JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

A Unique *cis*-3-Hydroxy-L-proline Dehydratase in the Enolase Superfamily

Xinshuai Zhang,[‡] Ritesh Kumar,[‡] Matthew W. Vetting,[§] Suwen Zhao,[∥] Matthew P. Jacobson,[∥] Steven C. Almo,[§] and John A. Gerlt*[,†,‡]

[†]Departments of Biochemistry and Chemistry and [‡]Institute for Genomic Biology, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States

[§]Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461, United States

[∥]Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143, United States

Supporting Information

ABSTRACT: The genome of *Labrenzia aggregata* IAM 12614 encodes an uncharacterized member of the muconate lactonizing enzyme (MLE) subgroup of the enolase superfamily (UniProt ID A0NXQ8). The gene encoding A0NXQ8 is located between genes that encode members of the proline racemase superfamily, 4*R*-hydroxyproline 2-epimerase (UniProt ID A0NXQ7; 4HypE) and *trans*-3-hydroxy-L-proline dehydratase (UniProt ID A0NXQ9; *t*3LHypD). A0NXQ8 was screened with a library of proline analogues; two reactions were observed with *cis*-3-hydroxy-L-proline (*c*3LHyp), competing 2-epimerization to *trans*-3-hydroxy-D-proline (1,1-proton transfer) and dehydration to Δ^1-pyrroline-2-carboxylate (β-elimination; *c*3LHyp dehydratase), with eventual total dehydration. The genome context encoding A0NXQ8 both (1) confirms its novel *c*3LHyp dehydratase function and (2) provides evidence for metabolic pathways that allow *L. aggregata* to utilize several isomeric 3- and 4-hydroxyprolines as sole carbon sources.

The members of the enolase superfamily (ENS) are functionally diverse, catalyzing reactions initiated by abstraction of the α-proton of a carboxylate substrate.[1] The reactions include 1,1-proton transfer (racemization and epimerization) and β-elimination (dehydration and deamination).[2] The UniProt database (release 2014_10) contains sequences for 83 038 members of the ENS. Perhaps 50% of these have unknown functions because they, or likely orthologues, have not been functionally characterized.

The MLE subgroup of the ENS includes the *cis*,*cis*-muconate cycloisomerase (MLE), *o*-succinylbenzoate synthase (OSBS), *N*-succinyl-amino acid racemase (NSAR), dipeptide epimerase, and 4*R*-hydroxyproline betaine 2-epimerase functions.[3,4] In this manuscript, we report the identification of the novel *cis*-3-hydroxy-L-proline dehydratase (*c*3LHypD) function, the first example of an amino acid dehydratase in the ENS.

Sequence−function relationships in enzyme superfamilies can be visualized using sequence similarity networks (SSNs).[5] When the SSN for the MLE subgroup is displayed at a BLAST *e*-value threshold of 10^{-50} (Figure 1A), several clusters are observed. One of the divergent clusters, designated the "Roseo" cluster (boxed), includes an uncharacterized member (UniProt ID A0NXQ8) from the bacterium *Labrenzia aggregata* IAM 12614.

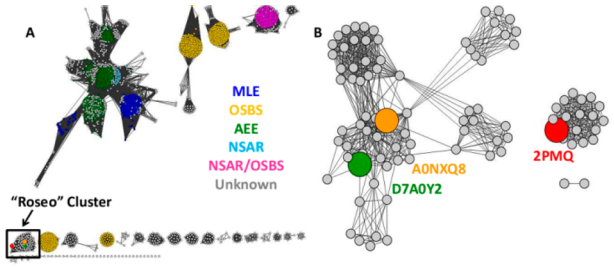

Figure 1. (A) SSN of the MLE subgroup displayed at a BLASTP *e*-value of 10^{-50}; the "Roseo" cluster is boxed. (B) SSN of the "Roseo" cluster displayed at a BLASTP *e*-value of 10^{-80}; 2PMQ, A0NXQ8, and D7A0Y2 are denoted by the red node, orange node, and green node, respectively.

Recently, our laboratories assigned the *trans*-4-hydroxy-L-proline betaine 2-epimerase function to Uniprot ID Q0FPQ4 (PDB 2PMQ) in the "Roseo" cluster.[4] A0NXQ8 shares 34% sequence identity with Q0FPQ4. When the BLAST *e*-value threshold is decreased to 10^{-80}, the "Roseo" cluster segregates into several presumed isofunctional families (Figure 1B), with A0NXQ8 and Q0FPQ4 located in different clusters, suggesting they have different *in vitro* activities and *in vivo* functions.

Genome neighborhood context can provide useful information for deducing the functions of uncharacterized bacterial enzymes by providing clues about the type of substrate; activity screening of a focused physical library then may allow the substrate to be identified.[3] The genome context for A0NXQ8 reveals that its gene is located between genes encoding members of the proline racemase superfamily (UniProt IDs A0NXQ7 and A0NXQ9; Figure 2A). Their *in vitro* activities are 4*R*-hydroxyproline 2-epimerase (4HypE; A0NXQ7) and *trans*-3-hydroxy-L-proline dehydratase (*t*3LHypD; A0NXQ9).[6] The genome context does not provide sufficient information to assign a function to A0NXQ8; however, we used it to hypothesize that a substituted proline is the substrate.

Received: October 15, 2014
Published: January 21, 2015

DOI: 10.1021/ja5103986
J. Am. Chem. Soc. 2015, 137, 1388−1391

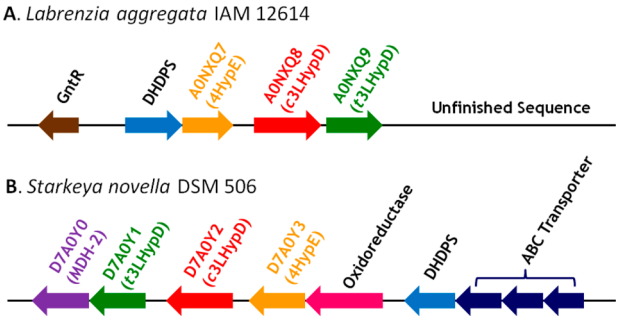

Figure 2. (A) Genome context of A0NXQ8 from *L. aggregata* IAM 12614. (B) Genome context of orthologous D7A0Y2 from *Starkeya novella* DSM 506. The genes encoding orthologues are highlighted with the same color. DHDPS, member of the dihydrodipicolinate synthetase superfamily.

We screened A0NXQ8 with a library of proline derivatives, including L-proline and several stereoisomers of 3-hydroxyproline, 4-hydroxyproline and their mono- and di-*N*-methylated derivatives (Figure S1). Only *cis*-3-hydroxy-L-proline (*c*3LHyp) was a substrate. The ^{1}H NMR spectrum of the reaction with *c*3LHyp performed in D_2O revealed (1) exchange of the α-proton with solvent, (2) 2-epimerization to *trans*-3-hydroxy-D-proline (*t*3DHyp), and (3) dehydration to Δ^1-pyrroline-2-carboxylate (Pyr2C). All three reactions are initiated by abstraction of the α-proton of the substrate, the common initial partial reaction for members of the ENS (Figure 3). As

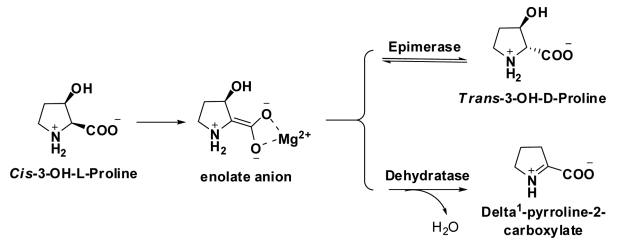

Figure 3. Proposed mechanism for the A0NXQ8-catalyzed reaction.

expected for irreversible dehydration and reversible epimerization reactions that share a common intermediate, only the Pyr2C dehydration product was observed when the reaction was complete.

The kinetic constants for disappearance of *c*3LHyp were measured using a polarimeter: k_{cat} = 23 ± 0.6 s^{-1}, K_M = 8.1 ± 0.9 mM, and k_{cat}/K_M = 2800 M^{-1} s^{-1}. The relative rates of epimerization and dehydration were determined by monitoring the reaction by ^{1}H NMR using several concentrations of *c*3LHyp: the resonances associated with C3 of *c*3LHyp, *t*3DHyp, and Pyr2C have different chemical shifts (Figure S3). As the reaction proceeds, the intensities of the signals associated with *c*3LHyp decrease as the intensities of the signals associated with the *t*3DHyp and Pyr2C products increase with no detectable lag (Figure 4 and Figure S4). The rate of epimerization is about 1.5-fold greater than that of dehydration at substrate concentrations both below and above the value of the K_M (Figure S5); thus, the k_{cat} for epimerization is ~14 s^{-1} and that for dehydration is ~9 s^{-1}. The initial linear dependence of product formation on time for both reactions (absence of a lag for dehydration) indicates that the substrate for the dehydration reaction is *c*3LHyp and not *t*3DHyp, the product of the 2-epimerization reaction.

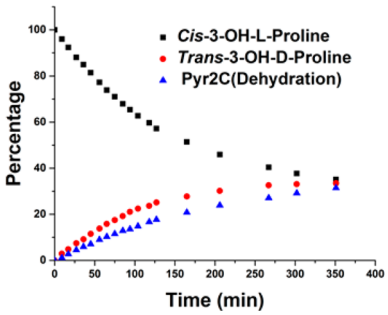

Figure 4. Initial progress of the A0NXQ8-catalyzed reaction with *c*3LHyp during the initial of the reaction showing formation of *t*3DHyp (epimerization) and Pyr2C (dehydration) monitiored by ^{1}H NMR; the ^{1}H NMR assay contained 10 mM *c*3LHyp, 10 mM $MgCl_2$, 50 mM sodium phosphate, pH 8.0, and 0.14 μM A0NXQ8. Eventually, only Pyr2C is observed.

We assign the *c*3LHyp dehydratase (*c*3LHypD) function to A0NXQ8. Four other reactions involving abstraction of the α-proton of a proline derivative have been described for the proline racemase superfamily (Pfam PF05544): proline racemase,[7] *trans*-4-hydroxy-L-proline 2-epimerase (4HypE),[8] *trans*-3-hydroxy-L-proline 2-epimerase (*t*3LHypE), and *trans*-3-hydroxy-L-proline dehydratase (*t*3LHypD).[6,9] To the best of our knowledge, A0NXQ8 is the first *c*3LHypD and the first amino acid dehydratase in the ENS.[3,4]

The crystal structure of A0NXQ8 was determined at 2.2-Å resolution with bound Mg^{2+} (PDB entry 4MGG). 4MGG has the ENS bidomain structure (Figure S7); as expected for a member of the MLE subgroup, the C-terminal $(\beta/\alpha)_7\beta$-barrel domain contains two conserved general acid/base residues, Lys 165 and Lys 265, located on the opposite faces of the active site at the ends of the second and sixth β-strands in the barrel domain.[3]

Mutagenesis studies were performed to investigate the reaction mechanism. Using *c*3LHyp, no activity was observed (change in optical rotation or exchange of the α-proton with solvent deuterium by ^{1}H NMR) for either the K165A or K265A mutant. With *t*3DHyp prepared using A0NXQ8 in H_2O (as a 1:1:1 mixture with *c*3LHyp and Pyr2C; *t*3DHyp is not commercially available), no dehydration or exchange of the α-proton with solvent deuterium was observed for either mutant (Figure S2). Therefore, the identities of the acid/base catalysts cannot be specified. Nevertheless, the reaction is expected to involve abstraction of the α-proton of *c*3LHyp (by either Lys 165 or Lys 265) to generate an enolate anion intermediate.

The promiscuity (competing dehydration and epimerization) likely is a consequence of divergent evolution of function in the enolase superfamily. Abstraction of the α-proton from the *cis*-3-hydroxy-L-proline substrate by one active site Lys results in formation of the enolate anion intermediate that partitions between general acid catalysis of departure of the 3-OH group to form Δ^1-pyrroline-2-carboxylate (β-elimination/dehydration) and protonation of carbon-2 to form *trans*-3-hydroxy-D-proline (1,1-proton transfer) by the second active site Lys. For dehydration of *c*3LHyp, departure of the 3-OH group from the intermediate would be general acid-catalyzed by the second Lys; for epimerization, the intermediate is protonated on the opposite face of carbon-2 by the same Lys (Figure 3). We have observed similar examples of competing dehydration and epimerization in acid sugar dehydratase in the ENS[10,11] and attributed these to the consequences of "hard-wired" active site

DOI: 10.1021/ja5103986
J. Am. Chem. Soc. 2015, 137, 1388−1391

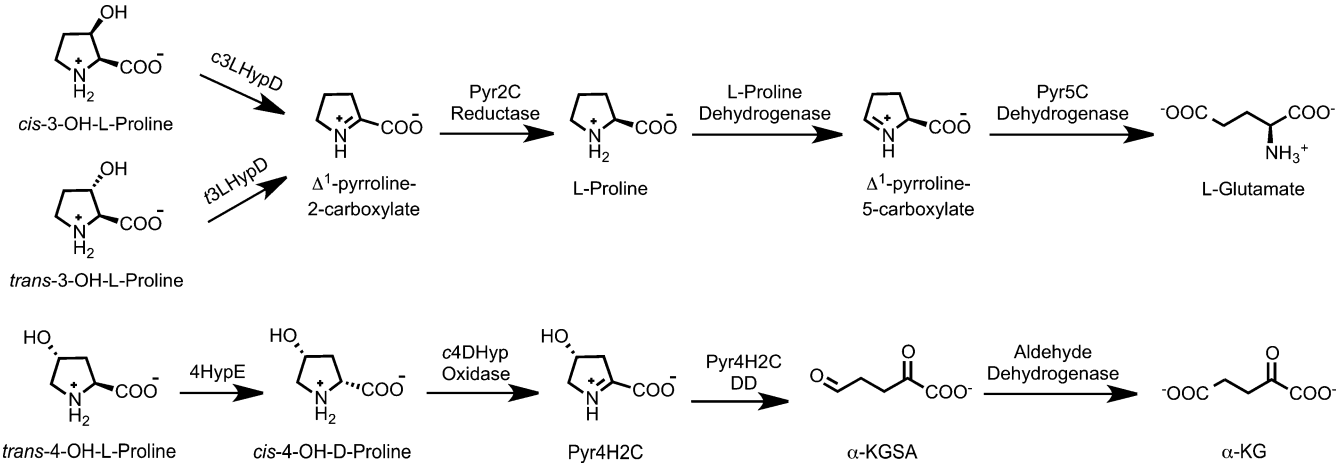

Figure 5. Pathways for degradation *trans*-4-hydroxy-L-proline, *trans*-3-hydroxy-L-proline, and *cis*-3-hydroxy-L-proline by *L. aggregata* IAM 12614 and *S. novella* DSM 506.

functional groups that can be accessed by substrates with different structures as determined by changes in the specificity determinants, thereby enhancing the evolution of new functions.[12]

L. aggregata IAM 12614 utilizes *c*3LHyp as carbon source. Transcriptomics using qRT-PCR revealed that the gene encoding A0NXQ8 is upregulated 10.5-fold when *L. aggregata* IAM 12614 is grown on *c*3LHyp as carbon source compared to growth on glucose (Table S1), implicating *c*3LHypD in the catabolic pathway.

The genome sequence of *L. aggregata* IAM 12614 is an incomplete draft sequence. To identify the catabolic pathway in which *c*3LHypD participates, we studied *S. novella* DSM 506 for which a complete genome sequence is available; it encodes an orthologue of A0NXQ8 (UniProt ID D7A0Y2; Figure 1B; A0NXQ8 and D7A0Y2 share 75% sequence identity).

The genome context of D7A0Y2 is shown in Figure 2B. Its gene is located between genes encoding orthologues of 4HypE (UniProt ID D7A0Y3) and a *t*3LHypD (UniProt ID D7A0Y1).[6] As expected from the sequence identity, D7A0Y2 also catalyzes the *c*3LHypD reaction. The gene encoding a member of the malate dehydrogenase-2 (MDH-2) superfamily (UniProt ID D7A0Y0) is adjacent to those encoding *t*3LHypD (D7A0Y1) and *c*3LHypD (D7A0Y2). We hypothesized that D7A0Y0 catalyzes the reduction of Pyr2C generated by dehydration of both *c*3LHypD and *t*3LHypD to L-proline;[6,13] *in vitro* assays of D7A0Y0 confirmed that it is a NADPH-dependent Pyr2C reductase, k_{cat} = 26 ± 1 s^{-1}, K_M = 2.8 ± 0.4 mM, and k_{cat}/K_M = 9100 M^{-1} s^{-1}. The L-proline product likely is catabolized to L-glutamate via the bifunctional L-proline dehydrogenase/Δ^1-pyrroline-5-carboxylate dehydrogenase that is encoded by the genome.[14]

Both *L. aggregata* IAM 12614 and *S. novella* DSM 506 encode pathways (Figure 5) for utilization of (1) *trans*-4-hydroxy-L-proline [catalyzed by orthologues of 4HypE (proline racemase superfamily), the FAD-dependent oxidoreductase (*c*4DHyp oxidase), and the member of the DHDPS superfamily (Δ^1-pyrroline 4-hydroxy 2-carboxylate dehydratase/deaminase, Pyr4H2C DD)],[4] (2) *trans*-3-hydroxy-L-proline [catalyzed by an orthologue of *t*3LHypD[6] and the Pyr2C reductase (vide supra)], and (3) *cis*-3-hydroxy-L-proline (catalyzed by *c*3LHypD and Pyr2C reductase). This work illustrates the use of genome context to facilitate the assignment of function to uncharacterized enzymes discovered in genome projects.

The X-ray coordinates and structure factors for *c*3HypD from *L. aggregata* IAM 12614 with Mg^{2+} in the active site have been deposited in the PDB (4MGG). This report describes the characterization of *in vitro* enzymatic activities of proteins with the following UniProt accession IDs: A0NXQ8, D7A0Y2, D7A0Y0, D7A0Y1, and D7A0Y3.

■ ASSOCIATED CONTENT

Supporting Information

A description of the experimental procedures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

j-gerlt@illinois.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by a program project grant and three cooperative agreements from the U.S. National Institutes of Health (P01GM071790, U54GM093342, U54GM074945, and U54GM094662). Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357. Use of the Lilly Research Laboratories Collaborative Access Team (LRL-CAT) beamline at Sector 31 of the Advanced Photon Source was provided by Eli Lilly Company, which operates the facility. We thank Dr. Xudong Guan, NMR specialist at the Institute for Genomic Biology, University of Illinois at Urbana−Champaign for assistance with the ^{1}H NMR kinetic experiments.

■ REFERENCES

(1) Babbitt, P. C.; Hasson, M. S.; Wedekind, J. E.; Palmer, D. R.; Barrett, W. C.; Reed, G. H.; Rayment, I.; Ringe, D.; Kenyon, G. L.; Gerlt, J. A. *Biochemistry* **1996**, *35*, 16489−16501.

(2) Gerlt, J. A.; Babbitt, P. C.; Rayment, I. *Arch. Biochem. Biophys.* **2005**, *433*, 59−70.

DOI: 10.1021/ja5103986
J. Am. Chem. Soc. 2015, 137, 1388−1391

(3) Gerlt, J. A.; Babbitt, P. C.; Jacobson, M. P.; Almo, S. C. *J. Biol. Chem.* **2012**, *287*, 29−34.
(4) Zhao, S.; Kumar, R.; Sakai, A.; Vetting, M. W.; Wood, B. M.; Brown, S.; Bonanno, J. B.; Hillerich, B. S.; Seidel, R. D.; Babbitt, P. C.; Almo, S. C.; Sweedler, J. V.; Gerlt, J. A.; Cronan, J. E.; Jacobson, M. P. *Nature* **2013**, *502*, 698−702.
(5) Atkinson, H. J.; Morris, J. H.; Ferrin, T. E.; Babbitt, P. C. *PLoS One* **2009**, *4*, e4345.
(6) Zhao, S.; Sakai, A.; Zhang, X.; Vetting, M. W.; Kumar, R.; Hillerich, B.; Francisco, B. S.; Solbiati, J.; Steves, A.; Brown, S.; Akiva, E.; Barber, A.; Seidel, R. D.; Babbitt, P. C.; Almo, S. C.; Gerlt, J. A.; Jacobson, M. P. *eLife* **2014**, DOI: 10.7554/eLife.03275.
(7) Buschiazzo, A.; Goytia, M.; Schaeffer, F.; Degrave, W.; Shepard, W.; Gregoire, C.; Chamond, N.; Cosson, A.; Berneman, A.; Coatnoan, N.; Alzari, P. M.; Minoprio, P. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1705−1710.
(8) Goytia, M.; Chamond, N.; Cosson, A.; Coatnoan, N.; Hermant, D.; Berneman, A.; Minoprio, P. *PLoS One* **2007**, *2*, e885.
(9) Visser, W. F.; Verhoeven-Duif, N. M.; de Koning, T. J. *J. Biol. Chem.* **2012**, *287*, 21654−21662.
(10) Palmer, D. R.; Hubbard, B. K.; Gerlt, J. A. *Biochemistry* **1998**, *37*, 14350−14357.
(11) Yew, W. S.; Fedorov, A. A.; Fedorov, E. V.; Almo, S. C.; Gerlt, J. A. *Biochemistry* **2007**, *46*, 9564−9577.
(12) Schmidt, D. M. Z.; Mundorff, E. C.; Dojka, M.; Bermudez, E.; Ness, J. E.; Govindarajan, S.; Babbit, P. C.; Minshull, J.; Gerlt, J. A. *Biochemistry* **2003**, *42*, 8387−8393.
(13) Muramatsu, H.; Mihara, H.; Kakutani, R.; Yasuda, M.; Ueda, M.; Kurihara, T.; Esaki, N. *J. Biol. Chem.* **2005**, *280*, 5329−5335.
(14) Srivastava, D.; Schuermann, J. P.; White, T. A.; Krishnan, N.; Sanyal, N.; Hura, G. L.; Tan, A.; Henzl, M. T.; Becker, D. F.; Tanner, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 2878−2883.

DOI: 10.1021/ja5103986
J. Am. Chem. Soc. 2015, 137, 1388−1391