioesterase |
| RPA1799 | 0.4 | 2e-01 | 1.8 | [1.7, 1.9] | 2.5 | 9e-05 | 2.1 | [2.0, 2.2] | -2.1 | 8e-06 | -0.3 | [-0.4, -0.2] | Putative branched-chain AA\(^e\) transporter, substrate-binding protein |
| RPA1791 | 1.9 | 2e-04 | 2.1 | [1.8, 2.3] | 2.1 | 2e-04 | 1.4 | [1.1, 1.6] | 0.1 | 5e-01 | 0.7 | [0.4, 1.0] | Putative branched-chain AA transporter, ATP-binding protein |
| RPA1792 | 2.2 | 1e-04 | 1.9 | [1.7, 2.2] | 2.5 | 7e-05 | 1.4 | [1.1, 1.6] | -0.3 | 8e-02 | 0.6 | [0.3, 0.9] | Putative branched-chain AA transporter, ATP-binding protein |
| RPA3011 | 2.5 | 2e-08 | 2.9 | [2.6, 3.3] | 2.6 | 1e-08 | 4.6 | [4.3, 4.9] | -0.1 | 7e-01 | -1.6 | [-2.1, -1.1] | Unknown protein |
| RPA3014 | 0.8 | 5e-03 | 2.1 | [1.6, 2.5] | 4.5 | 7e-09 | 2.2 | [1.9, 2.6] | -3.8 | 4e-06 | -0.2 | [-0.7, 0.4] | Two-component transcriptional regulator, LuxR family |
| RPA3101 | -0.1 | 9e-01 | 1.5 | [1.4, 1.7] | 0.2 | 5e-01 | 1.5 | [1.4, 1.6] | -0.3 | 3e-01 | 0.0 | [-0.1, 0.2] | Unknown protein |
| RPA3423 | 0.8 | 5e-04 | 1.6 | [1.3, 1.9] | 0.4 | 2e-02 | 1.6 | [1.3, 2.0] | 0.4 | 9e-02 | 0.0 | [-0.4, 0.4] | Unknown protein |
| RPA3893 | 4.0 | 4e-07 | 2.3 | [2.0, 2.6] | 3.6 | 6e-07 | 1.4 | [1.1, 1.8] | 0.4 | 6e-01 | 0.9 | [0.4, 1.3] | Putative carboxyesterase |
| RPA4092 | 2.8 | 5e-08 | 1.8 | [1.7, 2.0] | 2.7 | 9e-08 | 1.8 | [1.6, 2.2] | 0.1 | 8e-01 | 0.0 | [-0.4, 0.3] | Unknown protein |
| RPA4096 | 2.3 | 7e-06 | 2.6 | [2.0, 4.1] | 2.0 | 2e-05 | 3.1 | [2.7, 3.6] | 0.3 | 5e-01 | 0.5 | [-1.3, 1.0] | Possible multidrug efflux membrane fusion protein mexE |
| RPA4198 | 2.3 | 5e-05 | 3.8 | [3.5, 4.0] | 0.9 | 1e-02 | 1.2 | [1.1, 1.3] | 1.4 | 3e-06 | 2.6 | [2.3, 2.8] | Amidohydrolase 2 |

\(^a\) Log ratio: log\(_2\) abundance ratio.  
\(^b\) CI, confidence interval.  
\(^c\) ABC, ATP-binding cassette.  
\(^d\) AA, amino acid.
A probable locus was assigned to every enzyme in the non-\(\beta\)-oxidation route based on gene expression data and genome annotation.

Interestingly this locus showed insignificantly down-regulated mRNA abundance in the \(p\)-coumarate condition. It is unclear whether \(RPA1206\) has the substrate specificity for 4-hydroxybenzaldehyde. But given the expression data, this locus is a possible aldehyde dehydrogenase that could be used in the non-\(\beta\)-oxidation route (Fig. 6).

4-Hydroxybenzoate-CoA ligase is encoded by \(RPA0669\) (37) (Fig. 6). The substantial up-regulation of this enzyme in the \(p\)-coumarate condition suggests that a large flux of 4-hydroxybenzoate exists during \(p\)-coumarate catabolism. We have observed previously the accumulation of 4-hydroxybenzoate in culture media during the growth of \(R.\ palustris\) on \(p\)-coumarate (38). It is possible for the \(\beta\)-oxidation route to generate a flux of 4-hydroxybenzoate by hydrolyzing 4-hydroxybenzoyl-CoA and then religating 4-hydroxybenzoate with free CoA. But this is less plausible than the explanation with the non-\(\beta\)-oxidation route in which 4-hydroxybenzoate is an intermediate metabolite.

Taken in total, the gene expression data support that the anaerobic catabolism of \(p\)-coumarate in \(R.\ palustris\) proceeds through the non-\(\beta\)-oxidation route and then through the central benzoyl-CoA pathway (Fig. 1). Furthermore putative loci were identified for every enzyme in the non-\(\beta\)-oxidation route (Fig. 6). Formulation of such a testable pathway in its entirety demonstrates the value of the integrated transcriptomics and quantitative proteomics approach.

CONCLUSIONS

Quantitative proteomics measurements were conducted to provide a global view of the cellular pathways involved in \(p\)-coumarate catabolism for \(R.\ palustris\) by using metabolic stable isotope labeling, MudPIT analysis, and the ProRata data analysis program. By characterizing technical and biological replicates, over 1600 proteins were confidently quantified from \(R.\ palustris\) cultures grown with succinate, benzoyl, or \(p\)-coumarate as the sole carbon sources. The same \(R.\ palustris\) cultures were also examined by transcriptomics. The transcriptomics results and the quantitative proteomics results were integrated to reconstruct global gene expression profiles of \(R.\ palustris\) at the mRNA level and the protein level. This integrated gene expression data set provided evidence that the anaerobic degradation of \(p\)-coumarate proceeds through a non-\(\beta\)-oxidation route, rather than an alternative \(\beta\)-oxidation route, and then through the central benzoyl-CoA pathway. \(RPA1787, RPA1786,\) and \(RPA1206\) were hypothesized to be the probable \(p\)-coumarate-CoA ligase, \(p\)-coumaroyl-CoA hydratase/lyase, and 4-hydroxybenzaldehyde dehydrogenase, respectively, in the non-\(\beta\)-oxidation route.

Acknowledgments—Guruprasad Kora is acknowledged for assisting with Website construction for the proteomics data release. Oak Ridge National Laboratory is managed by the University of Tennessee-Battelle, L.L.C. for the Department of Energy under Contract DOE-AC05–00OR22725.

\* The experimental transcriptomics and proteomics work was supported by United States Department of Energy (Office of Biological and Environmental Research, Office of Science) Grant DE-FG02–05ER15707 and DE-FG02–01ER63241, respectively. The data analysis and integration work was supported by the “Exploratory Data Intensive Computing for Complex Biological Systems” project from United States Department of Energy (Office of Advanced Scientific Computing Research, Office of Science) and in part by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

†† The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

The microarray data are accessible at the NCBI Gene Expression Omnibus (GEO) database under accession number GSE6221.

** Present address: Dept. of Biomedical Informatics, Vanderbilt University, Nashville, Tennessee 37232.

§§ To whom correspondence may be addressed for questions on aromatic catabolism: Oak Ridge National Laboratory, P. O. Box 2008, Oak Ridge, TN 37831. Tel.: 865-576-2857; Fax: 865-574-6210; E-mail: pelletierda@ornl.gov.

To whom correspondence may be addressed for questions on aromatic catabolism: Dept. of Microbiology, University of Washington, Box 357242, 1959 N. E. Pacific St., Seattle, WA 98195. Tel.: 206-221-2848; Fax: 206-221-5041; E-mail: csh5@u.washington.edu.

To whom correspondence should be addressed for questions on proteomics: Oak Ridge National Laboratory, P. O. Box 2008, Oak Ridge, TN 37831. Tel.: 865-576-4968; Fax: 865-576-8559; E-mail: hettichrl@ornl.gov.
Anaerobic Catabolism of p-Coumarate in R. palustris

REFERENCES

1. Kirk, T. K. (1984) Degradation of lignin, in Microbial Degradation of Organic Compounds (Gibson, D. T., ed) pp. 399–437, Marcel Dekker, Inc., New York.

2. Sarkanen, K. V., and Ludwig, C. H. (1971) Definition and nomenclature, in Lignins: Occurrence, Formation, Structure and Reactions, pp. 1–18, John Wiley and Sons, New York.

3. Hartley, R. D., and Ford, C. W. (1989) Phenolic constituents of plant cell walls and wall biodegradability, in Plant Cell Wall Polymers: Biogenesis and Biodegradation (Lewis, N. G., and Paice, M. G., eds) pp. 137–145, American Chemical Society, Washington, D. C.

4. Diaz, E. (2004) Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility. Int. Microbiol., 7, 173–180

5. Egland, P. G., Pelletier, D. A., Dispensa, M., Gibson, J., and Harwood, C. S. (1997) A cluster of bacterial genes for anaerobic benzene ring biodegradation. Proc. Natl. Acad. Sci. U. S. A., 94, 6484–6489

6. Harwood, C. S., Burchhardt, G., Herrmann, H., and Fuchs, G. (1999) Anaerobic metabolism of aromatic compounds via the benzoyl-CoA pathway. FEMS Microbiol. Rev., 22, 439–458

7. Gibson, J., and Harwood, C. S. (2002) Metabolic diversity in aromatic compound utilization by anaerobic microbes. Annu. Rev. Microbiol., 56, 345–369

8. Tove, A., and Wood, J. M. (1970) Degradation of trans-ferulic acid by Pseudomonas aeruginosa. Biochemistry, 9, 337–343

9. Huang, Z. X., Dostal, L., and Rosazza, J. P. N. (1993) Mechanisms of ferulic acid conversions to vanillic acid and guaiacol by Rhodotorula rubra. J. Biol. Chem., 268, 23954–23958

10. Harwood, C. S. (2005) Functional genomic analysis of three nitrogenase isozymes in the photosynthetic bacterium Rhodopseudomonas palustris. Nat. Biotechnol., 23, 7784–7794

11. Gibson, J., Dispensa, M., Fogg, G. C., Evans, D. T., and Harwood, C. S. (1997) 4-Hydroxybenzoyl coenzyme A reductase (dehydroxylating) is required for anaerobic degradation of diverse aromatic-compounds by the photosynthetic bacterium Rhodopseudomonas palustris. Appl. Environ. Microbiol., 63, 252–262

12. Overhage, J., Priefert, H., and Steinbüchel, A. (1999) Biochemical and genetic analyses of ferulic acid catabolism in Pseudomonas sp strain HR199. Appl. Environ. Microbiol., 65, 4837–4847

13. Washburn, M. P., and Yates, J. R., III (2000) Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single-dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT. Int. J. Mass Spectrom., 219, 245–251

14. L. Harrison, F. H. (2005) Peripheral Pathways of Anaerobic Benzoate Degradation in Rhodopseudomonas palustris. Ph.D. thesis, University of Iowa