ACS SyntheticBiology

Letter

pubs.acs.org/synthbio

Refined Regio- and Stereoselective Hydroxylation of L-Pipecolic Acid by Protein Engineering of L-Proline *cis*-4-Hydroxylase Based on the X-ray Crystal Structure

Kento Koketsu,*[,†] Yasuhito Shomura,[‡] Kei Moriwaki,[†] Mikiro Hayashi,[†] Satoshi Mitsuhashi,[†] Ryotaro Hara,[§] Kuniki Kino,[§] and Yoshiki Higuchi[‡]

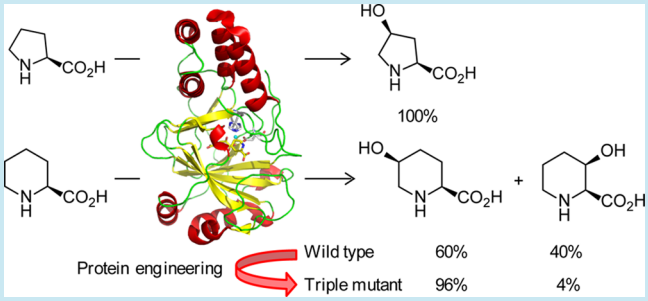
[†]Bioprocess Development Center, Kyowa Hakko Bio Co., Ltd., Tsukuba, Ibaraki 305-0841, Japan

[‡]Graduate School of Life Science, University of Hyogo, Koto 3-2-1, Kamigori-cho, Ako-gun, Hyogo 678-1297, Japan

[§]Department of Applied Chemistry, Faculty of Science and Engineering, Waseda University, Ohkubo 3-4-1, Shinjuku-ku, Tokyo 169-8555, Japan

Supporting Information

ABSTRACT: Enzymatic regio- and stereoselective hydroxylation are valuable for the production of hydroxylated chiral ingredients. Proline hydroxylases are representative members of the nonheme Fe^{2+}/α-ketoglutarate-dependent dioxygenase family. These enzymes catalyze the conversion of L-proline into hydroxy-L-prolines (Hyps). L-Proline *cis*-4-hydroxylases (*cis*-P4Hs) from *Sinorhizobium meliloti* and *Mesorhizobium loti* catalyze the hydroxylation of L-proline, generating *cis*-4-hydroxy-L-proline, as well as the hydroxylation of L-pipecolic acid (L-Pip), generating two regioisomers, *cis*-5-Hypip and *cis*-3-Hypip. To selectively produce *cis*-5-Hypip without simultaneous production of two isomers, protein engineering of *cis*-P4Hs is required. We therefore carried out protein engineering of *cis*-P4H to facilitate the conversion of the majority of L-Pip into the *cis*-5-Hypip isomer. We first solved the X-ray crystal structure of *cis*-P4H in complex with each of L-Pro and L-Pip. Then, we conducted three rounds of directed evolution and successfully created a *cis*-P4H triple mutant, V97F/V95W/E114G, demonstrating the desired regioselectivity toward *cis*-5-Hypip.



KEYWORDS: *hydroxyproline, hydroxy-L-pipecolic acid, hydroxylation, proline hydroxylase, protein engineering*

Hydroxy-L-prolines (Hyps) are nonproteinogenic hydroxylated derivatives of L-proline, which include four stereoisomers: *trans*-3-, *trans*-4-, *cis*-3- and *cis*-4-hydroxy-L-proline. Hyps have been used not only as dietary supplements[1] but also as versatile chiral building blocks for synthesizing pharmaceutical drugs such as carbapenem antibiotics, angiotensin-converting inhibitors and an antispastic agent.[2,3] Hydroxyl groups of the rigid ring of L-proline enables further modification, thereby expanding the versatility of Hyps as pharmaceutical ingredients toward, for example, the generation of halogenated analogs.[4] Efficient synthetic processes for the production of Hyps have been achieved by enzymatic regio- and stereoselective hydroxylation of L-proline using proline hydroxylases. Proline hydroxylases are members of the nonheme Fe^{2+}/α-ketoglutarate-dependent dioxygenase family. These enzymes introduce a hydroxyl group to the inactive carbon of the substrate, via the highly reactive ferryl (Fe^{IV}=O) species generated from Fe^{2+}, α-ketoglutarate (α-KG) and O_2 (Figure 1A).[5,6] Shibasaki and co-workers developed the process for *trans*-4-Hyp production using microbial L-proline-*trans*-4-hydroxylase (*trans*-P4H) isolated from *Dactylosporangium* sp. strain RH1,[7−9] while Mori and co-workers developed the process for *cis*-3-Hyp production using L-proline-*cis*-3-hydroxylases (*cis*-P3Hs), involving two homologues (type I and type II) isolated from *Streptomyces* sp. strain TH1.[10] Recently, Hara and Kino reported that the L-proline-*cis*-4-hydroxylases (*cis*-P4Hs), identified from the genomes of *Sinorhizobium meliloti* and *Mesorhizobium loti*, convert L-proline into *cis*-4-L-Hyp[11] (Figure 1B). These *cis*-P4Hs are named SmP4H and MlP4H, respectively, and display similar catalytic properties.

Hydroxy-L-pipecolic acids (Hypips) are the hydroxylated derivatives of L-pipecolic acid (L-Pip), a nonproteinogenic 6-membered cyclic amino acid. Similar to Hyps, Hypips are expected to prove chiral building blocks in the synthesis of novel pharmaceutical drugs. In previous reports, it was shown that proline hydroxylases can catalyze the hydroxylation of L-Pip as well as L-Pro.[12,13] For example, *trans*-P4H could convert L-Pip into *trans*-5-Hypip while *cis*-P3H converts L-Pip into *cis*-3-Hypip. Conversely, regio- and stereoselective hydroxylation for other Hypip isomers has not yet been developed. One of the isomers, *cis*-5-Hypip, is expected to be a novel chiral material for drug synthesis such as other Hypip isomers. In a previous

Received: May 11, 2014
Published: August 29, 2014

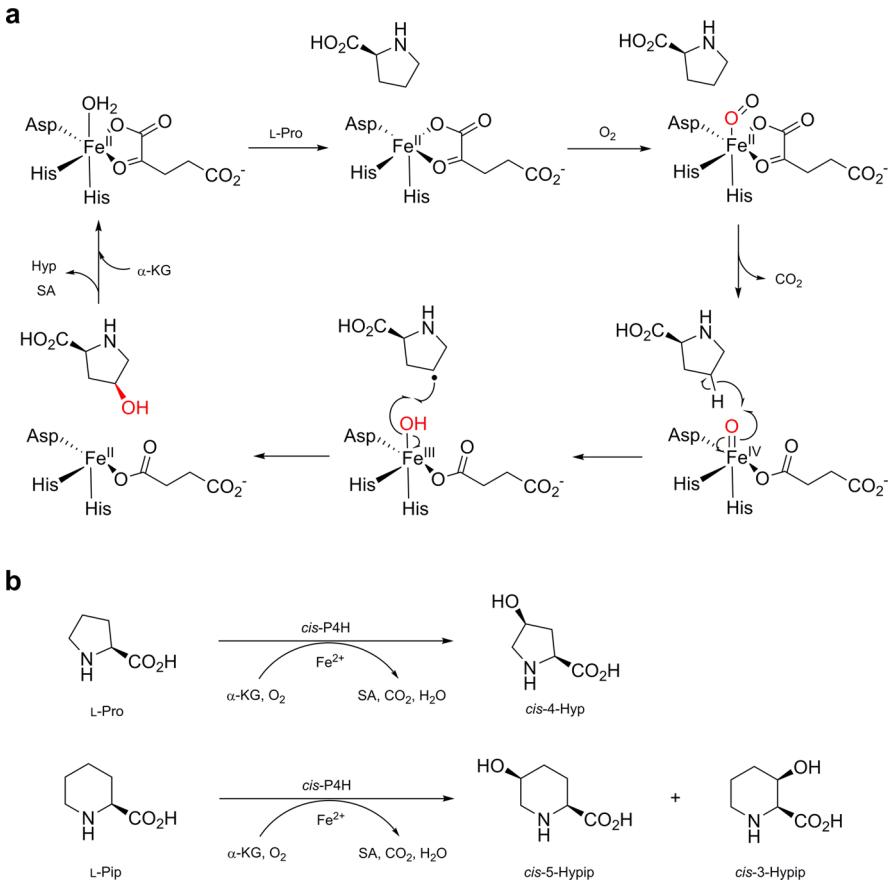


Figure 1. *cis*-Proline-4-hydroxylase (*cis*-P4H)-catalyzed reactions. (A) Proposed catalytic mechanism of *cis*-P4H. *cis*-P4H decomposes α-KG into SA and carbon dioxide by utilizing dioxygen to generate a highly reactive ferryl (Fe^{IV}=O) species. This potent oxidant abstracts a hydrogen radical from C4 carbon of L-proline generating a Fe^{III}—OH intermediates and the L-proline radical. Radical coupling of OH group of a Fe^{III}—OH species and the L-proline radical forms *cis*-4-Hyp and regenerates the starting Fe^{II} oxidation state. (B) *cis*-P4Hs from *Sinorhizobium meliloti* or *Mesorhizobium loti* catalyze regio- and stereoselective hydroxylation of L-Pro generating *cis*-4-Hyp and also catalyze hydroxylation of L-Pip providing two regioisomers, *cis*-5-Hypip and *cis*-3-Hypip. *cis*-4-Hyp, *cis*-4-hydroxy-L-proline; α-KG, α-ketoglutarate; SA, succinic acid.

report, it was demonstrated that *cis*-P4H from *S. meliloti* catalyzes the hydroxylation of L-Pip, thereby generating two regioisomers, *cis*-5-Hypip and *cis*-3-Hypip, at an approximate 1:1 ratio in *in vitro* experiments.[13] Since the production of pure isomers is required, the simultaneous generation of two isomers is undesirable because the undesired *cis*-3-Hypip must be removed during the purification step and its generation is waste of the substrate. Therefore, to produce *cis*-5-Hypip exclusively, protein engineering of *cis*-P4H is required. For protein engineering to be carried out, structural information on the target protein is critical; however, the X-ray structure of *cis*-P4H has never been solved.

In this study, we first solved the structure of MlP4H in complex with Co^{2+}, α-KG as a cofactor and L-Pro or L-Pip as substrates. Based on the resulting structural information, we carried out directed evolution of SmP4H by *in vivo* production assays. After a series of screening assays, we successfully created a *cis*-P4H triple mutant capable of selectively converting L-Pip into *cis*-5-Hypip in a regio- and stereoselective manner, with minimal *cis*-3-Hypip production. Although wild-type SmP4H was observed to produce *cis*-5-Hypip at a purity of 60% *in vivo*, the SmP4H triple mutant improved selective *cis*-5-Hypip production to greater than 95% purity. In addition, we verified the *in vitro* activity of the generated SmP4H mutants.

RESULTS AND DISCUSSION

To evaluate the *in vivo* hydroxylation of L-Pip by SmP4H and MlP4H, we constructed the expression plasmids pTrc99A-SmP4H and pTrc99A-MlP4H, respectively (Supporting Information Table S1 and S2). *E. coli* W3110 strains[14] transformed with pTrc99A-SmP4H or pTrc99A-MlP4H were grown in culture media containing 1.0 g/L L-Pip at 30 °C for 24 h and the supernatants of the resulting cultures were analyzed by HPLC after derivatization with 9-fluorenylmethyl chloroformate (Fmoc-Cl).[15] Based on HPLC and LC-MS analyses, we confirmed the production of the desired *cis*-5-Hypip, and the *cis*-5/*cis*-3 ratio was observed to be approximately 2:1 and 1:7 in the reactions by SmP4H and MlP4H, respectively (Supporting Information Figures S1, S2, and Table S3). This remarkable difference in regioselectivity of hydroxylation for L-Pip between SmP4H and MlP4H was quite surprising, considering that these *cis*-P4Hs have a relatively high sequence similarity (identity, 66%; similarity, 91%). To develop selective production of *cis*-5-Hypip over *cis*-3-Hypip, protein engineering of *cis*-P4H was required. However, the X-ray crystal structure of the target enzyme has never solved. Thus, to understand the molecular mechanisms of regioselectivity of L-Pip hydroxylation by *cis*-P4Hs, and to apply this information to our protein engineering efforts, we attempted to solve the crystal structure

dx.doi.org/10.1021/sb500247a | *ACS Synth. Biol.* 2015, 4, 383−392

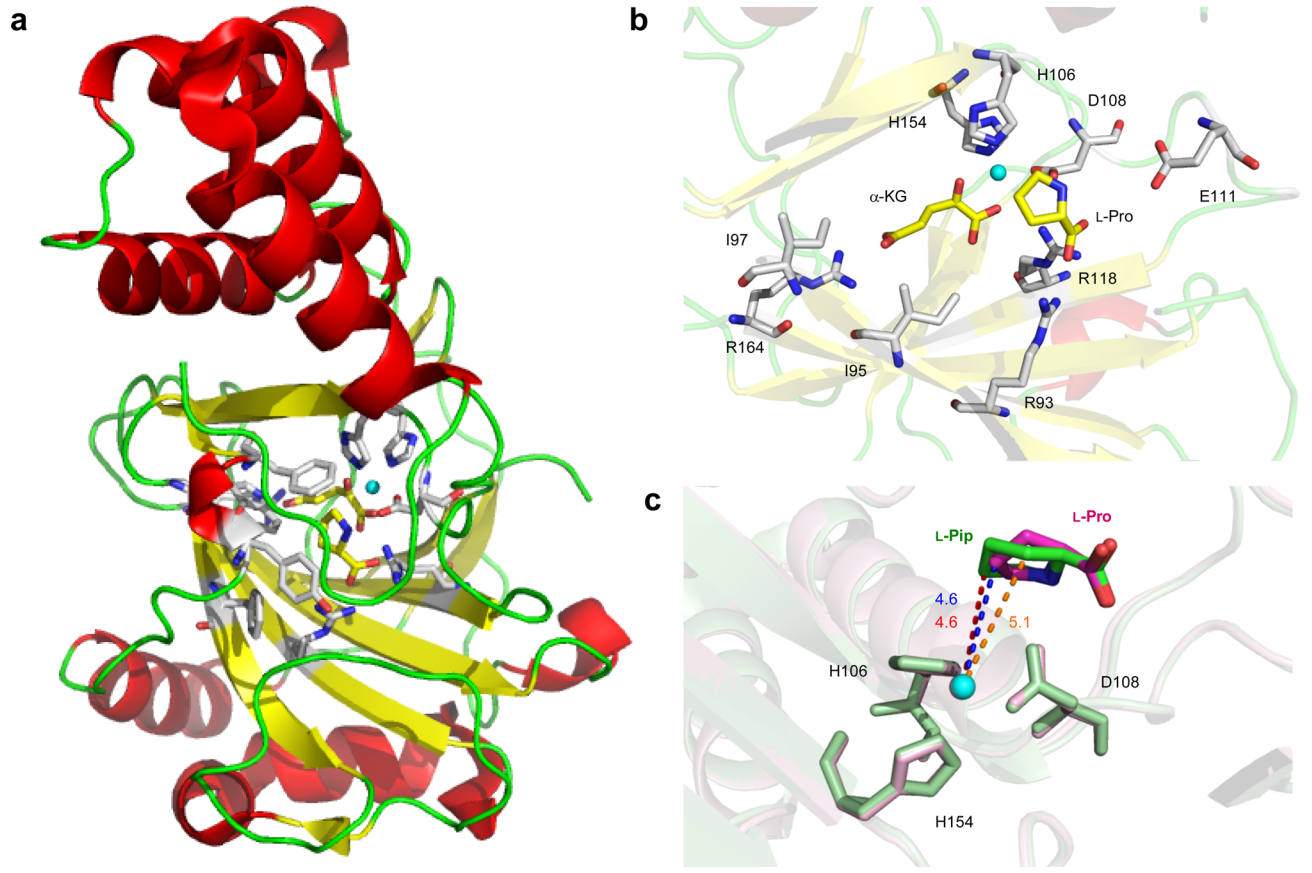


Figure 2. Crystal structure of MlP4H in complex with Co^{2+} or α-KG and L-Pro or L-Pip. (A) Overall structure of MlP4H. The structure shows a distorted jelly roll β-sheet core (yellow), which is sandwiched by the N-terminal and C-terminal α-helical domains (red). Co^{2+}, α-KG, and L-Pro are positioned in the center of the β-sheet core. (B) View of the MlP4H active site. Active site Co^{2+} (cyan) is coordinated by H106, H154, and D108 (silver sticks). L-Pro and α-KG (yellow sticks) are also shown. (C) A superposed image of L-Pro and L-Pip bound MIP4H. The C4 carbon of L-Pro positions at a distance of 4.6 Å from the Co^{2+} atom. The C5 and C3 carbons of L-Pip locate at distances of 4.6 and 5.1 Å, respectively, from the Co^{2+} atom.

of *cis*-P4H in complex with Fe^{2+}, α-KG as cofactors, and L-Pro or L-Pip as substrates.

The protein production of MlP4H and SmP4H was performed with exactly the same method, but crystals could only be obtained for the former. The crystals were prepared in the presence of L-Pro or L-Pip as a substrate, with α-KG as a cofactor, and Co^{2+} as a mimic of the catalytic metal center because of its stability under aerobic conditions and its enzymatic inactivity under anaerobic conditions. The crystal structure of L-Pro-bound MlP4H was first determined by the single-wavelength dispersion method with the bound Co^{2+} at the active site used as the anomalous scatter, while the other two structures (L-Pip-bound form and the L-Pro-bound form at a higher resolution) were determined by molecular replacement using the first L-Pro-bound structure as a search model.

The overall structure of MlP4H (Figure 2A) was closely similar to that of *cis*-P3H (type II) from *Streptomyces* sp. TH1, which has been previously reported.[16] However, to the best of our knowledge, this is the first example in complex with Co^{2+} (as a substitute for Fe^{2+}), α-KG as the cofactor, and L-Pro as the substrate, within the class of microbial L-proline hydroxylases (the crystal structure of *cis*-P3H did not include L-Pro as a substrate nor α-KG as a cofactor in the active site). The active site was composed of a distorted jelly roll β-sheet core, which was sandwiched by the N-terminal and C-terminal α−helical domains. Co^{2+} was coordinated by residues H106, H154, and D108, at distances of 2.14, 2.10, and 2.03 Å, respectively, which comprise the HxD/H motif conserved among Fe^{2+}/α-KG-dependent dioxygenases (Figure 2B). α-KG forms a salt bridge with R164, while the oxygen atoms of the keto acid moiety coordinate with Co^{2+}. L-Pro locates in the center of the jelly roll β-sheet core near Co^{2+} and the opposite face of L-Pro is surrounded by aromatic amino acid residues, Y32, F35, W40, and F42. The amino group of L-Pro forms a salt bridge with D108 at 2.84 Å and appears to form a hydrogen bond with E111 via a water molecule. The carboxylic group of L-Pro forms a salt bridge with R93 and R118 at distances of 2.76 and 2.92 Å, respectively. The *cis*-face of the C4 carbon of L-Pro is properly oriented toward Co^{2+}, which in nature exists as Fe^{IV}=O during the hydroxylation reaction, to generate the enantiopure *cis*-4-Hyp.

We also obtained the crystal structure in complex with L-Pip instead of L-Pro, and this provided us with an answer to controlling the regioselectivity of hydroxylation by *cis*-P4Hs, so as to selectively generate *cis*-5-Hypip over *cis*-3-Hypip (Figure 2c). L-Pip has a 6-membered ring and exists in a chair conformation in the active site of the enzyme. Because the C5 or C3 carbons of L-Pip lie in close proximity to the active site containing Co^{2+} at distances of 4.6 and 5.1 Å, respectively, it was apparent that hydroxylation would therefore occur on both carbons. It should also be noted that *cis*-4-Hypip has never been produced by *cis*-P4Hs owing to the length of the distance of the

dx.doi.org/10.1021/sb500247a | *ACS Synth. Biol.* 2015, 4, 383−392

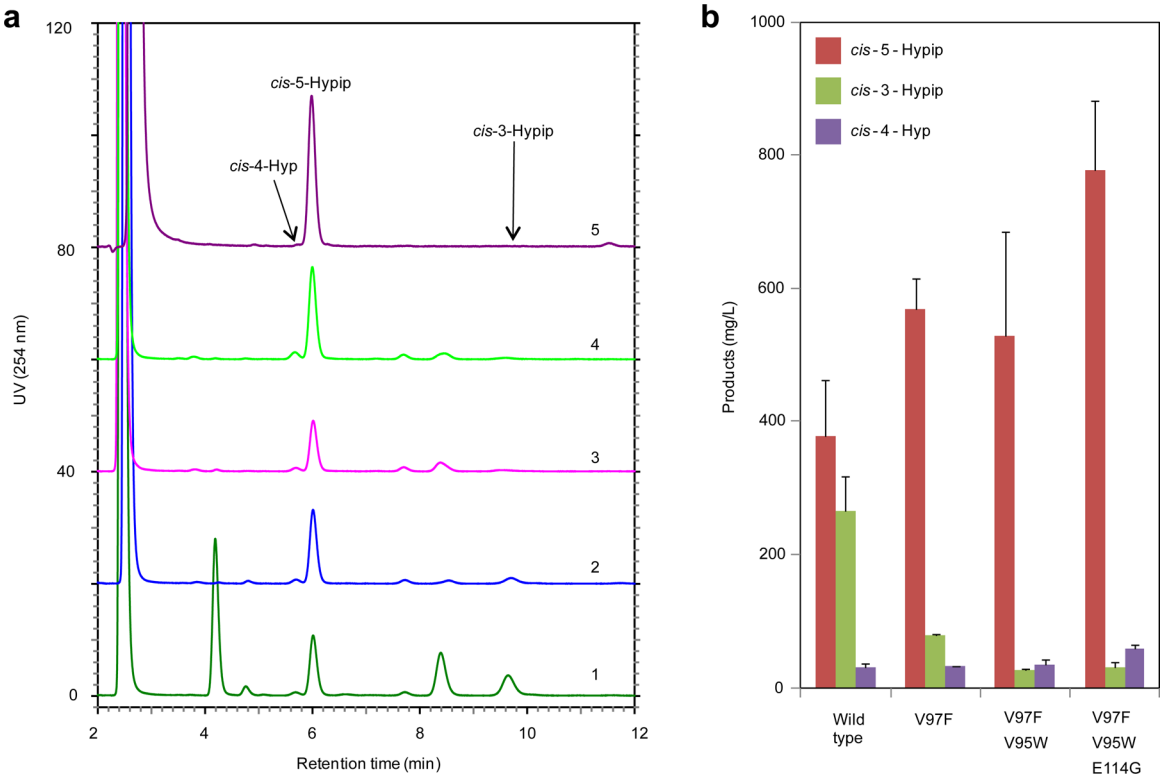


Figure 3. *In vivo* production of *cis*-5-Hypip and *cis*-3-Hypip using *E. coli* for expression of SmP4H and its mutant variants, V97F, V97F/V95W and V97F/V95W/E114G. *E. coli* W3110 *ilvG*$^+$ *pyrE*$^+$ strains harboring pTrS31-SmP4H or its variants are cultured at 30 °C for 48 h in 2.0 g/L L-Pip-containing medium. Hydroxylated products involved in the culture supernatants were analyzed. (A) HPLC chromatogram of Fmoc derivatives. Lane 1, SmP4H (wild-type); lane 2, SmP4H (V97F); lane 3, SmP4H (V97F/V95W); lane 4, SmP4H (V97F/V95W/E114G); lane 5, standards of *cis*-5-Hypip. *cis*-4-Hyp, *cis*-5-Hypip, and *cis*-3-Hypip are eluted at 5.7, 6.0, and 9.7 min, respectively. (B) The bar graph representation of *cis*-5-Hypip, *cis*-3-Hypip, and *cis*-4-Hyp production by wild-type SmP4H and its variants *in vivo*. The values of the Hypip productivities are listed in Supporting Information Table S7. Error bars represent the standard deviation of three independent experiments.

C4 carbon from the Co^{2+} atom. On the basis of this structural information, we hypothesized that the hydroxylation ratio could be changed by altering the orientation and proximity of L-Pip in the active site by effective mutations of amino acid residues within the active site. Although we were unable to obtain the SmP4H crystal structure in complex with L-Pip, reliable modeling was possible owing to the high sequence similarity between SmP4H and MlP4H (Supporting Information Figure S3). Thus, we initiated protein engineering of SmP4H to create a L-Pip *cis*-5-hydroxylase with higher selectivity toward *cis*-5-Hypip over *cis*-3-Hypip.

Among the amino acid residues located near L-Pip in the active site, R93 forms a tight interaction with the carboxyl moiety of L-Pip. We thought that changing the amino acid residues around R93 may affect the reactivity of L-Pip by SmP4H, and we first selected the V95 and V97 residues and substituted these with alanine residues. The expression plasmids of three SmP4H variants, V95A, V97A, and V95A/V97A, were constructed as pTrc99A-SmP4H (V95A), pTrc99A-SmP4H (V97A), and pTrc99A-SmP4H (V95A/V97A), respectively. *E. coli* W3110 strains harboring the constructed plasmids were cultivated in media containing 1.0 g/L L-Pip at 30 °C for 30 h, and the resulting culture supernatants were analyzed by HPLC after derivatization with Fmoc-Cl. Surprisingly, the strain expressing the SmP4H (V97A) variant produced 221 mg/L of *cis*-5-Hypip, which is higher than the 89 mg/L produced by the wild-type, despite their similar *cis*-5/*cis*-3 ratio of around 2:1 (Supporting Information Table S3). Conversely, the V95A mutant and V95A/V97A double mutant did not show any increase in hydroxylation activity nor any improvement in the *cis*-5/*cis*-3 ratio.

Because the V97A variant showed a good response, for the first round of protein engineering, we constructed various mutants of the V97 residue by substitution with other amino acid residues and then examined the productivity and *cis*-5/*cis*-3 ratio of Hypips by these mutants (Supporting Information Table S4). Based on *in vivo* production assays, two distinct mutant types were obtained: One was the V97C variant, which, as did the V97A mutant, showed an approximately 4-fold increase in hydroxylation activity, yielding 255 mg/L of *cis*-5-Hypip and 85 mg/L of *cis*-3-Hypip, respectively. Another distinct mutant type was the V97F and V97Y variants. The V97F mutant improved the regioselectivity of hydroxylation: the *cis*-5/*cis*-3 ratio improved from 1.4 (for the wild-type enzyme) to 5.3, whereas the total Hypip amount produced was comparable at 112 mg/L. The V97Y mutant showed better regioselectivity (*cis*-5/*cis*-3 = 9.0), but productivity was rather decreased compared with that of the V97F mutant.

Although the V97F mutant demonstrated higher selectivity of C5-hydroxylation, the selectivity and hydroxylation activity were not sufficient; thus, further protein engineering was required. Considering that improvements in regioselectivity would likely prove more difficult than attempts to increase the activity, we selected the V97F variant as the base enzyme for the next round of protein engineering rather than the V97A and

dx.doi.org/10.1021/sb500247a | *ACS Synth. Biol.* 2015, 4, 383−392

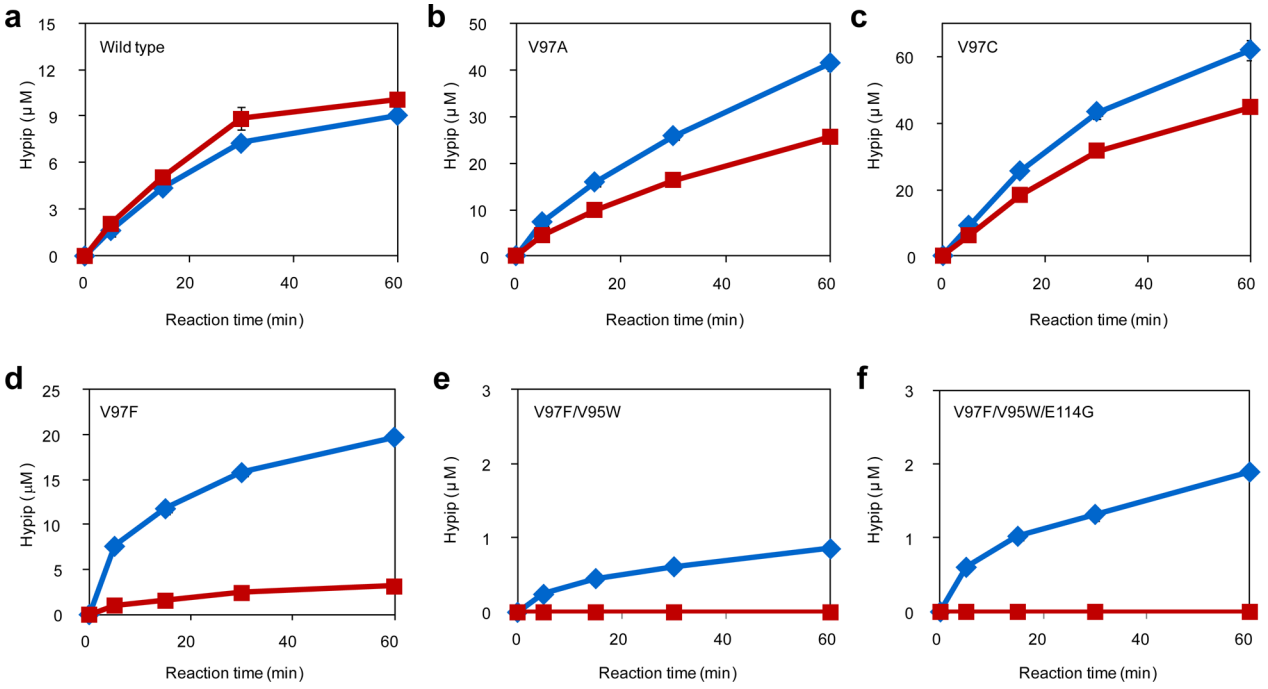


Figure 4. *In vitro* hydroxylation reactions for L-Pip catalyzed by SmP4H and its variants. *cis*-5-Hypip (blue) and *cis*-3-Hypip (red) production are indicated. (A) SmP4H, (B) V97A variant, (C) V97C variant, (D) V97F variant, (E) V97F/V95W variant, (F) V97F/V95W/E114G variant. In the reactions catalyzed by the double and triple mutants, *cis*-3-Hypip production was below the level of detection. Values were obtained from three independent experiments.

V97C variants. The V97Y mutant was an alternative candidate for this purpose, but the activity of this mutant was too low. We adopted a stepwise strategy for generating the desired hydroxylase from the V97F variant: (1) Improving the regioselectivity to >95% purity of *cis*-5-Hypip over *cis*-3-Hypip, even if the enzyme activity is reduced. (2) Increasing the activity based on the regioselectivity-improved variant. At this point, we changed the host to *E. coli* W3110 *ilvG*$^+$*pyrE*$^+$ and the vector to pTrS31. This vector includes a constitutive Trp promoter as used in the production of *trans*-4-Hyp in a previous study.[7] These changes resulted in increased regioselectivity and productivity (Supporting Information Table S5). Although the reason for this enhanced regioselectivity remains unclear, regarding to increased productivity, we assumed that improved growth of W3110 *ilvG*$^+$*pyrE*$^+$ strain should increase the supply of α-KG, a cosubstrate of *cis*-P4Hs. To demonstrate this hypothesis, we analyzed the amount of α-KG involved in the culture supernatants of W3110 and W3110 *ilvG*$^+$*pyrE*$^+$ strains, which do not have the plasmids. As expected, W3110 *ilvG*$^+$*pyrE*$^+$ strain accumulated α-KG at the amount of 0.53 ± 0.02 mM, while W3110 strain did not accumulate the detectable amount of α-KG (Supporting Information Table S6).

In the second round of protein engineering, we used saturated mutagenesis of eight amino acid residues (I94, V95, S98, E99, N100, A101, C102, I103) selected near the V97 residue in the active site. Mutagenesis libraries were constructed and the resulting colonies were inoculated into culture media containing 2.0 g/L L-Pip. This was followed by HPLC analysis according to the procedure described above. After a series of screening assays of 100 colonies per mutant at each of the eight amino acid residues, we discovered that the V95W variant (GTC → TGG) of SmP4H (V97F) significantly reduced the amount of *cis*-3-Hypip (<5%) production compared with that of *cis*-5-Hypip (Figure 3 and Supporting Information Table S7). Whereas the hydroxylation activity was decreased compared with the V97F single mutant, this V97F/V95W double mutant achieved the desired *cis*-5/*cis*-3 ratio. We therefore moved to the second stage of our strategy, which involved increasing the hydroxylation activity of the regioselectivity-improved mutant, SmP4H (V97F/V95W).

In the third round of protein engineering, we constructed a random mutagenesis library of SmP4H (V97F/V95W) using the error-prone PCR method. As a rough but simple screening procedure, we used thin layer chromatography (TLC)-based screening. Colonies harboring mutant library plasmids were cultivated in L-Pip-containing medium using 96-deep well plates, and the subsequent level of Hypip production was estimated by TLC with visualization using ninhydrin spray. After screening 2000 colonies, 14 positive clones were obtained and further tested using the same HPLC-based assay described above. One of the 14 clones showed approximately 1.5 fold higher production of *cis*-5-Hypip at the same *cis*-5/*cis*-3 ratio observed for the initial V97F/V95W double mutant (Figure 3 and Supporting Information Table S7). The responsible mutation, E114G, was identified by sequence analysis (GAG → GGG). Finally, we compared the Hypips productivity and the regioselectivity between SmP4H (wild-type) and three constructed mutants using pTrS31 vector and W3110 *ilvG*$^+$*pyrE*$^+$ strain or W3110 strain as hosts. As shown in Table S7, W3110 strains harboring pTrS31-SmP4H (or its mutants) resulted in less growth compared to W3110 *ilvG*$^+$*pyrE*$^+$ strains and the productivities of Hypips were also low. On the other hand, W3110 *ilvG*$^+$*pyrE*$^+$ strains harboring pTrS31-SmP4H (or its mutants) showed higher growths and relatively high productivities of Hypips. Particularly, W3110 *ilvG*$^+$*pyrE*$^+$ strain harboring pTrS31-SmP4H (V97F/V95W/E114G) demonstrated the highest growth and the productivity

dx.doi.org/10.1021/sb500247a | *ACS Synth. Biol.* 2015, 4, 383−392

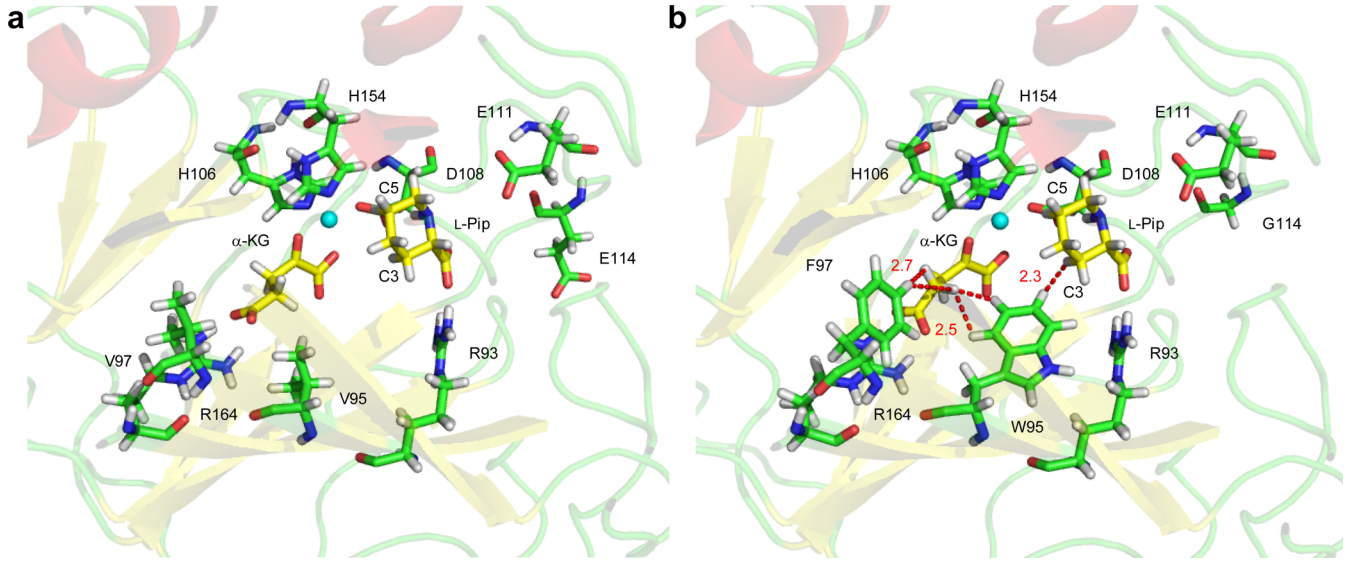


Figure 5. Active site views of the energy minimized homology models of SmP4H and its triple mutant. Modeling procedure was described in method. (A) Homology model of SmP4H (wild-type) constructed based on the MlP4H crystal structure. Instead of I95, I97, and E114 residues of MlP4H, V95, V97, and G114 exist in the active site of SmP4H. Other amino acids displayed in the figure are conserved between MlP4H and SmP4H. (B) Homology model of the SmP4H triple mutant (V97F/V95W/E114G) constructed based on the MlP4H crystal structure. The hydrogen atom of the C4 carbon of F97 locates near the hydrogen atom of the C4 carbon of α-KG at a distance of 2.7 Å. The hydrogen atoms of the C4 and C5 carbons of W95 locate near the C4 carbon of α-KG at distances of 2.5 and 2.4 Å, respectively. The hydrogen atom of the C6 carbon of W95 positions close to the hydrogen of the C3 carbon of L-Pip substrate at a distance of 2.3 Å.

of *cis*-5-Hypip in a regioselective manner (Supporting Information Table S7).

To characterize the created SmP4H variants in detail, *in vitro* hydroxylation assays were carried out. Wild-type SmP4H and five variants thereof (V97A, V97C, V97F, V97F/V95W, and V97F/V95W/E114G) were expressed in *E. coli* BL21 (DE3) as hexahistidine-tagged proteins using the pET-21d(+) vector. The wild-type and single mutant variant proteins were successfully expressed and purified at sufficient yields (~1.4 mg from 100 mL culture); however, the double and triple mutants gave low yields due to the low solubility (Supporting Information Figure S4). Due to substitution of valine residues with bulky aromatic amino acid residues in the active site, the stabilities of these mutants might be decreased. The *in vitro* enzyme assays were then carried out according to methods previously reported.[11] The results of the *in vitro* reactions were comparable to those of the *in vivo* reactions (Figure 4). The specific activities of the wild-type enzyme were 0.29 (*cis*-5) and 0.34 (*cis*-3) at a 1:1 ratio of *cis*-5/*cis*-3. The V97A and V97C mutants showed greatly increased hydroxylation activities; however, a significant amount of *cis*-3-Hypip was also produced (V97A: 1.05 (*cis*-5), 0.65 (*cis*-3). V97C: 1.69 (*cis*-5), 1.22 (*cis*-3)). Conversely, as observed in the *in vivo* assays, the V97F variant markedly improved the ratio of *cis*-5/*cis*-3 production (0.79 (*cis*-5) and 0.10 (*cis*-3)), and the addition of the V95W mutation further improved this ratio (0.03 (*cis*-5) and 0 (*cis*-3)), despite the significantly lowered activity. Although the double mutants showed reduced activity, the addition of the E114G mutation improved the activity approximately 2-fold. The *in vitro* activities of the double and triple mutants were rather low compared with those of *in vivo* production, possibly owing to the stability conferred to the enzymes by the cellular environment contributing to higher *cis*-5-Hypip productivity. Based on the *in vitro* enzymatic assays, we confirmed the successful generation of a *cis*-5-selective L-Pip hydroxylase by protein engineering of SmP4H.

The positive effect of the obtained mutations was considered from the energy minimized homology model of SmP4H and its triple mutant (Figure 5). The V97 is located near α-KG and V95 is located near both of α-KG and the C3 carbon of L-Pip (Figure 5A). As shown in Figure 5B, a distance between the hydrogen atom of the C4 carbon of F97 and the hydrogen atom of the C4 carbon of α-KG are 2.7 and 3.3 Å, respectively, and a distance between the hydrogen atom of F97 and the hydrogen atom of the C3 carbon of L-Pip is over 8.5 Å. Distances between the hydrogen atoms of the C4 and C5 carbons of W95 and the hydrogen atom of the C4 carbon of α-KG are 2.5 and 2.4 Å, respectively. Taking into account these observations, substitution of V97 with a bulky Phe residue seems to affect α-KG rather than L-Pip, and substitution of V95 residue with a bulky Trp residue appears to affect both L-Pip and α-KG. The mechanism of the positive effect of the F97 mutation supposed to arise from changing the binding mode of α-KG. Actually, α-KG binding mode of the energy minimized homology model of triple mutant appears to be slight different compared to that of wild type. W95 is also assumed to affect α-KG binding mode, synergistically. The shift of α-KG binding state may allow the change of the highly reactive Fe^{IV}=O species during the hydroxylation reaction; however, we could not demonstrate this hypothesis at this point. Additionally, W95 should affect the hydroxylation reaction of C3 carbon of L-Pip due to a close distance at 2.3 Å between the hydrogen atom of the C6 carbon of W95 and the hydrogen atom of the C3 carbon of L-Pip. We hypothesized that dual effect by F97 and W95 mutations allows the remarkable improvement of the regioselectivity of L-Pip hydroxylation by SmP4H. These substitutions are also likely to occupy the majority of the space in the active site of SmP4H and result in destabilization of the enzyme. The E114 residue is positioned near the E111 residue, which interacts with the amino group of L-Pip opposite to V95 and V97. Substituting E114 with the glycine residue may partly liberate the constricted state of the active site of the

dx.doi.org/10.1021/sb500247a | *ACS Synth. Biol.* 2015, 4, 383−392

Table 1. Data Collection, Phasing, and Refinement Statistics

data set	pro-bound	pip-bound	pro-bound (for SAD phasing)
		Crystal Params.	
space group	$P2_122_1$	$C222_1$	$P2_122_1$
cell dimensions (Å)			
a	55.39	112.68	57.43
b	73.61	123.40	73.30
c	95.04	56.98	94.88
		Data Collection	
wavelength (Å)	0.9000	0.9000	1.5900
resolution range (Å)[a]	19.96−1.80 (1.90−1.80)	19.87−1.30 (1.37−1.30)	19.34−2.80 (2.95−2.80)
total reflections	252 853 (30 583)	654 893 (83 927)	145 483 (21 149)
unique reflections[a]	36 529 (5251)	97 048 (14 030)	10 341 (1481)
R_{merge}[a,b]	0.068 (0.535)	0.055 (0.481)	0.099 (0.452)
$I/\sigma(I)$[a,c]	6.9 (1.5)	6.7 (1.6)	7.4 (1.7)
completeness[a]	0.999 (1000)	1.000 (1.000)	0.997 (1.000)
multiplicity[a]	6.9 (5.8)	6.7 (5.9)	14.1 (14.3)
Refinement			SAD phasing
resolution range (Å)	19.96−1.80	19.87−1.30	No. of Co sites 1
R-factor/free *R*-factor[d,e]	0.190/0.216	0.135/0.162	mean figure of merit after density modification
atoms in an asymmetric unit			acentric 0.69
protein	2263	2445	centric 0.64
ligand	19	68	
water	137	342	
deviations from ideal geometry			
bonds distances (Å)	0.011	0.009	
angles (deg)	1.468	1.415	
Ramachandran plot[f]			
favored	0.989	0.989	
allowed	0.011	0.011	

[a]Values in parentheses are for highest resolution shells. [b]$R_{merge} = \sum_{hkl}\sum_i(|I_i(hkl) - \langle I(hkl)\rangle|)/\sum_{hkl}\sum_i I_i(hkl)$. [c]Signal-to-noise ratio of intensities. [d]$R = \sum(|F_0 - F_c|)/\sum F_0$. [e]Five percent of reflections were randomly chosen for calculation of the free *R* value. [f]Values from Rampage.

V97F/V95W double mutant, thereby resulting in the observed increase in hydroxylation efficiency.

In summary, we solved the crystal structure of *cis*-P4H in complex with Co^{2+}, α-KG, L-Pro and L-Pip. This is the first example of a microbial proline hydroxylase structure in complex with substrates, and it may provide further insights into the regio- and stereoselectivity of the hydroxylation reaction. Based on the insights obtained from the X-ray crystal structure data, we applied protein engineering together with directed evolution to generate a refined regio- and stereoselective L-Pip hydroxylase from *cis*-P4H. For this purpose, we adopted a stepwise strategy comprising:[1] first increasing enzyme regioselectivity, and then[2] improving enzyme activity. This strategy resulted in the generation of the desired L-Pip hydroxylase for the selective production of *cis*-5-Hypip over *cis*-3-Hypip. The activity of the triple mutant may not prove sufficient for industrial scale production, but further improvements, especially in terms of enzyme stability, can still be made by use of protein engineering.[17]

METHODS

Protein Crystallization and Structure Determination. The gene for MlP4H (Accession number: BAB52605) was amplified from pEMlP4H[10] by PCR with the primers 5′-AAGGAAGGCGCCACAACGCGGATATTGGGTGTGGTC-3′ (*Ehe*I site is underlined) and 5′-AAGGAAGGTACCTC-AATAAGTCATGACCTCGCCAGC-3′ (*Kpn*I site is underlined) and inserted into the *Ehe*I and *Kpn*I sites of a pASK-IBA5plus expression vector (IBA GmbH, Göttingen, Germany). As a result, the MlP4H protein used for crystallization included an extra (Met)-Ser-Ala-Trp-Ser-His-Pro-Gln-Phe-Gly-Lys-Gly-Ala strep-tag II peptide at its N-terminus, and the original start codon of the wild-type was replaced by the underlined alanine, followed by the second codon of the wild-type. The resulting expression vector, designated pASK-MlP4H, was used for the transformation of *E. coli* pG-KJE8/BL21 (Takara Bio, Otsu, Japan). The cells were cultured at 37 °C in Luria−Bertani (LB) medium supplemented with 100 μg/mL ampicillin and 50 μg/mL chloramphenicol. After the culture reached an OD_{600} of 0.6, expression was induced by the addition of anhydrotetracycline to 0.2 μg/mL, and a further, overnight incubation was carried out at 24 °C. The cells were harvested by centrifugation at 10 000*g*, and the pellets were frozen in liquid nitrogen and stored at −80 °C. Protein purification was initiated by resuspension of the thawed cells in phosphate-buffered saline followed by cell disruption by sonication on ice. All the subsequent purification procedures were performed at 4 °C. Cell debris and the membrane fraction were removed by ultracentrifugation at 185 000*g* for 45 min and the supernatant was loaded onto Strep-Tactin Superflow resin (Qiagen, Venlo, Netherlands). The resin was washed with 7.5 volumes of phosphate-buffered saline and the target protein was eluted with four volumes of 25 mM Tris-HCl (pH 7.4) supplemented with 2.5 mM dethiobiotin. The protein was further purified with a HiTrap Q HP anion exchange column (GE Healthcare, Buckinghamshire HP7 9NA, England) using a 0−400 mM NaCl gradient in 25 mM Tris-HCl (pH 7.4) and a Superdex 200 10/30 gel filtration column (GE Healthcare)

dx.doi.org/10.1021/sb500247a | ACS Synth. Biol. 2015, 4, 383−392

equilibrated with 10 mM Tris-HCl (pH 7.4) containing 50 mM NaCl.

Crystallization experiments were performed at 288 K with the sitting-drop vapor diffusion method using a CrystalQuick 96-well plate (Greiner bio-one, Frickenhausen, Germany).The crystallization drop of the L-Pro-bound form was prepared by mixing 1 μL of the protein solution containing 28 mg/mL protein, 2 mM $CoCl_2$, 10 mM α-KG, and 20 mM L-Pro with 1 μL of the reservoir solution containing 0.1 M Bis-Tris propane (pH 8.5), 0.2 M sodium malonate, and 25% (v/v) PEG 3350. The drop of the L-Pip-bound form was prepared by mixing 1 μL of the protein solution containing 28 mg/mL protein, 2 mM $CoCl_2$, 10 mM α-KG, and 20 mM L-Pip with 1 μL of the reservoir solution containing 0.1 M CAPS (pH 10.5), 0.1 M lithium sulfate, and 1.8 M ammonium sulfate. After reaching their maximum sizes, crystals of the L-Pro-bound and L-Pip-bound forms were soaked in the cryobuffers containing 2 mM $CoCl_2$, 8 mM α-KG, 16 mM L-Pro, 0.1 M Bis-Tris propane (pH 8.5), 0.2 M sodium malonate, 28% (v/v) PEG3350, and 10% (v/v) ethylene glycol, and 2 mM $CoCl_2$, 8 mM α-KG, 16 mM L-Pip, 0.1 M CAPS (pH 10.5), 0.1 M lithium sulfate, 2.0 M ammonium sulfate, and 16% Glycerol, respectively, followed by flash cooling with liquid nitrogen.

The diffraction data were collected at the SPring-8 beamline BL44XU facility (Osaka University). Crystals were maintained at 90 K by a nitrogen gas stream. All data were integrated and scaled with iMosflm[18] and SCALA with the CCP4 interface.[19] Phasing and automated model-building were performed with the PHENIX software suite[20] using single-wavelength anomalous dispersion data (λ = 1.5900 Å). Subsequent manual model-building and structure refinement were carried out with the programs Coot[21] and Refmac.[22] The phases for the diffraction data comprising the L-Pro-bound and L-Pip-bound forms (λ = 0.9000 Å) were determined by molecular replacement using MolRep.[19] The ϕ/ψ torsion angles of the protein backbone were assessed by Rampage.[23] Statistics for data collection, phasing, and refinement are summarized in Table 1. The atomic coordinates and experimental data have been deposited in the Protein Data Bank, under the accession codes 4P7W (L-Pro-bound form) and 4P7X (L-Pip-bound form).

Homology Modeling. The homology models of SmP4H and its variants were first prepared by the SWISS-MODEL server,[24] and the energy minimization was performed using the Amber 14 program package.[25] The topology files were created with ff4SB and gaff force field. The proteins were solvated with the TIP3PBOX water model with leaving some water molecules observed in the MlP4H structure. The systems were neutralized by the addition of sodium ions. The distances between Fe^{2+} and its ligand atoms were restrained with experimental values during the energy minimization.

Plasmid Construction and Generation of the Mutant Libraries. The SmP4H gene (accession number: CAC47686) was amplified using primer 1 and primer 2 from pESmP4H[11] (Supporting Information Table S1 and S2). The amplified gene was digested with *Eco*RI and *Hind*III and introduced into the pTrc99A vector (Invitrogen, Carlsbad, CA, U.S.A.) digested with the corresponding restriction enzymes and subsequently dephosphorylated with shrimp alkaline phosphatase (Takara Bio, Otsu, Japan) to obtain pTrc99A-SmP4H. To generate various point mutants of SmP4H, fusion PCR was carried out using the primers listed in Supporting Information Table S2. For example, construction of pTrc99A-SmP4H (V97A) was conducted as follows: in the first PCR, the left fragment was amplified from pTrc99A-SmP4H using primer 1 and primer 3, while the right fragment was amplified using primers 2 and 4. After purification, both fragments were mixed and overlap extension PCR was carried out using primers 1 and 2. Fused PCR fragments were digested with *Eco*RI and *Hind*III and introduced into the pTrc99A vector, as described above.

***In Vivo* Hydroxylation of L-Pip.** Single colonies from *E. coli* W3110 strains harboring pTrc99A-SmP4H (wild-type and mutants) were inoculated into 5 mL LB medium containing 100 mg/L ampicillin and cultured at 30 °C overnight. An inoculum of the resulting overnight culture (100 μL) was transferred into the production medium (5 mL) comprising 20 g/L glucose, 2 g/L $MgSO_4$, 16 g/L KH_2PO_4, 14 g/L K_2HPO_4, 2 g/L NH_4SO_4, 1 g/L citric acid, 5 g/L casamino acids (Difco, Franklin Lakes, NJ, U.S.A.), 50 mg/L $FeSO_4{\cdot}7H_2O$, 10 mg/L thiamine·HCl, and 10 mg/L $MnSO_4.5H_2O$ (pH 7.2). L-Pip was added to the medium to a final concentration of 1.0 g/L (7.7 mM). After 4 h of cultivation at 30 °C, cultures were induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and further incubated at 30 °C for 30 or 48 h (Supporting Information Tables S3−S5). The amounts of Hypips involved in culture supernatants were analyzed by HPLC after the derivatization with Fmoc-Cl. Analytical methods are described in the Supporting Information.

For comparison of the Hypips productivity and *cis*-5/*cis*-3 ratio between SmP4H wild type enzyme and three generated mutants, V97F, V97F/V95W and V97F/V95W/E114G, *in vivo* assays were conducted with the pTrS31 vector and *E. coli* W3110 $ilvG^+pyrE^+$ strain as a host. In this case, the high $FeSO_4$-containing medium was used. The medium constituents were 40 g/L glucose, 0.5 g/L $MgSO_4$, 2 g/L K_2HPO_4, 10 g/L NH_4SO_4, 2 g/L NaCl, 1 g/L yeast extract, 15 mg/L $CaCl_2$, and 278 mg/L $FeSO_4{\cdot}7H_2O$. L-Pip was added to the medium to a final concentration of 2.0 g/L (15.4 mM). After the pH was adjusted at 7.2, 2.5% of $CaCO_3$ was added into the medium. Single colonies from the strains harboring pTrS31-SmP4H (wild-type and mutants) were inoculated into 5 mL LB medium containing 100 mg/L ampicillin and cultured at 30 °C overnight. An inoculum of the resulting overnight culture (100 μL) was transferred into the high $FeSO_4$-containing medium (5 mL). After cultivation at 30 °C for 48 h, the amounts of Hypips involved in culture supernatants were analyzed by HPLC after the derivatization with Fmoc-Cl as described above (Figure 3 and Supporting Information Table S7).

Construction and Screening of the SmP4H Mutant Library. To obtain the SmP4H V97F/V95W mutant, saturated mutagenesis[26] was carried out. I94, V95, S98, E99, N100, A101, C102, I103, T117, and L119 were selected as the target amino acid residues. Overlap extension PCRs were performed using the various primer sets listed in Supporting Information Table S2 with pTrS31-SmP4H (V97F) as a template. Each primer contained an NNK codon at the selected amino acid residue. Amplified fragments were purified, mixed, and overlap extension PCR was carried out using primer 29 and primer 30. Connected PCR fragments were digested with *Hind*III and *Sac*I and introduced into the pTrS31 vector (*Hind*III/*Sac*I) as described above. The plasmids were transformed into *E. coli* W3110 $ilvG^+pyrE^+$ and inoculated onto a LB plate containing 100 mg/L ampicillin. From these plates we obtained colonies of strains harboring the pTrS31-SmP4H (V97F/V95X) mutant library. A hundred colonies harboring the pTrS31-SmP4H (V97F/V95X) mutant library were inoculated into 5 mL LB

medium containing 100 mg/L ampicillin and then incubated at 30 °C overnight. An inoculum of the resulting overnight culture (100 μL) was transferred into high $FeSO_4$-containing medium (5 mL). After 48 h cultivation at 30 °C, culture supernatants were analyzed by HPLC as described before.

To generate the triple mutant (V97F/V95W/E114G), random mutagenesis was conducted. Error-prone PCR was performed using the Diversify random mutagenesis kit (Clontech CA, U.S.A.) with pTrS31-SmP4H (V97F/V95W) as a template. According to the kit protocol, Error-prone PCR was conducted at the condition introducing 2−3 mutations/1 kbp. Amplified PCR products were digested with *Hind*III and *Sac*I and introduced into the pTrS31 vector. The plasmid mixture was then transformed into *E. coli* HST08 (Takara Bio) by electroporation and the cells were spread onto an LB plate (100 mg/L ampicillin). After incubation at 37 °C for 16 h, ~10^5 colonies were obtained. Plasmids containing random mutants were purified from the mixture of colonies. Purified plasmids were introduced into *E. coli* W3110 *ilvG*$^+$*pyrE*$^+$ strain and spread onto an LB plate (100 mg/L ampicillin) followed by incubation at 30 °C for 16 h. About 2000 colonies were randomly selected and inoculated into a 96 deep-well plate (Greiner) containing 0.5 mL LB medium (100 mg/L ampicillin). After incubation at 30 °C for 18 h, a 50 μL aliquot of each culture was transferred into a 96 deep-well plate (Greiner) containing 0.5 mL high iron-containing medium (100 mg/L ampicillin), and further incubated at 30 °C for 48 h. After centrifugation at 3500*g* for 15 min, the amount of *cis*-5-Hypip in the culture supernatants was roughly estimated by TLC. A Funasel SF cellulose thin-layer plate was used (Funakosi, Tokyo, Japan) with a solvent system of *n*-BuOH/AcOH/H_2O = 3:1:1. A 1 μL aliquot of each supernatant was spotted onto the TLC plate. After the TLC plate was developed, *cis*-5-Hypip and L-Pip spots were visualized using ninhydrin spray (Wako Chemicals, Osaka, Japan). Positive strains were further tested by HPLC as described above. Mutants obtained from the first and second rounds of screening were then sequenced to identify the mutations of interest.

■ ASSOCIATED CONTENT

Supporting Information

Product analyses and *in vitro* assays are described in the supporting methods. Strains, plasmids, and primers used in this study (Tables S1 and S2). LC-MS analysis of Fmoc-Hypips (Figure S1). HPLC chromatogram of Fmoc-Hypips generated by wild-type SmP4H and MlP4H (Figure S2). Alignment between *cis*-P4Hs (Figure S3). Results of *in vivo* assays (Tables S3−S7). SDS-PAGE analysis (Figure S4). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Email: kento.koketsu@kyowa-kirin.co.jp.

Author Contributions

K. Koketsu and K. Moriwaki generated the SmP4H mutants and assayed all enzyme activities. Y. Shomura solved the MlP4H structures and carried out homology modeling of SmP4H and its mutants. S. Mitsuhashi, M. Hayashi, R. Hara, K. Kino, and Y. Higuchi provided guidance with the experimental design. K. Koketsu and Y. Shomura wrote the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Satoshi Koizumi and Mr. Koichiro Miyake for useful advice. We gratefully acknowledge Ms. Risa Miyashita for her technical assistance. We also thank laboratory members for their advice. We thank the beamline staff of BL44XU at SPring-8 for their support and the provision of beam time. The MX225-HE detector of the beamline is financially supported by Academic Sinica and the National Synchrotron Radiation Research Center (Taiwan, ROC).

■ REFERENCES

(1) Wu, G., Bazer, F. W., Burghardt, R. C., Johnson, G. A., Kim, S. W., Knabe, D. A., Li, P., Li, X., McKnight, J. R., Satterfield, M. C., and Spencer, T. E. (2011) Proline and hydroxyproline metabolism: Implications for animal and human nutrition. *Amino Acids 40*, 1053−1063.

(2) Remuzon, P. (1996) *Trans*-4-hydroxy-L-proline, a useful and versatile chiral starting block. *Tetrahedron 52*, 13803−13835.

(3) Kalamkar, N. B., Kasture, V. M., and Dhavale, D. D. (2010) Total synthesis of natural *cis*-3-hydroxy-L-proline from D-glucose. *Tetrahedron Lett. 51*, 6745−6747.

(4) Pandey, A. K., Naduthambi, D., Thomas, K. M., and Zondlo, N. J. (2013) Proline editing: a general and practical approach to the synthesis of functionally and structurally diverse peptides. Analysis of steric versus stereoelectronic effects of 4-substituted prolines on conformation within peptides. *J. Am. Chem. Soc. 135*, 4333−4363.

(5) Clifton, I. J., McDonough, M. A., Ehrismann, D., Kershaw, N. J., Granatino, N., and Schofield, C. J. (2006) Structural studies on 2-oxoglutarate oxygenases and related double-stranded β-helix fold proteins. *J. Inorg. Biochem. 100*, 644−669.

(6) Ogawa, J., and Shimizu, S. (2002) Industrial microbial enzymes: Their discovery by screening and use in large-scale production of useful chemicals in Japan. *Curr. Opin. Biotechnol. 13*, 367−375.

(7) Shibasaki, T., Mori, H., Chiba, S., and Ozaki, A. (1999) Microbial proline 4-hydroxylase screening and gene cloning. *Appl. Environ. Microbiol. 65*, 4028−4031.

(8) Shibasaki, T., Mori, H., and Ozaki, A. (2000) Enzymatic production of *trans*-4-hydroxy-L-proline by regio- and stereospecific hydroxylation of L-proline. *Biosci. Biotechnol. Biochem. 64*, 746−750.

(9) Shibasaki, T., Hashimoto, S., Mori, H., and Ozaki, A. (2000) Construction of a novel hydroxyproline-producing recombinant *Escherichia coli* by introducing a proline 4-hydroxylase gene. *J. Biosci. Bioeng. 90*, 522−525.

(10) Mori, H., Shibasaki, T., Uozaki, Y., Ochiai, K., and Ozaki, A. (1996) Detection of novel proline 3-hydroxylase activities in *Streptomyces* and *Bacillus* spp. by regio- and stereospecific hydroxylation of L-proline. *Appl. Environ. Microbiol. 62*, 1903−1907.

(11) Hara, R., and Kino, K. (2009) Characterization of novel 2-oxoglutarate dependent dioxygenases converting L-proline to *cis*-4-hydroxy-L-proline. *Biochem. Biophis. Res. Commun. 379*, 882−886.

(12) Shibasaki, T., Sakurai, W., Hasegawa, A., Uosaki, Y., Yoshida, M., and Ozaki, A. (1999) Substrate selectivities of proline hydroxylases. *Tetrahedron Lett. 40*, 5227−5230.

(13) Klein, C., and Hüttel, W. (2011) A simple procedure for selective hydroxylation of L-proline and L-pipecolic acid with recombinantly expressed proline hydroxylases. *Adv. Synth. Catal. 353*, 1375−1383.

(14) Hayashi, K., Morooka, N., Yamamoto, Y., Fujita, K., Isono, K., Choi, S., Ohtsubo, E., Baba, T., Wanner, B. L., Mori, H., and Horiuchi, T. (2006) Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. *Mol. Syst. Biol. 2* (2006), 0007.

(15) Strieker, M., Kopp, F., Mahlert, C., Essen, L.-O, and Marahiel, M. A. (2007) Mechanistic and structural basis of stereospecific Cβ-hydroxylation in calcium-dependent antibiotic, a daptomycin-type lipopeptide. *ACS Chem. Biol. 2*, 187−196.

(16) Clifton, I. J., Hsueh, L.-C., Baldwin, J. K., Harlos, K., and Schofield, C. J. (2001) Structure of proline 3-hydroxylase. Evolution of

the family of 2-oxoglutarate dependent oxygenases. *Eur. J. Biochem. 268*, 6625−6636.
(17) Reetz, M. T. (2013) Biocatalysis in organic chemistry and biotechnology: Past, present, and future. *J. Am. Chem. Soc. 135*, 12480−12496.
(18) Leslie, A. G. W., and Powell, H. R. (2007) Processing diffraction data with Mosflm. *Evol. Methods Macromol. Crystallogr. 245*, 41−51.
(19) Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D, Biol. Crystallogr. 67*, 235−242.
(20) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D, Biol. Crystallogr. 66*, 213−221.
(21) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D, Biol. Crystallogr. 60*, 2126−2132.
(22) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D, Biol. Crystallogr. 53*, 240−255.
(23) Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Cα geometry: ϕ, ψ, and Cβ deviation. *Proteins 50*, 437−450.
(24) Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T. G., Bertoni, M., Bordoli, L., and Schwede, T. (2014) SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res. 42*, W252−258.
(25) Case, D. A., Babin, V., Berryman, J. T., Betz, R. M., Cai, Q., Cerutti, D. S., Cheatham, T. E., III, Darden, T. A., Duke, R. E., Gohlke, H., Goetz, A. W.; Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolossváry, I., Kovalenko, A., Lee, T. S., LeGrand, S., Luchko, T., Luo, R., Madej, B., Merz, K. M., Paesani, F., Roe, D. R., Roitberg, A., Sagui, C., Salomon-Ferrer, R., Seabra, G., Simmerling, C. L., Smith, W., Swails, J., Walker, R. C., Wang, J., Wolf, R. M., Wu, X., and Kollman, P. A. (2014) *AMBER 14*; University of California, San Francisco.
(26) Siloto, R. M. P., and Weselake, R. J. (2012) Site saturation mutagenesis: Methods and applications in protein engineering. *Biocatal. Agric. Biotechnol. 1*, 181−189.