ELSEVIER

Biochimica et Biophysica Acta 1275 (1996) 154–156

BBA Biochimica et Biophysica Acta

Short sequence-paper

Structure of chromosomal DNA coding for *Pseudomonas putida* S-1 salicylate hydroxylase [1]

Kenzi Suzuki [a,*], Mitsuo Mizuguchi [a], Kuniharu Ohnishi [b], Eiji Itagaki [a]

[a] *Department of Chemistry, Faculty of Science, Kanazawa University, Kakuma-machi, Kanazawa 920-11, Japan*

[b] *Department of Microbiology, Faculty of Pharmaceutical Sciences, Hokuriku University, Kanagawa-machi, Kanazawa 920-11, Japan*

Received 12 March 1996; revised 7 May 1996; accepted 8 May 1996

Abstract

A gene coding for the salicylate hydroxylase has been isolated from chromosomal DNA of *Pseudomonas putida* S-1 and sequenced. The DNA fragment contained an open reading frame of 1266 bp encoding a polypeptide of 421 amino acid residues. The predicted amino acid sequence of the protein gave a good agreement with the sequences determined with the peptides isolated from the enzyme but methionine residue in the amino terminal was deleted in the N-terminal sequence of the enzyme protein. The nucleotide and amino acid sequences of the salicylate hydroxylase shared several common characteristics with those of the enzyme encoded on the plasmid DNA of *P. putida* PpG7; homology of nucleotide sequence is 58% and that of amino acid sequence is 56%. We could find two large conserved regions of the amino acid sequence at or near FAD- and NADH-binding regions. The FAD-binding site locates on the amino terminal and a lysine residue, functioning as an NADH-binding site (K. Suzuki, M. Mizuguchi, T. Gomi, and E. Itagaki, 1995, J. Biochem. 117,579–585), locates as Lys^{163}.

Keywords: Monooxygenase; Salicylate hydroxylase; Flavoprotein; Chromosomal DNA; (*P. putida* S-1)

Salicylate hydroxylase (salicylate, NADH: oxygen oxido-reductase (1-hydroxylating, decarboxylating), EC 1.14.13.1) is a mono-oxygenase catalyzing the decarboxylative hydroxylation of salicylate to produce catechol in consumption with NADH and O_2 [1]. The enzyme is induced in *Pseudomonas putida* S-1 by addition of salicylate as a sole carbon and energy source into the culture medium. It had been purified, crystallized, and characterized [2–4]. The enzyme is a monomeric flavoprotein with the molecular weight of 54000 containing one molecule of FAD as a prosthetic group [5]. The enzyme was also isolated from other bacterial sources including *P. putida* PpG7 [6–11]. We have studied on the reaction mechanisms [12] and chemical modifications of Arg and Lys residues of the enzyme [13,14]. To progress the studies, precise information about the enzyme structure are required strongly. Until now, *P. putida* PpG7 plasmid encoded salicylate hydroxylase has been sequenced; the structural analysis, however, was not reported [15]. The *nahG* gene encoding the salicylate hydroxylase was reported to be loaded on *sal* operon of a plasmid NAH7.

To understand the structure and function and the induction mechanism of the enzyme, we have amplified and characterized the genomic DNA of the enzyme of *P. putida* S-1 using PCR and cassette-primer PCR techniques [16]. Amplification of the chromosomal DNA with the primers described in Fig. 1 gave DNA fragments near 540bp and in the second PCR, the product of 470bp containing the sequences of the primers was isolated. Using the product as a probe, 4.2-kbp DNA fragment was obtained from the endonuclease (*Sal*I and *Hin*dIII)-digested chromosomal DNA. Transformed *Escherichia coli* cells with a vector ligated the DNA fragment exhibited a low activity of salicylate hydroxylase and the addition of salicylate to the cells induced the high enzyme activity. The fragment containing the *sal* gene of the salicylate hydroxylase was sequenced (Fig. 1). An ORF with 1266 nucleotides starting at ATG codon and ending at TAA codon is present in the sequence. The ORF encodes a protein of 421 amino acid residues with the molecular weight of 45288, which is smaller than the value (54000)

* Corresponding author. Fax: +81 762 645742; e-mail: kenzi@cacheibm.s.kanazawa-u.ac.jp.

[1] The sequence data reported in this paper have been submitted to the DDBJ/EMBL/GenBank Data Library under the accession number D67098.

PII S0005-2728(96)00069-2

```
   1 ATGAGCAAATCTCCCCTGCGTGTAGCTGTC ATTGGCGGAGGCATCGCTGGTACTGCCCTG
     MetSerLysSerProLeuArgValAlaVal IleGlyGlyGlyIleAlaGlyThrAlaLeu
  61 GCCCTCGGCCTCAGCAAATCCTCTCACGTC AATGTCAAACTGTTTGAAACTGCCCCTGCC
     AlaLeuGlyLeuSerLysSerSerHisVal AsnValLysLeuPheGluThrAlaProAla
 121 TTCGGCGAAATCGGTGCCGGCGTTTCCTTC GGCGTCAACGCTGTAGAAGCTATCCAACGC
     PheGlyGluIleGlyAlaGlyValSerPhe GlyValAsnAlaValGluAlaIleGlnArg
 181 CTGGGTATTGGCGAACTTTACAAAAGCGTT GCAGACAGCACCCCAGCACCTTGGCAAGAC
     LeuGlyIleGlyGluLeuTyrLysSerVal AlaAspSerThrProAlaProTrpGlnAsp
 241 ATCTGGTTTGAATGGCGTCATGCGCATGAT GCTTCGCTTGTAGGCGCCACCGTTGCGCCG
     IleTrpPheGluTrpArgHisAlaHisAsp AlaSerLeuValGlyAlaThrValAlaPro
 301 GGTATTGGCCAGTCATCCATCCATCGTGCA GACTTCATCGACATGCTCGAAAAGCGTTTG
     GlyIleGlyGlnSerSerIleHisArgAla AspPheIleAspMetLeuGluLysArgLeu
 361 CCTGCGGGCATCGCATCCCTGGGTAAGCAT GTCGTTGACTACACCGAAAACGCTGAAGGG
     ProAlaGlyIleAlaSerLeuGlyLysHis ValValAspTyrThrGluAsnAlaGluGly
 421 GTGACGCTCAATTTCGCAGATGGGAGCACC TACACTGCTGACGTAGCGATCGCTGCAGAC
     ValThrLeuAsnPheAlaAspGlySerThr TyrThrAlaAspValAlaIleAlaAlaAsp
 481 GGCATCAAGTCCTCCATGCGAAATACGCTG CTGCGTGCCGCCGGCCATGATGCCGTCCAT
     GlyIleLysSerSerMetArgAsnThrLeu LeuArgAlaAlaGlyHisAspAlaValHis
 541 CCGCAGTTCACCGGGACATCCGCCTACCGC GGGCTTGTAGAGACCTCTGCCCTTCGCGAG
     ProGlnPheThrGlyThrSerAlaTyrArg GlyLeuValGluThrSerAlaLeuArgGlu
 601 GCCTACCAAGCGGCATCACTGGACGAGCAT TTGCTCAATGTGCCGCAAATGTACTTGATC
     AlaTyrGlnAlaAlaSerLeuAspGluHis LeuLeuAsnValProGlnMetTyrLeuThr
 661 GAAGACGGCCACGTACTGACCTTCCCGGTT AAAAAGGGGAAGTTGATCATTATCGTGGCG
     GluAspGlyHisValLeuThrPheProVal LysLysGlyLysLeuIleIleIleValAla
 721 TTCGTGTCTGATCGCAGCGTCGCCAAACCG CAGTGGCCATCCGACCAACCTTGGGTTCGT
     PheValSerAspArgSerValAlaLysPro GlnTrpProSerAspGlnProTrpValArg
 781 CCCGCCACCACAGACGAGATGCTGCACCGA TTTGCAGGCGCCGGAGAGGCAGTAAAAACC
     ProAlaThrThrAspGluMetLeuHisArg PheAlaGlyAlaGlyGluAlaValLysThr
 841 CTCCTGACCAGCATCAAGAGCCCAACCCTC TGGGCCCTTCATGACTTTGACCCGCTGCCC
     LeuLeuThrSerIleLysSerProThrLeu TrpAlaLeuHisAspPheAspProLeuPro
 901 ACCTATGTGCATGGTCGCGTAGCACTGATT GGCGATGCTGCGCACGCCATGCTCCCACAC
     ThrTyrValHisGlyArgValAlaLeuIle GlyAspAlaAlaHisAlaMetLeuProHis
 961 CAAGGCGCAGGAGCAGGTCAGGGCCTTGAG GATGCTTACTTCATGGCCGAACTGCTCGGC
     GlnGlyAlaGlyAlaGlyGlnGlyLeuGlu AspAlaTyrPheMetAlaGluLeuLeuGly
1021 AACCCTCTTCACGAAGCTAGCGATATTCCA GCTCTCTTGGAGGTGTATGACGACGTTCGC
     AsnProLeuHisGluAlaSerAspIlePro AlaLeuLeuGluValTyrAspAspValArg
1081 AGGGGCCGCGCCTCCAAGGTTCAGCTGACC TCGCGTGAAGCAGGCGAACTCTATGAATAT
     ArgGlyArgAlaSerLysValGlnLeuThr SerArgGluAlaGlyGluLeuTyrGluTyr
1141 AGAACACCAGGTGTTGAACGCGATACCGCC AAGCTGAAGGCTTTGCTTGAGAGCCGTATG
     ArgThrProGlyValGluArgAspThrAla LysLeuLysAlaLeuLeuGluSerArgMet
1201 AACTGGATCTGGAACTACGACCTGGGTGCC GAGGCTCGTCTGGCAGTTAAACCCGCCCTC
     AsnTrpIleTrpAsnTyrAspLeuGlyAla GluAlaArgLeuAlaValLysProAlaLeu
1261 GCATAA
     Ala *
```

Fig. 1. Nucleotide sequence of the salicylate hydroxylase gene sal from *Pseudomonas putida* S-1 and the predicted amino acid sequence. The oligonucleotide primers used for isolation of the clone were primer 1; CC(AG)CTGCG(GC)GT(CG)GC(CG)AT(CT)GG(CT)GG, primer 2: C C (C G) G T (A G) A A (C T) T G (A C G T) G G (A G) T G (CG)AC(CG)GC(AG)TC, primer 3: GG(CT)GG(CT)GG(CT)AT(CT) GC(CG)GG, and primer 4: TT(AG)AT(AG)CC(AG)TC(CG)GC (CG)GC. These primers were designed from the Edman degradation data of the amino terminal peptide and two tryptic peptides for the NADH-binding site of salicylate hydroxylase [14]. Chromosomal DNA of *P. putida* S-1 was digested with *Sal*I and *Eco*RI, and 4.2-kbp DNA fragment was isolated and ligated to pUC18 plasmid vector. After the amplification, one plasmid was cloned with the probes as pSAH1. It was digested with *Pma*CI and *Hin*dIII and the nucleotide sequence of the open reading frame was sequenced in both directions using *Taq* cycle sequencing kit (Takara Shuzou). Sequence data were organized and analyzed using the GENETYX (version 9) program (Software Development Co. Ltd.). The positions corresponding to the oligonucleotide primers used for PCR were underlined in the figure. The sequences have been deposited in the DDBJ/ EMBL/GenBank Data Library under accession number D67098.

estimated by SDS gel electrophoresis and sedimentation experiments of purified salicylate hydroxylase [5]. The reason why the difference occurred was unknown. The sequence shown in Fig. 1 is supported by the amino acid composition, except cysteine residue, of the purified enzyme reported previously [5]. The nucleotide sequence has the homology of 58% to that of *nahG* of *P. putida* PpG7 [15]. The G + C content of the sal coding region is 57%, which is close to the value of 65% of *nahG* of *P. putida* PpG7 [15]. The preferential usage of C- and G-ending codons of *P. putida* S-1 enzyme is 61%. The value of *nahG* was reported to be 76% [15]. Hydrophobic and hydrophilic amino acid residues are 56% and 27%, respectively: almost the same values are seen in the salicylate hydroxylase of *P. putida* PpG7 [15].

The previous analytical data about the amino acid composition of the enzyme showed the presence of cysteine residues [5], but the present predicted sequence exhibited the absence of the residue. Re-examination of the analysis of the residue with 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB) method and also amino acid analysis of the performic acid-treated enzyme protein resulted in the confirmation of the sequence data. The amino terminal sequence of the enzyme protein was determined as SKSPLRVAVIGGGIAGTALALG. The sequence indicated the deletion of methionine residue from the amino terminal of the predicted sequence of salicylate hydroxylase in Fig. 1, indicating that the residue is processed off during the synthesis of the enzyme. The amino terminal residue of the plasmid-encoded salicylate hydroxylase of *P. putida* PpG7 was reported to be lysine [11].

Thus, the enzyme consisted of 420 amino acid residues with a molecular weight of 45 157. Homologies of the amino acid sequence predicted from the sal gene to the sequence of the product of *nahG* gene and to that of *p*-hydroxybenzoate hydroxylase from *P. fluorescens* are 56% and 25%, respectively (Fig. 2).

A long range homology was found between the amino acid sequences of *P. putida* S-1 and *P. putida* PpG7; 26 amino acid residues of residues 309-334 and residues 311-336, respectively. The regions were also highly homologous with the region of residues 283-304 of the sequence of *p*-hydroxybenzoate hydroxylase, in which the region constructs the part of the substrate binding pocket [17]. The region of residues 7-24 of *P. putida* S-1 enzyme, RVAVIGGGIAGTALALGL, contains the consensus sequence of ADP binding site of FAD and is homologous to those of salicylate hydroxylase of *P. putida* PpG7 [15] and of ruburedoxin reductase and of toluene reductase of *P. putida* F1 [18]. The region makes up bab-fold structure to bind ADP [19]. The consensus sequence of the second FAD binding region of flavoprotein is found in the sequence of residues 302-312 of *P. putida* S-1, which is highly conserved in that of *P. putida* PpG7, residues 304-314, and is similar to the sequence of *P. fluorescens* *p*-hydroxybenzoate hydroxylase, residues 276-286 (Fig. 2).

Our previous study of the chemical modification of salicylate hydroxylase revealed the presence of a lysine residue in the binding site of NADH and determined the amino acid sequence around the residue [14]. The sequence was found on the predicted amino acid sequence of the

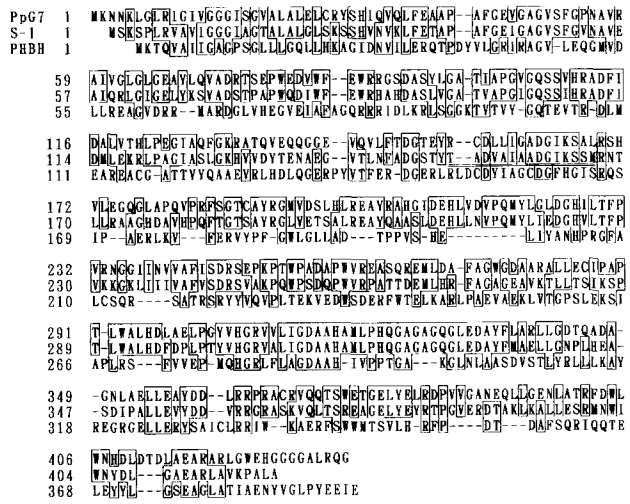

Fig. 2. Comparison of the deduced amino acid sequences for *P. putida* S-1 and *P. putida* PpG7 [15] salicylate hydroxylase and *P. fluorescens* *p*-hydroxybenzoate hydroxylase [20]. Residues that are identical in two or all three sequences are boxed and gapped position is shown as –. PpG7; salicylate hydroxylase of *P. putida* PpG7, S-1: salicylate hydroxylase of *P. putida* S-1, and PHBH: *p*-hydroxybenzoate hydroxylase of *P. fluorescens*. Double underlining indicates the peptide containing the lysine residue essential for binding of NADH.

protein, in which the lysine residue is located at residue 163 (Fig. 2). The residue is also conserved in the enzyme from *P. putida* PpG7.

The amino acid and nucleotide sequences obtained in this study are useful to disclose the functionally essential residues in the catalytic cycle of the salicylate hydroxylase by the mutagenesis and crystallography.

Acknowledgements

We thank Dr. S. Sato, Tokyo Medical and Dental University, for sequence analysis of the amino terminal peptide of salicylate hydroxylase.

References

[1] Katagiri, M., Takemori, S., Suzuki, K. and Yasuda, H. (1966) J. Biol. Chem. 241, 5675–5677.
[2] Takemori, S., Yasuda, H., Mihara, K., Suzuki, K. and Katagiri, M. (1969) Biochim. Biophys. Acta 191, 58–68.
[3] Takemori, S., Yasuda, H., Mihara, K., Suzuki, K. and Katagiri, M. (1969) Biochim. Biophys. Acta 191, 69–76.
[4] Suzuki, K., Takemori, S. and Katagiri, M. (1969) Biochim. Biophys. Acta 191, 77–85.
[5] Takemori, S., Hon-nami, K., Kawahara, F. and Katagiri, M. (1974) Biochim. Biophys. Acta 342, 137–144.
[6] White-Stevens, R.H. and Kamin, H. (1972) J. Biol. Chem. 247, 2358– 2370.
[7] Presswood, R.P. and Kamin, H. (1976) in Flavins and Flavo-proteins (Singer, T.P., ed.), pp. 145–154. Elsevier, Amsterdam.
[8] Tu, S.-C., Romero, F.A. and Wang, L.-H. (1981) Arch. Biochem. Biophys. 209, 423–432.
[9] Wang, L.-H. and Tu, S.-C. (1984) J. Biol. Chem. 259, 10682–10688.
[10] Sze, I. and Dagley, S. (1984) J. Bacteriol. 159, 353–359.
[11] You, I. -S., Murray, R.I., Jollic, D. and Gunsalus, I.C. (1990) Biochem. Biophys. Res. Commun. 169, 1049–1054.
[12] Takemori, S., Nakamura, M., Suzuki, K., Katagiri, M. and Nakamura, T. (1972) Biochim. Biophys. Acta 284, 382–393.
[13] Suzuki, K. and Ohnishi, K. (1990) Biochim. Biophys. Acta 1040, 327–336.
[14] Suzuki, K., Mizuguchi, M., Gomi, T. and Itagaki, E. (1995) J. Biochem. 117, 579–585.
[15] You, I. -S., Ghosal, D. and Gunsalus, I.C. (1991) Biochemistry 30, 1635–1641.
[16] Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Science 239, 487–491.
[17] Schreuder, H.A., Van der Laan, J.M., Hol, W.G.J. and Drenth, J. (1991) in Chemistry and Biochemistry of Flavo-enzyme (Muller, F. ed.), Vol. II, pp. 31–64, CRC Press, Boca Raton, FL.
[18] Zylstra, G.J. and Gibson, D.T. (1989) J. Biol. Chem. 264, 14940–14948.
[19] Wierenga, R.K., Terpstra, P. and Hol, W.G.J. (1986) J. Mol. Biol. 187, 101–107.
[20] Weijer, W.J., Hofsteenge, J., Vereijken, J.M., Jekel, P.A. and Beintema, J.J. (1982) Biochim. Biophys. Acta 704. 385–388.