THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 287, NO. 26, pp. 21654–21662, June 22, 2012

Identification of a Human *trans*-3-Hydroxy-L-proline Dehydratase, the First Characterized Member of a Novel Family of Proline Racemase-like Enzymes[S]

Received for publication, March 16, 2012 Published, JBC Papers in Press, April 23, 2012, DOI 10.1074/jbc.M112.363218

Wouter F. Visser[1], **Nanda M. Verhoeven-Duif, and Tom J. de Koning**

From the Department of Metabolic Diseases, University Medical Center Utrecht/Wilhelmina Children's Hospital Utrecht, 3508 AB Utrecht, The Netherlands

Background: The activity of a subset of proline racemase-like proteins was unresolved.
Results: Cys-Thr-type proline racemase-like enzymes, including human C14orf149, exhibit *trans*-3-hydroxy-L-proline dehydratase activity, whereas a T273C mutant of human C14orf149 exhibits epimerase activity.
Conclusion: Human C14orf149 is a previously uncharacterized *trans*-3-hydroxy-L-proline dehydratase.
Significance: These findings reveal a previously unknown pathway in human metabolism and will facilitate elucidation of the activity of related enzymes.

A family of eukaryotic proline racemase-like genes has recently been identified. Several members of this family have been well characterized and are known to catalyze the racemization of free proline or *trans*-4-hydroxyproline. However, the majority of eukaryotic proline racemase-like proteins, including a human protein called C14orf149, lack a specific cysteine residue that is known to be critical for racemase activity. Instead, these proteins invariably contain a threonine residue at this position. The function of these enzymes has remained unresolved until now. In this study, we demonstrate that three enzymes of this type, including human C14orf149, catalyze the dehydration of *trans*-3-hydroxy-L-proline to Δ^1-pyrroline-2-carboxylate (Pyr2C). These are the first enzymes of this subclass of proline racemase-like genes for which the enzymatic activity has been resolved. C14orf149 is also the first human enzyme that acts on *trans*-3-hydroxy-L-proline. Interestingly, a mutant enzyme in which the threonine in the active site is mutated back into cysteine regained 3-hydroxyproline epimerase activity. This result suggests that the enzymatic activity of these enzymes is dictated by a single residue. Presumably, human C14orf149 serves to degrade *trans*-3-hydroxy-L-proline from the diet and originating from the degradation of proteins that contain this amino acid, such as collagen IV, which is an important structural component of basement membrane.

Although prokaryotic proline racemase proteins have been known for a long time and have been well characterized, the first eukaryotic proline racemase was discovered only recently (1). It was initially discovered as an important factor for the mitogenic properties of *Trypanosoma cruzi*. Interestingly, the active site of the protein is required for its mitogenic properties (1, 2). Proline racemases have been evolutionarily preserved among trypanosomes and continue to attract considerable interest due to their potential as parasite-specific drug targets for the treatment of infections (3). The catalytic mechanism of these eukaryotic enzymes is essentially identical to that of the prokaryotic proline racemase proteins, which have been extensively studied (2, 4). Unlike alanine racemase and serine racemase, known proline racemases do not require pyridoxal-5-phosphate or any other cofactor. Instead, they operate by a two-base mechanism, in which two cysteine residues constitute the proton acceptor and donor that abstract and donate a proton from the proline α-carbon, respectively, thereby inversing its stereochemical configuration (2, 4). Interestingly, several studies (2, 5) have identified proline racemase-like genes that exhibit strong sequence similarity to proline racemases, but that appear to be missing a cysteine residue that is critical for the racemase activity. All known proteins of this type contain a threonine residue at this position instead. It has been confirmed that several of these proteins are devoid of racemase activity (5), as predicted, but their function has remained unresolved to date.

Most mammals contain a member of this subclass of proline racemase-like genes. The human gene is identified as *C14orf149*. Because considerable amounts of D-proline are present in human plasma, urine, and cerebrospinal fluid (6–8) and because D-serine and D-aspartate perform signaling functions in humans (9–11), it is conceivable that humans also synthesize and utilize D-proline. This prompted us to investigate the enzymatic activity of human C14orf149. To this end, we purified recombinant human C14orf149 and demonstrate that it is not a racemase but catalyzes the degradation of *trans*-3-hydroxy-L-proline (*trans*-3-L-Hyp) to Δ^1-pyrroline-2-carboxylate (Pyr2C).[2] In line with the finding that C14orf149 lacks one of the two cysteine residues that act as the proton acceptor and donor, the enzyme presumably operates by abstracting a proton from the *trans*-3-hydroxy-L-proline substrate, without donat-

[S] This article contains supplemental Figs. S1 and S2.

[1] To whom correspondence should be addressed: UMC Utrecht, Huispost KC 02.069.1, Postbus 85090, 3508 AB Utrecht, The Netherlands. Tel.: 31-88-7555318; Fax: 31-88-7554295; E-mail: w.visser-2@umcutrecht.nl.

[2] The abbreviations used are: Pyr2C, Δ^1-pyrroline-2-carboxylate; NIFE, *N*-(4-nitrophenoxycarbonyl)-L-phenylalanine 2-methoxyethyl ester; CV, cone voltage; Ecoll, collision energy; UPLC, ultra-high pressure liquid chromatography.

ing a proton from the other side of the ring. Interestingly, a mutant enzyme in which the Cys-273 residue was introduced regained racemase activity, catalyzing the conversion of *trans*-3-hydroxy-L-proline to *cis*-3-hydroxy-D-proline. Considering the evolutionary relationship between the proline racemase-like enzymes, our findings suggest that a mutation of one specific residue occurred during evolution, changing the function of the enzyme from a racemase to a dehydratase, and thus giving rise to a novel family of proline racemase-like proteins in humans and other animals. This hypothesis is strengthened by our observation that two other enzymes of this type, originating from *Mus musculus* and *Saccoglossus kowalevskii*, catalyze the same reaction. Presumably, the biological purpose of C14orf149 and its orthologues is to degrade *trans*-3-hydroxy-L-proline obtained through the diet and from the degradation of proteins that contain this amino acid, most notably collagen IV.

EXPERIMENTAL PROCEDURES

Materials—[D_7]-DL-Proline was purchased from Cambridge Isotope Laboratories (Andover, MA). UPLC grade solvents were from BioSolve (Valkenswaard, The Netherlands). Unless specified otherwise, all other chemicals came from Sigma-Aldrich (Zwijndrecht, The Netherlands).

Cloning and Mutagenesis—A construct containing the human *C14orf149* gene was obtained by PCR from IMAGE clone 4864867 (Promega, Leiden, The Netherlands) using primers 5′-AAAACTCgAgCCgACgCAgACCATggAgAg-3′ (forward) and 5′-AAAAgAATCCAAgTCACTTgAgAAgAAATCCATCCC-3′ (reverse) and ligated into a pGEM-T vector (Promega). This vector was sequenced using the BigDye Terminator method starting from the T7 and SP6 sites in the pGEM-T vector. A documented SNP (rs11555240) was found to occur in the sequence, resulting in a single amino acid residue change (R9W). Sequencing of the IMAGE vector revealed that the SNP was already present in this vector. Because it is a documented SNP and the amino acid residue is not evolutionarily conserved, the assumption was made that it does not fundamentally alter the enzymatic activity of the enzyme. The *C14orf149* construct was cut from the pGEM-T vector using XhoI and NotI restriction sites and ligated into a pGEX-6P1 vector (GE Healthcare, Diegem, Belgium), resulting in the vector pDA7. The T273C mutation was introduced in pDA7 by site-directed mutagenesis, using primers 5′-CAggTTgACAgAAgTCCCTgCggCTCAggAgTgACAgCC-3′ and 5′-ggCTgTCACTCCTgAgCCgCAgggACTTCTgTCAACCTg-3′, yielding the vector pDA31. The *M. musculus C14orf149* and *S. kowalevskii LOC100374014* genes were excised from synthetic plasmids (GenScript, Piscataway, NJ), using EcoRI and XhoI restriction sites, and ligated into a pGEX-6P1 expression vector, resulting in pDA36 and pDA37, respectively.

Expression and Purification—The pDA7, pDA31, pDA36, and pDA37 plasmids were used to transform the Rosetta *Escherichia coli* strain (Merck, Schiphol-Rijk, The Netherlands). The cells were grown in liquid LB medium containing the appropriate antibiotics to an A_{600} of 0.5. The cells were then transferred to 1 liter of Terrific Broth with additional glycerol (per liter: 12 g of Tryptone, 24 g of yeast extract, and 8 ml of glycerol) and appropriate antibiotics, at a density of $A_{600} = 0.1$. The cells were allowed to grow to a density of $A_{600} = 0.6$ at 22 °C, at which point 0.1 mM isopropyl-1-thio-β-D-galactopyranoside was added. After 5 h of incubation at 22 °C, the cells were harvested by centrifugation and washed once with ice-cold PBS. The pellet was resuspended in ice-cold PBS. The cells were lysed by sonication, and the lysate was cleared by centrifugation at 10,000 × *g* for 30 min. The protein was purified from the supernatant using glutathione-agarose 4B (GE Healthcare), according to the manufacturer's instructions. The *S. kowalevskii* LOC100374014 protein product was concentrated 10-fold by ultrafiltration using a 10-kDa cutoff filter (Millipore). The purified proteins were stored at −80 °C in a solution containing 50 mM Tris, pH 8.0, 15 mM reduced glutathione, and 20% w/v glycerol.

Substrate Specificity Assay—To check the substrate specificity, reaction mixtures containing 20 μg of the purified protein, 50 mM Tris·HCl, pH 8.0, 150 mM sodium chloride, and 1 mM substrate in a total volume of 100 μl were incubated at 37 °C. 5-μl samples of the reaction mixtures were taken at regular time intervals and immediately added to a vial containing an ice-cold mixture of 25 μl 50 μM [D_7]-DL-proline and 600 μl of acetonitrile to terminate the reaction and to deproteinize the sample. The terminated reactions were incubated on ice for 20 min and centrifuged (5 min at 13,000 rpm, at 4 °C). The supernatant was transferred to a glass vial and evaporated to dryness under N_2 flow, at 40 °C. The residue was subjected to derivatization with a chiral reagent to enable the resolution of the different isomers. For this purpose, the residue was dissolved in 50 μl of water and mixed with 35 μl of 0.15 M sodium tetraborate and 50 μl of (*S*)-NIFE solution (2.5 mg/ml in acetonitrile). After 10 min of incubation at room temperature, 10 μl of 4 M HCl and 355 μl of water were added, and the samples were vacuum-filtered through a 0.2-μm filter and analyzed using UPLC-MS/MS as follows. A Waters Acquity UPLC instrument (Milford, MA) was fitted with a 100 × 2.1-mm BEH-C18 1.7 μm column. A step gradient profile was used, consisting of 83% solvent A for 4.5 min followed by 75% A for 3 min, where solvent A is 0.1% (v/v) formic acid, and solvent B is acetonitrile. The MS/MS instrument (Waters Xevo) was configured for positive electrospray ionization and operated in multiple reaction monitoring mode. The following multiple reaction monitoring settings for the different compounds were used: hydroxyproline (all isomers), 381.35 > 132.1, CV = 22.0 V, Ecoll = 14.0 V; GABA, 353.36 > 120.1, CV = 22.0 V, Ecoll = 22.0 V; Pyr2C, 114.05 > 96.0, CV = 22.0 V, Ecoll = 14.0 V; proline, 365.3 > 120.1, CV = 24.0 V, Ecoll = 22.0 V; [D_7]-proline, 372.3 > 120.1, CV = 24.0 V, Ecoll = 22.0 V. Mass spectra obtained by direct infusion were recorded using a constant cone voltage of 24 V. The concentrations of the different compounds were calculated using a calibration curve of the different compounds prepared in water with [D_7]-L-Pro as the internal standard.

Colorimetric Enzyme Assay—Reaction mixtures containing 25 mM Tris·HCl pH 8.0, 1 mg/ml purified enzyme, and 25 mM *trans*-3-hydroxy-L-proline in a total volume of 50 μl were incubated at 37 °C for 60 s. The reactions were terminated by the addition of 900 μl of ice-cold 3% trichloroacetic acid. Then, 50

μl of a solution of 25 mg/ml 2-aminobenzaldehyde in ethanol was added to each terminated reaction, and the mixtures were incubated at 60 °C for 30 min. The absorbance was measured at 425 nm. For establishing the pH dependence, the Tris buffer was replaced with a mixture of 25 mM ammonium acetate and 25 mM potassium phosphate, the pH of which was adjusted using HCl and NaOH.

Preparation of Pyr2C—A reaction mixture containing 25 mM ammonium bicarbonate, pH 8.3, 1 mM D-proline, 0.25 mg of porcine D-amino acid oxidase, and 897 units of bovine catalase in a total volume of 100 μl was incubated for 60 min, at 37 °C. After this time, the reaction was terminated by the addition of 700 μl of ice-cold acetonitrile and incubated on ice for 30 min. The precipitate was removed by centrifugation (5 min, 20,000 × *g*), and the supernatant, containing Pyr2C, was used for experiments.

Tissue Distribution—A panel of glycerol-3-phosphate dehydrogenase-normalized cDNA samples prepared from human tissues was obtained from Clontech. PCR reactions were conducted with the following primers for C14orf149: 5′-ATGCTTATACCAAGGAACCAACC-3′ (forward) and 5′-ACTGTAAGGCAATTCGGGCTG-3′ (reverse). The primers included with the cDNA samples were used for glycerol-3-phosphate dehydrogenase.

Phylogenetic Analysis—The phylogenetic tree was estimated using maximum likelihood (phyML v3.0.1 (12)) using the BLOSUM62 substitution matrix and the nearest neighbor interchange algorithm. Protein sequences were initially aligned using Muscle (13) and then corrected manually.

RESULTS AND DISCUSSION

Inventory of Eukaryotic Proline Racemase-like Proteins—Several sequence motifs have been identified in proline racemase-like proteins that can be used to recognize proteins of this class in databases (5, 14). However, because it cannot be excluded that some homologs may have unusual sequence variations within these motifs, we made an inventory of eukaryotic proline racemase-like genes using pBLAST (15). The National Center for Biotechnology Information (NCBI) RefSeq database was searched using the *T. cruzi* proline racemase A (XP_811287.1) as the template, and the results were limited to eukaryotic organisms. Six sequences had low query coverage (<60%) and appeared to lack one of the two cysteine residues involved in the active site. These were assumed to be only partial sequences and were omitted. The resulting inventory of proline racemase-like proteins is listed in Table 1. These sequences were aligned (supplemental Fig. S2). A phylogenetic tree (Fig. 1) was derived from this alignment. The two cysteine residues that were previously found to be essential for the proline racemase activity are located at positions 145 and 364. The amino acids that are present at these two positions are included for each entry in Table 1. Interestingly, on the basis of these residues, three distinct types can be recognized: (i) proteins with a conserved pair of cysteine residues (Cys-Cys type); all known (hydroxy)proline racemases are of this type; (ii) proteins in which the Cys at position 145 is replaced with Ser (Ser-Cys type); and (iii) proteins in which the Cys at position 364 is replaced with Thr (Cys-Thr type). Only a minority of proline racemase-like proteins are of the Cys-Cys type known to catalyze racemization reactions: 19 out of 68 protein sequences (28%). 14 (21%) are of the Ser-Cys type, and 35 (51%) are of the Cys-Thr type. This suggests that the majority of the proline racemase-like proteins

TABLE 1

Inventory of eukaryotic proline racemase-like proteins obtained using pBLAST

The columns Cys-145 and Cys-364 list the active site residues at position 145 and 364. Underlining indicates a Cys in the active site.

Accession	Organism	Active site	
		Cys-145	Cys-364
XP_002913068.1	*Ailuropoda melanoleuca*	Cys	Thr
XP_003224770.1	*Anolis carolinensis*	Cys	Thr
XP_002374027.1	*Aspergillus flavus*	Cys	Thr
XP_002384351.1	*Aspergillus flavus*	Ser	Cys
XP_002384294.1	*Aspergillus flavus*	Ser	Cys
XP_002385107.1	*Aspergillus flavus*	Ser	Cys
XP_748142.1	*Aspergillus fumigatus*	Cys	Thr
XP_659725.1	*Aspergillus nidulans*	Cys	Thr
XP_001390789.1	*Aspergillus niger*	Cys	Thr
XP_001392593.2	*Aspergillus niger*	Cys	Cys
XP_001397532.2	*Aspergillus niger*	Ser	Cys
XP_003189630.1	*Aspergillus oryzae*	Cys	Thr
XP_001827113.2	*Aspergillus oryzae*	Ser	Cys
XP_001827065.1	*Aspergillus oryzae*	Ser	Cys
XP_001826937.2	*Aspergillus oryzae*	Ser	Cys
XP_001218441.1	*Aspergillus terreus*	Cys	Thr
NP_001029558.1	*Bos taurus*	Cys	Thr
XP_002604808.1	*Branchiostoma floridae*	Cys	Thr
XP_002604810.1	*Branchiostoma floridae*	Cys	Cys
XP_002604812.1	*Branchiostoma floridae*	Cys	Cys
XP_002604811.1	*Branchiostoma floridae*	Cys	Cys
XP_002604809.1	*Branchiostoma floridae*	Cys	Cys
XP_002587801.1	*Branchiostoma floridae*	Cys	Cys
XP_002807287.1	*Callithrix jacchus*	Cys	Thr
XP_547837.1	*Canis familiaris*	Cys	Thr
XP_002129543.1	*Ciona intestinalis*	Cys	Cys
XP_002124652.1	*Ciona intestinalis*	Cys	Cys
XP_694147.1	*Danio rerio*	Cys	Thr
XP_001492119.2	*Equus caballus*	Cys	Thr
XP_421428.2	*Gallus gallus*	Cys	Thr
XP_381905.1	*Gibberella zeae*	Ser	Cys
XP_388061.1	*Gibberella zeae*	Cys	Cys
NP_653182.1	*Homo sapiens*	Cys	Thr
XP_002160119.1	*Hydra magnipapillata*	Cys	Thr
XP_001094065.1	*Macaca mulatta*	Cys	Thr
XP_001377162.1	*Monodelphis domestica*	Cys	Thr
NP_080314.1	*Mus musculus*	Cys	Thr
XP_003042302.1	*Nectria haematococca*	Cys	Thr
XP_003049610.1	*Nectria haematococca*	Cys	Thr
XP_003052914.1	*Nectria haematococca*	Ser	Cys
XP_003040311.1	*Nectria haematococca*	Ser	Cys
XP_003042353.1	*Nectria haematococca*	Ser	Cys
XP_003044531.1	*Nectria haematococca*	Cys	Cys
XP_003043876.1	*Nectria haematococca*	Cys	Cys
XP_003043068.1	*Nectria haematococca*	Cys	Cys
XP_001627872.1	*Nematostella vectensis*	Cys	Thr
XP_001622857.1	*Nematostella vectensis*	Cys	Cys
XP_001266333.1	*Neosartorya fischeri*	Cys	Thr
XP_003267893.1	*Nomascus leucogenys*	Cys	Thr
XP_002723072.1	*Oryctolagus cuniculus*	Cys	Thr
XP_5009980.3	*Pan troglodytes*	Cys	Thr
XP_002556931.1	*Penicillium chrysogenum*	Cys	Thr
XP_002558663.1	*Penicillium chrysogenum*	Ser	Cys
XP_002148816.1	*Penicillium marneffei*	Cys	Thr
XP_002153067.1	*Penicillium marneffei*	Ser	Cys
XP_001792880.1	*Phaeosphaeria nodorum*	Cys	Cys
NP_001125373.1	*Pongo abelii*	Cys	Thr
NP_001101501.1	*Rattus norvegicus*	Cys	Thr
XP_002734182.1	*Saccoglossus kowalevskii*	Cys	Thr
XP_002734190.1	*Saccoglossus kowalevskii*	Cys	Cys
XP_002734189.1	*Saccoglossus kowalevskii*	Cys	Cys
XP_790644.1	*Strongylocentrotus purpuratus*	Cys	Thr
XP_787917.2	*Strongylocentrotus purpuratus*	Cys	Cys
XP_002200408.1	*Taeniopygia guttata*	Cys	Thr
XP_002487532.1	*Talaromyces stipitatus*	Ser	Cys
XP_811287.1	*Trypanosoma cruzi*	Cys	Cys
XP_803162.1	*Trypanosoma cruzi*	Cys	Cys
NP_001008128.1	*Xenopus (Silurana) tropicalis*	Cys	Thr

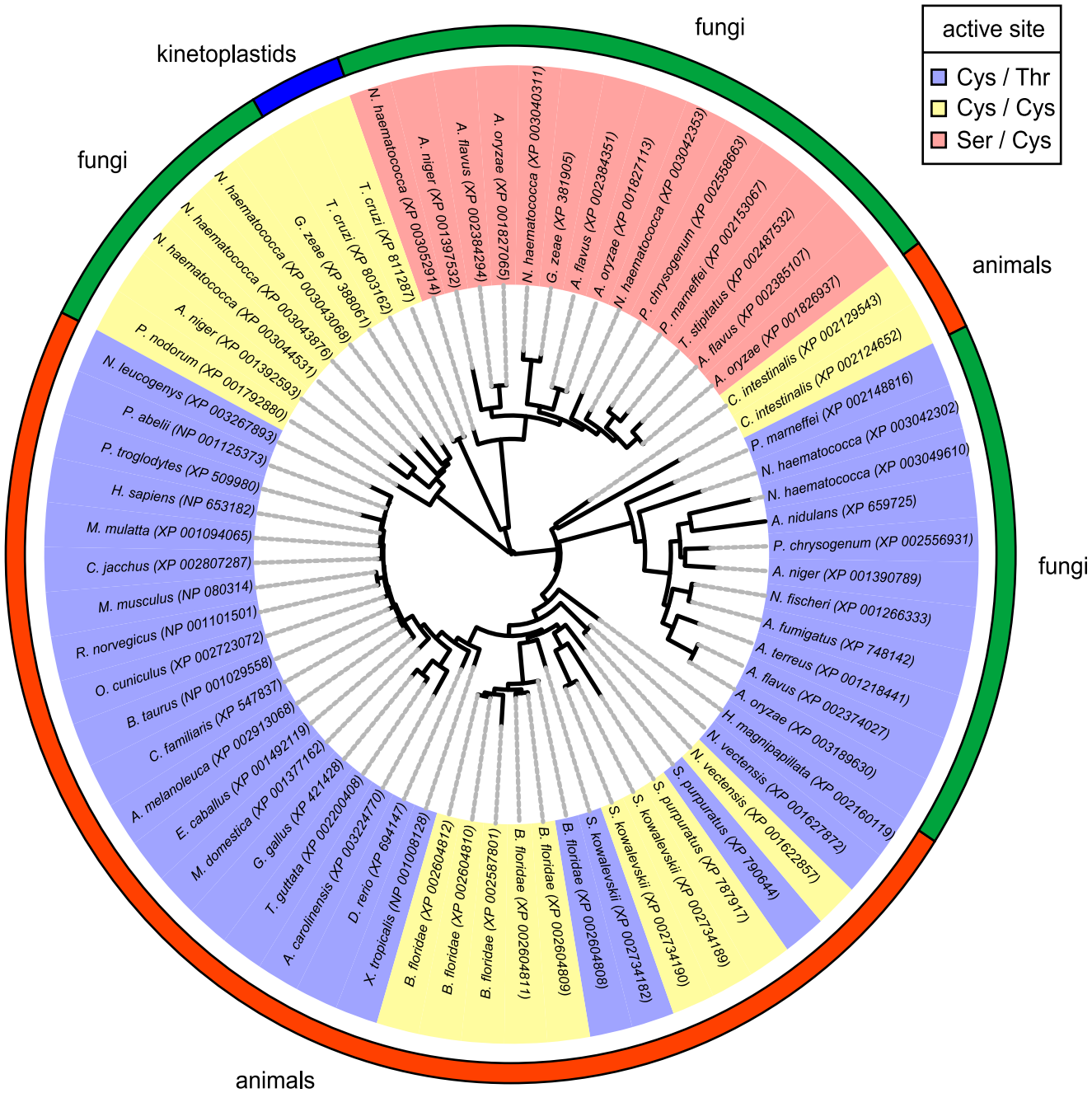

FIGURE 1. **Phylogenetic tree of proline racemase-like proteins.** The tree was drawn using iTol (27).

in this inventory do not catalyze a racemization reaction. A few general patterns can be recognized in the phylogenetic tree (Fig. 1). Nearly all animals contain a single gene of the Cys-Thr type, except the marine animals, which contain both a Cys-Thr and a Cys-Cys protein. The kinetoplastid *T. cruzi* contains two members of the Cys-Cys type. Proteins of the Ser-Cys type appear to be unique to fungi. In addition, both Cys-Thr-type and Cys-Cys-type proteins occur in fungi, often in the same species. In this inventory, the only exceptions to these patterns are *Ciona intestinalis* (a sea squirt), which despite being a marine animal appears to contain only two proteins of the Cys-Cys variant, and *Hydra magnipapillata* (a hydra), which contains only a Cys-Thr-type protein.

Cloning and Recombinant Expression of Human C14orf149—Interestingly, like that of most animals, the human genome also contains a gene encoding a proline racemase-like protein of the Cys-Thr type. It is annotated as a "probable proline racemase" or "hydroxyproline-2-epimerase," although the enzyme is of the Cys-Thr class and therefore lacks one of the cysteine residues known to be critical for the racemization/epimerization reaction. To investigate the enzymatic activity of this protein, we expressed a GST fusion construct of the human protein in *E. coli*. The protein was purified to >95% purity (Fig. 2). Under denaturating SDS-PAGE conditions, the protein migrates as a monomer of the expected molecular mass, 67 kDa (the GST tag contributes 29 kDa).

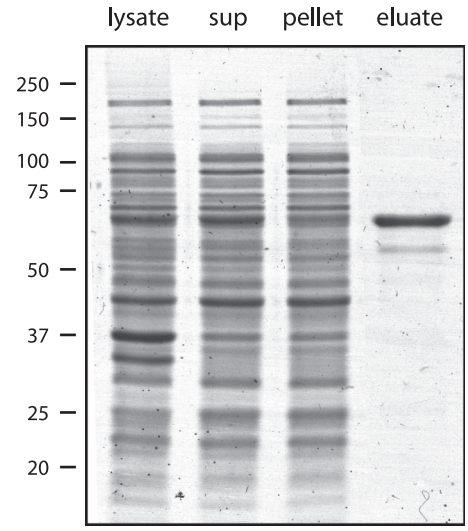

FIGURE 2. **Purification of GST-C14ORF149.** Shown is a Coomassie Brilliant Blue-stained SDS-PAGE gel loaded with different fractions obtained during purification of GST-C14ORF149. The *lysate lane* is the *E. coli* lysate obtained after disrupting the cells by sonication. The *sup* and *pellet lanes* are the supernatant and pellet fractions obtained by centrifugation of the lysate (30 min at 10,000 × *g*). The *eluate lane* is the purified protein fraction after elution from Sepharose 4B beads.

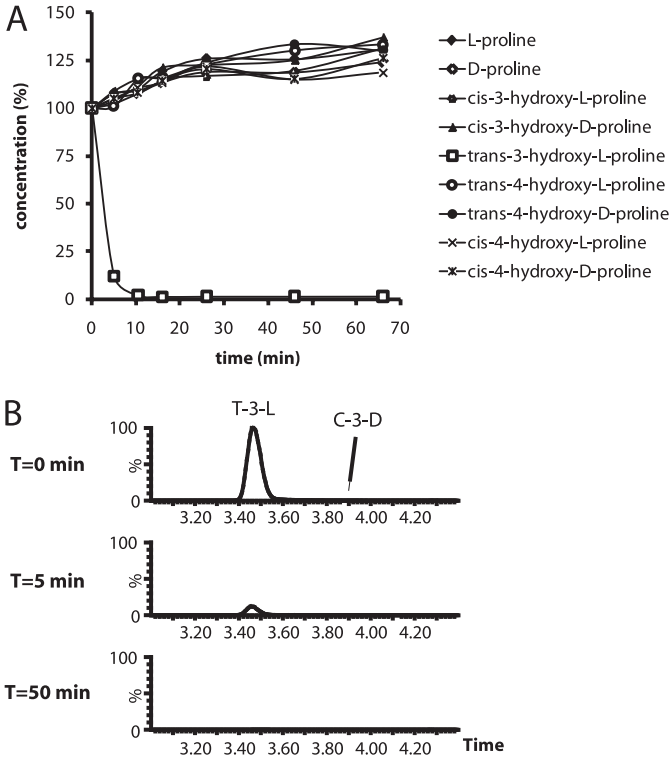

FIGURE 3. *A*, substrate utilization of C14orf149. The concentrations are expressed as the percentage of the amount initially present in the reaction mixture. The experiment was conducted three times. The data shown are from a representative experiment. *B*, C14orf149 does not catalyze the racemization of its substrate, *trans*-3-hydroxy-L-proline. The chromatographic elution profile of hydroxyproline is shown for three samples taken at the indicated time points. The expected elution times for the substrate (*trans*-3-hydroxy-L-proline, *T-3-L*) and the product (*cis*-3-hydroxy-D-proline, *C-3-D*) of the racemization reaction are indicated. The vertical scale for each of the three chromatograms is identical.

Substrate of Human C14orf149—Goytia *et al.* (14) identified three residues that are involved in substrate recognition, allowing them to distinguish between enzymes that act on proline from those that act on hydroxyproline, at least in a limited set of prokaryotic and kinetoplastid proline racemase-like proteins. However, in the majority of the eukaryotic members of this family examined here, including human C14orf149, different amino acids are present at the corresponding positions, which makes predictions regarding the substrate of these enzymes difficult. Therefore, the activity of purified GST-C14orf149 was experimentally evaluated with a wide range of isomers of proline and hydroxyproline. Interestingly, only *trans*-3-hydroxy-L-proline was rapidly consumed upon incubation with GST-C14orf149, whereas no reaction was observed with any of the other potential substrates (Fig. 3*A*). A slight increase in the concentrations of these substrates was observed, which might be attributed to slight evaporation of the reaction mixtures. To test whether C14orf149 can also act on *trans*-3-hydroxy-L-proline when it is present in a peptide, three different fragments of collagen known to contain this amino acid were incubated with C14orf149. However, in this case, no reaction was observed (supplemental Fig. S1).

Product of trans-3-Hydroxy-L-proline Degradation by C14orf149—When free *trans*-3-hydroxy-L-proline was incubated with C14orf149, the expected product of the 2-epimerase reaction, *cis*-3-hydroxy-D-proline, was not formed (Fig. 3*B*), indicating that the enzyme does not catalyze an epimerase reaction. To investigate the nature of the product of *trans*-3-hydroxy-L-proline degradation, the reaction mixture was monitored by mass spectrometry. The reaction was terminated at different time points and infused into the MS instrument directly. MS scans were made in the range $m/z = 2–650$. In the reaction mixture containing *trans*-3-hydroxy-L-proline as the substrate, the formation of a new peak was observed at $m/z = 114$ (Fig. 4*A*). This peak did not appear in reactions containing any of the other proline or hydroxyproline isomers (not shown) or in control reactions performed in the absence of enzyme (not shown). Because the (*S*)-NIFE reagent used for the chiral derivatization contributes a mass of 250 Da to the product and its derivatization products generally fragment to yield an ion with $m/z = 120.1$, which was not observed, the product with $m/z = 114$ did not appear to have reacted with the (*S*)-NIFE reagent. This implied that it did not contain an available amino group. To gain further information with respect to the identity of this compound, its mass was determined with 10-ppm accuracy using an LTQ orbitrap XL mass spectrometer (Thermo scientific, Breda, The Netherlands). The m/z of the compound was found to be 114.05445, which is consistent with the ion $[C_5H_7O_2N + H^+]$ (expected $m/z = 114.055504$). These results led us to hypothesize that the product of the reaction could be Δ^1-pyrroline-5-carboxylate (Pyr5C) or Pyr2C, both known intermediates of proline and hydroxyproline metabolism (16). These two compounds can be distinguished on the basis of their reaction with hydrogen peroxide (17). Pyr2C reacts with H_2O_2 to yield γ-aminobutyric acid (GABA) in nearly quantitative yield, whereas the reaction with Pyr5C yields small amounts of glutamate, but no GABA. Therefore, a sample of the reaction mixture was incubated with hydrogen peroxide, and the formation of GABA and glutamate was measured by UPLC-MS/MS. As shown in Fig. 4*B*, the reaction with H_2O_2 exclusively yielded

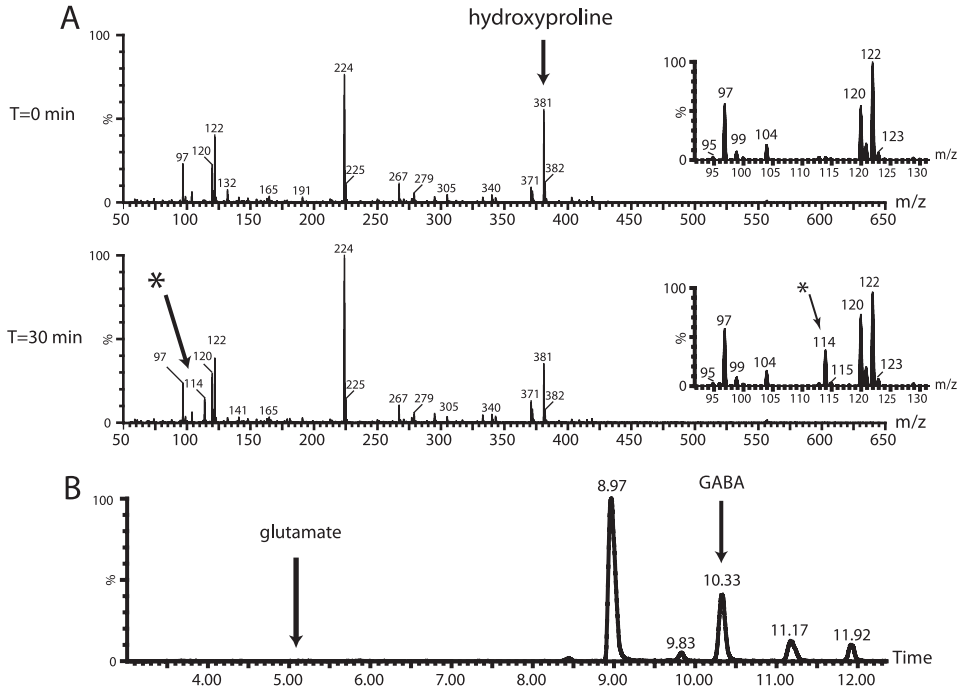

FIGURE 4. **Identification of reaction product of *trans*-3-hydroxy-L-proline degradation by C14orf149.** *A*, mass spectra of a reaction mixture containing C14orf149 and *trans*-3-hydroxy-L-proline, immediately after the addition of substrate and after 30 min of incubation. A newly formed peak at $m/z = 114$ is marked with an *asterisk*. Isomers of hydroxyproline have $m/z = 381$. The prominent peak at $m/z = 224$ is a side-product of the derivatization reaction. The *inset* shows the area around $m/z = 114$ in more detail. *B*, formation of GABA, but not glutamate, after incubation of the product with H_2O_2. Shown is a parent scan chromatogram from the reaction mixture after incubation with H_2O_2, showing the parents of $m/z = 120.1$. The expected retention times for GABA and glutamate are indicated.

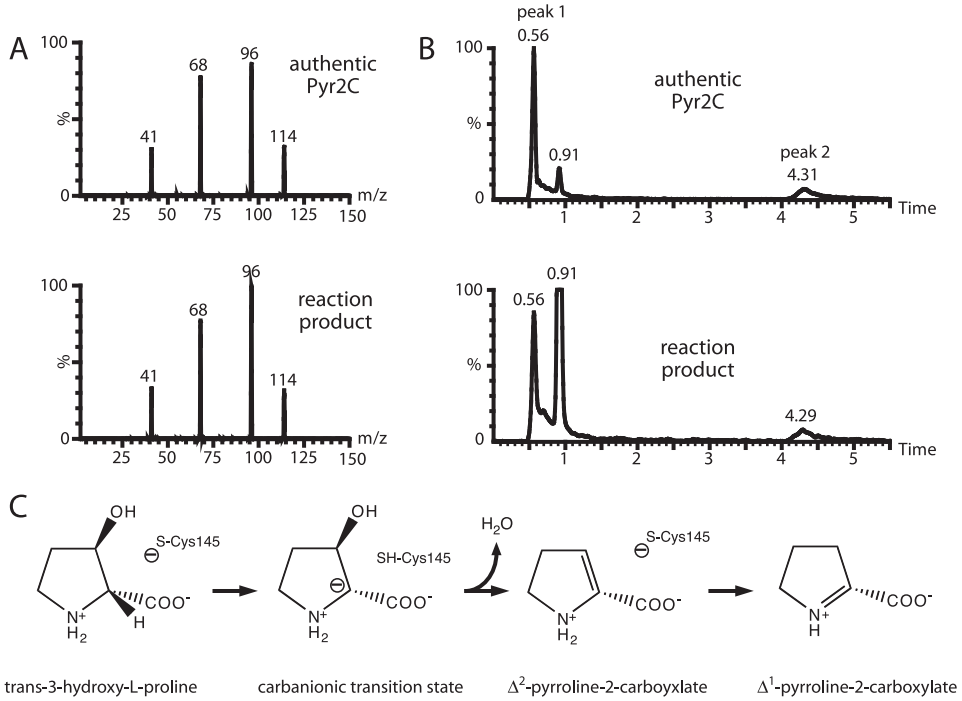

FIGURE 5. **Confirmation of enzymatic product as Δ^1-pyrroline-2-carboxylate.** *A*, fragmentation spectra of authentic Δ^1-pyrroline-2-carboxylate and that of the $m/z = 114$ product of *trans*-3-hydroxy-L-proline degradation by C14orf149. The spectra were recorded at a collision energy of 15 eV and cone voltage of 24 V. *B*, chromatograms of authentic Δ^1-pyrroline-2-carboxylate and of the product of *trans*-3-hydroxy-L-proline degradation by C14orf149. The chromatograms were recorded using the $114 > 96$ multiple reaction monitoring transition. *C*, the proposed reaction mechanism for the dehydration of *trans*-3-hydroxy-L-proline catalyzed by C14orf149.

GABA, suggesting that the product of the degradation of *trans*-3-hydroxy-L-proline by C14orf149 is Pyr2C. To further confirm the identity of the product, authentic Pyr2C was prepared by the reaction of D-proline and D-amino acid oxidase (18, 19). Fragmentation mass spectra of the ion at 114 *m*/*z* were recorded. As shown in Fig. 5*A*, the fragmentation spectra

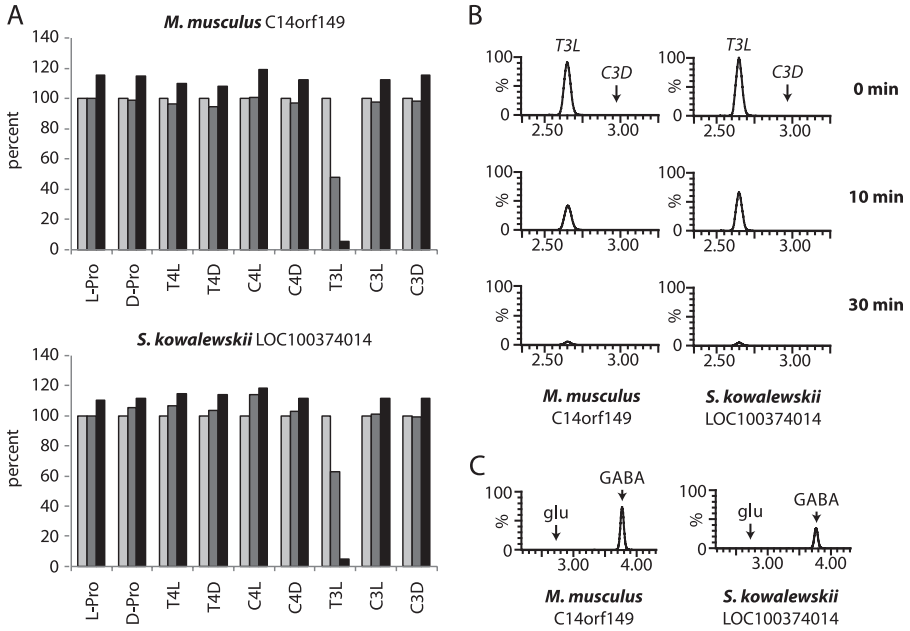

FIGURE 6. **Enzymatic activity of *M. musculus* C14orf149 and *S. kowalevskii* LOC100473014.** *A*, consumption of different potential substrates. Samples were taken at $T = 0$ (*light gray bars*), $T = 5$ min (*dark gray bars*), and $T = 10$ min (*black bars*). The isomers of hydroxyproline are abbreviated, indicating their configuration as follows: *T/C* (*trans*/*cis*), *3/4* (3-hydroxy or 4-hydroxy), and *D/L* (dextro/levo). For example, *T3L* is *trans*-3-hydroxy-L-proline. *B*, these enzymes do not catalyze the racemization of the substrate *trans*-3-hydroxy-L-proline. The chromatographic elution profile of hydroxyproline is shown for three samples taken at the indicated time points. The expected elution times for the substrate (*trans*-3-hydroxy-L-proline, *T3L*) and the product (*cis*-3-hydroxy-D-proline, *C3D*) of the racemization reaction are indicated. The vertical scale for each of the three chromatograms is identical. *C*, formation of GABA, but not glutamate, after incubation with H_2O_2. The expected retention times for these amino acids are indicated.

are highly similar, confirming the identity of the product. The chromatographic profiles of the authentic Pyr2C and the C14orf149 reaction mixtures were also compared, with the MS instrument set up to monitor the $114.05 > 96.0$ fragmentation reaction. As shown in Fig. 5*B*, the chromatographic elution profile of authentic Pyr2C matches that of the product of the C14orf149 reaction. The chromatogram exhibits two peaks with retention times of 0.56 and 4.31 min, which might be due to the spontaneous interconversion between Pyr2C and 5-amino-2-oxo-pentanoic acid. A third peak is observed, which elutes at 0.91 min, but this peak is present even in blank samples containing only water and derivatization agent and therefore unrelated to Pyr2C or the product of the C14orf149 reaction. Taken together, these experiments suggest that human C14orf149 operates according to the mechanism depicted in Fig. 5*C*, catalyzing the dehydration of *trans*-3-hydroxy-L-proline to Δ^1-pyrroline-2-carboxylate.

Other Cys-Thr-type Enzymes—To test whether other proline racemase-like proteins of the Cys-Thr type also catalyze a dehydratase reaction, we investigated the enzymatic activity of *M. musculus* C14orf149 (NCBI accession ID NP_080314.1) and *S. kowalevskii* LOC100374014 (XP_002734182.1). The *M. musculus* protein is orthologous to the human protein, whereas the *S. kowalevskii* protein is only distantly related (Fig. 1). GST fusion proteins were expressed in *E. coli*, purified, and incubated with different isomers of proline and hydroxyproline. As observed for the human enzyme, only *trans*-3-hydroxy-L-proline was consumed by these enzymes (Fig. 6*A*), demonstrating that they can utilize the same substrate. The expected product of the 2-epimerization reaction, *cis*-3-hydroxy-D-proline, was again not formed in these reactions (Fig. 6*B*). Instead, as found for human C14orf149, the product formed was Pyr2C, as evidenced by the appearance of GABA upon incubation of the reaction mixture with H_2O_2 (Fig. 6*C*).

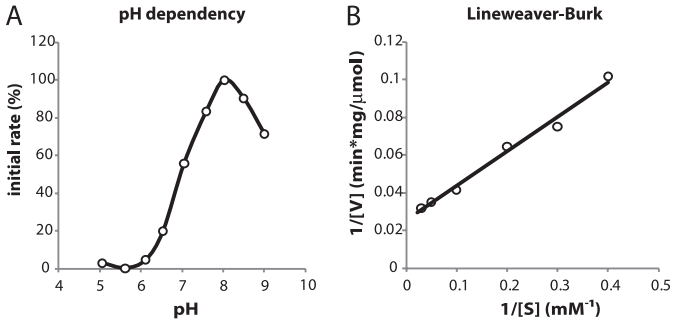

FIGURE 7. Kinetic parameters of C14orf149 (*A*) pH curve (*B*) Lineweaver-Burk plot.

Enzymatic Parameters of Human C14orf149—Accurate quantification of Pyr2C using UPLC-MS/MS was hindered by its elution in two asymmetric peaks. Therefore, a colorimetric assay was developed to enable accurate kinetic measurements. The assay is based on the reaction of *o*-aminobenzaldehyde with Pyr2C (18, 19), which yields a colored reaction product (λmax = 425 nm, $\epsilon = 1045\ \text{mol}^{-1}\text{cm}^{-1}$ in 3% trichloroacetic acid solution). Using this assay, the kinetic parameters of C14orf149 were established. Fig. 7*A* shows the pH dependence of the enzyme. The Michaelis-Menten parameters of the enzyme were obtained by measuring the initial reaction rate at different substrate concentrations. A Lineweaver-Burk plot calculated from these data is shown in Fig. 7*B*. From these results, the K_m and V_{max} were calculated and found to be 7.23 mM and 39.5 μmol/min/mg, respectively.

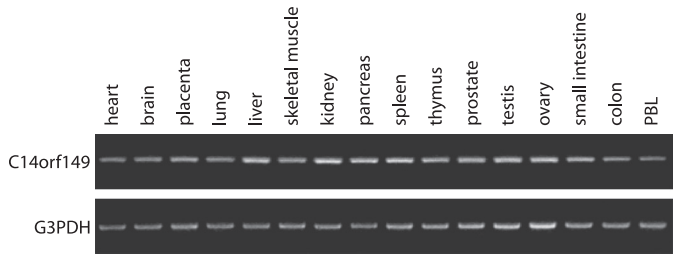

FIGURE 8. **Ubiquitous tissue distribution of C14orf149.** Shown are specific PCR products for C14orf149 (*upper panel*) and glycerol-3-phosphate dehydrogenase (*G3PDH*) (*lower panel*). *PBL*, peripheral blood leukocytes.

Tissue Distribution of Human C14orf149—The distribution of C14orf149 transcripts was examined in a panel of cDNA samples from 16 different human tissues (Fig. 8) using PCR. A product of the expected size was observed in all tissue samples, suggesting that human C14orf149 is ubiquitously expressed.

Racemase Activity after Introduction of a T273C Mutation—We hypothesized that the conversion of *trans*-3-hydroxy-L-proline to Δ^1-pyrroline-2-carboxylate occurs by abstraction of a hydrogen from the α-carbon of the substrate followed by spontaneous hydrolysis and tautomerization of the intermediate product (Fig. 5*C*) (20). It is then conceivable that upon reintroducing the second cysteine in the enzyme, lost in the evolutionary process, the enzyme might regain racemase activity. Therefore, we constructed a T273C mutant of human C14orf149 and incubated the purified, N-terminally GST-tagged mutant enzyme with *trans*-3-hydroxy-L-proline. As shown in Fig. 9*A*, the T273C mutant indeed yields the isomer expected from racemization, *cis*-3-hydroxy-D-proline. No Pyr2C could be detected in the reaction mixture, either by UPLC-MS/MS or spectrophotometrically (after the addition of *o*-aminobenzaldehyde), indicating that the mutant enzyme has lost its original dehydratase activity. The substrate is consumed by the mutant enzyme at a much lower rate (3.48 μmol/mg/min, or 8.8% of the wild type). Surprisingly, the mutant enzyme also catalyzed the racemization of L-proline to D-proline (Fig. 9*B*), albeit at a very low rate of 0.57 μmol/mg/min.

Conclusion—On the basis of two specific residues in the active site, located at positions 145 and 364 in our sequence alignment, three types of eukaryotic proline racemase-like genes can be discerned. The first type contains cysteine residues at both the 145 position and the 364 position (Cys-Cys type). All characterized members of this type are racemases that act on proline or hydroxyproline. The second type contains a Ser residue at the 145 position, rather than Cys (Ser-Cys type). Enzymes of this type seem to occur only in fungi, and their enzymatic activity has not been elucidated so far. In the third type, the residue at the 364 position is Thr, rather than Cys (Cys-Thr type). Like nearly all animals, humans contain a proline racemase-like gene of the Cys-Thr type, which is encoded by the C14orf149 gene. We show for the first time that this protein specifically catalyzes the dehydration of *trans*-3-hydroxy-L-proline. The product of the reaction is Δ^1-pyrroline-2-carboxylate (Fig. 5*C*). We propose to rename the *C14orf149* gene to L3HYPDH (*trans*-L-3-hydroxy-proline dehydratase). In addition to the human protein, we characterized two other enzymes of the same type, originating from *M. musculus*

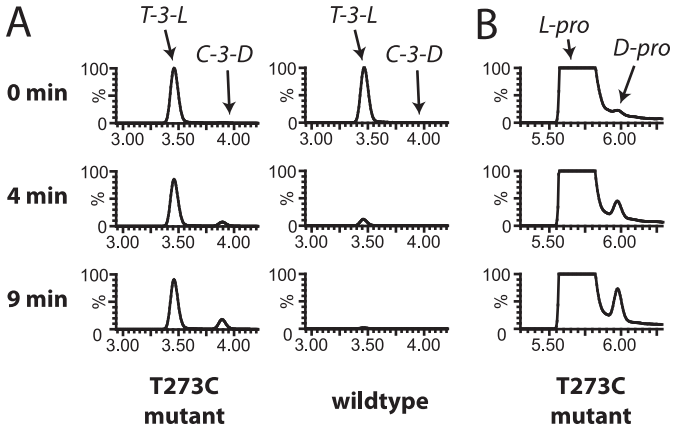

FIGURE 9. **Characterization of racemase activity of a T273C mutant enzyme.** *A*, *trans*-3-hydroxy-L-proline racemase activity. The mutant and wild-type enzymes were incubated with *trans*-3-hydroxy-L-proline (*T-3-L*), and the formation of *cis*-3-hydroxy-D-proline (*C-3-D*) was monitored by taking samples at regular intervals. The retention times of these isomers are indicated. The vertical scale is identical for all chromatograms. The experiment was conducted three times; a representative experiment is shown. *B*, the racemization of L-proline. Samples were taken from a reaction mixture that initially contained 1 mM L-proline at the indicated time points and analyzed. The vertical scale has been adjusted to show the formation of a small amount of D-proline, and the shown area of the L-proline peak is therefore not proportional to its concentration.

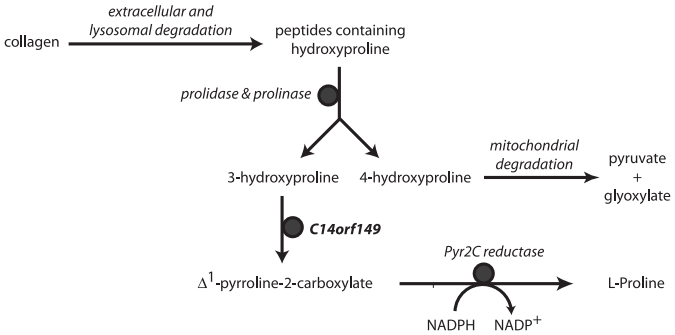

FIGURE 10. **Proposed model for role of C14orf149 in hydroxyproline metabolism.** Both *trans*-4-hydroxy-L-proline and *trans*-3-hydroxy-L-proline result from collagen breakdown, but they are degraded via different pathways. Pyr2C reductase (*Pyr2C reductase*) activity has been demonstrated in mammals (28, 29), but the gene encoding the enzyme has not yet been identified.

(C14orf149) and *S. kowalevskii* (LOC100374014). Both were found to possess the same enzymatic activity, which offers additional support for the hypothesis that Cys-Thr-type proteins catalyze dehydration reactions. To the best of our knowledge, this is the first study describing the enzymatic activity of proline racemase-like proteins of this type. It is also the first identification of a human protein that acts on *trans*-3-hydroxy-L-proline. It appears likely that C14orf149 serves to catalyze the *trans*-3-hydroxy-L-proline released by the degradation of proteins that contain this amino acid (Fig. 10) and from dietary sources. Small amounts of *trans*-3-hydroxy-L-proline have been shown to occur in different types of collagen, including collagen I, II, III, and V/XI, but it was found to be particularly abundant in collagen IV, in which *trans*-3-hydroxy-L-proline can be as much as 10% of the total hydroxyproline content (21). The *trans*-3-hydroxy-L-proline is introduced in this protein post-translationally by prolyl-3-hydroxylase enzymes (22, 23). Collagen IV is an important structural component of basement membrane, and the presence of *trans*-3-hydroxy-L-proline in

this protein is critical for its conformation and function (24). In this context, it is of interest to note that two independent studies (25, 26) have linked the genomic locus containing *C14orf149* (14q23.1) to the occurrence of aneurysms. Possibly, alterations in *trans*-3-hydroxy-L-proline metabolism lead to defects in the vascular basement membrane structure by affecting the level or stability of collagen IV. Now that the enzymatic activity of three different members of the Cys-Thr subclass of proline racemase-like proteins have been identified, it remains to be tested whether other members of this class have similar enzymatic activity. In these studies, different isomers of hydroxyproline and proline need to be included because it is currently difficult to predict the substrate specificity of proline racemase-like enzymes from sequence information alone. The enzymatic activity of the Ser-Cys class of proteins, unique to fungi, has not been explored at all to date. In light of the results presented in this study, it is tempting to speculate that proteins of this class catalyze the dehydration of the D-enantiomers of (hydroxy)proline. Future studies will have to reveal the true enzymatic activity of this previously unrecognized subclass of proline racemase-like proteins. The proline racemase from the *T. cruzi* parasite has been considered an attractive pharmacological target for the treatment of infections because it was believed that humans do not possess an enzyme with the same activity. However, it is conceivable that drugs that inhibit the *T. cruzi* racemase also interfere with the activity of human C14orf149, and we believe that this should be taken into consideration in future drug development. The origin of the appreciable quantities of D-proline observed in human tissues and body fluids (6–8) currently remains unresolved because C14orf149 has no detectable proline racemase activity. Because recent findings have shown that humans synthesize and utilize several other D-amino acids, including D-serine and D-aspartate, research into resolving the origin and possible function of D-proline in humans is of significant interest.

REFERENCES

1. Reina-San-Martín, B., Degrave, W., Rougeot, C., Cosson, A., Chamond, N., Cordeiro-Da-Silva, A., Arala-Chaves, M., Coutinho, A., and Minoprio, P. (2000) A B-cell mitogen from a pathogenic trypanosome is a eukaryotic proline racemase. *Nat. Med.* **6,** 890–897
2. Buschiazzo, A., Goytia, M., Schaeffer, F., Degrave, W., Shepard, W., Grégoire, C., Chamond, N., Cosson, A., Berneman, A., Coatnoan, N., Alzari, P. M., and Minoprio, P. (2006) Crystal structure, catalytic mechanism, and mitogenic properties of *Trypanosoma cruzi* proline racemase. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 1705–1710
3. Chamond, N., Cosson, A., Coatnoan, N., and Minoprio, P. (2009) Proline racemases are conserved mitogens: characterization of a *Trypanosoma vivax* proline racemase. *Mol. Biochem. Parasitol* **165,** 170–179
4. Rudnick, G., and Abeles, R. H. (1975) Reaction mechanism and structure of the active site of proline racemase. *Biochemistry* **14,** 4515–4522
5. Chamond, N., Grégoire, C., Coatnoan, N., Rougeot, C., Freitas-Junior, L. H., da Silveira, J. F., Degrave, W. M., and Minoprio, P. (2003) Biochemical characterization of proline racemases from the human protozoan parasite *Trypanosoma cruzi* and definition of putative protein signatures. *J. Biol. Chem.* **278,** 15484–15494
6. Visser, W. F., Verhoeven-Duif, N. M., Ophoff, R., Bakker, S., Klomp, L. W., Berger, R., and de Koning, T. J. (2011) A sensitive and simple ultra-high-performance-liquid chromatography-tandem mass spectrometry based method for the quantification of D-amino acids in body fluids. *J. Chromatogr. A* **1218,** 7130–7136
7. Nagata, Y., Masui, R., and Akino, T. (1992) The presence of free D-serine, D-alanine, and D-proline in human plasma. *Experientia* **48,** 986–988
8. Pätzold, R., Schieber, A., and Brückner, H. (2005) Gas chromatographic quantification of free D-amino acids in higher vertebrates. *Biomed. Chromatogr.* **19,** 466–473
9. Fuchs, S. A., Berger, R., Klomp, L. W., and de Koning, T. J. (2005) D-Amino acids in the central nervous system in health and disease. *Mol. Genet. Metab.* **85,** 168–180
10. Kim, P. M., Duan, X., Huang, A. S., Liu, C. Y., Ming, G. L., Song, H., and Snyder, S. H. (2010) Aspartate racemase, generating neuronal D-aspartate, regulates adult neurogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 3175–3179
11. Wolosker, H., Sheth, K. N., Takahashi, M., Mothet, J. P., Brady, R. O., Jr., Ferris, C. D., and Snyder, S. H. (1999) Purification of serine racemase: biosynthesis of the neuromodulator D-serine. *Proc. Natl. Acad. Sci. U.S.A.* **96,** 721–725
12. Guindon, S., and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52,** 696–704
13. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32,** 1792–1797
14. Goytia, M., Chamond, N., Cosson, A., Coatnoan, N., Hermant, D., Berneman, A., and Minoprio, P. (2007) Molecular and structural discrimination of proline racemase and hydroxyproline-2-epimerase from nosocomial and bacterial pathogens. *PLoS One* **2,** e885
15. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25,** 3389–3402
16. Adams, E., and Frank, L. (1980) Metabolism of proline and the hydroxyprolines. *Annu. Rev. Biochem.* **49,** 1005–1061
17. Mestichelli, L. J., Gupta, R. N., and Spenser, I. D. (1979) The biosynthetic route from ornithine to proline. *J. Biol. Chem.* **254,** 640–647
18. Meister, A. (1954) The α-keto analogues of arginine, ornithine, and lysine. *J. Biol. Chem.* **206,** 577–585
19. Wellner, D., and Scannone, H. (1964) Oxidation of L-proline and L-3,4-dehydroproline by D-amino acid oxidase. *Biochemistry* **3,** 1746–1749
20. Radhakrishnan, A. N., and Meister, A. (1957) Conversion of hydroxyproline to pyrrole-2-carboxylic acid. *J. Biol. Chem.* **226,** 559–571
21. Gryder, R. M., Lamon, M., and Adams, E. (1975) Sequence position of 3-hydroxyproline in basement membrane collagen: Isolation of glycyl-3-hydroxyprolyl-4-hydroxyproline from swine kidney. *J. Biol. Chem.* **250,** 2470–2474
22. Tiainen, P., Pasanen, A., Sormunen, R., and Myllyharju, J. (2008) Characterization of recombinant human prolyl 3-hydroxylase isoenzyme 2, an enzyme modifying the basement membrane collagen IV. *J. Biol. Chem.* **283,** 19432–19439
23. Vranka, J. A., Sakai, L. Y., and Bächinger, H. P. (2004) Prolyl 3-hydroxylase 1, enzyme characterization and identification of a novel family of enzymes. *J. Biol. Chem.* **279,** 23615–23621
24. Krane, S. M. (2008) The importance of proline residues in the structure, stability, and susceptibility to proteolytic degradation of collagens. *Amino Acids* **35,** 703–710
25. Mineharu, Y., Inoue, K., Inoue, S., Kikuchi, K., Ohishi, H., Nozaki, K., Hashimoto, N., and Koizumi, A. (2008) Association analyses confirming a susceptibility locus for intracranial aneurysm at chromosome 14q23. *J. Hum. Genet.* **53,** 325–332
26. Ozturk, A. K., Nahed, B. V., Bydon, M., Bilguvar, K., Goksu, E., Bademci, G., Guclu, B., Johnson, M. H., Amar, A., Lifton, R. P., and Gunel, M. (2006) Molecular genetic analysis of two large kindreds with intracranial aneurysms demonstrates linkage to 11q24–25 and 14q23–31. *Stroke* **37,** 1021–1027
27. Letunic, I., and Bork, P. (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Research* **39,** W475-W478
28. Garweg, G., von Rehren, D., and Hintze, U. (1980) L-Pipecolate formation in the mammalian brain: regional distribution of Δ^1-pyrroline-2-carboxylate reductase activity. *J. Neurochem.* **35,** 616–621
29. Meister, A., Radhakrishnan, A. N., and Buckley, S. D. (1957) Enzymatic synthesis of L-pipecolic acid and L-proline. *J. Biol. Chem.* **229,** 789–800