Supplementary Information

A pyridoxal phosphate-dependent enzyme that oxidizes an unactivated carbon-carbon bond

Yi-Ling Du,1 Rahul Singh,2 Lona M. Alkhalaf,1 Eugene Kuatsjah,3 Hai-Yan He,1 Lindsay D. Eltis,2 Katherine S. Ryan1,*

1Department of Chemistry, 2Department of Microbiology and Immunology, 3Genome Sciences and Technology Program, The University of British Columbia, Vancouver, BC, Canada

Supplementary Results

**TABLE OF CONTENTS**

<table>
<thead>
<tr>
<th>Figure / Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supplementary Figure 1</td>
<td>LC-MS/MS analysis and $^1$H-NMR on 6</td>
<td>3</td>
</tr>
<tr>
<td>Supplementary Figure 2</td>
<td>NMR data on 7</td>
<td>4</td>
</tr>
<tr>
<td>Supplementary Figure 3</td>
<td>HPLC analysis of Ind4 products</td>
<td>7</td>
</tr>
<tr>
<td>Supplementary Figure 4</td>
<td>Effect of EDTA and hydroxylamine on Ind4 activity</td>
<td>8</td>
</tr>
<tr>
<td>Supplementary Figure 5</td>
<td>Biochemical analysis of Ind4 and its variants</td>
<td>9</td>
</tr>
<tr>
<td>Supplementary Figure 6</td>
<td>Anaerobic assay and steady-state kinetics for Ind4</td>
<td>11</td>
</tr>
<tr>
<td>Supplementary Figure 7</td>
<td>Substrate specificity of Ind4</td>
<td>13</td>
</tr>
<tr>
<td>Supplementary Figure 8</td>
<td>Isotope tracking for interrogating deamination and reamination</td>
<td>15</td>
</tr>
<tr>
<td>Supplementary Figure 9</td>
<td>Biochemical analysis of the Ind4-Ind5/6 coupled reactions</td>
<td>16</td>
</tr>
<tr>
<td>Supplementary Figure 10</td>
<td>Distribution of Ind456 genes</td>
<td>19</td>
</tr>
<tr>
<td>Supplementary Figure 11</td>
<td>Analysis of Pel4 and Pel5</td>
<td>22</td>
</tr>
<tr>
<td>Supplementary Figure 12</td>
<td>Ind4 reaction with L-canavanine</td>
<td>25</td>
</tr>
<tr>
<td>Supplementary Figure 13</td>
<td>UV-Vis spectra of Ind4 reaction with L-arginine or L-canavanine</td>
<td>26</td>
</tr>
<tr>
<td>Supplementary Figure 14</td>
<td>Transient kinetic analysis of the reaction of Ind4 with L-arginine</td>
<td>28</td>
</tr>
<tr>
<td>Supplementary Figure 15</td>
<td>Global analysis of the reaction of Ind4 with L-arginine</td>
<td>29</td>
</tr>
<tr>
<td>Supplementary Figure 16</td>
<td>Mechanistic proposal for the Ind4 reaction</td>
<td>30</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>31</td>
</tr>
</tbody>
</table>
Supplementary Results

Supplementary Figure 1

**Supplementary Figure 1.** Structure of compound 6. (a) LC-MS/MS analysis of authentic 4-guanidinobutyric acid (left) and purified 6 (right) showing the identical ion fragmentation patterns. The diamond (◊) indicates the [M+H]+ ion. Additionally, purified 6 gave [M+H]+ ion at m/z 146.0933 in HR-ESI-MS, corresponding to the molecular formula C₅H₁₁N₃O₂ (calcd for C₅H₁₂N₃O₂ at 146.0930). (b) ¹H NMR data on 6 in D₂O at 400 MHz: δ 3.14 (t, \( J = 7.1 \) Hz, 2H), 2.21 (t, \( J = 7.3 \) Hz, 2H), 1.77 (m, 2H). These results demonstrate that 6 is 4-guanidinobutyric acid. Note: the signal at 1.89 ppm is from the methyl group of acetic acid that was used during the purification process.
Supplementary Results

Supplementary Figure 2

a

[Chemical structure and NMR spectrum]

b

[Chemical structure and NMR spectrum]
Supplementary Results

c

d
Supplementary Results

Supplementary Figure 2. NMR data on 7 in D$_2$O. (a) $^1$H NMR at 600 MHz. Note: the signal at 1.89 ppm is from the methyl group of acetic acid that was used during the purification process. (b) $^{13}$C NMR at 150 MHz, (c) COSY, (d) HSQC, and (e) HMBC spectra of 7 at 600 MHz. (f) NMR data assignment of 7. Purified 7 gave [M+H]$^+$ ion at $m/z$ 144.0770 in HR-ESI-MS, corresponding to the molecular formula C$_5$H$_9$N$_3$O$_2$ (calcld for C$_5$H$_{10}$N$_3$O$_2$ at 144.0773). These results demonstrate that 7 is 4-guanidino-3,4-dehydrobutyric acid.
Supplementary Results

Supplementary Figure 3.

Supplementary Figure 3. HPLC analysis of Ind4 products at 330 nm. (a) HPLC analysis of Ind4-catalyzed oxidation of 5 after pre-column derivatization with DNS-Cl (left) or o-phenylenediamine (OPD) (right) over time, showing a decrease in derivatized 5 (open diamond) and an increase in derivatized 8 (red diamond) and 9 (black diamond). (b) LC-MS analysis (selected ion monitoring) of the Ind4-catalyzed oxidation at 120 min after pre-column derivatization with OPD reveals the production of 8 (m/z 246 [M+H]⁺ of OPD-8) and 9 (m/z 244 [M+H]⁺ of OPD-9).
Supplementary Results

Supplementary Figure 4.

**Supplementary Figure 4.** Effect of EDTA and the PLP-enzyme inhibitor hydroxylamine on Ind4 activity. The reaction mixture (100 µl) contained 5 µM Ind4, 1 mM L-arginine (L-Arg) in 50 mM HEPES buffer (pH 7.5) with or without hydroxylamine (2 mM) / EDTA (10 mM), and was incubated at 28 °C for 1.5 h. Reaction mixtures were then subjected to HPLC analysis after pre-column derivatization by o-phthaldialdehyde (OPA) / mercaptopropionic acid (MPA) (see Online Methods for details). For OPA/MPA derivatization of ammonium, 0.5 mM (NH₄)₂SO₄ in 50 mM HEPES (pH 7.5) was used to react with OPA/MPA. Detection wavelength: 338 nm. The results show that hydroxylamine eliminates Ind4 activity, whereas EDTA does not affect turnover (After 1.5 h of incubation, both L-arginine + Ind4 and L-arginine + Ind4 + EDTA consumed 64% of the total amount of L-arginine).
Supplementary Results

Supplementary Figure 5.
Supplementary Results

Supplementary Figure 5. Biochemical analysis of Ind4 and its variants. (a) Homology-model of Ind4 constructed by SWISS-MODEL, based on the template human kynurenine aminotransferase I (hKAT-I) (PDB ID 1w7m), which shares 17% identity on the amino acid level with Ind4. The residues that form an active site in hKAT-I, are highly conserved among Ind4 and its homologs (see alignment in Supplementary Figure 10). In hKAT-I, PLP is covalently linked to K247 (K217 in Ind4), whereas R398 (R350 in Ind4) is thought to stabilize the substrate α-carboxylate moiety, resulting in the positioning of the substrate α-amino group just above the PLP C4’ reactive center. (b) SDS-PAGE analysis of Ind4 and its mutants, R350K and K217A, purified from E. coli. The PLP content of these proteins was determined to be 1.01 ± 0.05, 0.98 ± 0.05 and 0.42 ± 0.04 mol/mol of subunit, respectively. (c) In vitro biochemical assays for K217A and R350K in the presence of catalase and L-Arg show no activity for both of them. (d) UV-Vis spectra of reconstituted K217A, R350K, and Ind4 (see Online Methods for details on reconstitution protocol). Both Ind4 and R350K show an absorption peak at 417 nm that is typical of an enzyme-bound PLP cofactor in the form of a ketoenaminic internal aldimine, whereas K217A has absorption bands at 328 nm and 388 nm, which can be attributed to the unsubstituted and hydrated aldehyde forms of free PLP, suggesting a non-covalent binding of PLP to K217A. Non-covalent binding of PLP to the lysine substituted variants has also been described for other enzymes. These results confirm that the conserved K217 is involved in Schiff base formation with PLP.
Supplementary Figure 6. Anaerobic assay and steady-state kinetics for Ind4. (a) MS analysis of the Ind4-catalyzed reaction performed under air and argon shows the Ind4-catalyzed reaction is strictly oxygen-dependent, as no product can be detected when the reaction is performed under argon. Detected ions correspond to compounds 5 (m/z 175 [M+H]+); 6 (m/z 146 [M+H]+); 7 (m/z 144 [M+H]+); 8 (m/z 174 [M+H]+); and 9 (m/z 172 [M+H]+). Ind4 (final concentration: 5 µM) was injected into a vial containing 1 mM 5 in 3 ml of 50 mM phosphate buffer (pH 7.5), which had been previously deoxygenated by bubbling with argon for 30 min. The reaction mixture was kept under argon flow and incubated at room temperature for 3 h before the addition of 3
Supplementary Results

ml of methanol to stop the reaction, followed up by LC-MS analysis. As a control, a similar reaction was performed under air. Note: The reactions were performed in sodium phosphate buffer and analyzed by LC-MS with trifluoroacetic acid (TFA) in the mobile phase. Therefore, the m/z 159.1 that is present in both assays could be the [M+2Na]^+ for TFA. (b) Monitoring of the O₂ consumption of Ind4-catalyzed reactions using an O₂ electrode. The arrow indicates the time point when enzyme(s) were injected. The rate of O₂ consumption is 4.2 ± 0.1 µM/min with catalase (1 mg/ml) and 8.4 ± 0.2 µM/min without catalase. Assays were conducted in triplicate, and data are reported as mean ± standard deviation, with n = 3. (c) and (d) State-steady kinetics of the Ind4-catalyzed reaction. Rates were determined in terms of oxygen consumption, as described in the Online Methods. Assays were conducted in triplicate, and data are reported as mean ± standard deviation, with n = 3.
Supplementary Results

Supplementary Figure 7.

a

b

Supplementary Results

**Supplementary Figure 7.** Substrate specificity of Ind4. (a) Incubation of Ind4 with various amino acids failed to give any product, except with L-arginine. (-) indicates non-enzyme control, (+) indicates the presence of Ind4. TLC was developed with butanol/acetic acid/water (6:2:1). Visualization of the spots on TLC plate was performed by spraying with ninhydrin reagent (0.5% (w/v) ninhydrin and 1% (v/v) acetic acid in butanol). (b) MS analysis of Ind4 reaction mixtures (Ind4 + amino acid). All the reactions contain 5 µM Ind4 and 1 mM amino acid substrate in 50 mM Tris-HCl buffer (pH 7.5), and are incubated at 28 °C for 4 h. The ions at m/z 144 correspond to the [M+Na]^+ for Tris, and the ions at m/z 159 correspond to the [M+2Na]^+ for trifluoroacetate. The y-axis indicates intensity of the mass signal, which is shown as the percentage of the total ion count (TIC) (100% is the maximum TIC shown in each figure).
Supplementary Results

Supplementary Figure 8.

[Chemical structures and annotations, with m/z values marked for each compound.]
Supplementary Results

**Supplementary Figure 8.** An isotope tracking strategy for interrogating whether a deamination and reamination process is involved in the Ind4-Ind56-catalyzed dehydroarginine construction. (a) Four products would be expected from isotopic scrambling at the α-amino group from dehydroarginine construction with deamination and reamination, whereas only two products would be expected without deamination and reamination. Note that the Ind4-56 enzymes are removed by an Amicon Ultracentrifugal filter (10,000 MWCO) prior to addition of Ind3 (see Online Methods). (b) LC-MS analysis of indolmycin B from a 1:1 mixture of L-arginine and \(^{15}\)N-L-arginine, showing only \(m/z\) 357 and 360. TIC: total ion count. The mass signal for compound 4 is shown (right). Representative data from three replicates are shown.
Supplementary Results

Supplementary Figure 9.

a

b

Supplementary Figure 9 (continued)

c

d
Supplementary Results

Supplementary Figure 9. The Ind4-Ind56 coupled reaction catalyzes the production of 1. (a) LC-MS analysis of the Ind4-Ind56 coupled reaction. NADH consumption was followed at wavelength 340 nm, and NAD⁺ and 1 production are monitored at m/z 664 ([M+H]+ for NAD⁺) and m/z 173 ([M+H]+ for 1). (b) State-steady kinetics of the Ind4-Ind56 coupled reaction by following NADH consumption. Assays were conducted in triplicate, and data are reported as mean ± standard deviation, with n = 3. (c) and (d) LC-MS monitoring of the enzyme-free filtrate from the Ind4 reaction. The Ind4 reaction mixture (4 ml of 10 µM Ind4, 2 mM L-arginine (L-Arg), 500 units catalase in 50 mM Tris-HCl buffer, pH 7.5) was incubated for 1 h, and enzyme was quickly removed by ultracentrifugation (×3 times) using an Amicon Ultra Centrifugal filter (Millipore, 10,000 MWCO). The filtrate was subjected to LC-MS analysis immediately. The time point (0 min) indicates the time point for the first injection, and the same sample was auto-injected every 15 min. Detection wavelength for HPLC is 315 nm (only compounds 10 and 9 absorb at 315 nm). The identity of peaks in (e) was confirmed by their MS data. Multiple peaks for both 10 and 9 could be due to the imine-enamine or ketone-enol tautomerization, respectively. (d) Selected ion monitoring of 10 and 9 over time. TIC indicates total ion count. The stability of compound 10 over several hours contrasts with our observations for compound 12, which cannot be detected by MS (see Fig. 4a). The relative stability of compound 10 is likely due to its extended conjugation. (e) Selected ion monitoring (m/z 408, [M+H]+ of DNS-L/D-arginine) of Ind4 and Ind4-Ind56 reactions by LC-MS. Percent of arginine remaining after 2 h was calculated by comparing integrated peak areas for DNS-Arg. (f) Plots of NADH consumption (based on the HPLC peak areas from (a)) and production of 1. The concentration of 1 was calculated based on the consumption of indolmycenic acid (2) after incubation of Ind3, 2, ATP and Mg²⁺ with the samples from (a).
Supplementary Results

Supplementary Figure 10.

\[ \text{ind4} \rightarrow \text{ind5} \rightarrow \text{ind6} \rightarrow \text{Streptomyces griseus ATCC 12648} \]

\[ \text{Streptomyces sp. NRRL F-5135} \]

\[ \text{Pseudoalteromonas luteoviolacea H11} \]

\[ \text{pel4} \rightarrow \text{pel5} \rightarrow \text{Paenibacillus elgii B69} \]

\[ \text{Paenibacillus sp. MSt1} \]

\[ \text{Paenibacillus ehimensis} \]
Supplementary Figure 10. Distribution of Ind456 genes. (a) Ind456 genes distributed in phylogenetically distinct genera, which includes *Streptomyces* sp. NRRL F-5135 (accession number: NZ_JOCR01000013.1), *Pseudoalteromonas luteoviolacea* HI1 (NZ_JVIC01000004.1), *Paenibacillus ehimensis* (NZ_JFHX01000026.1), *Paenibacillus* sp. MST1 (NZ_JNVM01000031.1), and *Paenibacillus elgii* B69.
Supplementary Results

(NZ_AFHW01000184.1)\(^8\). (b) Sequence alignment of Ind4 and its homologs by Clustal Omega\(^9\). The residues that form a putative active site in the hKAT-I are highly conserved among Ind4 and its homologs (see also Supplementary Fig. 5).
Supplementary Results

I. $8 + \text{NADH}$
II. $8 + \text{Pel5}$
III. $8 + \text{NADH} + \text{Pel5}$
IV. $8 + \text{NADH} + \text{Pel5} + \text{Ind6}$
V. $8 + \text{NADH} + \text{Ind56}$
Supplementary Results

**Supplementary Figure 11.** *In vitro* biochemical characterization of Ind4/Ind5 homologs Pel4/Pel5. **(a)** Generation of Pel4 and Pel5 in an *E. coli* host. **(b)** MS analysis of reaction mixtures of Pel4 with L-Arg demonstrates that Pel4 is functionally equivalent to Ind4. **(c)** D-dehydroarginine (dhArg) production by the Pel4-Pel5 coupled reaction. **(d)** MS analysis of reaction mixtures shows a new mass signal at m/z 176 in the presence of Pel5. The mass signal m/z 176 is consistent with the [M+H]+ ion of the 2-hydroxy acid derivative of L-arginine (13). These data suggest that Pel5 might have a broader substrate tolerance than Ind5, and can catalyze the NADH-dependent reduction of 8 to give 13. **(e)** Confirmation of Pel5’s 2-hydroxy acid dehydrogenase activity. We enzymatically synthesized 8 (see Online Methods) and incubated it with Pel5 (or Pel5 + Ind6) and NADH. MS analysis of reaction mixtures shows that compound 13 was detected in both reactions. However, replacing Pel5 with Ind56 in the same reaction failed to give 13, supporting that only Pel5 catalyzes the reduction of 8 to 13. **(f)** Scheme for the reactions catalyzed by Pel5/Ind5. The sequence homology of Ind5/Pel5 with 2-hydroxy acid dehydrogenases, and the ability of Pel5 to reduce compound 8 to a 2-hydroxy acid further supports the idea that the imine of compound 10 and the imine acid derivative of L-arginine (12) are the substrates of Pel5 – and, by extension, Ind5 – rather than the corresponding enamines. Ind5 and its homologs could thus have evolved from the 2-hydroxyacid dehydrogenase family.
Supplementary Results

Supplementary Figure 12.

Supplementary Figure 12. MS analysis of the reaction mixture of Ind4 with (a) arginine analog L-canavanine, (b) L-arginine methyl ester (top) and L-citrulline (lower), in the presence of catalase. The results show that only the reaction with L-canavanine (m/z 177, [M+H]+), an arginine analog with an oxygen in place of C5, gives new products, whose mass signals are consistent with the corresponding keto acid (m/z 176, [M+H]+) and its decarboxylated product (m/z 148, [M+H]+). This result suggests that Ind4 can oxidatively deaminate L-canavanine, but cannot catalyze C4-O5 bond oxidation for this substrate. (c) Structures of L-arginine and its analogs. (d) Steady-state kinetic study of Ind4-catalyzed deamination of L-canavanine, as monitored by O2 consumption, gives a $K_m = 940 \pm 40$ µM, $k_{cat} = 4.7 \pm 0.1$ min⁻¹ for L-canavanine. Assays were conducted in triplicate, and data are reported as mean ± standard deviation, with n = 3.
Supplementary Results

Supplementary Figure 13

a

b

Supplementary Results

**Supplementary Figure 13.** UV-Vis electronic absorption spectroscopy of the reaction of Ind4 with L-arginine or L-canavanine. (a) UV-Vis spectra of Ind4 reaction with L-arginine (L-Arg) (top panel) or L-canavanine (L-Can) (bottom panel) in the presence or absence of O₂. The samples were prepared by mixing equal volumes of 60 µM Ind4 with 600 µM of substrates. The spectra were recorded before or after 1, 2, 5 min of mixing with the substrates. (b) The equilibrium between quinonoid and ketimine intermediates under anaerobic condition.
Supplementary Results

**Supplementary Figure 14**

(a) Transient kinetic analysis of the reaction of Ind4 with L-arginine. (a) The data was collected by mixing equal volumes of Ind4 (60 µM) and L-arginine (L-Arg) (50 µM) solutions using a stopped-flow apparatus. Spectra were recorded after mixing: air saturated (left), anaerobically prepared (middle) or anaerobically prepared Ind4: L-arginine complex with air saturated buffer (right). Representative data from three replicates are shown. (b) Traces extracted from the multi-wavelength absorption data at each of three indicated wavelengths for each of the three reactions shown in (a). The black, red and blue traces are data from the left, middle, and right reactions, respectively. The $k_{obs}$ from the exponential fit for change in absorbance at 326 nm (black trace), which shows a lag phase of $\sim$7 s, was 3.60 ± 0.03 min$^{-1}$. Assays were conducted in triplicate, and data are reported as mean ± standard deviation, with $n = 3$. 

Supplementary Results

Supplementary Figure 15

**Supplementary Figure 15.** Global analysis of the reaction of Ind4 with L-arginine using SVD and multi-step models. The data was collected by mixing equal volumes of Ind4 (60 µM) and L-arginine (50 µM) solutions using a stopped-flow apparatus. The left figure shows fitted spectra shown in Fig 5a (top row left) but after excluding the species D. The inset figure shows the region between 475-650 nm. The figure on right shows the fitted spectra obtained by fitting the data shown in Fig. 5a (top row left) but with using a three-step model for first 2 s of the reaction \((k_1 \ 132 \pm 6 \text{ min}^{-1}, \ k_2 \ 34 \pm 0.2 \text{ min}^{-1})\). Assays were conducted in triplicate, and data are reported as mean ± standard deviation, with \(n = 3\).
Supplementary Figure 16. Mechanistic proposal for the Ind4 reaction. The L-arginine substrate is coupled to the PLP cofactor through a Schiff base interaction (EA). Deprotonation at the alpha carbon then shuttles electrons into the cofactor, giving a quinonoid intermediate (Q1) (511 nm). Reaction with $O_2$ leads to formation of a hydroperoxide intermediate (H1). Subsequent elimination of $H_2O_2$ generates a PLP derivative of an enamine (PE), which can be deprotonated at C4, with electron density again shuttling through a conjugated chain into the cofactor, giving a second quinonoid intermediate (Q2) (567 nm). Activation of oxygen, formation of the hydroperoxide (H2), and elimination of $H_2O_2$ can then generate a product-PLP complex (PP) (326 nm). Finally, the active site lysine can displace the aldimine, releasing the product 10 and regenerating the enzyme-PLP Schiff base (IA) (417 nm).
Supplementary Results

References: