Pergamon

Tetrahedron Letters 40 (1999) 5227–5230

TETRAHEDRON LETTERS

Substrate Selectivities of Proline Hydroxylases

Takeshi Shibasaki,[a] Wataru Sakurai, Atsuhiro Hasegawa, Youichi Uosaki, Hideo Mori, Mayumi Yoshida and Akio Ozaki*

Tokyo Research Laboratories, Kyowa Hakko Kogyo Co. Ltd. 3-6-6, Asahimachi, Machida, Tokyo 194-8533, Japan
a) Present address: Technical Research Laboratories, Kyowa Hakko Kogyo Co. Ltd. 1-1, Kyowamachi, Hofu, Yamaguchi 747-8522, Japan

Received 8 April 1999; revised 6 May 1999; accepted 7 May 1999

Abstract: Substrate selectivities of microbial proline 4-hydroxylase and proline 3-hydroxylases, all of which were purified from recombinant *Escherichia coli*, were investigated. L-2-Azetidine carboxylate, 3,4-dehydro-L-proline and L-pipecolinic acid were hydroxylated by those enzymes in regio- and stereospecific manner.

Keywords: Enzymes and enzyme reactions; Hydroxylation; Regiospecificity; Stereospecificity.

3-Hydroxyproline and 4-hydroxyproline have been found as constituents of collagen[1] and several secondary metabolites including etamycin[2] and telomycin.[3] Among the naturally occuring three diastereomers of 4-hydroxyproline and two diastereomers of 3-hydroxyproline, *trans*-4-hydroxy-L-proline is the most abundant as a constituent of collagen. The biosynthesis of hydroxyprolines has been extensively studied in mammalian system, in which *cis*-3-hydroxy-L-proline and *trans*-4-hydroxy-L-proline are formed through the post-translational hydroxylation of peptidyl L-proline of procollagen catalysed by procollagen-proline 3-dioxygenase (prolyl 3-hydroxylase, EC1.14.11.7) and procollagen-proline dioxygenase (prolyl hydroxylase, EC1.14.11.2), respectively.[4] In contrast, proline hydroxylases which hydroxylate free L-proline occur in microorganisms.[5] Proline 4-hydroxylase was reported in the biosynthesis of *trans*-4-hydroxy-L-proline in etamycin producing *Streptomyces*[5a] and 3-hydroxylases were reported in the biosynthesis of *cis*-3-hydroxy-L-proline in *Streptomyces* including telomycin producer and *Bacillus*.[5b] Both mammalian prolyl hydroxylases and bacterial proline hydroxylases require 2-oxoglutarate and dioxygen as cosubstrates and ferrous ion as a cofactor. These 2-oxoglutarate dependent dioxygenases have commonly a relatively loose substrate specificity for hydroxylation. In the case of proline 4-hydroxylase of *S. griseoviridis* P8648,[5a] several proline analogues were shown to react as substrates, however, the structures of the products were not confirmed except in the case of 3,4-dehydro-L-proline **3** which was epoxidised to give *trans*-3,4-epoxy-L-proline **4**. The reactivities of other proline analogues were only shown through the increase of the turnover rate of 2-oxoglutarate.[6]

Here we report the studies on the substrate selectivities of proline 3-hydroxylases and proline 4-hydroxylase. Two genes encoding proline 3-hydroxylase typeI and typeII, respectively, from *Streptomyces* sp. TH1[5b] and a gene encoding proline 4-hydroxylase from *Dactylosporangium* sp. RH1 were cloned in *Escherichia coli* and overexpressed. Proline 3-hydroxylase typeI and typeII are homologous enzymes having 76% homology in amino acid sequences. All of the hydroxylases were purified to homogenuity from corresponding recombinant *E. coli* cells by means of DEAE Sepharose and Phenyl Sepharose HP chromatographies. Specific activties of the purified enzymes were 9674 U/mg (nmole/min/mg) for proline 4-hydroxylase, 7150 U/mg for proline 3-hydroxylase typeI, and 12700 U/mg for typeII. A number of anlogues were examined as substrates to reveal the substrate spectrum and the patterns of substrate recognition of the enzymes. Assays were basically done by measuring both the decrease of analogues and the appearance of a new

0040-4039/99/$ - see front matter
PII: S0040-4039(99)00944-2

peak with HPLC equipped with a chiral column SUMICHIRAL OA5000 by post column derivatization method.

(1) Peptides containing L-proline (L-pro-gly; gly-L-pro; L-pro-L-ala; L-pro-L-pro-gly-L-phe-L-ser-L-pro), amino and carboxyl derivatives of L-proline (L-prolinamide; *N*-acetyl-L-proline; *N*-Boc-L-proline; *N*-Fmoc-L-proline), 2-pyrrolidinemethanol, and pyrrolidine showed no activity for either 4-hydroxylase or two 3-hydroxylases. The results suggest that free imino and carboxyl moiety of L-proline are necessary to the action of both 4-hydroxylase and 3-hydroxylases. (2) L-2-Azetidinecarboxylic acid **1**, 3,4-dehydro-L-proline **3** and L-pipecolinic acid **6**, four to six membered ring analogues of L-proline, showed the reactivities for either 3-hydroxylases or 4-hydroxylase. Relative activities of proline hydroxylases and apparent *Kms* for these analogues are shown in Table 1. On the other hand, 3-azetidinecarboxylic acid and piperidine showed no reactivity. The results suggest that proline hydroxylases recognize an α-imino carboxylic acid structure independent of the ring size. (3) D-Proline, 3,4-dehydro-D-proline and D-pipecolinic acid did not react as substrates suggesting that the enzymes recognize L-configured imino acids.

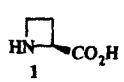
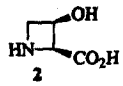
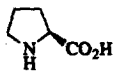
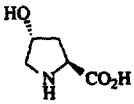
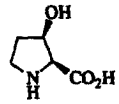
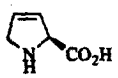
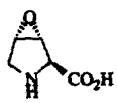
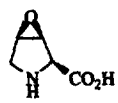
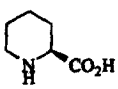
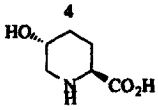
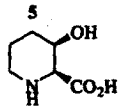
Substrate	4-Hydroxylase Products	4-Hydroxylase Relative Activity (Apparent *Km* in mM)	3-Hydroxylase Products	3-Hydroxylase Relative Activity (Apparent *Km* in mM) TypeI / TypeII
 **1**	Not Detected		 **2**	3 (NT) / 40 (2.1)
		100 (0.2)		100 (0.2) / 100 (0.4)
 **3**	 **4**	66 (5.1)	 **5**	30 (NT) / 57 (8.4)
 **6**	 **7**	109 (5.7)	 **8**	15 (3.8) / 3 (>100)

Table1: Substrate specificity of proline 4-hydroxylase, proline 3-hydroxylase typeI and typeII. Relative activities are presented assuming that the activity on L-proline is 100. Apparent *Kms* were calculated based on the observation of the decrease of the substrates. NT, not tested.

Although these results suggested that both proline 4-hydroxylase and 3-hydroxylases recognized α-imino carboxylic acid structure in L-configulation, regio- and stereochemistry of hydroxylation reaction was unclear. Moreover, even the decrease of substrates in the reaction cannot be regarded as the evidence of hydroxylation. Then, we isolated the reaction products of the analogues and identified their structures (Table 1).

Hydroxylation of L-2-azetidinecarboxylic acid **1**: A new peak was detected in HPLC analysis after the 3-hydroxylase typeII reaction on **1**[8a] suggesting the production of a new imino acid. The structure of the isolated product **2** corresponding to this peak was elucidated as *cis*-3-hydroxy-2-azetidinecarboxylic acid mainly by NMR spectral analyses.[9] Hydroxylation was defined to occur at C-3 because methine-methine-methylene skeleton was assigned by ^{1}H-^{1}H-COSY spectrum. The stereochemistry was elucidated as follows. The vicinal coupling constants for azetidines were reported to be 7-9 Hz in *cis*-orientation and 3-6 Hz in *trans*-orientation.[10] For the compound **2**, the coupling constant between H-2 and H-3 could not be determined because their resonances were not well resolved (δ4.92 and δ4.88, respectively). While, the coupling constants between H-3 and H-4a, and between H-3 and H-4b were analyzed to be 6.8 and 4.6 Hz, respectively. These coupling constants revealed that H-3 and H-4a were oriented in a *cis* relationship and H-3 and H-4b in a *trans* relationship (Fig. 1). Furthermore, NOEs were observed between H-4a and H-2 as well as H-3 while no NOE was observed between H-4b and H-2 or H-3, indicating H-2, H-3, and H-4a were spatially near to each other

and on the same side as shown in Fig 1. A diastereomer of **2**, *trans*-3-hydroxy-L-2-azetidinecarboxylic acid, was not detected in NMR spectral analyses. Based on these results including that absolute configulation of **1** is L, the structure of the isolated compound **2** was determined to be *cis*-3-hydroxy-L-2-azetidinecarboxylic acid. A peak having the identical retention time to **2** was also observed in the reaction with 3-hydroxylase typeI although the reactivity was very weak. On the contrary, no decrease of **1** was observed in the reaction with proline 4-hydroxylase of *Dactylosporangium* sp. RH1 suggesting that the enzyme was inactive on **1**, although proline 4-hydroxylase of *S. griseoviridis* P8648 was reported to show the reactivity on **1** in terms of the turnover of 2-oxoglutarate.[6]

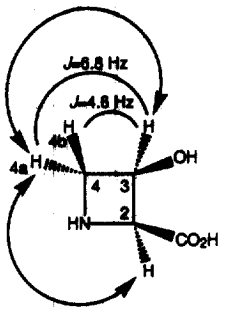


Figure 1: Coupling constants and NOEs of **2** observed in D_2O.

Epoxidation of 3,4-dehydro-L-proline **3**: Proline 4-hydroxylase from *S. griseoviridis* P-8648 was reported to catalyse the epoxidation of **3** to *trans*-3,4-epoxy-L-proline **4**.[6] Similarly, the product in the reaction on **3** with 4-hydroxylase from *Dactylosporangium*[8b] was confirmed to be identical to the synthetic compounds **4** by means of HPLC and TLC analyses,[11] and **5**, a diastereomer of **4**, was not detected. In the 3-hydroxylase typeII reaction on **3**, an another new peak was detected in HPLC analysis suggesting the production of a new imino acid. The structure of the product isolated[8c] was elucidated as 3,4-epoxy-proline from the spectral analyses,[12] however, the stereochemistry was not confirmed. Then the isolated product was compared with the synthetic *cis*-3,4-epoxy-L-proline **5**[6] as well as **4**. The product showed a spot with identical Rf in TLC and a peak with identical retention time in HPLC to those of **5**,[11] and **4**, a diastereomer of **5**, was not detected. The epoxidation product **5** was also detected in typeI 3-hydroxylase reaction using **3** as a substrate.

Hydroxylation of L-pipecolinic acid **6**: The hydroxylation product of **6** was isolated from the 4-hydroxylase reaction mixture.[8d] Spectral data of the isolated product[13] were identical with those reported for *trans*-5-hydroxy-D-pipecolinic acid[14a] and the specific rotation was opposite. In conclusion, the structure of the product was confirmed as *trans*-5-hydroxy-L-pipecolinic acid **7**. A diastereomer of **7**, *cis*-5-hydroxy-L-pipecolinic acid, was not detected in spectral analyses. On the other hand, spectral data of the isolated hydroxylation product of **6**[15] from the 3-hydroxylase typeI reaction mixture[8e] were identical with those reported for *cis*-3-hydroxy-L-pipecolinic acid **8**.[14b] The coupling constant between H-2 (δ3.37) and H-3 (δ4.24) was 2.9 Hz, while the reported coupling constant between H-2 and H-3 was 7.1 Hz in *trans*-3-hydroxy-L-pipecolinic acid.[14c] Therefore, the structure of the hydroxylation product of **6** with 3-hydroxylase typeI was concluded as *cis*-3-hydroxy-L-pipecolinic acid **8**. A diastereomer of **8**, *trans*-3-hydroxy-L-pipecolinic acid, was not detected in spectral analyses. Formation of **8** was also observed in the reaction with 3-hydroxylase typeII.

In summary, the substrate recognition and regio- and stereospecificity of hydroxylation by both L-proline 4-hydroxylase and L-proline 3-hydroxylases is rigid, in the sense that these proline hydroxylases recognize α-imino-L-carboxylic acid structures and hydroxylation occurs with high regularity, although the enzymes showed relatively loose substrate spectrum with regard to the ring size. These results suggest that these proline hydroxylases have a common hydroxylation mechanism, however, no homologies were observed in amino acid sequences between 3- and 4-hydroxylase except that a His-1 motif, which is considered to be a part of an active site of 2-oxoglutarate dependent dioxygenases, was conserved as reported in other 2-oxoglutarate dependent dioxygenases.[16] Proline hydroxylases would be of some utilities as enzyme catalysts to functionalise proline analogues, since these functionalised compounds are useful as chiral building blocks.[17]

Acknowledgement: We thank Ruriko Nishimura-Fujii, Takamichi Yamamoto and Yuko Uosaki for capable technical assistance.

References and Notes:

1. Kuttan, R.; Radhakrishnan, A. N. *Adv. Enzymol.* **1973**, 37, 273-347.

2. Sheehan, J. C.; Zahau, H. G.; Lawson, W. B. *J. Am. Chem. Soc.* **1958**, 80, 3349-3355.
3. Irreverre, F.; Morita, K.; Ishii, S.; Witkop, B. *Biochem. Biophys. Res. Commun.* **1962**, 9, 69-71.
4. Cardinale, G. J.; Udenfriend, S. *Adv. Enzymol.* **1974**, 41, 245-300.; Schomburg, D.; Stephan, D. (ed.) Oxidoreductases *in Enzyme Handbook*, vol. 8, Springer-Verlag; Berlin, **1994**, pp. 1.13-1.97.
5. (a) Onishi, M.; Okumura, Y.; Okamoto, R.; Ishikura, T. *Biochem. Biophys. Res. Commun.* **1984**. 120, 45-51.; Lawrence, C. C.; Sobey, W. J.; Field, R. A.; Baldwin, J. E.; Schofield, C. J. *Biochem. J.* **1996**, 313, 185-191.; (b) Mori, H; Shibasaki, T.; Uozaki, Y; Ochiai, K.; Ozaki, A. *Appl. Environ. Microbiol.* **1996**, 62, 1903-1907.; Mori, H; Shibasaki, T.; Yano, Y; Ozaki, A. *J. Bacteriol.* **1997**, 179, 5677-5683.
6. Baldwin, J. E.; Field, R. A.; Lawrence, C. C.; Lee, V.; Robinson, J. K.; Schofield, C. J. *Tetrahedron Lett.* **1994**, 4649-4652.
7. Ozaki, A.; Shibasaki, T.; Mori, H. *Biosci. Biotech. Biochem.* **1995**, 59, 1764-1765.
8. Reactions were done under the following conditions. The substrates added were consumed completely in all reactions. (a) 50 mM 2-(*N*-morpholino)ethanesulfonic acid (MES), pH 6.0, 10 mM 2-oxoglutarate, 2.5 mM **1**, 2 mM $FeSO_4$, 1 g/l of bovine serum albumin (BSA) and 4624 U/ml 3-hydroxylase type II in 30 ml, 40°C, 1h. **2** (2.5 mg) was isolated (isolation yield 28%).; (b) 50 mM MES, pH 6.5, 20 mM 2-oxoglutarate, 2.5 mM **3**, 2 mM $FeSO_4$ and 4611 U/ml 4-hydroxylase, in 20 ml, 35°C, 1h. **4** (1.5 mg) was isolated (23%).; (c) 50 mM MES, pH 6.0, 10 mM 2-oxoglutarate, 2.5 mM **3**, 2 mM $FeSO_4$, 1 g/l BSA and 9103 U/ml 3-hydroxylase type II in 40 ml, 40°C, 1h. **5** (4.4 mg) was isolated (35%).; (d) 50 mM MES, pH 6.5, 20 mM 2-oxoglutarate, 8 mM **6**, 2 mM $FeSO_4$ and 3904 U/ml 4-hydroxylase, in 30 ml, 35°C, 1h. **7** (13.9 mg) was isolated (40%).; (e) 50 mM MES, pH 6.0, 10 mM 2-oxoglutarate, 2.5 mM **6**, 2 mM $FeSO_4$, 1 g/l BSA and 19468 U/ml 3-hydroxylase type I in 20 ml, 35°C, 1h. **8** (2.7 mg) was isolated (38%).
9. ^{1}H NMR (500 MHz, D_2O) δ 3.92 (1H, dd, *J*=4.6, 11.6 Hz), 4.39 (1H, dd, *J*=6.8, 11.6 Hz), 4.88 (1H, m), 4.92 (1H, m); ^{13}C NMR (125 MHz, D_2O) δ 54.0(t), 64.0(d), 67.5(d) 170.5(s); HRFABMS *m/z* 118.0500 $(M+H)^+$ for $C_4H_7O_3N$+H; $[\alpha]_D^{30}$ = -49.9 (*c* 0.12, H_2O).
10. Doemes, E.; Cromwell, N. H. *J. Org. Chem.* **1969**, 34, 310-317.
11. Retention time of **4** = 9.2 min. and **5** = 33.5 min. with HPLC system equipped with chiral column OA5000 by post column derivatization with NBD.[7]; Rf of **4** = 0.41, **5** = 0.32 (phenol / propanol / water : 8 / 2 / 1).
12. ^{1}H NMR (500 MHz, D_2O) δ 3.52 (1H, dd, *J*=1.0, 13.3 Hz), 3.77 (1H, d, *J*=13.3 Hz), 4.09 (1H, dd, *J*=1.0, 2.9 Hz), 4.22 (1H, dd, *J*=1.2, 2.9 Hz), 4.34 (1H, d, *J*=1.3 Hz); ^{13}C NMR (125 MHz, D_2O) δ 46.8(t), 56.0(d), 57.6(d), 62.1(d), 170.8(s); MS *m/z* 130 $(M+H)^+$.
13. ^{1}H NMR (500 MHz, D_2O) δ 1.64 (1H, m, *J*=3.6, 9.4, 11.3, 12.8 Hz), 1.81 (1H, m, *J*=3.6, 6.8, 10.7, 11.3 Hz), 2.13 (1H, m, *J*=3.6, 5.6, 8.1, 12.8 Hz), 2.34 (1H, m, *J*=3.6, 3.9, 5.6, 6.8 Hz), 2.88 (1H, ddd, *J*=1.0, 9.5, 12.3 Hz), 3.49 (1H, ddd, *J*=1.5, 4.2, 12.3 Hz), 3.66 (1H, dd, *J*=3.9, 10.7 Hz), 4.00 (1H, m, *J*=4.2, 8.1, 9.4, 9.5 Hz); MS *m/z* 145 $(M+H)^+$; $[\alpha]_D^{22}$ = -6.10 (*c* 0.09, H_2O).
14. Hosrau, S.; Fauchere, J. L.; Pappalardo, L.; Roumestant M. L.; Viallefont, P. *Tetrahedron: Asymmetry*, **1996**, 7, 2585-2593.; (b) Kusano, G.; Ogawa, H.; Takahashi, A.; Nozoe, S.; Yokoyama, K. *Chem. Pharm. Bull.*, **1987**, 35, 3482-3486.; (c) Greck, C.; Ferreira, F.; Genet, J.P. *Tetrahedron Lett.*, **1996**, 37, 2031-2034.
15. ^{1}H NMR (500 MHz, D_2O) δ 1.54 (1H, m), 1.77 (1H, m), 1.78 (1H, m), 1.92 (1H, m), 2.70 (1H, m, *J*=3.6, 7.4, 13.4 Hz), 3.11 (1H, m), 3.37 (1H, d, *J*=2.9 Hz), 4.24 (1H, m, *J*=2.9, 4.4, 6.6 Hz).; MS *m/z* 145 $(M+H)^+$.
16. Valegard, K.; van Scheltinga, A. C.; Lloyd, M. D.; Hara, T.; Ramaswamy, S.; Perrakis, A.; Thompson, A.; Lee, H. J.; Baldwin, J. E.; Schofield, C. J.; Hajdu, J.; Anderson, I. *Nature*, **1998**, 394, 805-809.; Myllyla, R.; Gunzler, V.; Kivirikko, K. I.; Kaska, D. D. *Biochem. J.*, **1992**, 286, 923-927.
17. Remuzon, R. *Tetrahedron.* **1996**, 52, 13803-13835.