Catalytic mechanism for the conversion of salicylate into catechol by the flavin-dependent monooxygenase salicylate hydroxylase

Débora M. A. Costa,[1,†] Stefanya V. Gómez,[2,†] Simara S. de Araújo,[1] Mozart S. Pereira,[2] Rosemeire B. Alves,[2] Denize C. Favaro,[3] Alvan C. Hengge,[4] Ronaldo A. P. Nagem,[1,] and Tiago A. S. Brandão[2,*]*

[1] Departamento de Bioquímica e Imunologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31270-901, Brazil.

[2] Departamento de Química, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31270-901, Brazil.

[3] Instituto de Química, Universidade Estadual de Campinas, Campinas, SP 13083-970, Brazil.

[4] Department of Chemistry and Biochemistry, Utah State University, Logan, UT 84322-0300, USA.

[†] D.M.A.C. and S.V.G. contributed equally to this work.

EMAIL ADDRESS: nagem@icb.ufmg.br (R.A.P.N.); tasb@ufmg.br (T.A.S.B.)

* CORRESPONDING AUTHORS: Tel.: +55 31 3409-5766; Fax: +55 31 3409-5700

Abbreviations: NahG, salicylate hydroxylase; 6xHis-NahG, recombinant NahG with 6xHis N-terminal tag; Sal, salicylate; S_EAr, aromatic electrophilic substitution.

Enzymes: Salicylate hydroxylase (NahG), EC **1.14.13.1**

Keywords: *Pseudomonas putida* G7; salicylate hydroxylase; NahG; oxidative decarboxylation; mechanism; crystal structure.

Database: Structural data are available in the PDB database under the accession number 6BZ5.

Conflicts of Interest. The authors declare no competing interests.

Abstract

Salicylate hydroxylase (NahG) is a flavin-dependent monooxygenase that catalyzes the decarboxylative hydroxylation of salicylate into catechol in the naphthalene degradation pathway in *Pseudomonas putida* G7. We explored the mechanism of action of this enzyme in detail using a combination of structural and biophysical methods. NahG shares many structural and mechanistic features with other versatile flavin-dependent monooxygenases, with potential biocatalytic applications. The crystal structure at 2.0 Å resolution for the *apo* form of NahG adds a new snapshot preceding the FAD binding in flavin-dependent monooxygenases. The k_{cat}/K_m for the salicylate reaction catalyzed by the *holo* form is greater than 10^5 M^{-1} s^{-1} at pH 8.5 and 25 ºC. Hammett plots for K_m and k_{cat} using substituted salicylates indicates a change in rate-limiting step. Electron-donating groups favor the hydroxylation of salicylate by a peroxyflavin to yield a Wheland-like intermediate, whereas the decarboxylation of this intermediate is faster for electron-withdrawing groups. The mechanism is supported by structural data and kinetic studies at different pHs. The salicylate carboxyl group lies near a hydrophobic region that aids decarboxylation. A conserved histidine residue is proposed to assist the reaction by general base/general acid catalysis.

1. Introduction

Salicylate hydroxylase (NahG, EC 1.14.13.1) is a flavin-dependent monooxygenase that catalyzes the decarboxylative hydroxylation of salicylate into catechol, CO_2, and water, with stoichiometric consumption of NADH and O_2 (Scheme 1) [1]. This reaction links the upper and lower pathways of naphthalene degradation by soil pseudomonads [2], a bacterial genus encompassing many species that can use naphthalene or salicylate as sole carbon sources [3-5].

A number of *nahG* genes and encoded proteins have been identified in several naphthalene-degrading strains, such as *P. putida* G7 [6, 7], which harbors the archetype catabolic plasmid NAH7 [8]; *P. putida* S-1 [9-11]; *P. putida* KF715 [12]; *P. stutzeri* AN10 [13], whose dissimilatory genes are chromosomally encoded [14]; and *Pseudomonas* sp. ND6 [15], among others [16]. Other salicylate hydroxylase genes situated outside, but in close proximity, to the transcriptional unit forming the naphthalene degradation lower pathway, *i.e. nahW* and *nahU*, have been reported for *P. stutzeri* AN10 [17, 18] and *Pseudomonas* sp. ND6 [19], respectively. Interestingly, the coexistence of the classic NahG with NahW was shown to improve cell growth in salicylate-supplemented media [17]. The structures and catalytic mechanisms of NahW and NahU are unknown. NahG is the most-studied salicylate hydroxylase, though poorly understood.

NahG hydroxylates and decarboxylates the substrate at the same aromatic carbon atom (*ipso* substitution) [20, 21], an exquisite feature that is remarkable from a synthetic perspective. NahG is one of the few examples of a flavin enzyme that catalyzes an *ortho* hydroxylation relative to the substrate's phenol group, which is usually required for efficient catalysis. Other monooxygenases exhibit different regioselectivities. For example, the oxidative decarboxylation of 6-hydroxynicotinate catalyzed by 6-hydroxynicotinate 3-monooxygenase (NicC) occurs *para*

to the substrate phenol group [22]. In the reaction catalyzed by 3-hydroxybenzoate 6-hydroxylase (3HB6H), which the substrate carboxylate group is *meta* to the phenol, the 3-hydroxybenzoate is converted to 2,5-dihydroxybenzoate through a hydroxylation followed by a deprotonation [23].

Catalysis in these enzymes involves a C(4a)-hydroperoxyflavin species, which provides a powerful peroxo electrophile tuned for oxygen insertion at nucleophilic carbons and soft centers. Examples include hydroxylation coupled with different kinds of substitution, epoxidation, Baeyer-Villiger oxidation, and oxidation of heteroatoms (B, S, Se, N, and P) [24-27]. Formation of the C(4a)-hydroperoxyflavin species from the oxidized flavin and NAD(P)H requires two reactions coupled with dynamics of the isoalloxazine group between two positions [28]. In an external position, aside from the hydroxylation site, the oxidized isoaloxazine group is reduced by NAD(P)H. Then, the reduced isoaloxazine swings to an internal position, where it reacts with molecular oxygen to yield the C(4a)-hydroperoxyflavin species. This reaction is fast in enzymes, with rate constants typically between 10^4 and 10^6 M^{-1} s^{-1} at 4 ºC [29-32]; in contrast, the rate constant for the corresponding reaction in water is about 250 M^{-1} s^{-1} at 30 ºC [33].

Herein, we explore the catalytic mechanism of NahG, that, like those other monooxygenases, remains a subject of intense debate [10, 20]. Particular goals are to identify the catalytic groups involved in catalysis, the reactivity of reaction intermediates, and the reaction pathway that affords the products. The mechanistic proposal resulting from this work is supported by kinetics and x-ray crystallography data that provide a close look into the active site of NahG. In addition, we report the first structure to date of an *apo* flavin enzyme, providing a new snapshot to the set of events that culminate with the conversion of the *apo* to the peroxyflavin-bound protein.

2. Experimental section

2.1. Materials

Synthesis of 4-nitrosalicylic acid was as reported in the literature [34]. All other compounds were analytical grade and used without further purification. Doubly de-ionized water with a conductivity of less than 17.3 μS.cm was used to prepare all aqueous solutions.

2.2. Cloning, protein expression and purification

The coding sequence for salicylate hydroxylase NahG (NCBI Gene ID: 3974216) was amplified from *P. putida* G7 NAH7 plasmid (NCBI Accession No. AB237655) and cloned as described in the Supporting Information. Protein expression in *Escherichia coli* Rosetta (DE3) cells was carried out at 18 ºC for 16 h by induction with IPTG. The recombinant histidine-tagged NahG was purified by nickel affinity chromatography prior to a final size exclusion chromatography. The 6xHis-tag was removed for binding and kinetic studies using TEV protease prior a second nickel affinity chromatography step. In the text, we refer to the histidine-tagged form as 6xHis-NahG, while the tag-free enzyme is termed NahG. The protein concentration was estimated using the molar extinction coefficient of 76,890 M^{-1} cm^{-1} at 280 nm, calculated with the ProtParam tool [35] from the amino acid composition, and corrected for the amount of FAD. The concentration of FAD was estimated at 450 nm (ε = 11,300 M^{-1} cm^{-1}) [21], and used to calculate its absorption contribution at 280 nm with ε = 23,000 M^{-1} cm^{-1}. Residual FAD (< 30%) from protein expression was found in all purified protein samples. Detailed information regarding NahG production is provided in the Supporting Information.

2.3. Dynamic light scattering and size-exclusion chromatography

The hydrodynamic radius of NahG molecules and polydispersity of the purified sample were estimated using dynamic light scattering (DLS) assays performed on a Zetasizer Nano ZS system (Malvern Instruments). Protein samples were centrifuged at 17,000 *g* for 10 min at 4 °C prior to any measurement. The final protein concentration for DLS measurements was about 1.5 mg/mL in buffer C. Each sample was measured in a 1 cm path-length quartz cuvette at 25 °C. Data were collected and analyzed using the Malvern Zetasizer software v. 6.01, in which the "Protein Utility" routine was used to estimate particle mass based on the hydrodynamic radius.

The oligomeric state of the enzyme was also estimated by size-exclusion chromatography, using a calibrated HiLoad 16/60 Superdex 200 column (GE Healthcare) previously equilibrated with buffer C (50 mM Tris-HCl buffer at pH 7.4, 50 mM NaCl) at a flow rate of 1.0 mL/min. The following calibration standards were used: β-amylase (200 kDa); alcohol dehydrogenase (150 kDa); albumin (66 kDa); carbonic anhydrase (29 kDa); and cytochrome C (12.4 kDa) (from Sigma-Aldrich).

2.4. Crystallization, data collection, and structure determination

Pure and monodisperse 6xHis-NahG samples at concentrations of 30 mg/mL and 65 mg/mL were used in crystallization trials. Initial conditions were identified by high-throughput robotic screening [36] (SaltRX, Crystal Screen, and Crystal Screen 2 from Hampton Research; Precipitant Synergy, Wizard I, and Wizard II from Emerald BioSystems; PACT and JCSG+ from Qiagen) using the sitting-drop vapor diffusion method at 18 °C.

Diffraction-quality crystals were obtained by the hanging-drop vapor-diffusion method at 18 °C after manual optimization of the condition 1.5 M ammonium sulfate, 0.1 M Tris at pH 8.5, and 12% v/v glycerol by grid screening of salt and precipitant concentrations, and buffer pH. Crystallization drops were formed by mixing equal volumes (1 μL) of protein and reservoir

solutions. The crystal dimensions were approximately 0.38 x 0.18 x 0.06 mm. An iodide derivative was obtained using the quick cryo-soaking approach [37-39]. Essentially, a native crystal was soaked for about 30 seconds in a cryoprotectant solution – formed by the reservoir solution supplemented with 10% (v/v) ethylene glycol – additionally containing 0.06 M sodium iodide. Prior to data collection, crystals were flash-frozen in a nitrogen stream.

X-ray diffraction data were collected at the Laboratório Nacional de Luz Síncrotron (LNLS, Campinas-SP, Brazil) at the W01B-MX2 beamline [40]. X-ray diffraction images were recorded on a Rayonix MarMosaic 225 CCD detector and processed using the HKL-2000 software suite [41]. The 6xHis-NahG 3D structure was solved by the Single-wavelength Anomalous Dispersion method. Eight iodide sites were located by SHELXC/D [42]. AutoSol in the PHENIX software suite [43, 44] was used to calculate initial electron density maps using these heavy atom sites and also to perform automated preliminary model building. Structure refinement was carried out using PHENIX [44] and iterative rounds of manual model adjustments were done using Coot [45]. Data collection parameters and final refinement statistics are listed in Table 1. The refined 6xHis-NahG structure was deposited in the Protein Data Bank (entry 6BZ5). Pictures were generated with the program UCSF Chimera [46].

A three-dimensional search using the DALI server (http://ekhidna.biocenter.helsinki.fi/dali_server/) [47] was performed to identify proteins structurally similar to NahG. The primary structures of selected flavoprotein monooxygenases were aligned, and their crystallographic structures were superimposed, allowing a comparison among the *apo* and *holo* forms of the related enzymes.

2.5. Fluorometric titrations

Experiments were carried out on a Cary Eclipse fluorometer fitted with a magnetic stirring mechanism and a Peltier single cell holder at 25 °C. Quartz cuvettes of 1 cm pathlength and 3 mL capacity were used. Solutions containing the enzyme were buffered with 55.6 mM HEPES, 181 mM NaCl and 1.11 mM EDTA (I = 0.2). The initial NahG concentration was 1.0 μM, which in the titration of *apo* NahG with FAD consisted of 26 % *holo* form; determined by UV-Vis at 450 nm using ε = 11,300 M^{-1} cm^{-1} [21]. Addition of 1.5 μM FAD to this solution afforded the *holo* NahG (1.0 μM) used in the salicylate binding studies. Titrant aliquots with known concentrations were added to the enzyme solution, mixed and measured. Fluorescence data were corrected for dilution that was lower than 5 % of the initial volume. The excitation wavelength was 292 nm using the following parameters: excitation slit of 2.5 nm (salicylate binding) and 10 nm (holoenzyme formation); emission slit of 2.5 nm (salicylate binding) and 10 nm (holoenzyme formation); and averaging time of 0.1 s.

2.6. Product identification by NMR

Reactions were followed at 25 ºC by ^{1}H NMR spectroscopy on Bruker spectrometers at 400 MHz (256 transients) and 600 MHz (128 transients), equipped with room temperature probes. The experiments were performed using 600 μL of aqueous solutions at pH 8.2 containing 10% D_2O. The reaction mixture consisted of 56 mM Hepes, 150 μM FAD, 1.0 mM NADH, and an equivalent amount of NaCl to give an ionic strength of 0.222. The initial concentration of salicylate and 3-methylsalicylate were 200 μM and 800 μM, respectively. The reactions were initiated by addition of NahG in the same buffer as described above to give a final concentration of 0.32 μM.

2.7. Kinetics

Reactions were followed on a Cary50 spectrophotometer at 25.0 ºC by monitoring NADH consumption at 340 nm. For substrates (5-NH_2 and 5-CH_3O substituted salicylates) that absorbed appreciably at this wavelength, reactions were followed by fluorimetry using the NADH emission at 472 nm (excitation at 340 nm). The initial rate (v) was calculated from $v = \Delta[NADH]/\Delta t$. The NADH concentration was determined using a molar absorptivity of 6220 M^{-1} cm^{-1} in the spectrophotometric studies, or a standard curve for NADH in the fluorimetric studies. Quartz cuvettes of 1 cm pathlength were used in all studies. The reaction mixture consisted of 56 mM Buffer, 1.1 mM EDTA, and 171 mM NaCl to give an ionic strength of 0.222. The following buffers were used according to the indicated pH ranges: Bis-Tris (pH 5.3-7.3); Hepes (pH 6.7-8.5); Bicine (pH 7.4-9.4); Ches (pH 8.6-9.9); and Caps (pH 9.7-10.3). Reactions were initiated by addition of enzyme (10 μL) using an add-mixer device. The steady-state kinetic parameters (K_m and V_{max}) for substrates and cofactors were determined by non-linear regression of rate *vs* [Substrate] or [Cofactor] using the Michaelis-Menten equation. The k_{cat} value was calculated from $k_{cat} = V_{max}/[E]_o$, where $[E]_o$ is the total enzyme concentration. Studies in different $[O_2]$ were carried under argon. Samples were prepared from dilution of an O_2-saturated solution with a previously degassed solution obtained by three freezing-thaw cycles under vacuum. The concentration of O_2 in the saturated solution was determined by the Winkler method [48].

2.8. Hydrogen peroxide determination

Peroxide concentrations produced by the uncoupled path were quantified in triplicate using a modification of the Trinder method [49]. NADH oxidation (400 μM) was monitored by UV-Vis spectroscopy at 25 °C under aerobic conditions in the presence of 20 μM FAD, 0.3 μM NahG, and 120-600 μM salicylates. The pH and ionic strength were kept constant as described above for the kinetic experiments. After complete oxidation of NADH, the reaction mixture (1 mL) was

mixed with a solution containing horseradish peroxidase (700 U/L), 20 mM phenol, and 5 mM 4-aminoantipyrine in 50 mM sodium phosphate, pH 7.2, 0.02% (w/v) sodium azide, and 0.5% (w/v) BSA (stored at 4 °C). After incubation at 37 °C for 10 min, the absorbance was measured at 505 nm. A standard curve for H_2O_2, previously standardized with a standard $KMnO_4$ solution, was used to calculate the amount of H_2O_2 formed in the reactions catalyzed by NahG. Control experiments showed that the presence of salicylates does not affect the determination of H_2O_2. In addition, benzoate was used as pseudosubstrate to evaluate the efficiency of the H_2O_2 determination. We measured the formation of 100 ± 3% H_2O_2, in agreement with previous studies in which benzoate stimulates NADH oxidation in salicylate hydroxylase without substrate oxygenation [50, 51].

3. Results and Discussion

3.1. The overall structure of 6xHis-NahG

The 6xHis-NahG exists as a monomer in solution according to size exclusion chromatography (Figure S1) and dynamic light scattering assays (Figure S2), in agreement with literature observations for the native enzyme [52]. The crystal structure of 6xHis-NahG has been determined at 2.0 Å resolution using SAD phasing, and represents the first structure in the *apo* form for a flavin-dependent monooxygenase. Refinement statistics are summarized in Table 1.

The asymmetric unit contains two monomers of *apo* 6xHis-NahG, and the final model consists of residues 7-248 and 251-424 (native NahG numbering) (Figure 1). The first six residues of the native protein, together with the earlier 16 residues for the 6xHis-affinity tag and TEV protease cleavage site, were not visible in the electron density map and therefore were not included in the final model [53]. Two other regions, corresponding to residues 249-250 and 425-434, are also lacking in the final 6xHis-NahG *apo* model due to weak electron density. As expected, NahG

exhibits an overall α/β fold similar to that observed for other flavin-dependent monooxygenases [26, 54], which consists of three domains. The central part of the first domain includes the nucleotide-binding motif or Rossman fold. The main feature of the second domain is a large β-sheet comprised by mostly antiparallel β strands that defines a significant portion of the substrate binding cavity, while the C-terminal domain is an α-helical region arranged next to the first two domains. In some flavoprotein monooxygenases the C-terminal domain is involved in oligomerization [55-58], which seems not to be the case for NahG, given that the enzyme is a monomer in solution.

Although cells were disrupted in a lysis buffer with an excess of salicylate, no electron density for the aromatic substrate was found in the active site. Despite numerous attempts, NahG crystals could not be obtained from cocrystallization trials with salicylate, FAD, or NADH. Soaking experiments were equally unsuccessful with these ligands. A number of variables and issues influence ligand co-crystallization and soaking which can hinder formation of protein-ligand complex crystals [59].

3.2. Comparison of NahG with others flavoprotein monooxygenases

According to the DALI server [47], the *apo* NahG model displays good structural agreement with the salicylate hydroxylase from *P. putida* S-1 (SALH, PDB entry 5EVY) [10], in addition to the homologues urato oxidase from *Klebsiella pneumoniae* (HpxO, PDB entry 3RP7) [60], 3-hydroxybenzoate 6-hydroxylase from *Rhodococcus jostii* RHA1 (3HB6H, PDB entry 4BJZ) [23], 6-hydroxynicotinate 3-monooxygenase from *P. putida* KT2440 (NicC, PDB entry 5EOW) [22], and flavin-containing monooxygenase from *Pseudomonas aeruginosa* (PhzS, PDB entry 2RGJ) [55]. The root mean square deviation for these enzymes with respect to NahG is less than 3.2 Å for at least 332 Cα atoms with Z-scores of ≥ 32.4 (Figure S3).

A comparison of NahG with the structurally related enzymes provides insight into the active site architecture, including the FAD and substrate-binding sites, which are in equivalent positions among all the aligned structures. A remarkable exception is made for the loop comprised by the residues 38-53 in the *apo* 6xHis-NahG, which is found in a distinct closed position compared to FAD-bound structures as shown in Figure 1 for the superimposition with the *holo* 3-hydroxybenzoate 6-hydroxylase (3HB6H, PDB entry 4BJZ, green) from *Rhodococcus jostii* RHA1 [23].

The N-terminal GXGXXG fingerprint sequence (G14-X-G16-XX-G19), comprising the secondary structure elements β1 and α1, represents the common βαβ-fold of the dinucleotide-binding domain, known to interact with the ADP moiety of the FAD cofactor [61]. The second FAD binding fingerprint consists of the GD sequence (G313-D314), in which the highly conserved aspartate residue forms hydrogen bonds with the O-3 hydroxyl group of the ribityl chain of the flavin moiety, while the glycine accommodates the PO_4 group of a FAD [62]. Besides, a strictly conserved DG sequence motif (D162-G163) on the large turn (residues 161-166) that connects strand β11 and helix α5 is involved in binding the pyrophosphate moieties of both FAD and NADH [63].

The putative substrate-binding pocket, inferred from superimposition of NahG structure to that of the ligand-bound structures of 4-hydroxybenzoate hydroxylase (PHBH, PDB entry 1DOD) [64] and *P. putida* S-1 SALH (PDB entry 5EVY) [10], is clearly divided into hydrophobic and hydrophilic regions. The aromatic residues F85, W87, F230, F243, and W293 are positioned on the antiparallel β-sheet opposite from the FAD isoalloxazine ring and likely form a hydrophobic environment facing the salicylate (Figure 2). In addition, a number of nonpolar amino acid residues, including A190, M194, M219, L221, L228, and V241, also compose the substrate-

binding pocket, whereas the charged residues D224, H226, R247, H322, E381, and R383 are distinctly grouped near the active site, laterally positioned on the *re*-side of the isoalloxazine and leading to the solvent-accessible protein surface. Curiously, in the second closest relative, urato oxidase (HpxO, PDB entry 3RP7) [60], the charged residues R204, D293, and H348 (equivalent to H226, H332 and E381, respectively, in NahG) take part in an unusual hydrogen-bonding network which is mostly buried in a hydrophobic environment, while the catalytic residue R204 is proposed to deprotonate the substrate [60]. Indeed, a remarkable feature of the *para*-hydroxybenzoate hydroxylase (PHBH, PDB entry 1PBE) [65] from *Pseudomonas fluorescens*, the best known member of this family of enzymes, is a hydrogen-bonding network connecting the phenolic group of the substrate *p*-hydroxybenzoate (*p*OHB) in the buried active site to the surface of the protein, which lowers the p*K*a of *p*OHB and favors proton exchange between the substrate and the external solvent and regulates the coordination of conformational states for efficient catalysis [56]. A similar mechanism of substrate activation has been proposed for the structural homologue 3-hydroxybenzoate 6-hydroxylase (3HB6H; PDB entry 4BJZ) [23], in which mutagenesis of the basic residue H213 has been shown to impair substrate hydroxylation without preventing substrate binding and C4a-hydroperoxyflavin formation. The role of H213 has been suggested to facilitate the deprotonation of the substrate's phenol group during the electrophilic aromatic substitution [23, 66].

Several reports on PHBH have concluded that the internal proton-transfer network and substrate activation through phenolic ionization constitute a "password" or "signal", for molecular recognition, in which the enzyme tests for the presence of a suitable aromatic substrate before allowing the FAD's isoalloxazine ring to assume the "out" conformation, which is solvent accessible and required for reduction by NADPH [67-69]. Furthermore, regions of positive

electrostatic potential in the active site would situate the anionic reduced flavin to the "in" position adjacent to the substrate, in addition to previously promoting the formation of the substrate phenolate [70]. Interestingly, a hydrogen-bonding network is not found in the active site of kynurenine 3-monooxygenase (KMO), in which conformational changes and flavin reduction are triggered by π-π interactions between the substrate and the loop above the *re*-side of the isoalloxazine group [71].

It has been established that the rearrangement of FAD between the "out" and "in" positions are required in the catalytic mechanism comprised of reductive and oxidative reactions [57, 69]., which are shared among NahG and other flavin-dependent monooxygenases. Unlike all known structures enzymes in this class, there is no ligand in the structure of NahG. Interestingly, a notable structural feature is a large loop connecting strand β2 and helix α2, which exhibits a conformation markedly different from that in other homologs (Figure 3). This segment of 16 residues spanning from E38 to G53 stretches over the FAD-binding cleft, instead of being laterally positioned on the *si*-side of the isoalloxazine ring, as observed in the homologs. This loop is suggested to control the tunnel access that allows the substrate to enter the active site of the HpxO homologue [60]. A similar flexible loop in the phenol hydroxylase from *Trichosporon cutaneum* moves during catalysis to accommodate changes in FAD conformation, shielding the catalytic site during hydroxyl transfer or opening a channel for substrates and products to enter or leave the active site [57].

3.3. Stoichiometry and binding constants

Formation of *holo* NahG and its interaction with salicylate (Sal) followed tight-binding profiles as observed by fluorimetric titrations at pH 8.0 and 25 °C (Figure 4). In *holo* NahG formation, the excitation of *apo* NahG at 292 nm gives the characteristic emission for the

tryptophan residues at 340 nm (F_{obs} at 340nm), which is quenched by FAD binding to the enzyme (Figure 4a). An event solely caused by *holo* NahG formation, since FAD shows no substantial effect in the F_{obs} at 340nm in absence of enzyme and the self-inner filtering effect due to FAD absorption accounts for less than 2% of the fluorescence quenching as estimated according to Birdsall *et al*. [72].

The tight-binding profile in *holo* NahG formation (Figure 4b) fits to Equation 1, in which the *holo* NahG molar fraction (χ_{bound}) is calculated by Equation 2. The F_{obs}, F_{free}, F_{bound} represent the fluorescence intensities, respectively, for the observed, free and fully bound enzyme, and $[L]_o$ and $[E]_o$ are the total concentrations of the ligand and enzyme, respectively. The data provides an estimated dissociation constant (K_d) of 2 x 10^{-9} M for the *holo* NahG, which is in the same range reported for the salicylate hydroxylase from *P. putida* S-1 (K_d = 45 x 10^{-9} M) [73].[§] The inflection point occurs near to equivalent concentrations of FAD and *apo* NahG as expected for a 1:1 complex.

$$F_{obs} = F_{free} + (F_{bound} - F_{free})\chi_{bound} \quad \textbf{(1)}$$

$$\chi_{bound} = \frac{[E_{bound}]}{[E]_o} = \frac{([L]_o + [E]_o + K_d) - \sqrt{([L]_o + [E]_o + K_d)^2 - 4[L]_o[E]_o}}{2[E]_o} \quad \textbf{(2)}$$

The interaction of salicylate with *holo* NahG was followed by excitation at 292 nm and emission at 410 nm (F_{obs} at 410nm, Figure 4c). The emission is associated with intramolecular proton transfer in the excited state (ESIPT), which is common for *ortho*-substituted phenols with acidic groups [75]. The fluorescence intensity for salicylate follows a tight-binding profile for a 1:1 complex (Figure 4d). The K_d calculated by non-linear curve fitting with Equation 1 is 2.2 x

[§] The salicylate hydroxylase was purified from a soil gram-negative bacterium, later identified as *Pseudomonas putida* strain S-1 [74].

10^{-7} M for the salicylate-*holo* NahG complex, about 10 times lower than reported by Suzuki *et al*. (K_d = 3.5 x 10^{-6} M) for the corresponding substrate-enzyme complex using the salicylate hydroxylase from *P. putida* S-1 [73]. The fluorescence increase is consistent with salicylate binding to NahG, which hinders the substrate from interaction with the solvent and results in efficient radiative decay.

3.4. Identification of reaction products

We probe the regioselectivity of the oxidative decarboxylation of salicylate catalyzed by *holo* NahG using 3-methylsalicylate as substrate. Comparing the ^{1}H NMR spectra (Figure 5) before and after partial substrate conversion shows that 3-methylsalicylate signals disappear at 2.14 ppm (3H, singlet), 6.79 ppm (1H, triplet), 7.26 ppm (1H, doublet) and 7.59 ppm (1H, doublet) and new signals for 3-methycatechol appeared at 2.13 ppm (3H, singlet) and 6.67-6.70 ppm (3H, multiplet) in agreement to a pure sample. The conversion of salicylate to catechol catalyzed by NahG shows a similar profile (Figure S4). These results are strong evidences of *ipso* substitution in the salicylate reaction catalyzed by NahG, as the hydroxylation of 3-methylsalicylate in a carbon site different from C-1 would not afford 3-methycatechol. Similar observation is reported for another salicylate hydroxylase from *Pseudomonas putida* S-1, which also catalyzes the conversion of 3-methylsalicylate into 3-methylcatechol [76]. The *ipso* substitution is also supported by computational studies [20] and by early experimental studies [21], which elegantly showed that salicylate hydroxylase catalyzes the conversions of 2,3- and 2,6-dihydroxybenzoates reactions to the same product pyrogallol, while 2,4- and 2,5-dihydroxybenzoates are converted to different products. These observations do not support hydroxylation at the C-3 atom of salicylate followed by decarboxylation of a 2,3-dihydroxybenzoate intermediate as previously reported [10].

In general, aromatic hydroxylations by flavin-monooxygenases occur via electrophilic aromatic substitution (S_EAr). The *ipso* substitution is favored by phenol groups at *ortho* and *para* positions of the aromatic ring [77]. Salicylate undergoes decarboxylation in the presence of salicylate hydroxylase, but there are several other examples, including NahG itself [78-80], which the step following the hydroxylation involves deformylation, denitrification, dehalogenation, among others [26, 77]. In some cases, the hydroxylation is followed by deprotonation, for instance, 3-hydroxybenzoate is converted to 2,5- or 3,4-dihydroxybenzoate in the presence of 3-hydroxybenzoate 6-hydroxylase (3HB6H) [30, 66] or 3-hydroxybenzoate hydroxylase (MHBH) [58], respectively. On the other hand, 3,4-dihydroxybenzoate can also be obtained from 4-hydroxybenzoate in the presence of 4-hydroxybenzoate hydroxylase (PHBH) [64]. The reasons that control the regioselectivity in these enzymes are still not well understood, nonetheless, hydroxylation seems to be the key step [81]. Literature reports show that mutation of nonpolar residues at the active-site [82] and the presence of co-substrates [23] may change the orientation and proximity of the substrate to the C4a-hydroxyperoxyflavin group in the hydroxylation step.

3.5. Kinetic Studies

The Michaelis-Menten parameters shown in Table 2 were determined at pH 8.0 and 25 °C for salicylate, dioxygen (O_2), and NADH. The $v/[E]_o$ (s^{-1}) measured at different concentrations of salicylate and dioxygen were determined in the presence of 200 μM NADH, 84 % of saturation based on the K_m value for NADH (Figure S5). The same experiments under different concentrations of dioxygen and NADH were carried out in the presence of 192 μM salicylate, which represents 87 % of saturation regard the K_m value for salicylate. The FAD and dioxygen concentrations were kept under saturation conditions, except when varied. The kinetic data under

increasing concentrations of FAD fitted to a 1:1 binding model with a K_d value of 150±30 nM (Figure S5c). Therefore, k_{cat}^{ap} represents the product of k_{cat} and the molar fractions for salicylate, dioxygen, FAD, and NADH. Because the C(4a)-hydroperoxyflavin may generate H_2O_2 without causing substrate oxygenation (uncoupled path) [50, 51], the k_{cat} was also corrected to the amount of H_2O_2 (5.5 %) determined after complete NADH oxidation in the presence of excess substrate.

The k_{cat} and K_m values for salicylate and NADH catalyzed by NahG at pH 8.0 and 25 °C (Table 2) have the same magnitude reported for the salicylate hydroxylase from *P. putida* S-1, which the k_{cat} is about 21 s^{-1} and K_m values are 1.6-1.9 μM for salicylate and 2.6-8.2 μM for NADH [21, 74, 83]. Stopped-flow kinetic data for salicylate hydroxylase from *P. putida* S-1 indicated that salicylate binds prior NADH, and its oxidation to afford catechol is rate-limiting [74].

3.6. Linear Free Energy Relationships

Further details about the mechanism of catalysis were determined in the presence of 200 μM NADH (84 % saturation) and increasing concentrations of substituted salicylates (Figure S6 and S7). The k_{cat}^{ap} values were corrected for the uncoupled path, estimated from H_2O_2 determinations after complete NADH oxidation in excess of each substrate. We observed that k_{cat}^{ap} varies systematically for salicylates substituted with electron-donating and electron-withdrawing groups (Table 3). This observation is consistent with formation and cleavage of covalent bonds affected by the electronic demand in the substrate and inconsistent with a rate-limiting step controlled only by diffusion of substrate, products, or cofactors.

The kinetic parameters for the substituted salicylates were fitted to Hammett relationships (Figure 6) as described by equation 3. Interested readers may find background information in chemistry textbooks [84-87] and reviews [88, 89]. The $\sigma_{carboxy}$ (x-axis) corresponds to the values of

σ_{para} and σ_{meta} [88] for the respective substituents at positions 4 and 5 of salicylate, ϱ is the sensitivity factor as commonly employed in linear free energy relationships.

$$\log k_{cat}^{ap\,(X)}(or\ K_m^{(X)}) = \sigma\rho + \log k_{cat}^{ap\,(H)}(or\ K_m^{(H)}) \qquad \textbf{(3)}$$

Linear Hammett relationships provide information for rate-limiting steps that are affected by the electronic demand in the substrate. The presence of two intersecting straight lines indicates a change in the rate-limiting step. This effect is found for both K_m and k_{cat}^{ap}, near $\sigma_{carboxy}$ = zero. Linear behaviors are evident for both parameters at $\sigma_{carboxy} > 0$ and for K_m at $\sigma_{carboxy} < 0$. The k_{cat}^{ap} for 5-amino and 5-methoxy salicylates do not correlate with the others and are lower than expected. These deviations are discussed later in the text. Poor relationships are observed with σ_{phenol} (*i.e.*, σ_{meta} and σ_{para} of the respective substituents at positions 4 and 5 of salicylate, Figure S8).

The Hammett observations fit the mechanism proposed in Scheme 2, an aromatic electrophilic substitution (S_EAr) through a Wheland-type intermediate, in which either hydroxylation or decarboxylation is rate-limiting, depending on the electronic character of the aryl substituent. The S_EAr mechanism proposal is also supported by a recent theoretical study [20].

Analysis of the Hammett plot for k_{cat}^{ap} considers only the rate-limiting step. The hydroxylation step is facilitated by electron donating groups ($\sigma_{carboxy} < 0$) that increase the nucleophilicity of the aryl nucleophile, but when $\sigma_{carboxy} > 0$ the hydroxylation step is slower and rate-limiting. The converse is observed when $\sigma_{carboxy} < 0$, in which the decarboxylation is the rate-limiting step and faster towards electron withdrawing groups.

The log K_m values correlate well with $\sigma_{carboxy}$ (Figure 6). In the hydroxylation step, the less steep ϱ of -2.6 for log k_{cat}^{ap} compared to the ϱ of -3.5 for log K_m represents the additive effects of electron-withdrawing groups to the formation of the ES complex. Because $K_m = (k_{off} + k_{hyd})/k_{on}$ in the hydroxylation step, when the equilibrium is shifted toward ES, k_{off} decreases and k_{on} increases with

electron-withdrawing groups, resulting in a lower K_m. When the decarboxylation step is rate limiting, $K_m = (k_{off} + k_{dec}K_{hyd})/k_{on}$ where $K_{hyd} = k_{hyd}/k_{-hyd}$, in which case ϱ is less steep than in the hydroxylation step because k_{dec} decreases with electron-donating groups in contrast to the effect of electron-withdrawing groups in formation of the ES complex.

The driving-force for decarboxylation of the Wheland-type intermediate is provided by proximity to the hydrophobic pocket formed by the residues F85, L228, F230, V241 and F243 (Figure 2). These residues are situated at the *si*-face of salicylate, which, upon hydroxylation by the C4a-hydroperoxyflavin at the *re*-face, pushes the carboxyl group of the Wheland-type intermediate towards to the hydrophobic pocket. The resulting electrostatic destabilization of the Wheland-type intermediate favors the more neutral transition state leading to catechol and carbon dioxide formation (Scheme 2). It is worth mentioning that such assistance of a hydrophobic pocket for decarboxylation is a common feature of decarboxylases [90-93].

NahG activity requires the presence of a hydroxyl group *ortho* to the carboxyl group, as shown by the fact that NahG does not catalyze the decarboxylation of benzoate. The hydroxyl group allows the generation of the aryl nucleophile in the S_EAr step, and, as a carbonyl group in the second step, assists the oxidative decarboxylation of salicylate as an electronic relay providing charge dispersion in the transition state. Thus, substituents simultaneously influence the hydroxyl and carboxyl groups on the salicylate. Therefore, it is not surprising that the k_{cat}^{ap} values for 5-amino and 5-methoxy are lower than predicted by the linear fit at $\sigma_{carboxy} < 0$ (Figure 6). In these cases, the rate-limiting step is decarboxylation, in which the amino and methoxy groups at position 5 of salicylate contribute by resonance in the Wheland-type intermediate and increase the electron density at the carbon neighboring the site of decarboxylation. This slows down the

decarboxylation because the restoration of the resonance in the transition state is impaired in higher extent compared to groups that show only electron-donation by inductive effects.

The Jaffé relationship (Equation 4) [89, 94] separates the electronic effects over both groups in the salicylate and for each rate-limiting step. Scheme 3 explains the effect of the aryl substituent on each step. In the hydroxylation step, ϱ_X and ϱ_Y arise from groups *para*-positioned relative to the carboxyl and phenol groups, which activate these positions by resonance. On the other hand, activation by resonance in the decarboxylation step is provided by groups *meta*-positioned relative to the carboxyl and carbonyl sites, which give ϱ_Y and ϱ_X as sensitivity factors, respectively. It is important to note that the distance between the substituents and the reaction centers in the decarboxylation step is shorter than in the hydroxylation step, and therefore, the influence of substituents on the ϱ values is expected to be higher in the decarboxylation step.

$$\frac{1}{\sigma_{\text{carboxy}}}\log\frac{k_{\text{cat}}^{\text{ap (X)}}}{k_{\text{cat}}^{\text{ap (H)}}} = \frac{\sigma_{\text{phenol}}}{\sigma_{\text{carboxy}}}\rho_Y + \rho_X \qquad \textbf{(4)}$$

Figure 7 results from application of Equation 4 to the data in Table 3. Well-defined linear relationships were obtained for each rate-limiting step, sorted according to the electronic character of the salicylate substituent, as discussed above for the Hammett relationships. The hydroxylation or decarboxylation step becomes rate-limiting when salicylate is substituted with electron-withdrawing or electron-donating groups, respectively.

The ϱ_X = -2.0 in the hydroxylation step results from the effect of electron-donating groups that increase the nucleophilicity of the salicylate C-1 atom . This effect is smaller than observed for non-enzymatic S_EAr's, where ϱ values range from -3 to -12 [95, 96]. The ϱ_Y of 0.1 is consistent with minimal to no charge formation on C-2 or phenolic oxygen atoms in the transition state (Scheme 2), a situation satisfied by a concerted proton transfer from the phenol group to a nearby general-base. The most likely general base in NahG is H226, which is structurally equivalent to

H211 and H213 in 6-Hydroxynicotinate 3-Monooxygenase (NicC) [22] and 3-Hydroxybenzoate 6-Hydroxylase (3HB6H) [23, 66], respectively. In both enzymes, these residues have been proposed to promote substrate activation by the same mechanism. As observed for H213A 3HB6H mutant [23, 66], the H226A mutation in NahG also resulted in an inactive enzyme for NADH oxidation and catalysis.

The ϱ_Y = 1.9 in the decarboxylation step is consistent with the effect of electron-withdrawing groups that weaken the bond between the C-1 atom and the carboxyl group and stabilize the developing charge in the transition state. In contrast, the ϱ_X = -2.3 denotes the effect of electron-donating groups over the carbonyl position of the Wheland-type intermediate, which is converted to catechol by general-acid catalysis of the protonated H226 generated in the hydroxylation step (Scheme 2).

3.7. pH-rate Profile

The k_{cat}^{ap} and K_m values for the oxidative decarboxylation of salicylate catalyzed by NahG were determined from pH 3.8 to 11 in the presence of 200 μM NAD^+ at 25 °C (Table S1). The K_m values are lower under acidic conditions (Figure S9), an effect that is not caused by protonation of salicylate to give salicylic acid, which has pK_a values of 3.0 and 13.7 at 25 °C [97] and is found in the monoanionic state over the entire pH range of these kinetic experiments. The effect on K_m is ascribed to protonation of active site groups, most likely histidine residues (*viz*. H110, H226 or H322) as revealed by inspection of the x-ray structure.

The log k_{cat}^{ap} follows a bell-shaped profile as a function of pH (Figure 8) and is uncorrected for the contribution of the uncoupled path, which routine H_2O_2 measurements between pH 5-10 showed to account for less than 6 % of the overall NADH consumption. A reaction pathway that accounts for the pH effect on the oxidative decarboxylation of salicylate is shown in Scheme 4.

On the basic limb of the bell, log k_{cat} decreases with a slope of -1 as result of a single deprotonation of ES_{AH_1} to afford the catalytically inactive ES_A species. On the acidic limb, the log k_{cat}^{ap} decreases as result of acid-base equilibria for two active species of the enzyme. The ES_{AH_3} deprotonation provides the ES_{AH_2} species that converts the substrate to product by k_i, whereas, under slightly basic conditions, the ES_{AH_2} deprotonates to afford the ES_{AH_1} species, which is the most active species and catalyzes the reaction through k_{max}. These rate constants correlate with k_{cat}^{ap} according to Equation 5, which gives Equation 6 when the molar fractions of the catalytic species ES_{AH_1} and ES_{AH_2} are expressed in relation to K_{ES1}, K_{ES2}, and K_{ES3}. The logarithmic form of Equation 6 was used to fit the data in Figure 8, which afforded the kinetic and equilibrium parameters shown in Table 4.

$$k_{cat}^{ap} = k_i \chi_{ES_{AH2}} + k_{max} \chi_{ES_{AH1}} \quad \textbf{(5)}$$

$$k_{cat}^{ap} = \frac{k_i}{\frac{[H^+]}{K_{ES1}} + 1 + \frac{K_{ES2}}{[H^+]} + \frac{K_{ES2}K_{ES3}}{[H^+]^2}} + \frac{k_{max}}{\frac{[H^+]^2}{K_{ES1}K_{ES2}} + \frac{[H^+]}{K_{ES2}} + 1 + \frac{K_{ES3}}{[H^+]}} \quad \textbf{(6)}$$

We attribute the pK_{ES3} of 10.6 to the hydroxyl peroxide in the C4a-hydroperoxyflavin, which is the electrophile in the hydroxylation step. Deprotonation of the C4a-hydroperoxyflavin would render it catallytically inactive. This pK_a is in the range observed for other peroxides. Specifically, the pK_a for H_2O_2 is 11.6 [97], and free C4a-hydroperoxyflavin has a pK_a of about 9.2 [98]. The acidity of enzyme-bound C4a-hydroperoxyflavin may vary. A pK_a > 9.4 has been estimated for the reaction catalyzed by *p*-hydroxyphenylacetate 3-hydroxylase [99], while a pK_a = 8.4 has been reported for cyclohexanone monooxygenase [100].

The rate increase from pH 4 to 8 is consistent with the role of basic residues in catalysis. We attribute the pK_a of 6.6 to the conjugate acid of H226, general-base in hydroxylation step, and

directly related to the formation of most active species ES_{AH_1}. The identity of the residue responsible for the pK_a of 4.2 is still uncertain. Our best assumption is H110, which is near the NADH- and FAD-binding domains and close to the cationic side chain of R111 that may alter the acidity of H110 from the pK_a of 6.5 for free histidine. Mutagenesis studies are underway to test this hypothesis and will be reported in a future communication.

4. Conclusions

The crystal structure for the FAD-unbound NahG from *P. putida* G7 adds a new snapshot for the events that culminate with conversion of the *apo* to the peroxyflavin-bound protein. The loop of 16 residues from E38 to G53 stretches over the FAD-binding cleft. Its closed conformation in the *apo* form must change to an open conformation for FAD binding. In the *holo* protein, the E38-G53 loop forms a lid over FAD and contributes to strong interactions responsible for the tight binding of FAD. The formation of the peroxyflavin species is fast and highly favorable in NahG. Our results provide compelling evidence and new information on the catalytic groups involved in catalysis, the reactivity of reaction intermediates, and the reaction pathway. Catalysis by salicylate hydroxylase (NahG), and very likely by other flavin-dependent monooxygenases as well, occurs through a S_EAr mechanism via a Wheland-like intermediate. Intermediate formation is favored by electron-donating groups in the hydroxylation of salicylate by a neutral peroxo moiety in the peroxyflavin, whereas the decarboxylation of the Wheland-like intermediate is favored electron-withdrawing groups. An important feature in the active site of NahG is a hydrophobic region surrounding the salicylate's carboxyl group, which provides

additional driving force for decarboxylation of the Wheland-like intermediate. These steps are aided by H226, that, in proximity to the phenol group, acts as a general base in the hydroxylation step and, presumably, as a general acid in the decarboxylation step. This sequence of events does not require proton exchange between the substrate and the external solvent as previously proposed for *para*-hydroxybenzoate hydroxylase [56, 65].

Acknowledgment. We thank Dr. Masataka Tsuda and Dr. Masahiro Sota for kindly providing us the *P. putida* G7 strain, Laboratório Nacional de Luz Síncrotron (LNLS, Campinas, Brazil), Laboratório Nacional de Biociências (LNBio, Campinas, Brazil), INCT-Catalysis, VALE S.A., and the Brazilian Foundations CNPq, Capes and FAPEMIG for financial support.

Author Contributions. Conceptualization, R.A.P.N. and T.A.S.B.; Methodology, R.B.A., D.C.F., R.A.P.N., and T.A.S.B.; Investigation, D.M.A.C., S.V.G., and M.S.P.; Writing - Original Draft, D.M.A.C., S.S.A., R.A.P.N., and T.A.S.B.; Writing - Review & Editing, S.S.A., R.B.A., D.C.F., A.C.H., R.A.P.N., and T.A.S.B.; Funding Acquisition, R.A.P.N. and T.A.S.B.; Resources, D.C.F., R.A.P.N., and T.A.S.B.; Supervision, R.B.A., D.C.F., R.A.P.N., and T.A.S.B.

Supporting Information. Experimental procedures for cloning, expression and protein purification. Size exclusion chromatogram and dynamic light scattering for *apo* NahG. Sequence alignment of NahG with others flavin-dependent monooxygenases. NMR spectra for the oxidative decarboxylation of 3-methylsalicylate catalyzed by NahG. Kinetic data for the reaction of NahG-catalyzed oxidative decarboxylation of substituted salicylates.

References

[1] M. Katagiri, S. Yamamoto, O. Hayaishi, Flavin Adenine Dinucleotide Requirement for Enzymic Hydroxylation and Decarboxylation of Salicylic Acid, J. Biol. Chem. 237(7) (1962) PC2413-PC2414.

[2] R.H. Peng, A.S. Xiong, Y. Xue, X.Y. Fu, F. Gao, W. Zhao, Y.S. Tian, Q.H. Yao, Microbial biodegradation of polyaromatic hydrocarbons, FEMS Microbiol. Rev. 32(6) (2008) 927-955.

[3] K. Lee, J.W. Park, I.S. Ahn, Effect of additional carbon source on naphthalene biodegradation by *Pseudomonas putida* G7, J. Hazard. Mater. 105(1-3) (2003) 157-167.

[4] A.C. Grimm, C.S. Harwood, Chemotaxis of *Pseudomonas* spp. to the polyaromatic hydrocarbon naphthalene, Appl. Environ. Microbiol. 63(10) (1997) 4111-4115.

[5] K. Das, A.K. Mukherjee, Differential utilization of pyrene as the sole source of carbon by *Bacillus subtilis* and *Pseudomonas aeruginosa* strains: role of biosurfactants in enhancing bioavailability, J. Appl. Microbiol. 102(1) (2007) 195-203.

[6] I.-S. You, R.I. Murray, D. Jollic, I.C. Gunsalus, Purification and characterization of salicylate hydroxylase from *Pseudomonas putida* PpG7, Biochem. Biophys. Res. Commun. 169 (1990) 1049-1054.

[7] I.S. You, D. Ghosal, I.C. Gunsalus, Nucleotide sequence analysis of the *Pseudomonas putida* PpG7 salicylate hydroxylase gene (nahG) and its 3'-flanking region, Biochemistry 30(6) (1991) 1635-1641.

[8] M. Sota, H. Yano, A. Ono, R. Miyazaki, H. Ishii, H. Genka, E.M. Top, M. Tsuda, Genomic and functional analysis of the IncP-9 naphthalene-catabolic plasmid NAH7 and its transposon Tn4655 suggests catabolic gene spread by a tyrosine recombinase, J. Bacteriol. 188(11) (2006) 4057-4067.

[9] K. Suzuki, M. Mizuguchi, K. Ohnishi, E. Itagaki, Structure of chromosomal DNA coding for *Pseudomonas putida* S-1 salicylate hydroxylase, Biochim. Biophys. Acta 1275(3) (1996) 154-156.

[10] T. Uemura, A. Kita, Y. Watanabe, M. Adachi, R. Kuroki, Y. Morimoto, The catalytic mechanism of decarboxylative hydroxylation of salicylate hydroxylase revealed by crystal structure analysis at 2.5 Å resolution, Biochem. Biophys. Res. Commun. 469(2) (2016) 158-163.

[11] S. Yamamoto, M. Katagiri, H. Maeno, O. Hayaishi, Salicylate Hydroxylase, a Monooxygenase Requiring Flavin Adenine Dinucleotide. I. Purification and General Properties, J. Biol. Chem. 240 (1965) 3408-3413.

[12] J. Lee, J. Oh, K.R. Min, Y. Kim, Nucleotide sequence of salicylate hydroxylase gene and its 5'-flanking region of *Pseudomonas putida* KF715, Biochem. Biophys. Res. Commun. 218(2) (1996) 544-548.

[13] R. Bosch, E. Garcia-Valdes, E.R. Moore, Complete nucleotide sequence and evolutionary significance of a chromosomally encoded naphthalene-degradation lower pathway from *Pseudomonas stutzeri* AN10, Gene 245(1) (2000) 65-74.

[14] R.A. Rossello-Mora, J. Lalucat, E. Garcia-Valdes, Comparative biochemical and genetic analysis of naphthalene degradation among *Pseudomonas stutzeri* strains, Appl. Environ. Microbiol. 60(3) (1994) 966-72.

[15] W. Li, J. Shi, X. Wang, Y. Han, W. Tong, L. Ma, B. Liu, B. Cai, Complete nucleotide sequence and organization of the naphthalene catabolic plasmid pND6-1 from *Pseudomonas* sp. strain ND6, Gene 336(2) (2004) 231-240.

[16] J.J. Dennis, G.J. Zylstra, Complete sequence and genetic organization of pDTG1, the 83 kilobase naphthalene degradation plasmid from *Pseudomonas putida* strain NCIB 9816-4, J. Mol. Biol. 341(3) (2004) 753-768.

[17] R. Bosch, E.R. Moore, E. Garcia-Valdes, D.H. Pieper, NahW, a novel, inducible salicylate hydroxylase involved in mineralization of naphthalene by *Pseudomonas stutzeri* AN10, J. Bacteriol. 181(8) (1999) 2315-2322.

[18] M.P. Lanfranconi, J.A. Christie-Oleza, C. Martin-Cardona, L.Y. Suarez-Suarez, J. Lalucat, B. Nogales, R. Bosch, Physiological role of NahW, the additional salicylate hydroxylase found in *Pseudomonas stutzeri* AN10, FEMS Microbiol. Lett. 300(2) (2009) 265-272.

[19] H. Zhao, D. Chen, Y. Li, B. Cai, Overexpression, purification and characterization of a new salicylate hydroxylase from naphthalene-degrading *Pseudomonas* sp. strain ND6, Microbiol. Res. 160(3) (2005) 307-313.

[20] X. Wang, Q. Hou, Y. Liu, Insights into the decarboxylative hydroxylation of salicylate catalyzed by the Flavin-dependent monooxygenase salicylate hydroxylase, Theor. Chem. Acc. 137(7) (2018) 89.

[21] S. Yamamoto, M. Katagiri, H. Maeno, O. Hayaishi, Salicylate Hydroxylase, a Monooxygenase Requiring Flavin Adenine Dinucleotide: I. Purification and General Properties, J. Biol. Chem. 240(8) (1965) 3408-3413.

[22] K.A. Hicks, M.E. Yuen, W.F. Zhen, T.J. Gerwig, R.W. Story, M.C. Kopp, M.J. Snider, Structural and Biochemical Characterization of 6-Hydroxynicotinic Acid 3-Monooxygenase, A Novel Decarboxylative Hydroxylase Involved in Aerobic Nicotinate Degradation, Biochemistry 55(24) (2016) 3432-3446.

[23] S. Montersino, R. Orru, A. Barendregt, A.H. Westphal, E. van Duijn, A. Mattevi, W.J.H. van Berkel, Crystal Structure of 3-Hydroxybenzoate 6-Hydroxylase Uncovers Lipid-assisted Flavoprotein Strategy for Regioselective Aromatic Hydroxylation, J. Biol. Chem. 288(36) (2013) 26235-26245.

[24] D.E.T. Pazmino, H.M. Dudek, M.W. Fraaije, Baeyer-Villiger monooxygenases: recent advances and future challenges, Curr. Opin. Chem. Biol. 14(2) (2010) 138-144.

[25] D. Holtmann, M.W. Fraaije, I.W. Arends, D.J. Opperman, F. Hollmann, The taming of oxygen: biocatalytic oxyfunctionalisations, Chem. Comm. 50(87) (2014) 13180-13200.

[26] M.M. Huijbers, S. Montersino, A.H. Westphal, D. Tischler, W.J. van Berkel, Flavin dependent monooxygenases, Arch. Biochem. Biophys. 544 (2014) 2-17.

[27] E. Romero, J.R. Gómez Castellanos, G. Gadda, M.W. Fraaije, A. Mattevi, Same Substrate, Many Reactions: Oxygen Activation in Flavoenzymes, Chem. Rev. 118(4) (2018) 1742-1769.

[28] D.P. Ballou, B. Entsch, L.J. Cole, Dynamics involved in catalysis by single-component and two-component flavin-dependent aromatic hydroxylases, Biochem. Biophys. Res. Commun. 338(1) (2005) 590-598.

[29] J. Sucharitakul, P. Chaiyen, B. Entsch, D.P. Ballou, Kinetic mechanisms of the oxygenase from a two-component enzyme, p-hydroxyphenylacetate 3-hydroxylase from *Acinetobacter baumannii*, J. Biol. Chem. 281(25) (2006) 17044-17053.

[30] J. Sucharitakul, C. Tongsook, D. Pakotiprapha, W.J.H. van Berkel, P. Chaiyen, The Reaction Kinetics of 3-Hydroxybenzoate 6-Hydroxylase from *Rhodococcus jostii* RHA1 Provide an Understanding of the para-Hydroxylation Enzyme Catalytic Cycle, J. Biol. Chem. 288(49) (2013) 35210-35221.

[31] B. Entsch, D.P. Ballou, V. Massey, Flavin-Oxygen Derivatives Involved in Hydroxylation by p-Hydroxybenzoate Hydroxylase, J. Biol. Chem. 251(9) (1976) 2550-2563.

[32] K. Maedayorita, V. Massey, On the Reaction-Mechanism of Phenol Hydroxylase - New Information Obtained by Correlation of Fluorescence and Absorbency Stopped Flow Studies, J. Biol. Chem. 268(6) (1993) 4134-4144.

[33] C. Kemal, T.W. Chan, T.C. Bruice, Reaction of $^3O_2$ with dihydroflavins. 1. $N^{3,5}$-Dimethyl-1,5-dihydrolumiflavin and 1,5-dihydroisoalloxazines, J. Am. Chem. Soc. 99(22) (1977) 7272-7286.

[34] Y.H. Zhang, J.Q. Yu, Pd(II)-catalyzed hydroxylation of arenes with 1 atm of O_2 or air, J. Am. Chem. Soc. 131(41) (2009) 14654-14655.

[35] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server, in: J.M. Walker (Ed.), The Proteomics Protocols Handbook, Humana Press 2005, pp. 571-607.

[36] J.R. Luft, R.J. Collins, N.A. Fehrman, A.M. Lauricella, C.K. Veatch, G.T. DeTitta, A deliberate approach to screening for initial crystallization conditions of biological macromolecules, J. Struct. Biol. 142(1) (2003) 170-179.

[37] Z. Dauter, M. Dauter, K.R. Rajashankar, Novel approach to phasing proteins: derivatization by short cryo-soaking with halides, Acta Crystallogr., Sect. D: Biol. Crystallogr. 56(Pt 2) (2000) 232-237.

[38] R.A. Nagem, Z. Dauter, I. Polikarpov, Protein crystal structure solution by fast incorporation of negatively and positively charged anomalous scatterers, Acta Crystallogr., Sect. D: Biol. Crystallogr. 57(Pt 7) (2001) 996-1002.

[39] R.A. Nagem, I. Polikarpov, Z. Dauter, Phasing on rapidly soaked ions, Methods Enzymol. 374 (2003) 120-137.

[40] B.G. Guimaraes, L. Sanfelici, R.T. Neuenschwander, F. Rodrigues, W.C. Grizolli, M.A. Raulik, J.R. Piton, B.C. Meyer, A.S. Nascimento, I. Polikarpov, The MX2 macromolecular crystallography beamline: a wiggler X-ray source at the LNLS, J. Synchrotron Radiat. 16(Pt 1) (2009) 69-75.

[41] Z. Otwinowski, W. Minor, Processing of X-ray Diffraction Data Collected in Oscillation Mode, in: C.W. Carter, R.M. Sweet (Eds.), Methods Enzymol., Academic Press, New York, 1997, pp. 307 - 326.

[42] G.M. Sheldrick, Experimental phasing with SHELXC/D/E: combining chain tracing with density modification, Acta Crystallogr., Sect. D: Biol. Crystallogr. 66(Pt 4) (2010) 479-485.

[43] T.C. Terwilliger, P.D. Adams, R.J. Read, A.J. McCoy, N.W. Moriarty, R.W. Grosse-Kunstleve, P.V. Afonine, P.H. Zwart, L.W. Hung, Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard, Acta Crystallogr., Sect. D: Biol. Crystallogr. 65(Pt 6) (2009) 582-601.

[44] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr., Sect. D: Biol. Crystallogr. 66(Pt 2) (2010) 213-221.

[45] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallogr., Sect. D: Biol. Crystallogr. 66(Pt 4) (2010) 486-501.

[46] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera - a visualization system for exploratory research and analysis, J. Comput. Chem. 25(13) (2004) 1605-1612.

[47] L. Holm, P. Rosenstrom, Dali server: conservation mapping in 3D, Nucleic Acids Res. 38(Web Server issue) (2010) W545-W549.

[48] H.A.C. Montgomery, N.S. Thom, A. Cockburn, Determination of dissolved oxygen by the winkler method and the solubility of oxygen in pure water and sea water, J. Appl. Chem. 14(7) (1964) 280-296.

[49] P. Trinder, Determination of glucose in blood using glucose oxidase with an alternative oxygen acceptor, Ann. Clin. Biochem. 6(1) (1969) 24-27.

[50] R.H. White-Stevens, H. Kamin, Uncoupling of oxygen activation from hydroxylation in a bacterial salicylate hydroxylase, Biochem. Biophys. Res. Commun. 38(5) (1970) 882-889.

[51] R.H. White-Stevens, H. Kamin, Studies of a flavoprotein, salicylate hydroxylase. I. Preparation, properties, and the uncoupling of oxygen reduction from hydroxylation, J. Biol. Chem. 247(8) (1972) 2358-2370.

[52] I.-S. You, R.I. Murray, D. Jollic, I.C. Gunsalus, Purification and characterization of salicylate hydroxylase from *Pseudomonas putida* PpG7, Biochem. Biophys. Res. Commun. 169 (1990) 1049-1054.

[53] A. Wlodawer, W. Minor, Z. Dauter, M. Jaskolski, Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures, FEBS J. 275(1) (2008) 1-21.

[54] W.J.H. van Berkel, N.M. Kamerbeek, M.W. Fraaije, Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts, J. Biotechnol. 124(4) (2006) 670-689.

[55] B.T. Greenhagen, K. Shi, H. Robinson, S. Gamage, A.K. Bera, J.E. Ladner, J.F. Parsons, Crystal structure of the pyocyanin biosynthetic protein PhzS, Biochemistry 47(19) (2008) 5281-5289.

[56] B. Entsch, L.J. Cole, D.P. Ballou, Protein dynamics and electrostatics in the function of p-hydroxybenzoate hydroxylase, Arch. Biochem. Biophys. 433(1) (2005) 297-311.

[57] C. Enroth, H. Neujahr, G. Schneider, Y. Lindqvist, The crystal structure of phenol hydroxylase in complex with FAD and phenol provides evidence for a concerted conformational change in the enzyme and its cofactor during catalysis, Structure 6(5) (1998) 605-617.

[58] T. Hiromoto, S. Fujiwara, K. Hosokawa, H. Yamaguchi, Crystal structure of 3-hydroxybenzoate hydroxylase from *Comamonas testosteroni* has a large tunnel for substrate and oxygen access to the active site, J. Mol. Biol. 364(5) (2006) 878-896.

[59] I. Muller, Guidelines for the successful generation of protein-ligand complex crystals, Acta Crystallogr., Sect. D: Biol. Crystallogr. 73(Pt 2) (2017) 79-92.

[60] K.A. Hicks, S.E. O'Leary, T.P. Begley, S.E. Ealick, Structural and Mechanistic Studies of HpxO, a Novel Flavin Adenine Dinucleotide-Dependent Urate Oxidase from *Klebsiella pneumoniae*, Biochemistry 52(3) (2013) 477-487.

[61] R.K. Wierenga, P. Terpstra, W.G.J. Hol, Prediction of the Ocurrence of the ADP-binding βαβ-fold in Proteins, Using an Amino Acid Sequence Fingerprint, J. Mol. Biol. 187 (1986) 101-107.

[62] G. Eggink, H. Engel, G. Vriend, P. Terpstra, B. Witholt, Rubredoxin Reductase of *Pseudomonas oleovorans*. Structural Relationship to Other Flavoprotein Oxidoreductases Based on One NAD and Two FAD fingerprints, J. Mol. Biol. 212 (1990) 135-142.

[63] M.H.M. Eppink, H.A. Schreuder, W.J.H. Van Berkel, Identification of a novel conserved sequence motif in flavoprotein hydroxylases with a putative dual function in FAD/NAD(P)H binding., Protein Sci. 6 (1997) 2454-2458.

[64] D.L. Gatti, B.A. Palfey, M.S. Lah, B. Entsch, V. Massey, D.P. Ballou, M.L. Ludwig, The mobile flavin of 4-OH benzoate hydroxylase, Science 266(5182) (1994) 110-114.

[65] H.A. Schreuder, P.A. Prick, R.K. Wierenga, G. Vriend, K.S. Wilson, W.G. Hol, J. Drenth, Crystal structure of the p-hydroxybenzoate hydroxylase-substrate complex refined at 1.9 A resolution. Analysis of the enzyme-substrate and enzyme-product complexes, J. Mol. Biol. 208(4) (1989) 679-696.

[66] J. Sucharitakul, D. Medhanavyn, D. Pakotiprapha, W.J. van Berkel, P. Chaiyen, Tyr217 and His213 are important for substrate binding and hydroxylation of 3-hydroxybenzoate 6-hydroxylase from *Rhodococcus jostii* RHA1, FEBS J. 283(5) (2016) 860-81.

[67] B.A. Palfey, G.R. Moran, B. Entsch, D.P. Ballou, V. Massey, Substrate recognition by "password" in p-hydroxybenzoate hydroxylase, Biochemistry 38(4) (1999) 1153-1158.

[68] K.K. Frederick, B.A. Palfey, Kinetics of proton-linked flavin conformational changes in p-hydroxybenzoate hydroxylase, Biochemistry 44(40) (2005) 13304-13314.

[69] B.A. Palfey, C.A. McDonald, Control of catalysis in flavin-dependent monooxygenases, Arch. Biochem. Biophys. 493(1) (2010) 26-36.

[70] G.R. Moran, B. Entsch, B.A. Palfey, D.P. Ballou, Electrostatic effects on substrate activation in para-hydroxybenzoate hydroxylase: studies of the mutant lysine 297 methionine, Biochemistry 36(24) (1997) 7548-7556.

[71] H.T. Kim, B.K. Na, J. Chung, S. Kim, S.K. Kwon, H. Cha, J. Son, J.M. Cho, K.Y. Hwang, Structural Basis for Inhibitor-Induced Hydrogen Peroxide Production by Kynurenine 3-Monooxygenase, Cell Chem. Biol. 25 (2018) 1-13.

[72] B. Birdsall, R.W. King, M.R. Wheeler, C.A. Lewis, Jr., S.R. Goode, R.B. Dunlap, G.C. Roberts, Correction for light absorption in fluorescence studies of protein-ligand interactions, Anal. Biochem. 132(2) (1983) 353-361.

[73] K. Suzuki, S. Takemori, M. Katagiri, Mechanism of the salicylate hydroxylase reaction. IV. Fluorometric analysis of the complex formation, Biochim. Biophys. Acta 191(1) (1969) 77-85.

[74] S. Takemori, M. Nakamura, K. Suzuki, M. Katagiri, T. Nakamura, Mechanism of the salicylate hydroxylase reaction. V. Kinetic analyses, Biochim. Biophys. Acta 284(2) (1972) 382-393.

[75] P.J. Kovi, C.L. Miller, S.G. Schulman, Biprotonic versus intramolecular phototautomerism of salicylic acid and some of its methylated derivatives in the lowest excited singlet state, Anal. Chim. Acta 61(1) (1972) 7-13.

[76] S. Takemori, H. Yasuda, K. Mihara, K. Suzuki, M. Katagiri, Mechanism of the salicylate hydroxylase reaction. 3. Characterization and reactivity of chemically or photochemically reduced enzyme-flavin, Biochim. Biophys. Acta 191(1) (1969) 69-76.

[77] B. Ricken, B.A. Kolvenbach, P.F. Corvini, Ipso-substitution - the hidden gate to xenobiotic degradation pathways, Curr. Opin. Biotechnol. 33 (2015) 220-7.

[78] K. Suzuki, M. Katagiri, Mechanism of salicylate hydroxylase-catalyzed decarboxylation, Biochim. Biophys. Acta, Enzymol. 657(2) (1981) 530-534.

[79] K. Suzuki, T. Gomi, E. Itagaki, Intermediate and mechanism of hydroxylation of o-iodophenol by salicylate hydroxylase, J. Biochem. 109(5) (1991) 791-797.

[80] K. Suzuki, T. Gomi, T. Kaidoh, E. Itagaki, Hydroxylation of o-halogenophenol and o-nitrophenol by salicylate hydroxylase, J. Biochem. 109(2) (1991) 348-353.

[81] A. Alfieri, F. Fersini, N. Ruangchan, M. Prongjit, P. Chaiyen, A. Mattevi, Structure of the monooxygenase component of a two-component flavoprotein monooxygenase, Proc. Natl. Acad. Sci. U.S.A. 104(4) (2007) 1177-82.

[82] A. Bregman-Cohen, B. Deri, S. Maimon, Y. Pazy, A. Fishman, Altering 2-Hydroxybiphenyl 3-Monooxygenase Regioselectivity by Protein Engineering for the Production of a New Antioxidant, ChemBioChem 19(6) (2018) 583-590.

[83] K. Suzuki, K. Ohnishi, Functional modification of an arginine residue on salicylate hydroxylase, Biochim. Biophys. Acta 1040(3) (1990) 327-36.

[84] A. Williams, Free energy relationships in organic and bio-organic chemistry, RSC, Cambridge, 2003.

[85] H. Maskill, Structure and reactivity in organic chemistry, Oxford University Press, Oxford, 1999.

[86] J. Clayden, N. Greeves, S.G. Warren, Organic chemistry, 2nd ed., Oxford University Press, Oxford, 2012.

[87] E.V. Anslyn, D.A. Dougherty, Modern physical organic chemistry, University Science, Sausalito, CA, 2006.

[88] C. Hansch, A. Leo, R.W. Taft, A survey of Hammett substituent constants and resonance and field parameters, Chem. Rev. 91 (1991) 165-195.

[89] H.H. Jaffé, A reexamination of the Hammett equation, Chem. Rev. 53(2) (1953) 191-261.

[90] M.C. Ho, J.F. Menetret, H. Tsuruta, K.N. Allen, The origin of the electrostatic perturbation in acetoacetate decarboxylase, Nature 459(7245) (2009) 393-397.

[91] T. Bock, E. Luxenburger, J. Hoffmann, V. Schutza, C. Feiler, R. Muller, W. Blankenfeldt, AibA/AibB Induces an Intramolecular Decarboxylation in Isovalerate Biosynthesis by *Myxococcus xanthus*, Angew. Chem. Int. Ed. 56(33) (2017) 9986-9989.

[92] S.L. Guimaraes, J.B. Coitinho, D.M. Costa, S.S. Araujo, C.P. Whitman, R.A. Nagem, Crystal Structures of Apo and Liganded 4-Oxalocrotonate Decarboxylase Uncover a Structural Basis for the Metal-Assisted Decarboxylation of a Vinylogous beta-Keto Acid, Biochemistry 55(18) (2016) 2632-2645.

[93] L.K. Jackson, H.B. Brooks, D.P. Myers, M.A. Phillips, Ornithine decarboxylase promotes catalysis by binding the carboxylate in a buried pocket containing phenylalanine 397, Biochemistry 42(10) (2003) 2933-2940.

[94] A.R. Fersht, A.J. Kirby, Structure and Mechanism in Intramolecular Catalysis. Hydrolysis of Substituted Aspirins, J. Am. Chem. Soc. 89(19) (1967) 4853-4857.

[95] H.C. Brown, Y. Okamoto, Directive Effects in Aromatic Substitution. 30. Electrophilic Substituent Constants, J. Am. Chem. Soc. 80(18) (1958) 4979-4987.

[96] Y. Yukawa, Y. Tsuno, Resonance Effect in Hammett Relationship. 3. The Modified Hammett Relationship for Electrophilic Reactions, Bull. Chem. Soc. Jpn. 32(9) (1959) 971-981.

[97] A.E. Martell, Z.M. Smith, R.J. Motekaitis, NIST Critical Stability Constants of Metal Complexes Database Version 8 (for Windows): NIST Standard Reference Database 46, NIST, Gaithersburg, 2004.

[98] G. Eberlein, T.C. Bruice, The chemistry of a 1,5-diblocked flavin. 2. Proton and electron transfer steps in the reaction of dihydroflavins with oxygen, J. Am. Chem. Soc. 105(22) (1983) 6685-6697.

[99] K. Thotsaporn, P. Chenprakhon, J. Sucharitakul, A. Mattevi, P. Chaiyen, Stabilization of C4a-Hydroperoxyflavin in a Two-component Flavin-dependent Monooxygenase Is Achieved through Interactions at Flavin N5 and C4a Atoms, J. Biol. Chem. 286(32) (2011) 28170-28180.

[100] D. Sheng, D.P. Ballou, V. Massey, Mechanistic Studies of Cyclohexanone Monooxygenase: Chemical Properties of Intermediates Involved in Catalysis, Biochemistry 40(37) (2001) 11156-11167.

[101] F.R. Carneiro, T.C. Silva, A.C. Alves, T. Haline-Vaz, F.C. Gozzo, N.I. Zanchin, Spectroscopic characterization of the tumor antigen NY-REN-21 and identification of heterodimer formation with SCAND1, Biochem. Biophys. Res. Commun. 343(1) (2006) 260-268.

[102] D.M.A. Costa, M.A.F. Costa, S.L. Guimarães, J.B. Coitinho, S.V. Gómez, T.A.S. Brandão, R.A.P. Nagem, A combined approach for enhancing the stability of recombinant cis-dihydrodiol naphthalene dehydrogenase from *Pseudomonas putida* G7 allowed for the structural and kinetic characterization of the enzyme, Protein Expression Purif. 132 (2017) 50-59.

Appendix A. Supplementary data

Catalytic mechanism for the conversion of salicylate into catechol by the flavin-dependent monooxygenase salicylate hydroxylase

Cloning of salicylate hydroxylase *nahG* gene from *P. putida* G7

The coding sequence for salicylate hydroxylase NahG (NCBI Gene ID: 3974216) was amplified from *P. putida* G7 NAH7 plasmid (NCBI Accession No. AB237655) using standard PCR conditions with forward 5'**CAT ATG** AAA AAC AAT AAA CTT GGC TTG CG3' and reverse 5'**GAA TTC** TCA CCC TTG ACG TAG CG 3' primers. The *Nde*I and *Bam*HI restriction enzyme cloning sites are shown in bold, respectively. The PCR product was purified, cloned into a pGEM-T vector (Promega), and subcloned into pET28a(TEV) expression vector [101]. The PCR-derived DNA was sequenced and verified to contain no errors. *Escherichia coli* strain DH5α was used as a recipient for transformations during plasmid construction and for plasmid propagation and storage. Plasmid DNA extraction from the cells was done using the AxyPrep Plasmid Miniprep Kit from Axygen Biosciences and the purification of PCR products or DNA fragments from agarose gel slice was made by Ilustra GFX PCR DNA and Gel Band Purification Kit from GE Healthcare Life Sciences

Protein expression and purification

E. coli Rosetta (DE3) competent cells (Novagen) were transformed with the recombinant expression vector pET28a(TEV)-*nahg*. A single colony from an agar plate was used to inoculate 20 mL of LB medium supplemented with 50 µg/mL kanamycin and 34 µg/mL chloramphenicol. After overnight growth at 37 ºC with agitation at 180 rpm, this primary culture was diluted 1:50 in 1 L LB medium supplemented with the same antibiotics. When cells reached an OD_{600} of 0.6 - 0.7, the temperature was reduced to 18 ºC, and 6xHis-NahG expression was induced by addition of 0.5 mM isopropyl β-D-thiogalactopyranoside. After 16 h with continuous shaking at 18 ºC,

cells were harvested by centrifugation at 4 °C for 10 min at 5,000 *g*, and the pellet was stored at -20 ºC.

Purification of the enzyme was done according to a combined approach used to stabilize another enzyme from the same naphthalene pathway [102]. Prior to cell lysis, the frozen pellet was resuspended in 30 mL of lysis buffer (50 mM Tris-HCl buffer at pH 7.4, 1% w/v sucrose, 1% v/v Tween 20, 1% v/v glycerol). Approximately 50 mg of salicylic acid was added to the resuspended cells, which were disrupted using an EmulsiFlex-C3 homogenizer (Avestin). The lysate was clarified by centrifugation at 10,000 *g* for 30 min at 4 ºC. The soluble fraction was immediately loaded onto a 5 mL Ni^{2+} affinity column HisTrap HP (GE Healthcare) connected to an ÄKTAprime chromatography system (GE Healthcare). Both the column and system were previously equilibrated with buffer A (50 mM Tris-HCl buffer at pH 7.4, 500 mM NaCl, 30 mM imidazole). After sample loading, the column was washed with buffer A to remove nonspecific binding proteins, and the target protein was eluted using an increasing imidazole gradient obtained with buffer B (50 mM Tris-HCl buffer at pH 7.4, 500 mM NaCl, 500 mM imidazole). Fractions containing the recombinant protein were pooled and buffer exchanged into buffer C (50 mM Tris-HCl buffer at pH 7.4, 50 mM NaCl) using a HiPrep 26/10 desalting column (GE Healthcare). The imidazole-free sample was then loaded onto a HiLoad 16/60 Superdex 200 (GE Healthcare) size-exclusion chromatography column equilibrated with buffer C for a final purification step. For binding and kinetic studies, samples were incubated with recombinant histidine-tagged TEV protease (5% w/w) at 30 ºC for 16 h for 6xHis-tag removal. A second Ni^{2+} affinity chromatography step was then used to further separate tag-free NahG from contaminants – including histidine-tagged TEV protease – prior to a final size-exclusion chromatography step. Whenever required, the resulting NahG samples were concentrated by ultrafiltration using a

Vivaspin sample concentrator (10 kDa MWC, GE Healthcare). Protein sample purity was checked by Coomassie blue stained 12% SDS-PAGE at all steps.

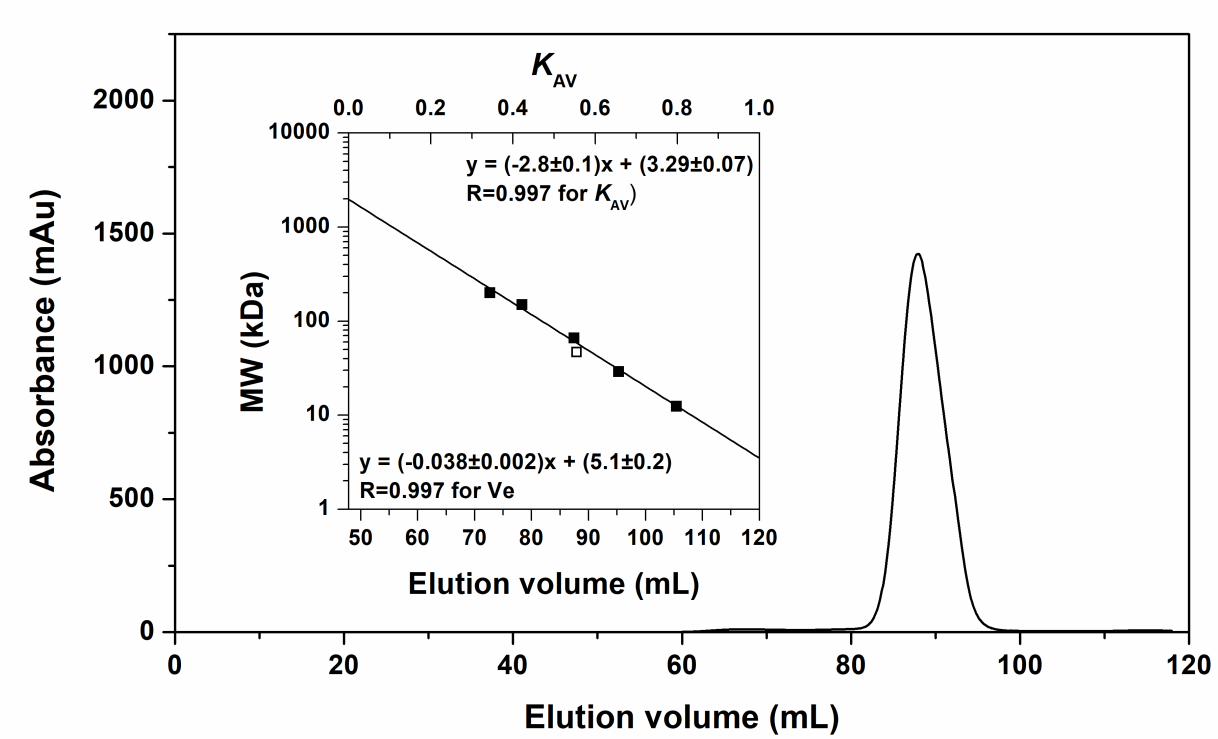


Fig. S1. Size exclusion chromatogram for apo NahG using the Hiload™ 16/60 Superdex 200 column (GE Healthcare) with a flow rate of 1.0 mL/min at 25 ºC. Inset: Column calibration curve for MW *versus* Ve or K_{AV}. The open square is the data point for apo NahG (47 kDa). Proteins were eluted isocratically at pH 7.4 with 50 mM Tris-HCl containing 50 mM NaCl. K_{AV} refers to the partition coefficient given by the equation [(Ve-Vo)/(Vc-Vo)], where Ve corresponds to the elution volume for each protein, and Vc and Vo to the respective total (120 mL) and void (47.9 mL using Blue Dextran) volumes for the column. The molecular weight standards are β-amylase (200 kDa), alcohol dehydrogenase (150 kDa), albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome c (12.4 kDa).

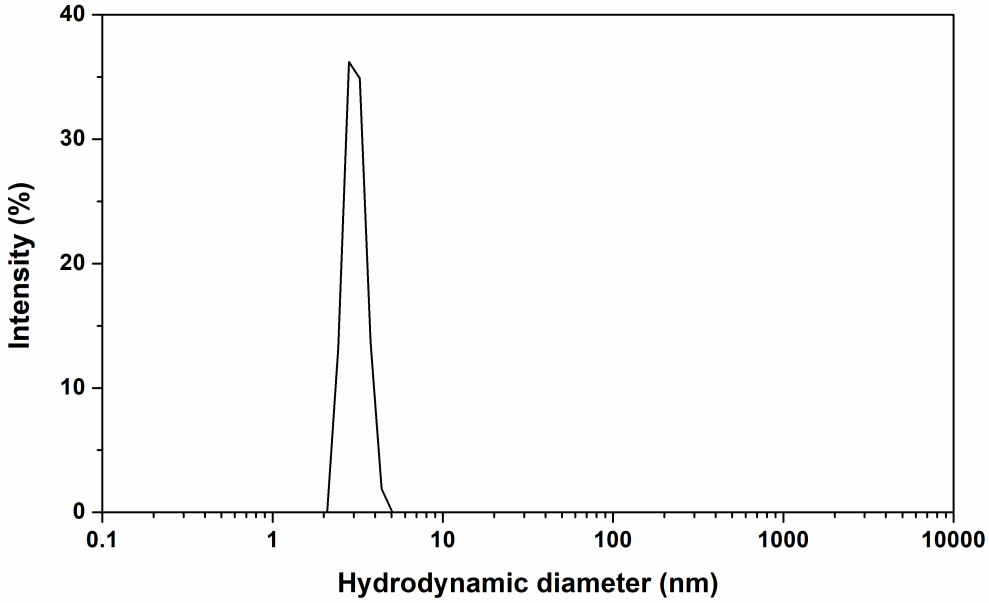


Fig. S2. Dynamic light scattering for apo NahG (1.5 mg/mL) at pH 7.4 with 50 mM Tris-HCl containing 50 mM NaCl. The hydrodynamic diameter for this particle in solution is 3.1 ± 0.3 nm and its estimated molecular weight is 49 ± 5 kDa, consistent with a NahG monomer.

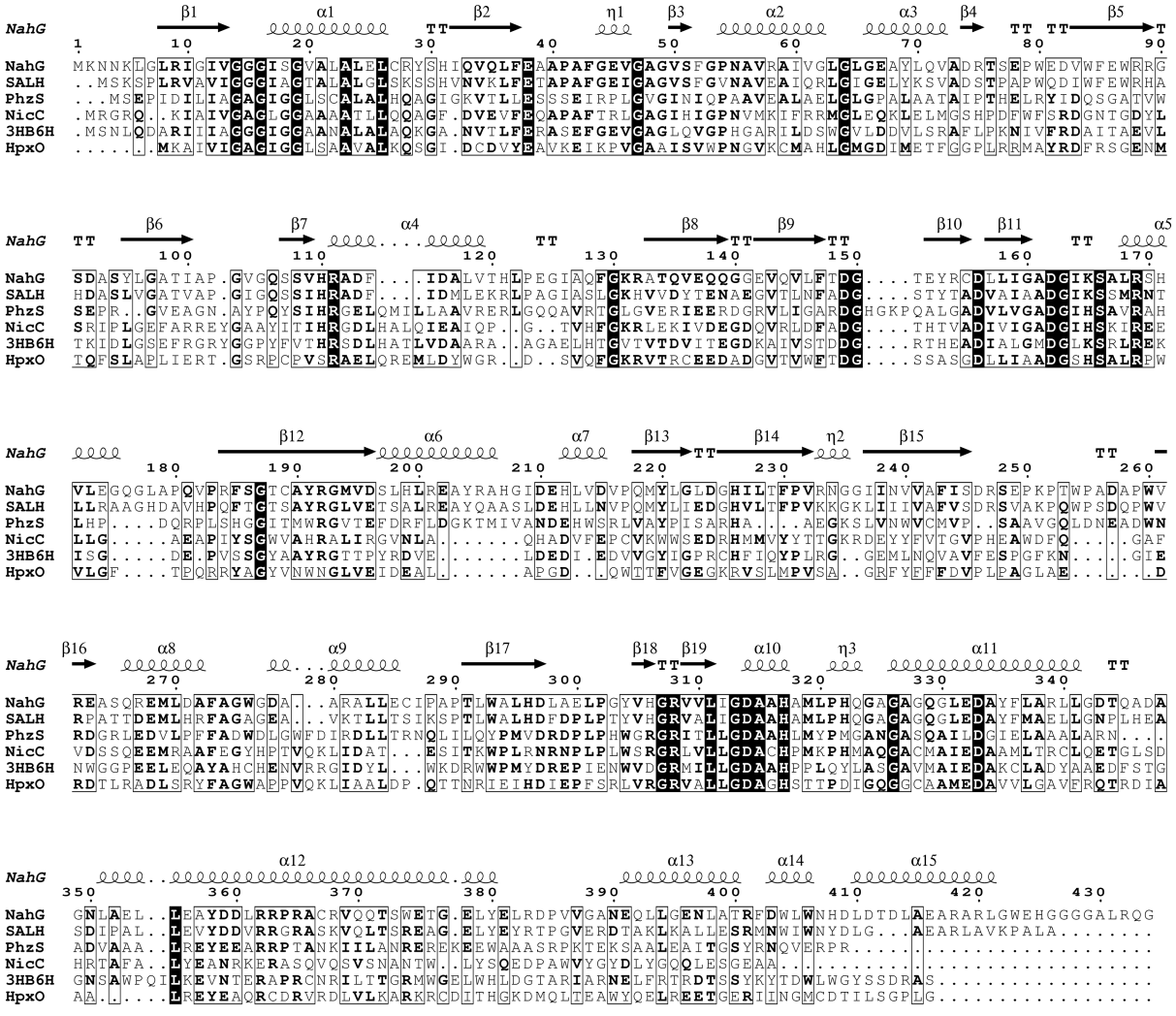


Fig. S3. Sequence alignment of NahG with other flavin-dependent monooxygenases. The five closest structural homologs to NahG are: Salicylate Hydroxylase (SALH); Urato Oxidase (HpxO); 3-Hydroxybenzoate 6-Hydroxylase (3HB6H); 6-Hydroxynicotinate 3-Monooxygenase (NicC); and Flavin-containing Monooxygenase (PhzS). The rmsd for each of these enzymes with respect to NahG is < 3.2Å for at least 332 aligned residues with Z scores ≥ 32.4. NahG secondary structure elements (helices and β-strands) are labeled, numbered and depicted in the first line of the figure. Identical residues are marked with black boxes while similar ones are grouped in white boxes.

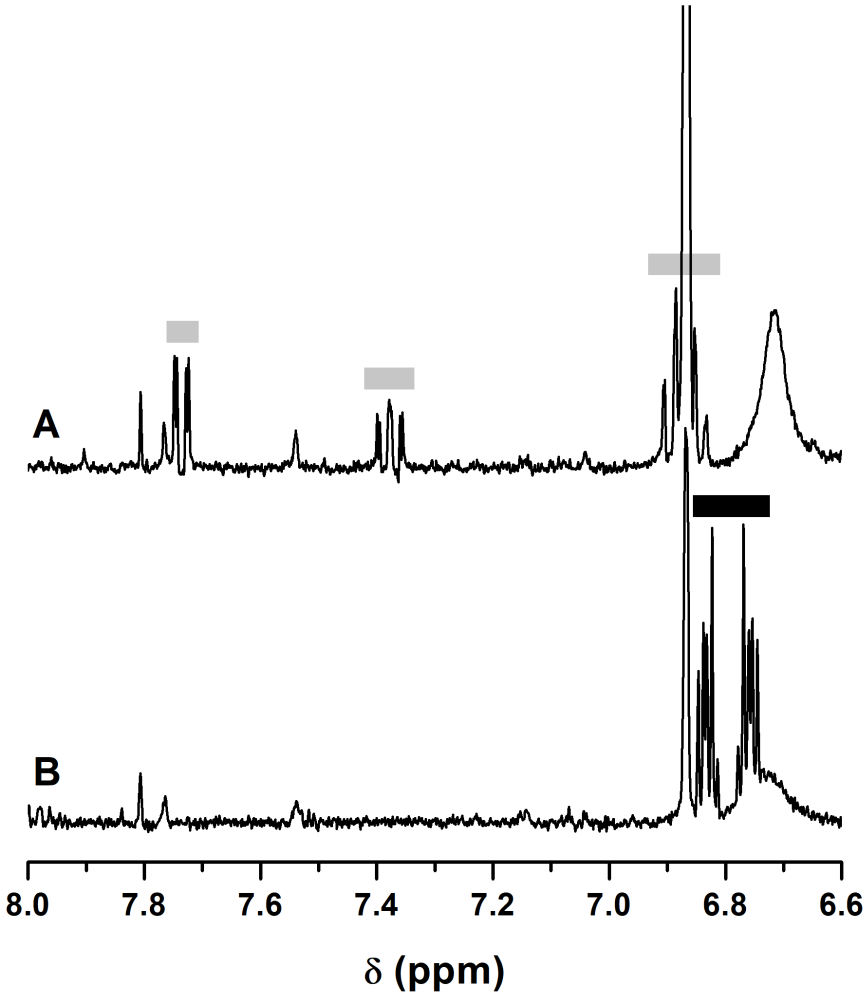


Fig. S4. NMR spectra at 400 MHz before (A) and after (B) the oxidative decarboxylation of salicylate catalyzed by NahG in the presence of NADH at 25 ºC and pH 8.2. The gray and black boxes indicate the ^{1}H signals for salicylate and catechol, respectively. In spectrum A the ^{1}H NMR signals for salicylate at 6.83-6.90 ppm (2H, multiplet) are superimposed with NADH nicotinamide's signal at 6.87 ppm.

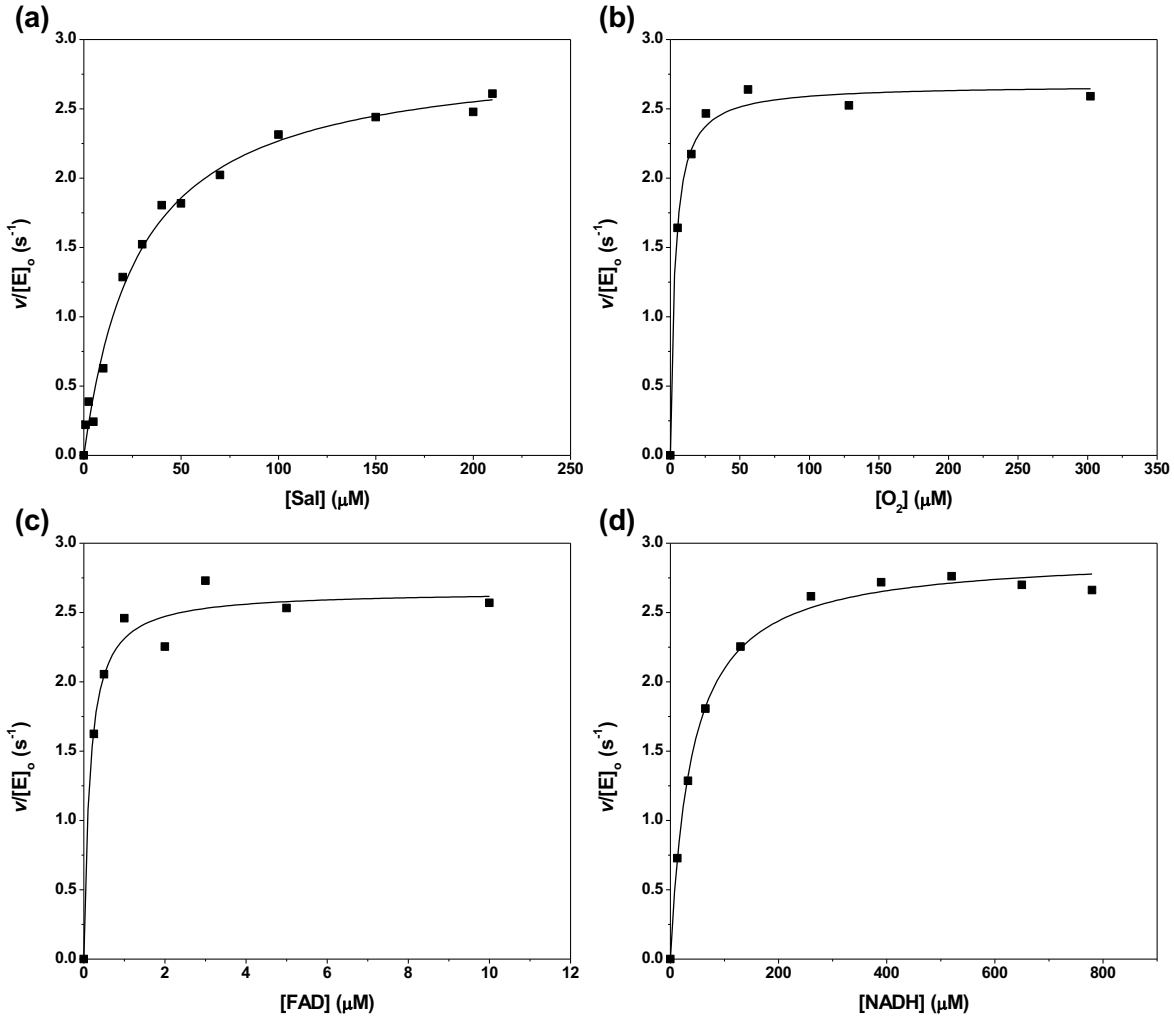


Fig. S5. Reaction rate as a function of concentration for salicylate (a), O_2 (b), FAD (c), and NADH (d) for conversion of salicylate (Sal) into catechol catalyzed by NahG at pH 8.5 and 25 ºC. The solid line represents nonlinear fits using the Michaelis-Menten equation and the parameters presented in Table 2.

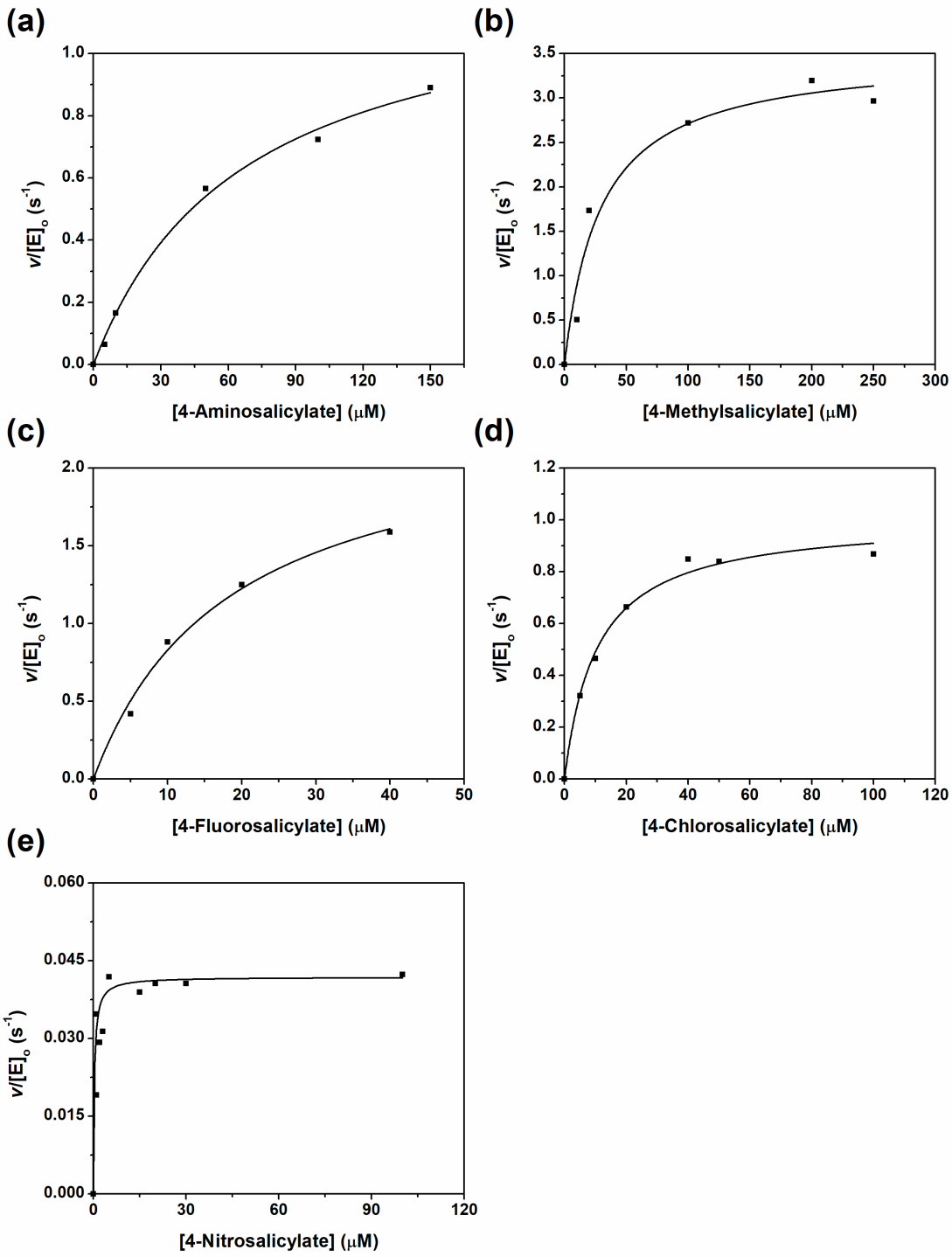


Fig. S6. Reaction rate as a function of substrate concentration for the reaction of 4-substituted salicylates catalyzed by NahG at pH 8.5, I = 0.222, and 25 ºC. The reaction mixture consisted of 200 μM NADH, 1.5 μM FAD, 0.192 μM NahG in 56 mM Hepes, 1.1 mM EDTA, and 171 mM NaCl. The solid lines represent nonlinear fits using the Michaelis-Menten equation and the parameters presented in Table 3.

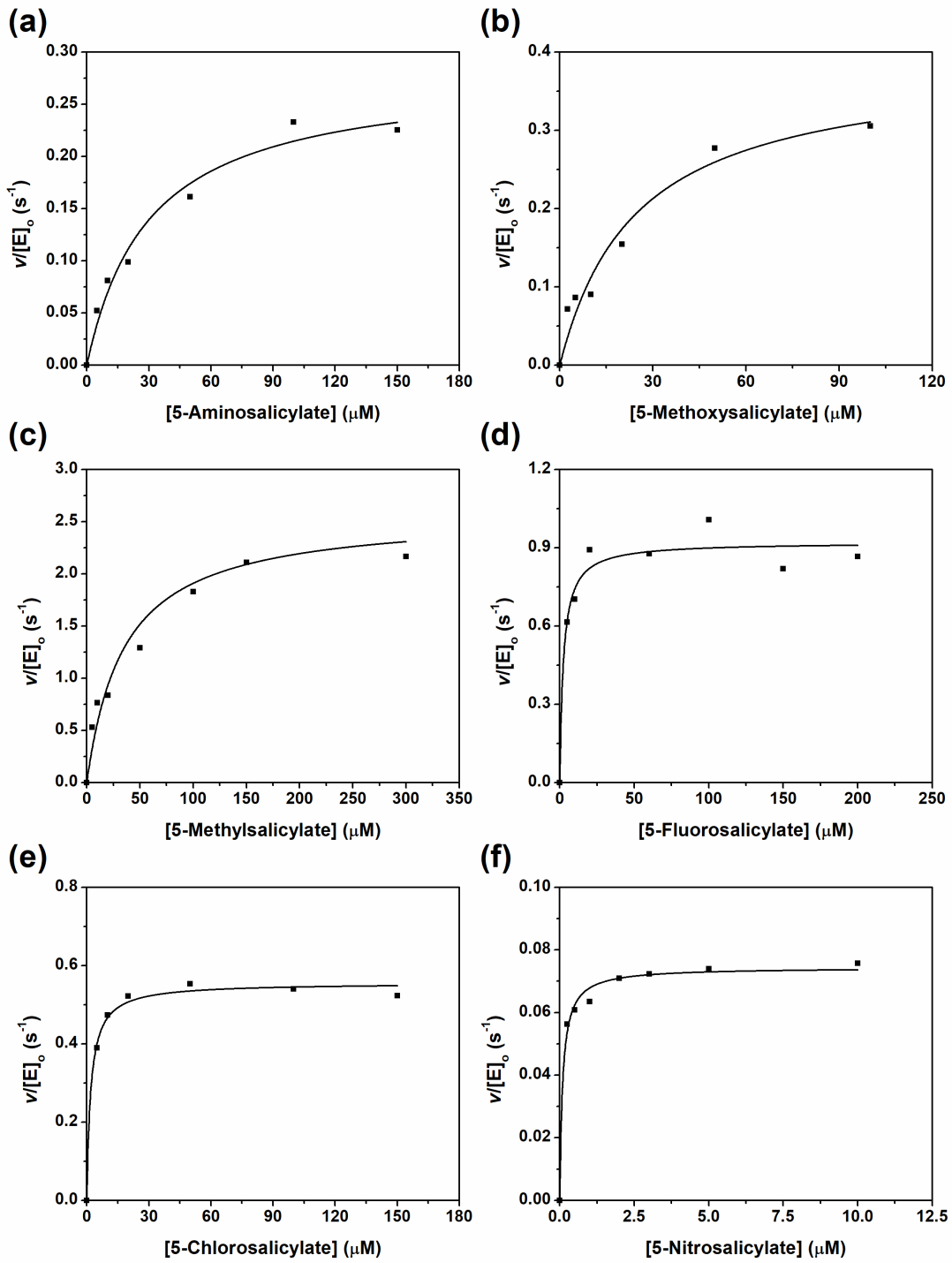


Fig. S7. Reaction rate as a function of substrate concentration for the reaction of 5-substituted salicylates catalyzed by NahG at pH 8.5, I = 0.222, and 25 ºC. The reaction mixture consisted of 200 μM NADH, 1.5 μM FAD, 0.192 μM NahG in 56 mM Hepes, 1.1 mM EDTA, and 171 mM NaCl. The solid lines represent nonlinear fits using the Michaelis-Menten equation and the parameters presented in Table 3.

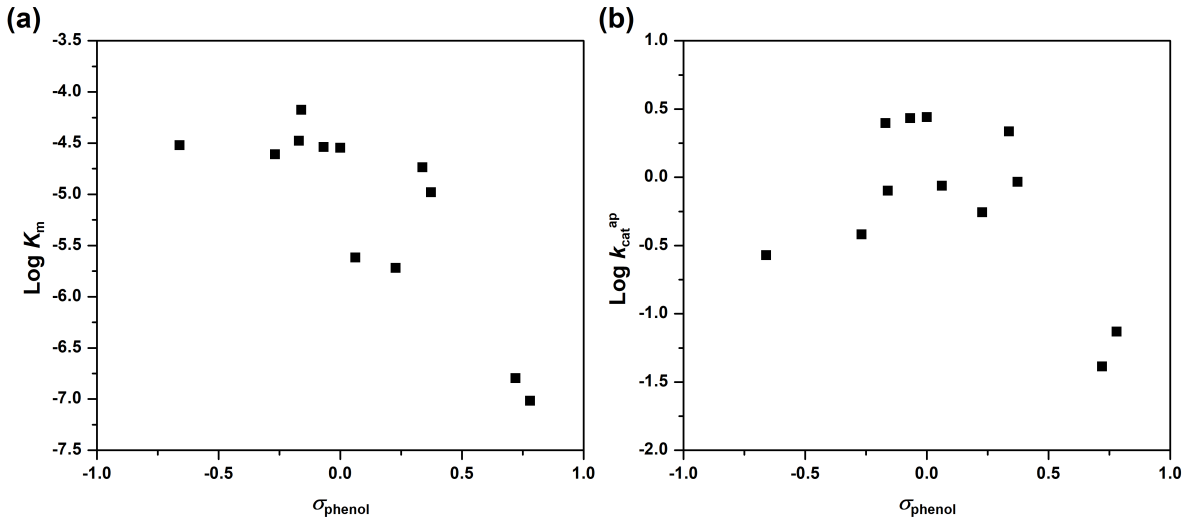


Fig. S8. Hammett relationships for K_m (a) and k_{cat} (b) for the reaction of substituted salicylates catalyzed by NahG at pH 8.5 and 25 ºC. The σ_{phenol} values are σ_m and σ_p for salicylates substituted at positions 4 and 5, respectively.

Table S1. Kinetic parameters for the oxidative decarboxylation of salicylate catalyzed by NahG in the presence of 200 μM NAD^+ and 1.5 μM FAD at 25.0 ºC; buffer conditions: 56 mM Buffer, 1.1 mM EDTA, and ionic strength of 0.222 (NaCl).

pH	K_m μM	V_{max} nM s^{-1}	$[E]_o$ nM	k_{cat}^{ap} s^{-1}	Buffer
3.80	0.33±0.05	93±2	512	0.18±0.01	Acetic acid
4.00	nd	106±5	512	0.21±0.01	Acetic acid
4.20	1.3±0.2	181±3	512	0.35±0.01	Acetic acid
4.32	0.9±0.1	195±5	512	0.38±0.01	Acetic acid
4.69	0.5±0.2	230±10	512	0.45±0.02	Acetic acid
5.07	0.6±0.3	210±10	512	0.40±0.02	Acetic acid
5.40	1.0±0.4	280±10	512	0.56±0.02	Acetic acid
5.70	3.3±0.6	310±10	512	0.61±0.02	Acetic acid
5.80	1.4±0.2	77±2	64	1.20±0.03	Bis-tris
6.31	2.5±0.1	101±1	64	1.58±0.01	Bis-tris
6.57	nd	276±5	128	2.15±0.04	Bis-tris
6.78	6.1±0.8	113±4	64	1.77±0.06	Bis-tris
6.86	nd	203±7	128	1.59±0.05	Bis-tris
7.15	nd	283±2	128	2.21±0.02	Hepes
7.58	16±1	249±4	64	3.89±0.06	Hepes
7.86	32±3	240±10	128	2.63±0.08	Hepes
8.20	41±2	202±4	64	3.15±0.07	Hepes
8.50	28±3	186±3	64	2.92±0.04	Hepes
8.60	36±4	320±10	128	2.49±0.09	Ches
9.46	46±3	382±9	128	2.99±0.07	Ches
9.74	nd	145±5	64	2.27±0.08	Ches
10.00	41±2	200±4	64	3.12±0.07	Ches
10.30	nd	1200±40	384	3.12±0.09	Caps
10.66	58±2	526±8	384	1.37±0.02	Caps
10.97	53±8	510±30	640	0.79±0.06	Caps

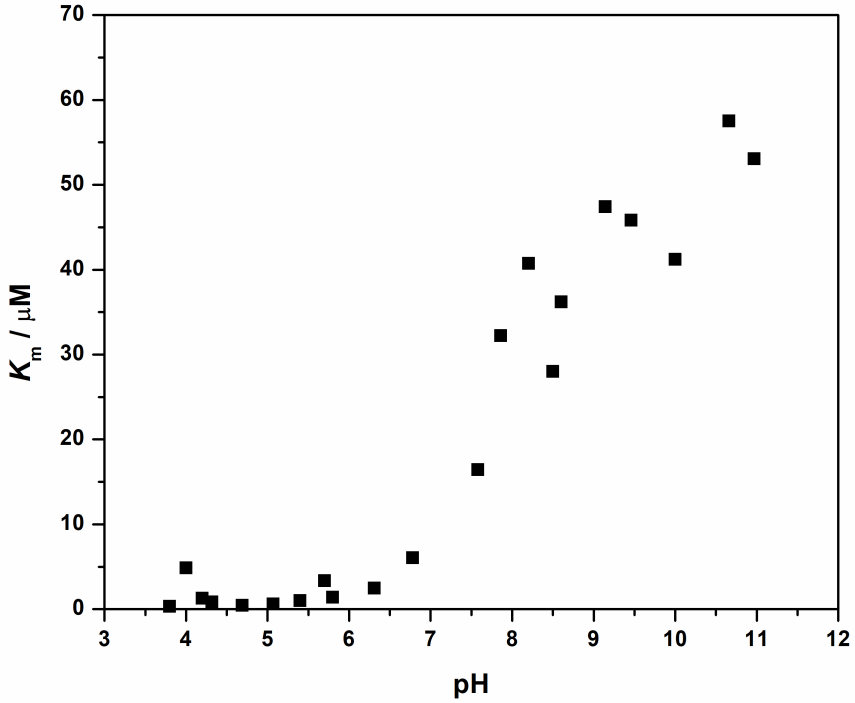


Fig. S9. Effect of pH on K_m for the oxidative decarboxylation of salicylate catalyzed by NahG in the presence of 200 μM NAD^+ and 1.5 μM FAD at 25 °C.

References

[1] F.R. Carneiro, T.C. Silva, A.C. Alves, T. Haline-Vaz, F.C. Gozzo, N.I. Zanchin, Spectroscopic characterization of the tumor antigen NY-REN-21 and identification of heterodimer formation with SCAND1, Biochem. Biophys. Res. Commun. 343(1) (2006) 260-268.

[2] D.M.A. Costa, M.A.F. Costa, S.L. Guimarães, J.B. Coitinho, S.V. Gómez, T.A.S. Brandão, R.A.P. Nagem, A combined approach for enhancing the stability of recombinant cis-dihydrodiol naphthalene dehydrogenase from *Pseudomonas putida* G7 allowed for the structural and kinetic characterization of the enzyme, Protein Expression Purif. 132 (2017) 50-59.

Table 1. Data collection and structural refinement statistics for iodine 6xHis-NahG crystal. Values in parentheses are for the highest resolution shell.

Data collection statistics	
Wavelength (Å)	1.459
Temperature (K)	100
Crystal-to-detector distance (mm)	110
Rotation range per image (°)	0.5
Total rotation range (°)	360
Space group	$P2_12_12_1$
Unit cell parameters (Å)	a = 81.62, b = 98.08, c = 130.14 Å
Resolution range (Å)	50.00 – 2.00 (2.03 – 2.00)
No. of observations	951,839 (33,675)
No. of unique reflections	133,525 (6,476)
Data completeness (%)	99.0 (95.6)
$\langle I/\sigma(I) \rangle$	25.4 (2.9)
Redundancy	7.1 (5.2)
R_{merge}[a]	0.064 (0.526)
Structural refinement statistics	
No. of reflections, working set	132,697
No. of reflections, test set	6,750
Final R_{factor}[b]	0.2109
Final R_{free}[c]	0.2336
No. of non-H atoms[d]	
Protein	6,434
Ion (I^-)	14
Water	462
R.m.s. deviations	
Bonds (Å)	0.002
Angles (°)	0.471
Average B factors (Å^2)	

Protein	47.5
Ion (I^-)	52.1
Water	50.6
Ramachandran plot	
Most favored (%)	98.07
Allowed (%)	1.93

[a] $R_{merge} = \Sigma_{hkl} \Sigma_i | I_i(hkl) - \langle I(hkl) \rangle | / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the i^{th} intensity measurement of reflection hkl and $\langle I(hkl) \rangle$ is its average. [b] R-factor $= \Sigma_{hkl} || F_{obs}| - k|F_{cal}|| / \Sigma_{hkl} |F_{obs}|$, where F_{obs} and F_{cal} are observed and calculated structure factors, respectively. [c] For R_{free}, the sum is extended over a subset of reflections (5%) excluded from all stages of refinement. [d] Double conformations counted twice.

Table 2. Kinetic parameters for salicylate (Sal), dioxygen (O_2), FAD, and NADH for conversion of salicylate into catechol catalyzed by NahG at pH 8.5 and 25 ºC.[a]

Parameters	Sal	O_2	FAD	NADH
K_m (μM)	28 ± 3	3.2± 0.4	0.15± 0.03	39 ± 3
k_{cat} (s^{-1})		3.3 ± 0.1[b]		

[a] In 56 mM Buffer HEPES, 1.1 mM EDTA and I = 0.222 (NaCl), [NahG] = 16-64 nM. Except where the concentration was varied, salicylate, FAD, and NADH concentrations were 192 μM, 31.5 μM and 200 μM, respectively. Errors are standard deviations.[b] Corrected for the uncoupled path (6 ± 2 %); uncorrected k_{cat} = 3.5 ± 0.1 s^{-1}.

Table 3. Kinetic parameters at pH 8.5 and 25 ºC for the NahG-catalyzed reaction of salicylates substituted at positions 4 and 5.[a]

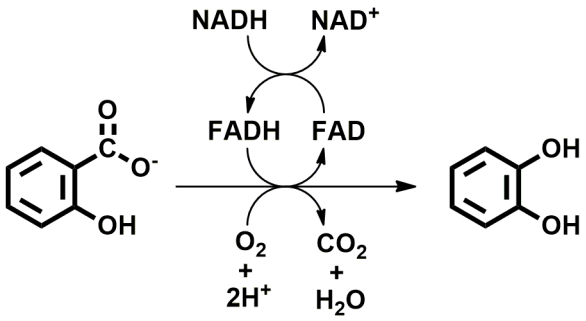
Salicylates	$\sigma_{carboxy}$ [b]	σ_{phenol} [b]	K_m (10^{-6} M)	k_{cat}^{ap} (s^{-1}) [c]	% Uncoupling
H	0	0	28 ± 3	2.76 ± 0.04	6 ± 2
4-NH_2	-0.66	-0.16	77 ± 3	0.80 ± 0.02	37 ± 2
4-CH_3	-0.17	-0.07	29 ± 5	2.7 ± 0.2	22 ± 2
4-F	0.06	0.34	18.4 ± 0.9	2.16 ± 0.06	8 ± 2
4-Cl	0.23	0.37	10.5 ± 0.7	0.93 ± 0.02	8 ± 2
4-NO_2	0.78	0.72	0.16 ± 0.01	0.041 ± 0.02	2 ± 2
5-NH_2	-0.16	-0.66	30 ± 3	0.28 ± 0.01	4 ± 1
5-OCH_3	0.12	-0.27	25 ± 3	0.37 ± 0.02	3 ± 2
5-CH_3	-0.07	-0.17	33 ± 5	2.6 ± 0.1	1 ± 1
5-F	0.34	0.06	2.4 ± 0.6	0.87 ± 0.03	6 ± 2
5-Cl	0.37	0.23	1.9 ± 0.2	0.55 ± 0.01	2 ± 2
5-NO_2	0.72	0.78	0.10 ± 0.01	0.074 ± 0.001	1 ± 2

[a] In 56 mM Hepes, 1.1 mM EDTA and I = 0.222 (NaCl). The kinetic parameters were obtained from the Michaelis-Menten plots shown in Figure S6 and S7 (refer to the Supporting Information). Errors are standard deviations. [b] The $\sigma_{carboxy}$ values are Hammett's σ_p and σ_m [88] for salicylates substituted at positions 4 and 5, respectively. The σ_{phenol} refers to σ_p and σ_m for salicylates substituted at the respective positions 5 and 4. [c] Corrected for uncoupling.

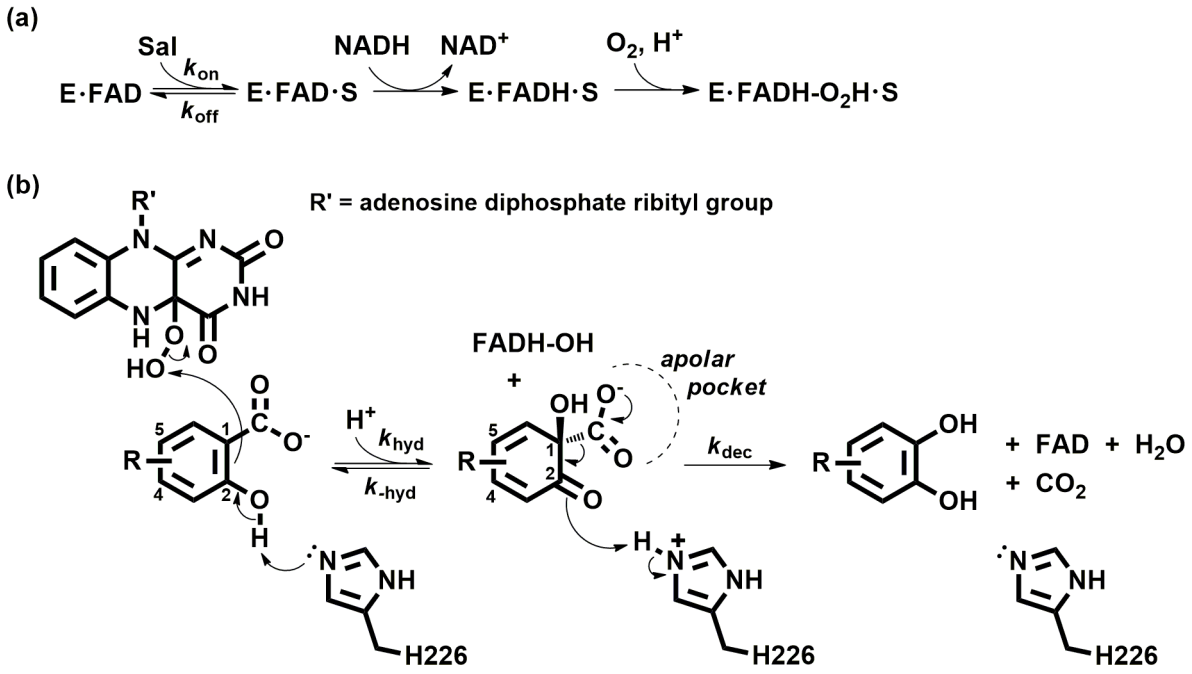
Table 4. Kinetic and equilibrium parameters for the oxidative decarboxylation of salicylate catalyzed by NahG at 25 °C. The parameters were obtained by the non-linear fit (solid line) of the data in Figure 8 using Equation 6.

Parameters	Values
k_i (s^{-1})	0.5 ± 0.1
k_{max} (s^{-1})	3.0 ± 0.3
pK_{ES1}	4.0 ± 0.2
pK_{ES2}	6.6 ± 0.2
pK_{ES3}	10.6 ± 0.1

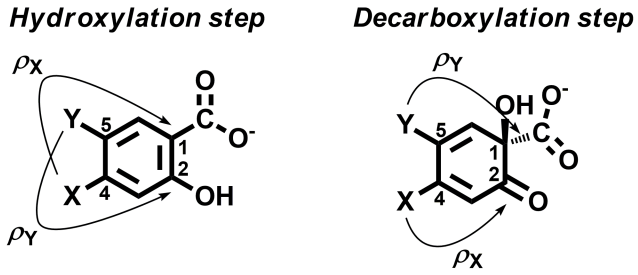
Scheme 1. Oxidative decarboxylation of salicylate catalyzed by NahG.



Scheme 2. Proposed mechanism for NahG catalyzed the oxidative decarboxylation of substituted salicylates.



Scheme 3. Guide for analysis of substituent effects on the oxidative decarboxylation of salicylate.



Scheme 4. Species in the reaction pathway of NahG.

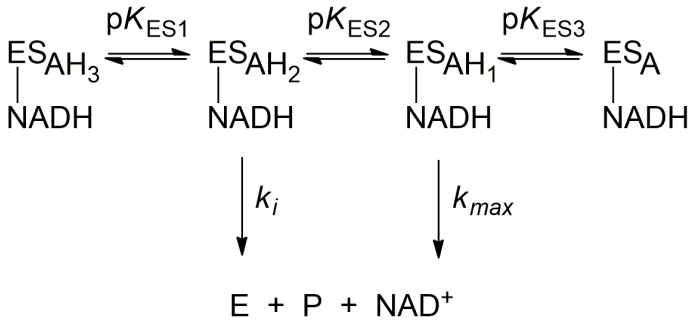


Figure Captions

Figure 1. Crystal structures of *apo* 6xHis-NahG (light blue) showing two monomers in the asymmetric unit. The structure at the bottom was superimposed with *holo* 3-hydroxybenzoate 6-hydroxylase (3HB6H, PDB entry 4BJZ, green) from *Rhodococcus jostii* RHA1 [23]. The region colored in blue (residues 38-53) for the *apo* 6xHis-NahG is found in a new position upon FAD binding as shown in dark green (residues 36-51) for the *holo* 3HB6H.

Figure 2. Active-site of 6xHis-NahG showing the putative positions of FAD isoalloxazine ring and salicylate (Sal) relative to the hydrophobic region formed by the residues F85, M219, H226, L228, F230, V241, and F243. The ligands FAD and Sal were modeled using superpositions with the respective structures of 3HB6H (PDB entry 4BJZ) [23] and PHBH bound to 2,4-dihydroxybenzoate (PDB entry 1DOD) [64], in which the 4-OH group was suppressed to give Sal. The hydrophobic region depicts only residues within 5 Å of distance from the plane formed by all Sal atoms.

Figure 3. FAD binding cleft showing three conformational states: (a) the *apo* state is represented by 6xHis-NahG, in which the loop from E38 to G53 is found in a closed conformation; (b) and (c), the equivalent loop lies in an open conformation upon FAD binding, and the isoalloxazine ring swings between the "out" (PDB entry 3C96 for the flavin-containing monooxygenase phzS from *Pseudomonas aeruginosa*) and "in" (PDB entry 4BJZ for the 3HB6H) states.

Figure 4. Formation of *holo* NahG and its interaction with salicylate (Sal) at pH 8.0, I = 0.222 (NaCl) and 25 °C: (a) and (c), fluorescence spectra of *apo* (0.74 μM) and *holo* NahG (1.0 μM) as a function of FAD and salicylate concentrations, respectively; (b) and (d), fluorescence intensity at the indicated wavelengths as a function of ligand concentration. Arrows in each fluorescence

spectra indicate the direction of an increasing ligand concentration. Solid lines are fits according to Equation 1 and the following parameters: (b) F_{free} = 566, F_{bound} = 203, $[E]_o$ = 0.74 μM and K_d = 2 x 10^{-9} M, (d) F_{free} = zero, F_{bound} = 285, $[E]_o$ = 1.0 μM and K_d = 2.0 x 10^{-7} M.

Figure 5. NMR spectra at 600 MHz before (A) and after (B) the partial conversion in the oxidative decarboxylation of 3-methylsalicylate catalyzed by NahG in the presence of NADH at 25 ºC and pH 8.2. The gray and black boxes indicate the ^{1}H signals for 3-methylsalicylate and 3-methylcatechol, respectively.

Figure 6. Hammett relationships for K_m (a) and k_{cat} (b) for the reaction of substituted salicylates catalyzed by NahG at pH 8.5 and 25 ºC. The $\sigma_{carboxy}$ values are σ_p and σ_m [88] for salicylates substituted at positions 4 and 5, respectively.

Figure 7. Jaffé relationships for the reaction of substituted salicylates catalyzed by NahG at pH 8.5 and 25 ºC.

Figure 8. Effect of pH on log k_{cat} for the oxidative decarboxylation of salicylate catalyzed by NahG at 25 °C. The solid line represents the nonlinear fit of the kinetic data using Equation 4 and the parameters presented in Table 4. The dotted and dashed lines represent, respectively, the individual influence of k_i and k_{max}.

Figure 1.

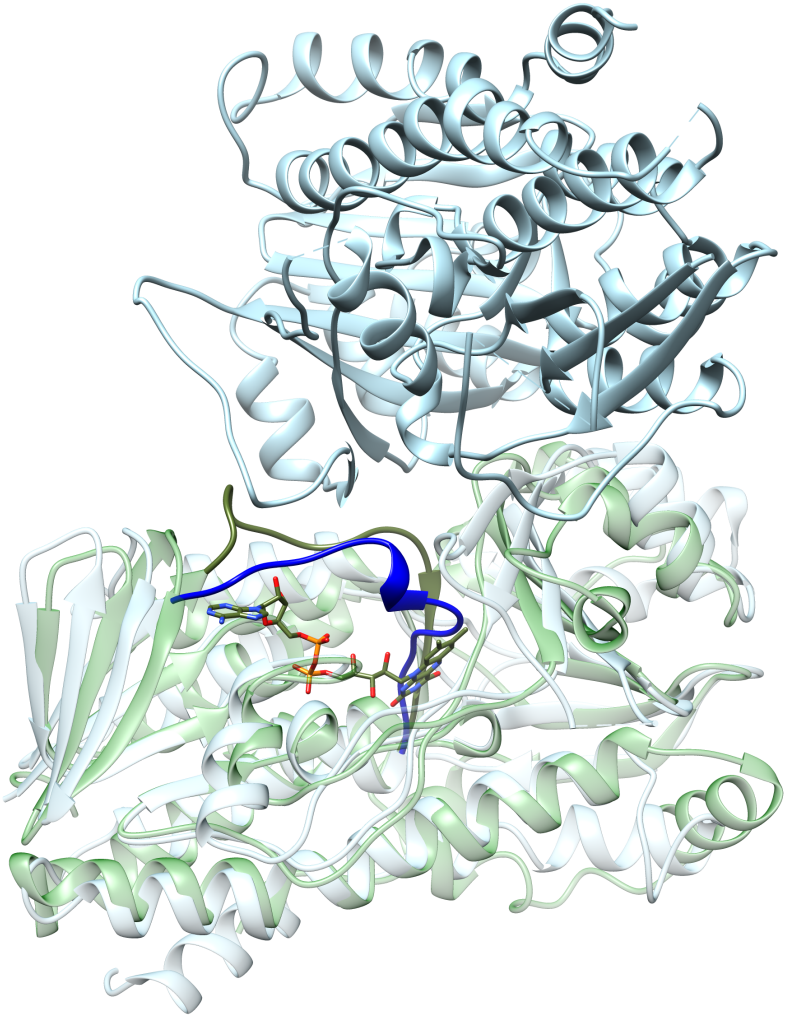


Figure 2.

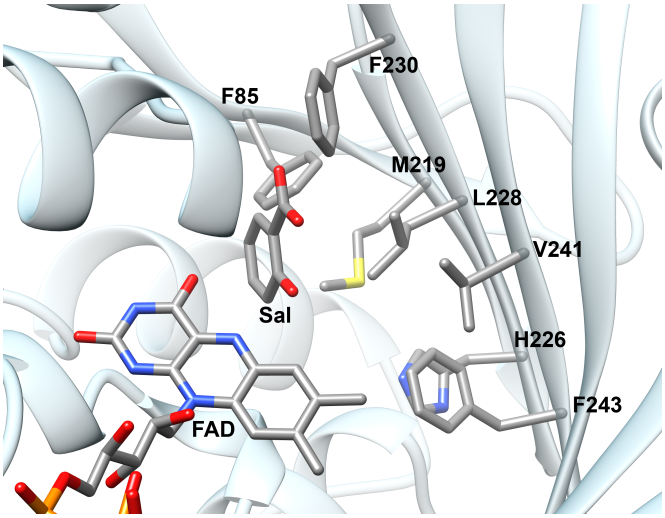


Figure 3.

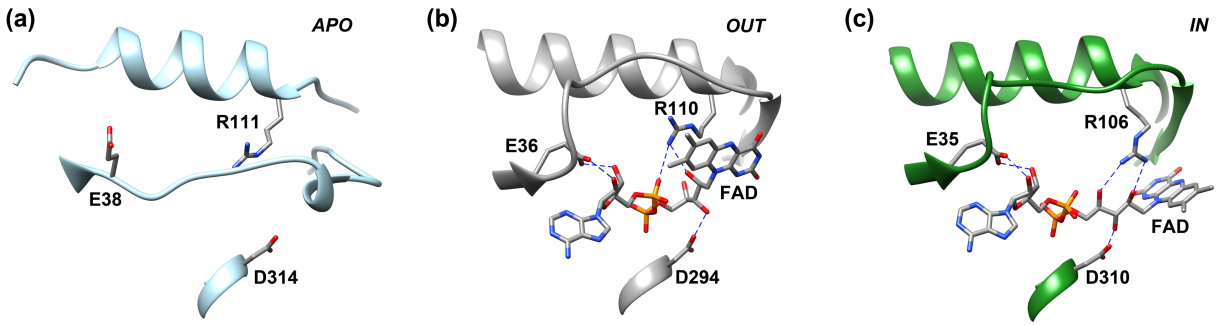

(a)
APO
R111
E38
D314
(b)
OUT
R110
E36
FAD
D294
(c)
IN
R106
E35
FAD
D310

Figure 4.

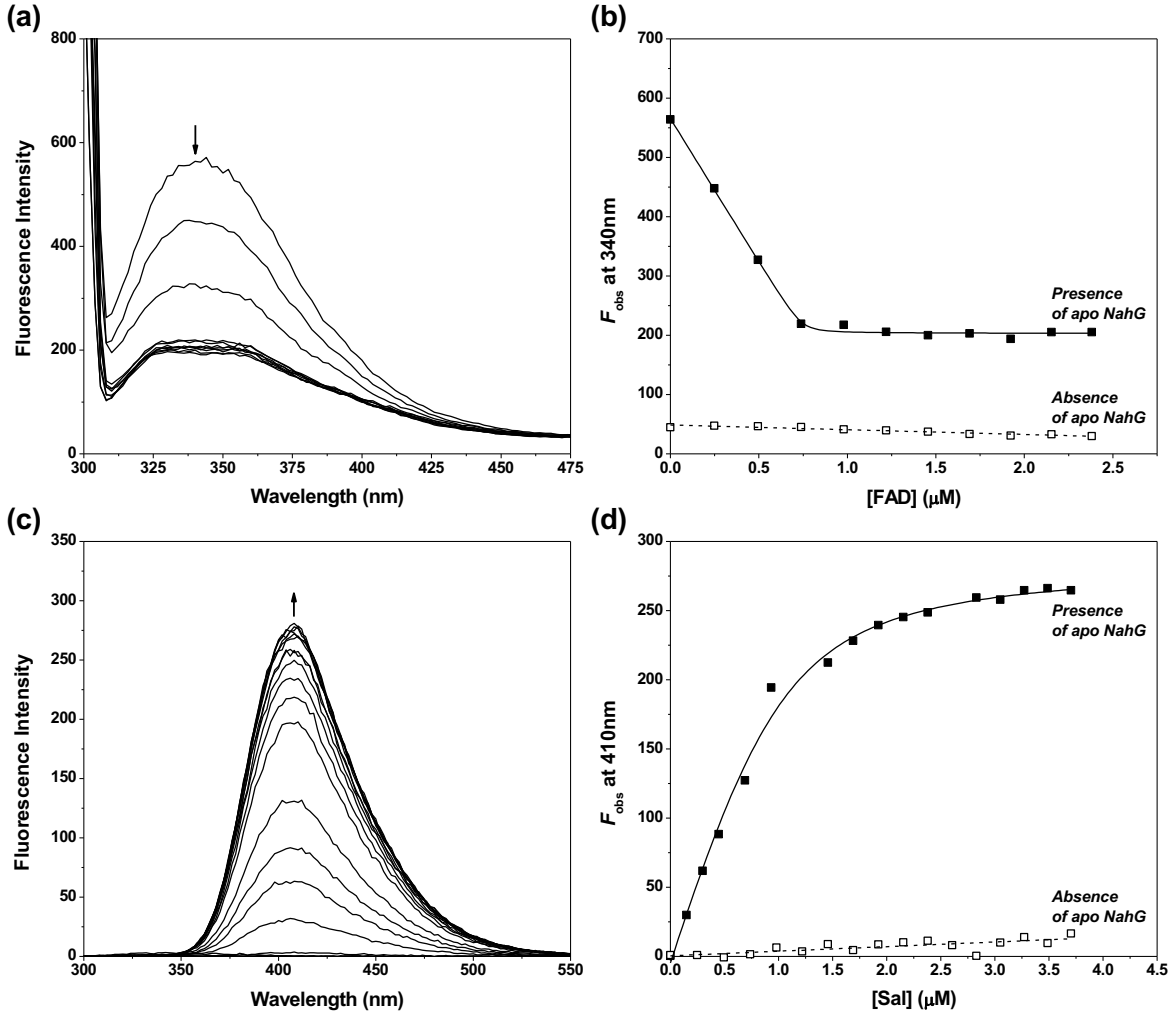

(a)
Fluorescence Intensity
Wavelength (nm)
(b)
F_{obs} at 340nm
[FAD] (μM)
Presence of apo NahG
Absence of apo NahG
(c)
Fluorescence Intensity
Wavelength (nm)
(d)
F_{obs} at 410nm
[Sal] (μM)
Presence of apo NahG
Absence of apo NahG

Figure 5.

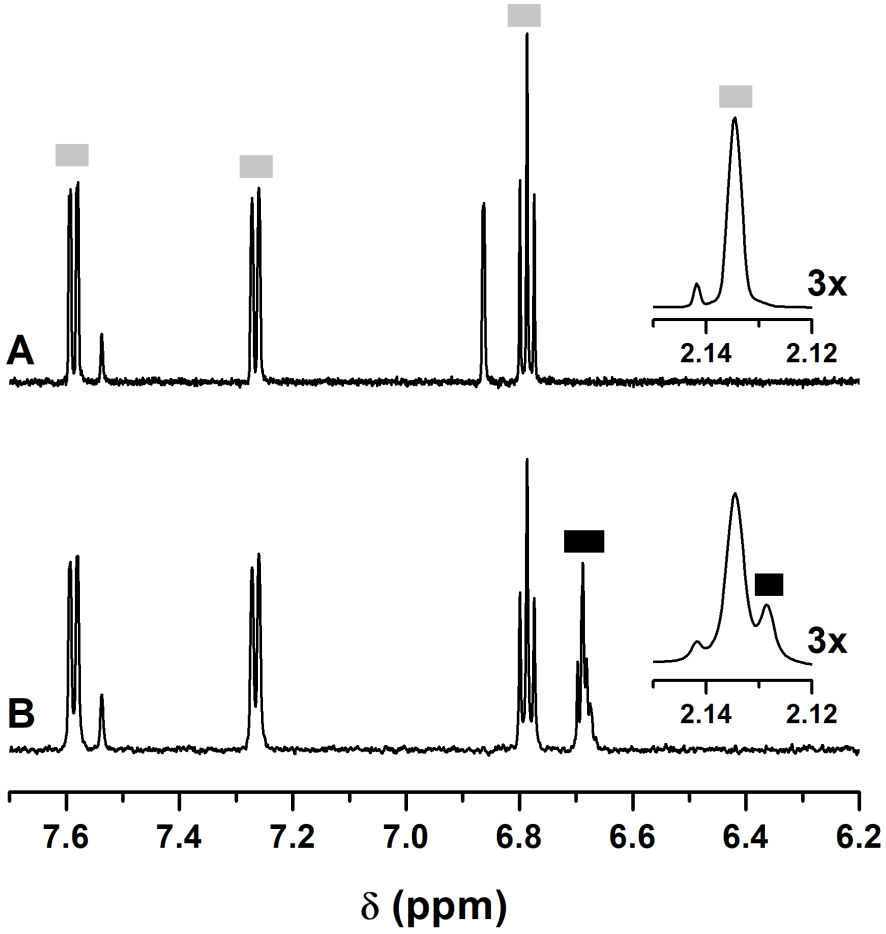


Figure 6.

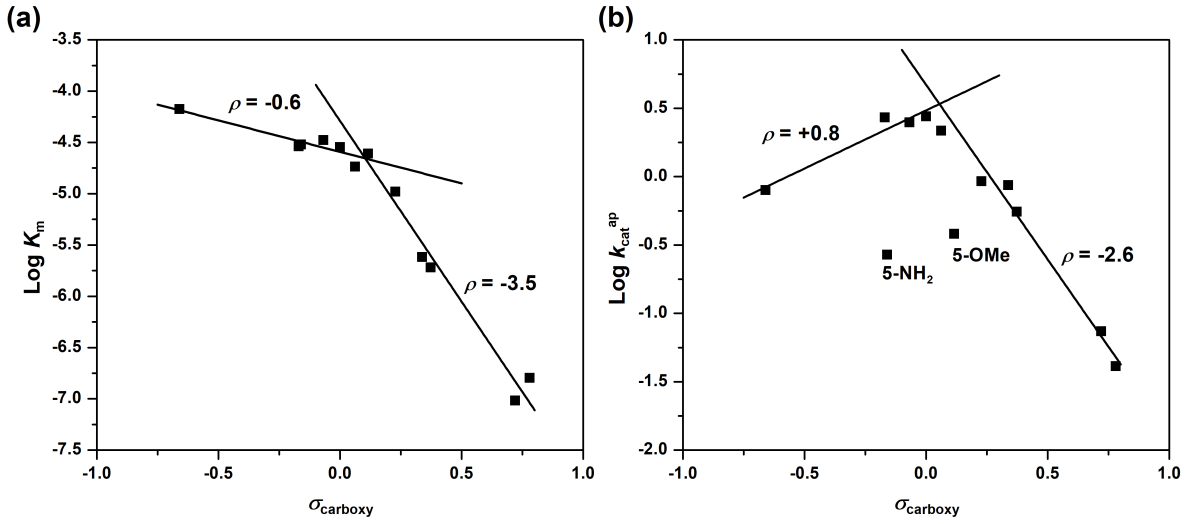

(a)
Log K_m
ρ = -0.6
ρ = -3.5
$\sigma_{carboxy}$
(b)
Log k_{cat}^{ap}
ρ = +0.8
5-NH$_2$
5-OMe
ρ = -2.6
$\sigma_{carboxy}$

Figure 7.

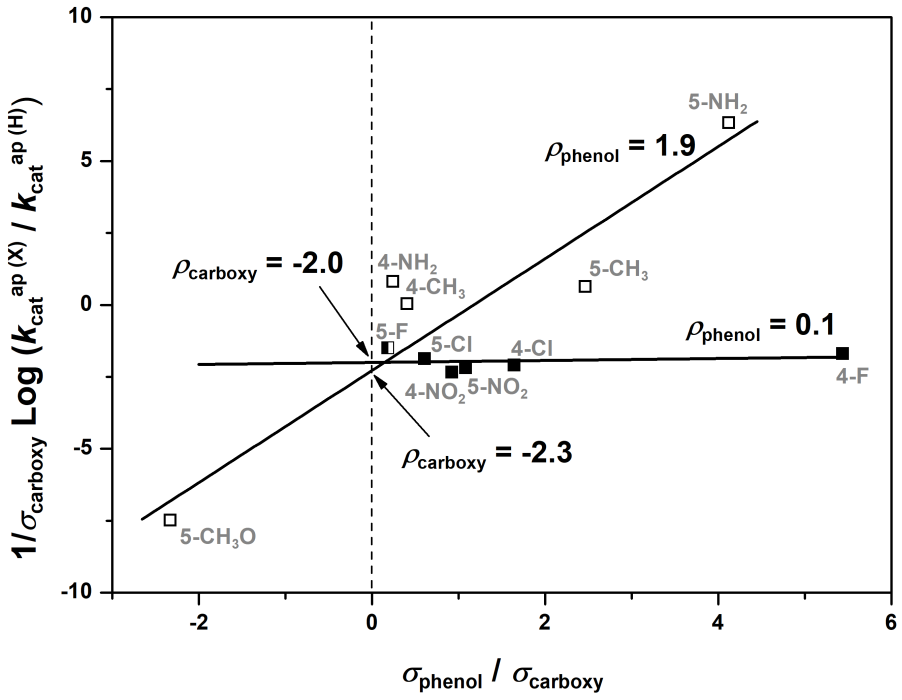

10
5
0
-5
-10
-2
0
2
4
6
1/σcarboxy Log (kcat ap (X) / kcat ap (H))
σphenol / σcarboxy
ρphenol = 1.9
ρcarboxy = -2.0
ρphenol = 0.1
ρcarboxy = -2.3
5-NH2
4-NH2
4-CH3
5-CH3
5-F
5-Cl
4-Cl
4-F
4-NO2
5-NO2
5-CH3O

Figure 8.