Functional characterization of a gene cluster involved in gentisate catabolism in *Rhodococcus* sp. strain NCIMB 12038

Ting-Ting Liu · Ying Xu · Hong Liu · Sha Luo · Ya-Jie Yin · Shuang-Jiang Liu · Ning-Yi Zhou

Received: 10 September 2010 / Revised: 22 November 2010 / Accepted: 23 November 2010 / Published online: 23 December 2010 © Springer-Verlag 2010

**Abstract** *Rhodococcus* sp. strain NCIMB 12038 utilizes naphthalene as a sole source of carbon and energy, and degrades naphthalene via salicylate and gentisate. To identify the genes involved in this pathway, we cloned and sequenced a 12-kb DNA fragment containing a gentisate catabolic gene cluster. Among the 13 complete open reading frames deduced from this fragment, three (narIKL) have been shown to encode the enzymes involved in the reactions of gentisate catabolism. NarI is gentisate 1,2-dioxygenase which converts gentisate to maleylpyruvate, NarL is a mycothiol-dependent maleylpyruvate isomerase which catalyzes the isomerization of maleylpyruvate to fumarylpyruvate, and NarK is a fumarylpyruvate hydrolase which hydrolyzes fumarylpyruvate to fumarate and pyruvate. The narX gene, which is divergently transcribed with narIKL, has been shown to encode a functional 3-hydroxybenzoate 6-monooxygenase. This led us to discover that this strain is also capable of utilizing 3-hydroxybenzoate as its sole source of carbon and energy. Both NarL and NarX were purified to homogeneity as His-tagged proteins, and they were determined by gel filtration to exist as a trimer and a monomer, respectively. Our study suggested that the gentisate degradation pathway was shared by both naphthalene and 3-hydroxybenzoate catabolism in this strain.

**Keywords** Catabolism · Gene cluster · Gentisate · *Rhodococcus* sp. strain NCIMB 12038

**Introduction**

Naphthalene is one of the common bicyclic aromatics and is often released into the environment as coal tar and coal tar products, such as creosote (Mueller et al. 1989). Two pathways have been well characterized for the microbial degradation of naphthalene, and salicylate is an intermediate in both cases. In the classical pathway, which was studied extensively in Gram-negative *Pseudomonas*, salicylate undergoes oxidative decarboxylation by salicylate 1-hydroxylase to produce catechol as the ring-cleavage substrate (Dunn and Gunsalus 1973; Yen and Serdar 1988). The alternative pathway was discovered in *Ralstonia* sp. strain U2 a decade ago and involves a salicylate 5-hydroxylase which yields the ring-cleavage substrate gentisate (Fuenmayor et al. 1998), which is further metabolized via a GSH-dependent maleylpyruvate isomerization pathway (Zhou et al. 2001).

The Gram-positive genus *Rhodococcus* possesses extreme capability to degrade a wide range of xenobiotics (Larkin et al. 2005), including the naphthalene utilization strains NCIMB 12038 (Allen et al. 1997; Larkin and Day 1986) and B4 (Grund et al. 1992). Naphthalene degradation in both strains was found to occur via the gentisate pathway (Allen et al. 1997; Grund et al. 1992). In strain NCIMB 12038, an 11,548-bp DNA fragment involved in naphthalene degradation was sequenced, and the narAaAb-encoded naphthalene dioxygenase and the narB-encoded cis-naphthalene dihydrodiol dehydrogenase were functionally characterized (Kulakov et al. 2000; Larkin et al. 1999). However, no genetic determinates for its gentisate degradation
have been identified to date. More recently, a putative gene cluster encoding a gentisate pathway in the naphthalene utilizer *Rhodococcus opacus* R7 has been proposed, but the biochemical functions of these genes have not been experimentally confirmed (Di Gennaro et al. 2010).

Here, we report the genetic and biochemical characterization of the gentisate catabolic pathway in *Rhodococcus* sp. strain NCIMB 12038, revealing the involvement of a mycothiol-dependent maleylpyruvate isomerase.

### Materials and methods

#### Bacterial strains, plasmids, primers, culture conditions, and chemicals

The bacterial strains, plasmids, and polymerase chain reaction (PCR) primers used in this study are listed in Table 1. Strain NCIMB 12038 was grown aerobically in lysogeny broth (LB) or M9 minimum media (Miller 1972) supplemented with naphthalene (0.5%, w/v), salicylate (2 mM) or 3-hydroxybenzoate (2.5 mM), or pyruvate (10 mM), as the sole carbon source at 30°C on a rotary shaker at 150 rpm. *Escherichia coli* strains were grown in LB at 37°C with 100 μg/ml ampicillin, 50 μg/ml kanamycin, or 34 μg/ml chloramphenicol, as necessary. Mycothiol (MSH) was prepared according to the procedures described previously (Feng et al. 2006).

#### Cloning of gentisate degradation genes and sequence analyses

Primers were designed based on a conserved region of the gentisate 1,2-dioxygenase (GDO) genes from *R. opacus* (accession no. BAD35142) and *Streptomyces* sp. strain WA46 (accession no. BAC78375) to amplify the putative gene encoding GDO from strain NCIMB 12038. A 599-bp fragment was PCR amplified from the genomic DNA of strain NCIMB 12038 and ligated into pGEM-T vector (Promega) for sequencing. The genome walking strategy (Siebert et al. 1995) was then employed to clone the flanking regions of the above fragment. Nucleotide sequences were determined by Invitrogen Technologies Co. (Shanghai,

### Table 1  
**Bacterial strains, plasmids, and primers used in this study**

<table>
<thead>
<tr>
<th>Strains and plasmids</th>
<th>Relevant genotype or characteristic(s)</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Rhodococcus</em> sp. strain NCIMB 12038</td>
<td>Naphthalene utilizer, wild type</td>
<td>NCIMB</td>
</tr>
<tr>
<td><em>Corynebacterium glutamicum</em> ATCC 13032</td>
<td>Gentisate utilizer, wild type</td>
<td>ATCC</td>
</tr>
<tr>
<td><strong>E. coli</strong> DH5α</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosetta (DE3) pLysS</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGEM-T</td>
<td>Ap′, cloning vector</td>
<td>Promega</td>
</tr>
<tr>
<td>pET5a</td>
<td>Ap′, expression vector</td>
<td>Novagen</td>
</tr>
<tr>
<td>pET28a</td>
<td>Km′, expression vector</td>
<td>Novagen</td>
</tr>
<tr>
<td>pZW102</td>
<td>NdeI-BamHI fragment containing <em>narX</em> inserted into pET5a</td>
<td>This study</td>
</tr>
<tr>
<td>pZW106</td>
<td>NdeI-EcoRI fragment containing <em>narI</em> inserted into pET5a</td>
<td>This study</td>
</tr>
<tr>
<td>pZW107</td>
<td>NdeI-EcoRI fragment containing <em>narL</em> inserted into pET5a</td>
<td>This study</td>
</tr>
<tr>
<td>pZW108</td>
<td>NdeI-EcoRI fragment containing <em>narK</em> inserted into pET5a</td>
<td>This study</td>
</tr>
<tr>
<td>pZW109</td>
<td>NdeI-BamHI fragment containing <em>narX</em> inserted into pET28a</td>
<td>This study</td>
</tr>
<tr>
<td>pZW110</td>
<td>NdeI-EcoRI fragment containing <em>narL</em> inserted into pET28a</td>
<td>This study</td>
</tr>
<tr>
<td>pZW111</td>
<td>NdeI-EcoRI fragment containing <em>ncgl2918</em> inserted into pET5a</td>
<td>This study</td>
</tr>
<tr>
<td><strong>Primer</strong></td>
<td>Sequence (5′–3′)</td>
<td>Purpose</td>
</tr>
<tr>
<td>h1</td>
<td>5′-TCTGGATCCATAATGCAGCAGAGC-3′</td>
<td>Forward primer for <em>narX</em> expression</td>
</tr>
<tr>
<td>h2</td>
<td>5′-AATTGAGTACCCACTACCATCCAGT-3′</td>
<td>Reverse primer for <em>narX</em> expression</td>
</tr>
<tr>
<td>g1</td>
<td>5′-AGATCTAGCCAGCCACCCGAA-3′</td>
<td>Forward primer for <em>narI</em> expression</td>
</tr>
<tr>
<td>g2</td>
<td>5′-GCTGAATTCCGGTGAAGCTTCAATGAG-3′</td>
<td>Reverse primer for <em>narI</em> expression</td>
</tr>
<tr>
<td>r1</td>
<td>5′-CCCTTGAATTCCATATGGAACGCTGCCACCT-3′</td>
<td>Forward primer for <em>narK</em> expression</td>
</tr>
<tr>
<td>r2</td>
<td>5′-AGCGAATCTGGTCATCCTATTTGACG-3′</td>
<td>Reverse primer for <em>narK</em> expression</td>
</tr>
<tr>
<td>m1</td>
<td>5′-ATACTATGACCAACCTTGAACGCTTC-3′</td>
<td>Forward primer for <em>narL</em> expression</td>
</tr>
<tr>
<td>m2</td>
<td>5′-GAGCAGATTCTAGCCACCCGATG-3′</td>
<td>Reverse primer for <em>narL</em> expression</td>
</tr>
</tbody>
</table>

*Letters in italics indicate restriction enzyme cutting site added for cloning.*
Cloning and expression of the narX, narI, narK, and narL genes in E. coli

Individual nar genes were amplified by PCR from strain NCIMB 12038 using Pyrobest polymerase (Takara, Dalian, China). Primers used for amplification of the target genes are listed in Table 1. PCR products were then cloned into pET5a individually to produce the plasmids listed in Table 1. The narX and narL genes were also ligated into pET28a to produce pZWT09 and pZWT10, respectively. Both procedures resulted in N-terminal six-His-tagged fusion proteins. The inserts of these clones were sequenced to ensure that no mutations had been incorporated during the PCR.

All of the pET5a and pET28a constructs were transformed into E. coli strain Rosetta (DE3) pLysS. The recombinant strains were grown in LB at 37°C to an optical density of 0.6 at 600 nm and were then induced with 0.1 mM of isopropyl-b-D-thiogalactopyranoside (IPTG) for an additional 12 h of cultivation at 16°C.

Preparation of cell extracts

Cells were harvested by centrifugation and disrupted by sonication at 4°C (3 s sonication followed by a 7-s break) in ice-cold phosphate buffer (50 mM, pH 7.4). After centrifugation at 13,000×g for 40 min at 4°C, the supernatant (cell extract) was collected for the enzyme assay. For protein purification, the buffer was changed to a binding buffer containing 50 mM sodium phosphate buffer (pH 8.0) and 300 mM NaCl.

Protein purification

All procedures were carried out at 4°C. Cell extracts containing His-tagged NarX or NarL were applied to nickel-nitrilotriacetic acid (Ni2+-NTA) agarose (Merck Biosciences, Darmstadt, Germany) according to the supplier’s recommendations. The column was then washed with binding buffer, and the protein was released with elution buffer (50 mM sodium phosphate buffer, 300 mM NaCl, 80 mM imidazole [pH 8.0]) at a flow rate of 1.0 ml/min.

Enzyme assays

All assays were carried out in 50 mM phosphate buffer (pH 7.4). The 3-hydroxybenzoate (3HBA) 6-monoxygenase activity was determined by measuring the decrease in absorbance at 340 nm due to the substrate-dependent oxidation of NADH or NADPH, the molar extinction coefficient of which was taken as 6,200 M⁻¹ cm⁻¹ (Liu et al. 2005; Wang et al. 1987). GDO was assayed by measuring the increase in absorbance at 330 nm due to the conversion of gentisate to maleylpyruvate, the molar extinction coefficient was taken as 13,000 M⁻¹ cm⁻¹ (Lack 1959). Maleylpyruvate was prepared from gentisate by GDO as described previously (Feng et al. 2006). Maleylpyruvate isomerase activity was monitored by measuring the change in absorbance at 330 nm due to maleylpyruvate disappearance in the presence of boiled cell extracts of strain NCIMB 12038 (or strain ATCC 13032) or MSH as a cofactor (Feng et al. 2006). Rates of isomerization of maleylpyruvate to fumarylpyruvate were calculated using a value of 2,400 M⁻¹ cm⁻¹ for the extinction change at 330 nm (Hagedorn et al. 1985). Fumarylpyruvate hydrolase was assayed by measuring the decrease at 340 nm due to fumarylpyruvate disappearance; the molar extinction coefficient of fumarylpyruvate was taken as 9,400 M⁻¹ cm⁻¹ (Tanaka et al. 1957). The protein concentration was determined according to the Bradford (1976) method, using bovine serum albumin as the standard. One unit of enzyme activity was defined as the amount required for the disappearance (or production) of 1 μmol of substrate (or product) per min at 28°C. Specific activities are expressed as units per milligram of protein.

HPLC analysis

HPLC analysis was performed on an Agilent series 1200 system equipped with a C18 reversed-phase column (5 μm, 4.6×250 mm, Agilent Technologies). The mobile phase was methanol/water (20:80, v/v) containing 100 mM ammonium acetate (pH 4.2) at a flow rate of 1 ml/min. Detection was performed by UV at 300 nm. The retention times for 3-hydroxybenzoate and gentisate were 8.3 and 3.8 min, respectively. The analysis of fumarate and pyruvate was performed as described previously (Zhou et al. 2001), the retention times were 15.1 and 9.2 min, respectively.

Molecular weight determination

The molecular weight and subunit composition of proteins were determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Gel filtration chromatography was performed to determine the native molecular masses of the purified proteins as described previously (Zhang et al. 2009).

Reverse transcription PCR

Total RNA of strain NCIMB 12038 was isolated using an RNeasy mini kit (QIAGEN). DNase digestion was carried out under the manufacturer’s instructions.
out to remove any DNA contamination. A reverse transcription-PCR (RT-PCR) was performed using Moloney murine leukemia virus reverse transcriptase (Promega), and the resulting cDNAs were amplified with gene-specific primers for each nar gene.

Nucleotide sequence accession number

The GenBank accession number for the nucleotide sequence reported in this paper is HM852512.

Results

Cloning of the gentisate degradation genes and DNA sequence analysis

Based on a conserved region of the gentisate 1,2-dioxygenase gene, a 599-bp PCR product was obtained from strain NCIMB 12038. Subsequently, a 12,052-bp DNA sequence flanking the GDO gene was obtained after several cycles of genome walking. Thirteen putative ORFs were identified and annotated on the basis of BLAST analysis of this genome walking. Thirteen putative ORFs were identified flanking the GDO gene was obtained after several cycles of sequence analysis.

Table 2 BLAST searching results and identified functions of genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Proposed function</th>
<th>Position in sequence</th>
<th>Product size (residues, kDa)</th>
<th>Related gene product (species, % amino acid identity, accession no.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>narX</td>
<td>3-Hydroxybenzoate 6-hydroxylase</td>
<td>4256–3054</td>
<td>400 (43.8)</td>
<td>Salicylate hydroxylase (Rhodococcus jostii RHA1, 92, BAH49794)</td>
</tr>
<tr>
<td>narR3</td>
<td>Transcriptional regulator</td>
<td>4433–5203</td>
<td>256 (28.2)</td>
<td>IclR family transcriptional regulator (R. jostii RHA1, 94, ABG93678)</td>
</tr>
<tr>
<td>narI</td>
<td>Gentisate 1,2-dioxygenase</td>
<td>5268–6347</td>
<td>359 (39.9)</td>
<td>Gentiopicroside 1,2-dioxygenase (Corynebacterium glutamicum ATCC 13032, 64, NP 602217)</td>
</tr>
<tr>
<td>narK</td>
<td>Fumarylpyruvate hydrolase</td>
<td>6354–7181</td>
<td>275 (29.7)</td>
<td>3-Maleylpyruvate isomerase (Rhodococcus opacus, 92, BAD89181); fumarylpyruvate hydrolase (C. glutamicum ATCC 13032, 58, NP 602216)</td>
</tr>
<tr>
<td>narL</td>
<td>3-Maleylpyruvate isomerase</td>
<td>7184–7912</td>
<td>239 (26.1)</td>
<td>Mycothiol-dependent maleylpyruvate isomerase (C. glutamicum ATCC 13032, 51, NP 602215)</td>
</tr>
</tbody>
</table>
enzyme activity by 10%. SDS-PAGE indicated an overexpressed polypeptide of ~40 kDa (data not shown), as expected from the deduced amino acid composition.

NarL acts as a mycothiol-dependent maleylpyruvate isomerase

According to the results of a BLAST search, NarL was presumed to be an MSH-dependent maleylpyruvate isomerase.

When cell extracts of IPTG-induced E. coli [pZWT07, a pET5a-based construct] carrying the narL gene were incubated with maleylpyruvate generated from gentisate by the action of NarI, no change in the absorption spectrum between 250 and 400 nm was observed, unless 10 μl of boiled cell extracts of strain NCIMB 12038 was added. The spectrum changed due to a shift in the λ\text{max} from 330 to 340 nm that corresponded to the conversion of maleylpyruvate to fumarylpyruvate (Fig. 3) as described previously (Zhou et al. 2001). In this case, the specific activity was 0.016 U/mg. The same transformation also took place when the boiled cell extracts of strain ATCC 13032 or MSH purified from strain ATCC 13032 were used with specific activities of 0.011 and 0.017 U/mg, respectively. To compare NarL with the maleylpyruvate isomerase Ncgl2918 from strain ATCC 13032, the cell extracts of E. coli [pZWT11] carrying ncl2918 were also investigated by

---

**Fig. 1** a Proposed metabolic pathway for naphthalene and 3HBA degradation in *Rhodococcus* sp. strain NCIMB 12038. b Organization of 3HBA-gentisate catabolic cluster in strain NCIMB 12038. The locations of amplified DNA fragments RTG, RTI, RTIK, and RTKL for RT-PCR are indicated by short lines. c Analysis of narX and narIKL transcription by RT-PCR. Lane M molecular size markers (100- to 2,000-bp), lanes 2, 6, 10, and 14 corresponding products RTG, RTI, RTIK, and RTKL, respectively, induced with naphthalene, lanes 4, 8, 12, and 16 indicated the amplified results with growth on pyruvate, lanes 1, 3, 5, 7, 9, 11, 13, and 15 corresponding negative controls without RT

---

**Fig. 2** SDS-PAGE of H6-NarX (a) and H6-NarL (b). a Lane 1 purified H6-NarX, lane 2 molecular mass standards (molecular masses [in kilodaltons] are indicated on the right). b Lane 1 molecular mass standards, lane 2 purified H6-NarL.

---

**Fig. 3** The spectral changes of isomerization of maleylpyruvate to fumarylpyruvate catalyzed by the mycothiol-dependent maleylpyruvate isomerase NarL.

---

© Springer
adding the boiled cell extracts from strain NCIMB 12038. Negl2918 also showed activity and its specific activity was ten times higher than that of NarL.

To purify this MSH-dependent maleylpyruvate isomerase, the narL gene was cloned into expression vector pET28a to produce pZWT10. The cell extracts of IPTG-induced E. coli [pZWT10] exhibited maleylpyruvate isomerase specific activity of 0.1 U/mg with MSH as a cofactor. However, the purified Hε-NarL exhibited specific activity as high as 7.9 U/mg. A single polypeptide band of ~30 kDa was detected by SDS-PAGE (Fig. 2b), and a protein with a molecular mass of 29 kDa was detected by gel filtration, indicating that this enzyme is a monomeric protein, the same as Negl2918 from strain ATCC 13032 (Wang et al. 2007).

NarK acts as a fumarylpyruvate hydrolase

SDS-PAGE analysis of the cell extracts of E. coli [pZWT08] carrying the narK gene showed a polypeptide of ~30 kDa, matching the predicted molecular mass of NarK deduced from its amino acid sequence. To investigate the role of NarK, the cell extracts were incubated with the product of the joint action of NarI and NarL on gentisate. It was evident from the absorption spectrum that the peak, with a λmax at 340 nm, decreased and then disappeared, corresponding to the hydrolysis of fumarylpyruvate. This indicated that NarK acts as a fumarylpyruvate hydrolase, although most of its close homologues in the GenBank database were annotated as maleylpyruvate isomerases. The measured specific activity of fumarylpyruvate hydrolase for NarK was 5.5 U/mg. The products of fumarylpyruvate hydrolase were confirmed as fumarate and pyruvate by HPLC.

Transcription of the nar genes in strain NCIMB 12038

To determine the transcription of the narXR3IKL cluster, RNA was extracted from pyruvate and naphthalene-grown cells of strain NCIMB 12038. The narI, narK, and narL genes were found to be transcribed as a single operon, only in the presence of naphthalene. However, transcription of the narX gene occurred under both cases, suggesting that the narX gene is constitutively transcribed (Fig. 1c).

Discussion

The genus *Rhodococcus* is regarded as a diverse group of bacteria that possesses the catabolic versatility to degrade a large number of organic compounds (Larkin et al. 2005). Naphthalene catabolism in *Rhodococcus* sp. strain NCIMB 12038 was reported to occur via the gentisate pathway in the 1980s (Larkin and Day 1986). However, genetic and biochemical level studies on the gentisate pathway have not been reported in the literature in this “historic” strain or in other Gram-positive naphthalene utilizers. In this study, we cloned and characterized three genes involved in gentisate catabolism via MSH-dependent maleylpyruvate isomerization, a pathway that differs from the gentisate pathway in the Gram-negative naphthalene utilizing *Ralstonia* sp. strain U2 (Zhou et al. 2001). The presence of a functional 3HBA monooxygenase-encoding gene in this cluster led us to discover that strain NCIMB 12038 could also grow on 3HBA. RT-PCR analysis showed that naphthalene induced the transcription of genes involved in gentisate degradation. Therefore, we proposed that the gentisate degradation pathway was shared by both naphthalene and 3HBA catabolism in this strain, as shown in Fig. 1a.

In Gram-negative bacteria, glutathione (GSH) is the most abundant thiol (Newton et al. 1996) and acts as a cofactor of maleylpyruvate isomerase in the gentisate pathway (Zhou et al. 2001), while mycothiol functions as a GSH analogue in actinomycete bacteria (Newton et al. 1996). To date,
Ngl2918 in Corynebacterium glutamicum strain ATCC 13032 is the only MSH-dependent maleylpyruvate isomerase identified (Feng et al. 2006). In Rhodococcus spp., there is some evidence that MSH-dependent maleylpyruvate isomerase is involved in the gentisate pathway. In a recent report on Rhodococcus jostii strain RHA1, an mshC (an essential gene for the biosynthesis of MSH)-deficient strain failed to grow when gentisate and 3HBA were used as the sole carbon sources (Dosanjh et al. 2008), suggesting that MSH is involved in the gentisate pathway in this strain. In Rhodococcus sp. strain R7, gene genI, annotated as 3-maleylpyruvate isomerase (Di Gennaro et al. 2010), shares 89% amino acid identity with narK, a fumarylpyruvate hydrolase described in the current study, while GenL was experimentally.

On alignment of nar homologues from different bacteria (Fig. 4), we found that three gentisate pathway-encoding clusters share the same organization among these strains, regardless of whether their products are MSH- or GSH-dependent. By contrast, gentisate degradation genes tend to be clustered with the 3HBA 6-monoxygenase gene in both Gram-positive and -negative strains, suggesting that a coevolutionary process may have occurred during the development of their biodegradation capacity.

Acknowledgements This work was supported by the National High Technology Research and Development Program of China (2007AA10Z402) and the Chinese Academy of Sciences (KSCX2-YW-G-052).

References


Springer

