Published in final edited form as:
Arch Biochem Biophys. 2013 October 1; 538(1): 34–40. doi:10.1016/j.abb.2013.07.024.

Structural basis of substrate selectivity of Δ¹-pyrroline-5-carboxylate dehydrogenase (ALDH4A1): Semialdehyde chain length

Travis A. Pemberton[a] and **John J. Tanner**[a,b,*]

[a]Department of Chemistry, University of Missouri-Columbia, Columbia, MO 65211, USA

[b]Department of Biochemistry, University of Missouri-Columbia, Columbia, MO 65211, USA

Abstract

The enzyme Δ^1-pyrroline-5-carboxylate (P5C) dehydrogenase (aka P5CDH and ALDH4A1) is an aldehyde dehydrogenase that catalyzes the oxidation of γ-glutamate semialdehyde to L-glutamate. The crystal structures of mouse P5CDH complexed with glutarate, succinate, malonate, glyoxylate, and acetate are reported. The structures are used to build a structure-activity relationship that describes the semialdehyde carbon chain length and the position of the aldehyde group in relation to the cysteine nucleophile and oxyanion hole. Efficient 4- and 5-carbon substrates share the common feature of being long enough to span the distance between the anchor loop at the bottom of the active site and the oxyanion hole at the top of the active site. The inactive 2- and 3-carbon semialdehydes bind the anchor loop but are too short to reach the oxyanion hole. Inhibition of P5CDH by glyoxylate, malonate, succinate, glutarate, and L-glutamate is also examined. The K_i values are 0.27 mM for glyoxylate, 58 mM for succinate, 30 mM for glutarate, and 12 mM for L-glutamate. Curiously, malonate is not an inhibitor. The trends in K_i likely reflect a trade-off between the penalty for desolvating the carboxylates of the free inhibitor and the number of compensating hydrogen bonds formed in the enzyme-inhibitor complex.

Keywords

X-ray crystallography; Aldehyde dehydrogenase; ALDH4A1; Proline catabolism; Substrate recognition

Introduction

Δ^1-Pyrroline-5-carboxylate (P5C)[1] dehydrogenase (P5CDH) is an NAD^+-dependent aldehyde dehydrogenase that catalyzes the last step of proline catabolism, the oxidation of γ-glutamate semialdehyde (GSA) to L-glutamate (Fig. 1, upper reactions) [1]. The aldehyde substrate for P5CDH in this pathway is generated by the hydrolysis of P5C, which is produced from the oxidation of proline by the flavoenzyme proline dehydrogenase

*Corresponding author at: Department of Chemistry, University of Missouri-Columbia, 601 S. College Ave., Columbia, MO 65211, USA. Fax: +1 573 882 2754. tannerjj@missouri.edu (J.J. Tanner).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.abb.2013.07.024.

[1]*Abbreviations used:* P5C, Δ^1-pyrroline-5-carboxylate; P5CDH, Δ^1-pyrroline-5-carboxylate dehydrogenase; GSA, γ-glutamate semialdehyde; PRODH, proline dehydrogenase; ALDH, aldehyde dehydrogenase; MmP5CDH, Δ^1-pyrroline-5-carboxylate dehydrogenase from *Mus musculus*.

(PRODH), known as proline oxidase in humans. Mammalian P5CDHs also function in hydroxyproline catabolism [2]. This pathway begins with the oxidation of hydroxyproline by hydroxyproline oxidase and ends with the P5CDH-catalyzed oxidation of 4-hydroxyglutamate semialdehyde to 4-erythro-hydroxy-L-glutamate (Fig. 1, lower reactions) [3]. In eukaryotes, P5CDH is a nuclear-encoded mitochondrial matrix enzyme, whereas in some bacteria P5CDH is combined with PRODH into a single bifunctional enzyme known as proline utilization A [1].

P5CDHs are of considerable biomedical importance. Mutations in the P5CDH gene (*ALDH4A1*) that abrogate enzyme activity cause the metabolic disorder type II hyperprolinemia [4–6]. Both proline catabolic genes are activated by the tumor suppressor p53 [7,8], indicating a role for the pathway in apoptosis, stress, and cancer [9–11]. In this context, the coordination of P5CDH, proline oxidase, and the proline biosynthetic enzyme P5C reductase is required for maintaining the proper balance of mitochondrial reactive oxygen species [12]. In the fungal pathogen *Cryptococcus neoformans*, P5CDH is required for optimal production of the major virulence factors [13]. And bacterial P5CDHs are being considered as components of vaccines against *Staphylococcus aureus* [14].

P5CDH belongs to the aldehyde dehydrogenase (ALDH) superfamily and is known as ALDH4A1. The ALDH superfamily comprises hundreds of distinct genes from all three domains of life, including 19 human ALDHs [15]. ALDHs share a common protein fold (Fig. 2) and catalytic mechanism for the oxidation of aldehydes to carboxylates. The mechanism [16–18] begins with the binding of the substrate aldehyde group in the oxyanion hole, which positions the C atom of the aldehyde for nucleophilic attack by the essential Cys residue. Nucleophilic attack results in formation of a hemithioacetal intermediate. Hydride transfer to $NAD(P)^+$ generates NAD(P)H and the thioacylenzyme intermediate. Finally, hydrolysis of the thioacylenzyme yields the carboxylate product and regenerates the Cys nucleophile.

P5CDH was first characterized in the late 1980s [19–22]. In this early work performed on placental and liver P5CDHs, several potential substrates were tested in steady-state kinetic assays in order to help establish P5C/GSA as the physiological substrate [19,21,22]. The human enzyme exhibits good activity with glutarate semialdehyde and adipate semialdehyde. The catalytic efficiencies (V/K_m) for these substrates at pH 7.0 are about 50% of that of the physiological substrate, P5C/GSA. The shorter semialdehyde, succinate semialdehyde, was also found to be a substrate, but the efficiency was somewhat lower at 16% of the value for P5C/GSA. Activity was undetectable with malonate semialdehyde, glyoxylate, and aspartate semialdehyde. These studies clearly suggested that semialdehyde chain length is an important factor in substrate selectivity, however; the structural basis for this aspect of substrate recognition has not been explored.

To better understand the basis of substrate selectivity by P5CDH, we have determined the high resolution crystal structures of *Mus musculus* P5CDH (MmP5CDH, 92% identical to human P5CDH) complexed with carboxylate ligands having chain lengths ranging from 5 carbons to 2 carbons: glutarate (**1** in Fig. 3), succinate (**2**), malonate (**3**), glyoxylate (**4**), and acetate (**5**). The structures explain the relationship between semialdehyde chain length and enzyme activity.

Experimental procedures

Crystallization and soaking

MmP5CDH was expressed, purified, and crystallized as described previously [23]. Briefly, the stock enzyme solution used for crystallization contained 10 mg/mL MmP5CDH in a

buffer of 50 mM Tris, 50 mM NaCl, 0.5 mM EDTA, 0.5 mM THP, and 5% glycerol at pH 7.5. This His-tag was not removed. Crystallization experiments were performed in sitting drops at room temperature with drops formed by mixing 1 μL of the enzyme solution and 1 μL of reservoir solution. The latter consisted of 15–25% (w/v) PEG 3350, 0.2 M Li_2SO_4, and 0.1 M Bis–Tris at pH 5.0–6.5. The space group is $P2_12_12_1$ with unit cell lengths of a = 85 Å, b = 94 Å, and c = 132 Å. The asymmetric unit contains two protein molecules, which form a dimer.

Crystals of Mm5CDH complexed with the anions in Fig. 3 were obtained by soaking the aforementioned crystals in a solution containing cryobuffer (24% PEG 3350, 0.1 M Bis–Tris at pH 6.5, 15–20% PEG 200) supplemented with high concentration of the ligand (440 mM acetate, 220 mM glyoxylate, 220 mM malonate, 270 mM succinate, or 270 mM glutarate). We note that the high concentration was necessary to displace the adventitious sulfate ion that binds to the active site. The crystals were soaked for about 15 min, harvested with Hampton loops, and flash-cooled in liquid nitrogen.

X-ray data collection and refinement

X-ray diffraction data were collected using a Rigaku rotating anode source with an R-AXIS IV++ detector. Each data set consisted of 240–260 images collected with an exposure time of 5 min per frame, oscillation width of 0.5°, and detector distance of 110 mm or 130 mm. The data were integrated in XDS [24] and scaled in SCALA [25]. Data processing statistics are listed in Table 1.

Refinement with PHENIX [26] was initiated from a 1.3 Å resolution structure of MmP5CDH (PDB code 4V9J). A common test set of reflections was used for cross validation, which was based on the one used previously for refinements of MmP5CDH [23,27]. COOT was used for model building [28]. The PHENIX elbow utility [29] was used to create ligand restraint files from ideal coordinates downloaded from PDB Ligand Expo [30]. Refinement statistics are listed in Table 1.

Steady-state inhibition kinetics

P5CDH activity was measured at 20 °C by monitoring the production of NADH at 340 nm as described previously for HsP5CDH [23]. The assay buffer contained 0.1 M sodium phosphate and 1 mM EDTA at pH 7.0. The enzyme concentration was 9.6 μg/ml (0.15 μM), except for the measurements performed in the presence of malonate, which were done with 3.2 μg/ml enzyme (0.05 μM). Inhibition of MmP5CDH by glyoxylate, malonate, succinate, glutarate, and L-glutamate was studied using succinate semialdehyde as the variable substrate at fixed NAD^+ concentration of 1.0 mM. (We note that P5C/GSA is not commercially available.) The substrate range was 10–400 μM when glyoxylate, succinate, or glutarate was the inhibitor, 10–300 μM with malonate as the inhibitor, and 10–350 μM with L-glutamate as the inhibitor. In each case, 10–12 substrate concentrations were used (Fig. S1). The inhibitors were present at the following concentrations: glyoxylate, 1 mM and 4 mM; malonate, 50 mM and 100 mM; succinate, 100 mM and 200 mM; glutarate, 100 mM and 200 mM; L-glutamate, 50 mM and 100 mM.

The kinetic data were analyzed by global fitting using Origin 9 software. Fitting the data sets to the competitive model (Eq. (1)) yielded R^2 values in the range 0.981–0.991 (Fig. S1), whereas the fits to the noncompetitive model (Eq. (2) with $\alpha = 1$) were noticeably worse ($R^2 < 0.956$). The noncompetitive model was thus rejected. Use of the more general mixed inhibition model (Eq. (2)) yielded fits that were indistinguishable (R^2 = 0.980–0.992) from those of the competitive model, and since the competitive model has one fewer parameter, it was selected for the final determination of kinetic parameters.

$$V = V_{\max}[S]/\{K_{\mathrm{m}}(1+[I]/K_{\mathrm{i}})+[S]\} \quad (1)$$

$$V = V_{\max}[S]/\{K_{\mathrm{m}}(1+[I]/K_{\mathrm{i}})+[S](1+[I]/\alpha K_{\mathrm{i}})\} \quad (2)$$

Results

5- and 4-Carbon carboxylate ligands: glutarate and succinate

The location of the aldehyde substrate-binding site has been deduced from crystal structures of P5CDHs complexed with the product L-glutamate [23,31]. L-glutamate binds in the cleft between the catalytic and NAD^+-binding domains, as demonstrated by glutarate bound to MmP5CDH (Fig. 2). The γ-carboxylate of L-glutamate represents the aldehyde group of GSA and binds in the oxyanion hole near the catalytic Cys residue, while the backbone of L-glutamate is anchored to a loop that connects the final strand of the catalytic domain to the NAD^+-binding domain. This loop is the second of two crossover peptides that connect the major domains of the enzyme and will be referred to as the anchor loop.

The canonical aldehyde-binding mode is exhibited by glutarate, the longest ligand used in the present study (Fig. 4A). One of the carboxylate groups of glutarate binds to the anchor loop by forming hydrogen bonds with Gly512 and Ser513. This carboxylate also forms a hydrogen bond with Ser349 of the catalytic loop. These interactions are identical to those formed by the α-carboxylate of L-glutamate (Fig. 4C), and by inference, GSA. The three methylene carbons of glutarate are sandwiched between Phe512 and Phe520 (Fig. 2, inset), which is also reminiscent of the L-glutamate complex. These residues are part of a more general aromatic box that is common to ALDHs [32]. The upper carboxylate of glutarate binds near catalytic Cys348 and thus represents the aldehyde group of glutarate semialdehyde (Fig. 4A). One of the O atoms of the upper carboxylate occupies the oxyanion hole and accepts hydrogen bonds from the side chain of Asn211 and the backbone of Cys348. The other O atom of the upper carboxylate extends toward a solvent cavity and represents the atom derived from nucleophilic attack of water on the thioacyl intermediate. The interactions of the upper carboxylate of glutarate are nearly identical to those of the γ-carboxylate of L-glutamate (Fig. 4C).

The position of the upper carboxylate is consistent with efficient nucleophilic attack by Cys348. In the glutarate complex, Cys348 has $\chi^1 = -65°$, which is the most favored rotamer for Cys but not the one that is populated during nucleophilic attack ($\chi^1 = 55°$). We note that the $\chi^1 = -65°$ rotamer is typically observed in P5CDH structures that lack NAD^+, as is the case here. Rotation of Cys348 to $\chi^1 = 55°$ brings the nucleophile into attack position and places the S atom 2.3 Å from the carboxylate C atom (Fig. 4C). For reference, this distance is also 2.3 Å in the MmP5CDH-L-glutamate complex. The structural similarity between the glutarate and L-glutamate poses is consistent with glutarate semialdehyde being a good substrate for human P5CDH [19,21,22].

The conformation of succinate is close to the canonical one. As observed for glutarate, the lower carboxylate group mimics the backbone of L-glutamate and is anchored to Gly512, Ser513, and Ser349 (Fig. 4B). The upper carboxylate, however, deviates from the canonical conformation. One of the O atoms of the upper carboxylate is in the oxyanion hole, as expected for a product, but the other one engages Ser349 rather than pointing to the water-filled cavity. As a result, when Cys348 is rotated into attack position ($\chi^1 = 55°$), the distance between the S atom and the carboxylate C atom is 2.6 Å, which is 0.3 Å longer than observed for L-glutamate and glutarate. The suboptimal placement of the upper carboxylate

of succinate is consistent with succinate semialdehyde being a less efficient substrate than GSA and glutarate semialdehyde [19,21,22].

3- and 2-Carbon ligands: malonate, glyoxylate, and acetate

The pose of malonate (Fig. 5A) deviates substantially from the canonical pose. Although malonate binds to the anchor loop and Ser349, it is too short to occupy the oxyanion hole. Rather, a water molecule fills the oxyanion hole. As a result, the upper carboxylate of malonate is too far from Cys348 for nucleophilic attack. Specifically, when Cys348 is rotated into the attack rotamer, the S-C distance is 3.2 Å. This result is consistent with the lack of enzymatic activity observed with malonate semialdehyde [21].

Electron density maps suggest that glyoxylate has two conformations that are related by a 180° rotation around the C–C bond axis (Fig. 5B). In both conformations, the carboxylate group is bound to the anchor residues. In conformation A (occupancy 0.6), the carbonyl is directed toward the substrate entrance channel and hydrogen bonds to Ser513. In conformation B (occupancy 0.4), the carbonyl points in the direction of the oxyanion hole. As observed with malonate, glyoxylate is too short to simultaneously occupy the anchor site and the oxyanion hole. This result is consistent with the lack of activity observed with glyoxylate [21].

The carboxylate group of acetate binds to the aldehyde backbone anchor residues (Fig. 5C). Obviously, if acetaldehyde binds similarly, the aldehyde group is out of reach of the Cys nucleophile. The fact that activity, albeit weak, is observed with acetaldehyde suggests that the acetate complex is not representative of enzyme-acetaldehyde complex.

Inhibition of Mm5CDH

The discovery that glyoxylate and malonate bind in the active site, yet neither glyoxylate nor malonate semialdehyde is a substrate for mammalian P5CDHs [21,22], suggested the possibility that glyoxylate and malonate could be inhibitors. This idea was tested with steady-state kinetic measurements. As a control, we verified that enzymatic activity was undetectable with glyoxylate as the substrate, in agreement with previous studies [21,22]. The analogous check with malonate semialdehyde was not performed due to the lack of a commercial source for this compound.

Glyoxylate is a sub-millimolar competitive inhibitor of MmP5CDH (Table 2). Initial rate data could be fit satisfactorily to the competitive inhibition model (Eq. (1)), which is consistent with glyoxylate binding in the aldehyde site, yielding K_i of 0.27 mM (Fig. S1A).

Surprisingly, malonate does not appear to be an inhibitor of MmP5CDH. Inhibition was not apparent even when malonate was present at 50–100 mM (Fig. S1B).

Inhibition studies using succinate, glutarate, and L-glutamate were also performed. All three compounds are millimolar competitive inhibitors (Fig. S1). The estimated K_i values are 58 mMfor succinate, 30 mM for glutarate, and 12 mM for L-glutamate. The K_i for L-glutamate agrees with value of 14 mM determined previously for human P5CDH [20].

Discussion

Pioneering studies of mammalian P5CDHs showed that the length of the semialdehyde chain is an important determinant of substrate selectivity [21]. The catalytic efficiency is highest for the 5-carbon substrates GSA and glutarate semialdehyde, somewhat lower for the 4-carbon succinate semialdehyde, and essentially zero with 3- and 2-carbon semialdehydes malonate semialdehyde and glyoxylate.

The structures reported here, along with that of the L-glutamate complex reported previously, provide a satisfying explanation for the substrate selectivity of mammalian P5CDH. Analysis of these structures suggests a structure-activity relationship consisting of four distances that describe the E–S complex (Fig. 6, Table 3). These parameters include the distance between the two terminal C atoms of the bound substrate (L_S), the nucleophilic attack distance (d_N), and the distances between the aldehyde carbonyl oxygen atom and the two hydrogen bond donors of the oxyanion hole (d_{OH1}, d_{OH2}). GSA is the most efficient substrate for P5CDH, and thus the distance parameters from the enzyme-L-glutamate complex represent the optimal case.

The distances describing the E–S complex for glutarate semialdehyde are nearly identical to those of GSA (Table 3), which is consistent with the observation that glutarate semialdehyde also has high catalytic efficiency. The somewhat higher efficiency of GSA is likely due to the fact that it forms an extra hydrogen bond that is not possible for glutarate semialdehyde (noted in Fig. 4C).

The parameters describing the E–S complex of succinate semialdehyde differ from those of the 5-carbon substrates GSA and glutarate semialdehyde. In particular, L_S is 1.7 Å shorter, d_N is 0.3 Å longer, and the oxyanion hole hydrogen bonding distances are 0.1–0.2 Å longer. These values reflect the fact that the aldehyde of succinate semialdehyde is not optimally positioned for nucleophilic attack and are consistent with the lower efficiency of succinate semialdehyde.

The deduced parameters for the inactive semialdehydes, malonate semialdehyde and glyoxylate, are clearly suboptimal. The chain lengths are only 2.5 and 1.5 Å for malonate semialdehyde and glyoxylate, respectively, and since these ligands bind to the anchor loop, the aldehyde groups are poorly positioned for nucleophilic attack. In particular, the attack distances are 3.2 and 4.4 Å for malonate semialdehyde and glyoxylate, respectively. Furthermore, the oxyanion hole distances are 3.6–5.7 Å, which is beyond optimal hydrogen bonding range. The parameters describing malonate semialdehyde and glyoxylate are thus consistent with the lack of activity observed with these semialdehydes. In summary, active semialdehyde substrates share the common feature of being long enough to span the distance between the anchor loop and the oxyanion hole, which is required to position the aldehyde group for nucleophilic attack by the catalytic Cys.

Absent structural information, the lack of catalytic activity observed with malonate semialdehyde and glyoxylate might lead one to suspect that these compounds and their respective dicarboxylates do not bind the enzyme active site. Our studies show otherwise; glyoxylate and malonate mimic actual substrates by binding to the anchor loop (Fig. 5A and B).

Glyoxylate was found to be a competitive inhibitor (K_i = 0.27 - mM). In fact, it is the most potent inhibitor reported to date for P5CDH. The pose of glyoxylate is similar to that of another known P5CDH inhibitor, L-proline. L-proline is a competitive inhibitor (with GSA) of human P5CDH with K_i of 3 mM [20], and we previously determined the structure of MmP5CDH complexed with L-proline [27]. Glyoxylate and L-proline bind to P5CDH similarly and form identical hydrogen bonds with the enzyme (Fig. 7). Note that the amine of proline and the aldehyde oxygen atom of glyoxylate (conformation A) form a hydrogen bond with the hydroxyl of Ser513. An analogous interaction is present in the L-glutamate complex (Fig. 4C) but absent in the glutarate, succinate, and malonate complexes.

The trends in K_i (i.e., affinity) of the ligands investigated here are perhaps unexpected but nonetheless consistent with the structures. Glyoxylate has the highest affinity of the ligands

tested. The extra hydrogen bond with Ser513 may partly account for this observation. Also, glyoxylate is accommodated in the active site in two orientations, suggesting a more favorable (or less unfavorable) entropy of binding. The desolvation of carboxylate versus aldehyde groups provides another possible explanation. The free energy of hydration of the acetate ion is approximately −80 kcal/mol, whereas that of acetaldehyde is only about −4 kcal/mol [33–35], implying that a larger desolvation penalty must be paid to bind a carboxylate group to the enzyme than an aldehyde group. Thus, binding a dicarboxylate ligand, such as malonate, presumably requires a larger desolvation penalty than glyoxylate, which contains only one carboxylate group. The higher affinity of glutarate compared to succinate likely reflects the better oxyanion hole hydrogen bonding of glutarate (Table 3). The higher affinity of L-glutamate compared to glutarate is probably due to the extra hydrogen bond between the amine of L-glutamate and Ser513. Comparing the K_i values for L-glutamate and glutarate suggests that this hydrogen bond contributes about −0.5 kcal/mol to the free energy of binding. Finally, it is puzzling that malonate does not bind to MmP5CDH in solution. This result perhaps reflects the fact that whereas two carboxylate groups must be desolvated, only the lower carboxylate forms compensating hydrogen bonds when bound to the enzyme (Fig. 5A). The upper carboxylate fails to engage the hydrogen bond donors of the oxyanion. These ideas about the trade-off between the desolvation penalty of the inhibitor and the compensating hydrogen bonds formed in the enzyme-inhibitor complex could potentially aid the discovery of new ALDH inhibitors [32,36].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health via Grant GM065546. We thank Dr. Dhiraj Srivastava for advice on expression, purification, and crystallization of MmP5CDH.

References

1. Tanner JJ. Amino Acids. 2008; 35:719–730. [PubMed: 18369526]
2. Adams E, Frank L. Annu. Rev. Biochem. 1980; 49:1005–1061. [PubMed: 6250440]
3. Valle D, Goodman SI, Harris SC, Phang JM, Clin J. Invest. 1979; 64:1365–1370.
4. Geraghty MT, Vaughn D, Nicholson AJ, Lin WW, Jimenez-Sanchez G, Obie C, Flynn MP, Valle D, Hu CA. Hum. Mol. Genet. 1998; 7:1411–1415. [PubMed: 9700195]
5. Phang, JM.; Hu, CA.; Valle, D. Metabolic and molecular basis of inherited disease. Scriver, CR.; Beaudet, AL.; Sly, WS.; Valle, D., editors. New York: McGraw Hill; 2001. p. 1821-1838.
6. Valle D, Goodman SI, Applegarth DA, Shih VE, Phang JM. J. Clin. Invest. 1976; 58:598–603. [PubMed: 956388]
7. Yoon KA, Nakamura Y, Arakawa H. J. Hum. Genet. 2004; 49:134–140. [PubMed: 14986171]
8. Polyak K, Xia Y, Zweier JL, Kinzler KW, Vogelstein B. Nature. 1997; 389:300–305. [PubMed: 9305847]
9. Phang JM, Donald SP, Pandhare J, Liu Y. Amino Acids. 2008; 35:681–690. [PubMed: 18401543]
10. Phang JM, Pandhare J, Liu Y. J. Nutr. 2008; 138:2008S–2015S. [PubMed: 18806116]
11. Phang JM, Liu W, Hancock C, Christian KJ. Front. Oncol. 2012; 2:60. [PubMed: 22737668]
12. Liang X, Zhang L, Natarajan SK, Becker DF. Antioxid. Redox Signal. 2013 http://dx.doi.org/10.1089/ars.2012.5074.
13. Lee IR, Lui EY, Chow EW, Arras SD, Morrow CA, Fraser JA. Genetics. 2013 http://dx.doi.org/10.1534/genetics.1113.150326.

14. Lijek RS, Luque SL, Liu Q, Parker D, Bae T, Weiser JN. Proc. Natl. Acad. Sci. USA. 2012; 109:13823–13828. [PubMed: 22869727]
15. Sophos NA, Vasiliou V. Chem. Biol. Interact. 2003:143–144. 5–22.
16. Perez-Miller SJ, Hurley TD. Biochemistry. 2003; 42:7100–7109. [PubMed: 12795606]
17. Steinmetz CG, Xie P, Weiner H, Hurley TD. Structure. 1997; 5:701–711. [PubMed: 9195888]
18. Farres J, Wang TT, Cunningham SJ, Weiner H. Biochemistry. 1995; 34:2592–2598. [PubMed: 7873540]
19. Farres J, Julia P, Pares X. Biochem. J. 1988; 256:461–467. [PubMed: 3223924]
20. Forte-McRobbie C, Pietruszko R. Biochem. J. 1989; 261:935–943. [PubMed: 2803253]
21. Forte-McRobbie CM, Pietruszko R. J. Biol. Chem. 1986; 261:2154–2163. [PubMed: 3944130]
22. Small WC, Jones ME. J. Biol. Chem. 1990; 265:18668–18672. [PubMed: 2211729]
23. Srivastava D, Singh RK, Moxley MA, Henzl MT, Becker DF, Tanner JJ. J. Mol. Biol. 2012; 420:176–189. [PubMed: 22516612]
24. Kabsch W. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:125–132. [PubMed: 20124692]
25. Evans P. Acta Cryst. 2006; D62:72–82.
26. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:213–221. [PubMed: 20124702]
27. Pemberton TA, Still BR, Christensen EM, Singh H, Srivastava D, Tanner JJ. Acta Crystallogr. D Biol. Crystallogr. 2012; 68:1010–1018. [PubMed: 22868767]
28. Emsley P, Cowtan K. Acta Cryst. 2004; D60:2126–2132.
29. Moriarty NW, Grosse-Kunstleve RW, Adams PD. Acta Crystallogr. D Biol. Crystallogr. 2009; 65:1074–1080. [PubMed: 19770504]
30. Feng Z, Chen L, Maddula H, Akcan O, Oughtred R, Berman HM, Westbrook J. Bioinformatics. 2004; 20:2153–2155. [PubMed: 15059838]
31. Inagaki E, Ohshima N, Takahashi H, Kuroishi C, Yokoyama S, Tahirov TH. J. Mol. Biol. 2006; 362:490–501. [PubMed: 16934832]
32. Riveros-Rosas H, Gonzalez-Segura L, Julian-Sanchez A, Diaz-Sanchez AG, Munoz-Clares RA. Chem. Biol. Interact. 2013; 202:51–61. [PubMed: 23219887]
33. Pearson RG. J. Am. Chem. Soc. 1986; 108:6109–6114.
34. Hawkins GD, Cramer CJ, Truhlar DG. J. Phys. Chem. B. 1997; 101:7147–7157.
35. Wang J, Wang W, Huo S, Lee M, Kollman PA. J. Phys. Chem. B. 2001; 105:5055–5067.
36. Koppaka V, Thompson DC, Chen Y, Ellermann M, Nicolaou KC, Juvonen RO, Petersen D, Deitrich RA, Hurley TD, Vasiliou V. Pharmacol. Rev. 2012; 64:520–539. [PubMed: 22544865]
37. Weiss M. J. Appl. Cryst. 2001; 34:130–135.
38. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Proteins. 2003; 50:437–450. [PubMed: 12557186]
39. DeLano, WL. The PyMOL User's Manual, DeLano Scientific. Palo Alto, CA, USA: 2002.

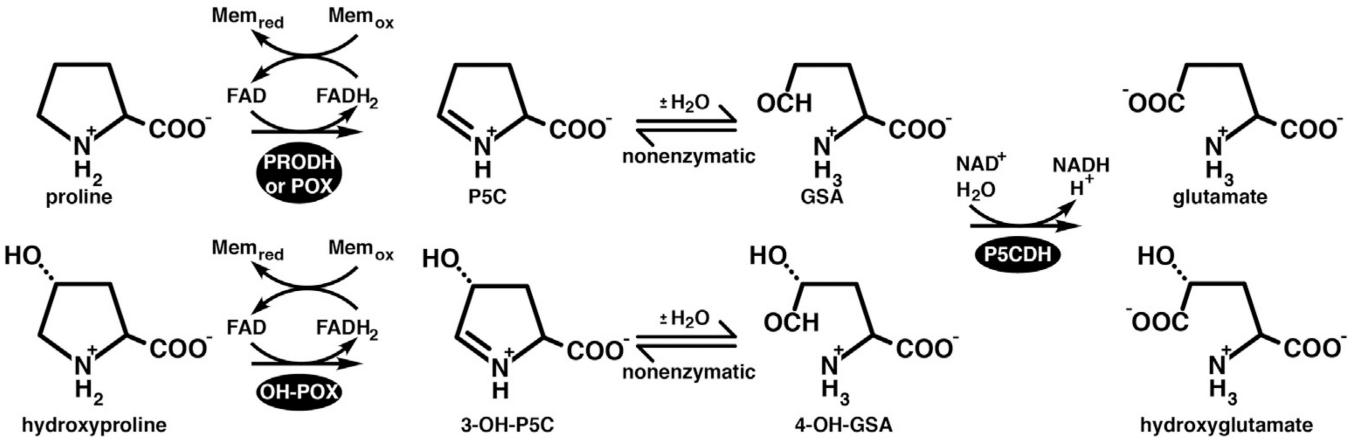

Fig. 1.
The reactions of proline (upper) and hydroxyproline (lower) catabolism in mammals.

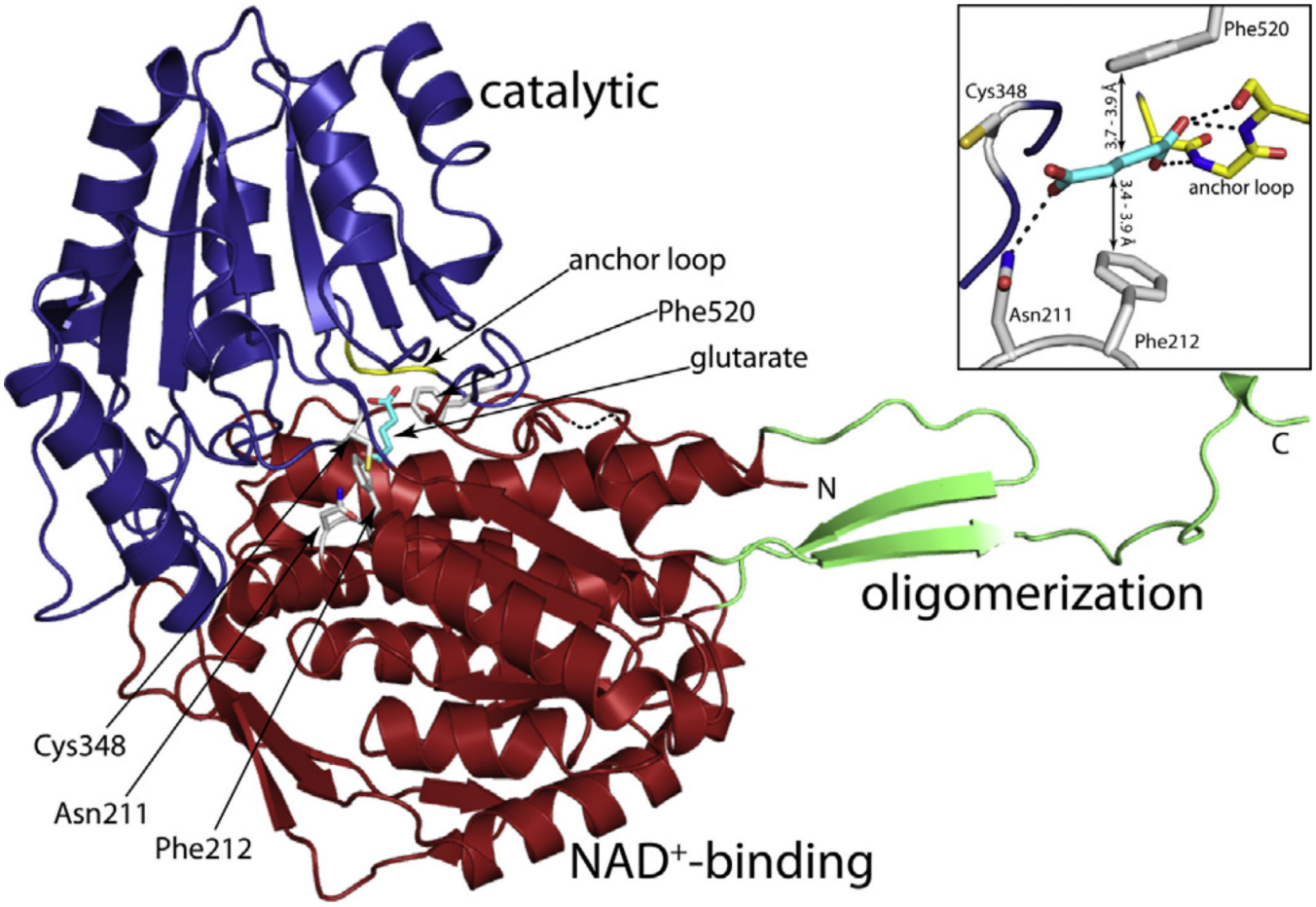

Fig. 2.

Structure of MmP5CDH complexed with glutarate (cyan). The NAD^+-binding, catalytic, and oligomerization domains are colored red, blue, and green, respectively. The side chains of selected active site residues are shown. The inset shows the essential elements of substrate recognition. This figure and others were created with PyMOL [39]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

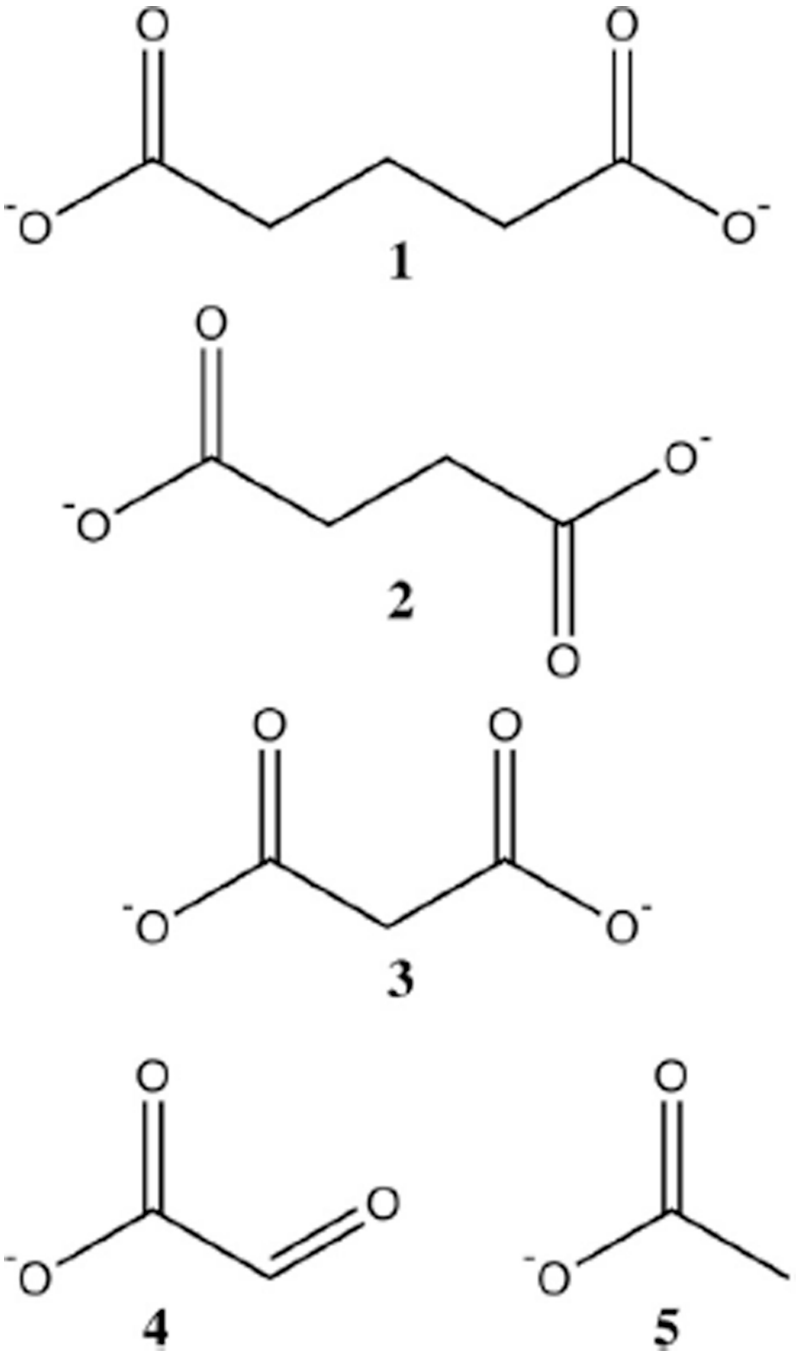

Fig. 3.
Ligands used in crystal structure determinations: (1) glutarate, (2) succinate, (3) malonate, (4) glyoxylate, and (5) acetate.

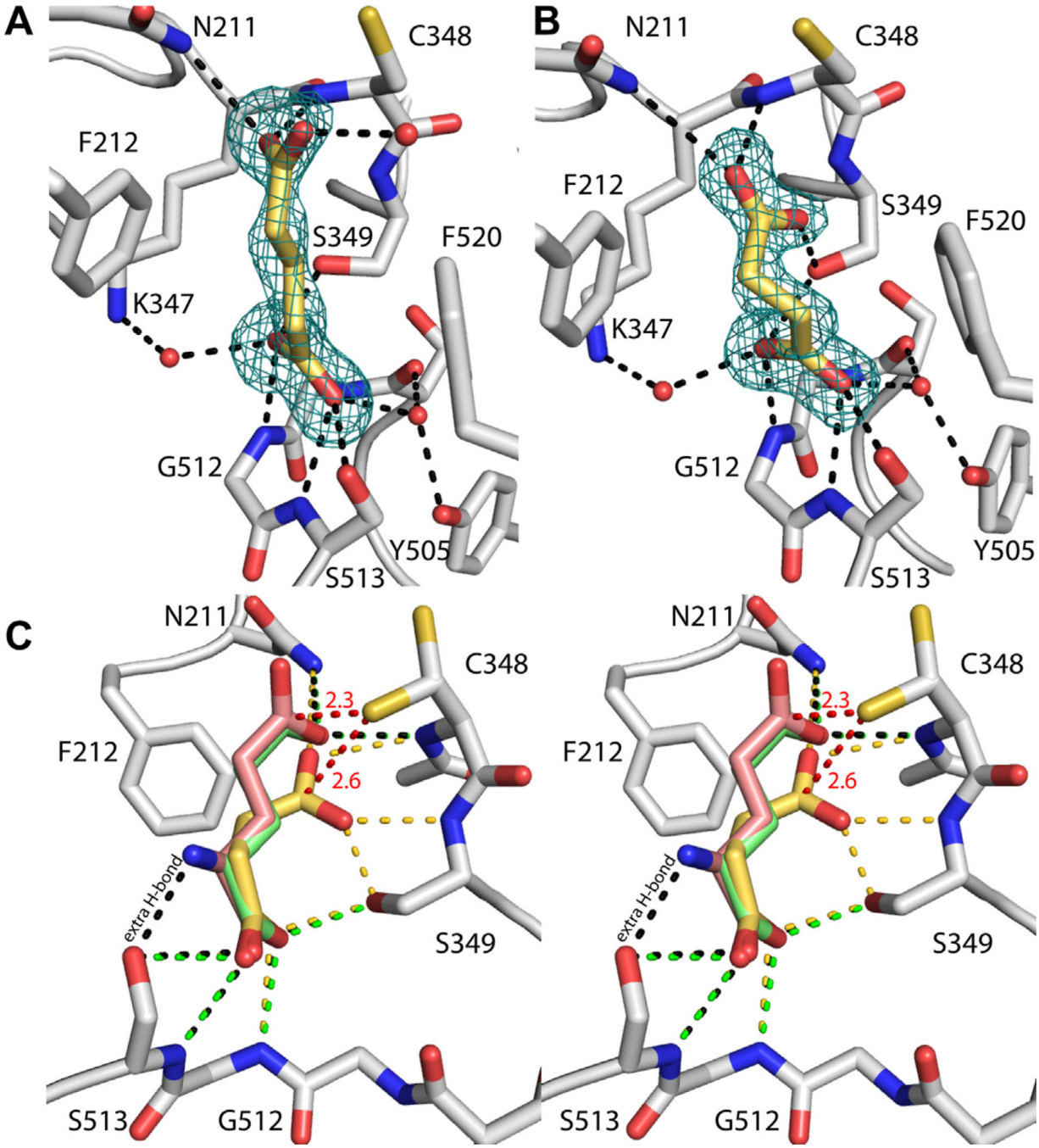

Fig. 4.

The binding of products to MmP5CDH. Electron density and interactions for (A) glutarate and (B) succinate. The cages represent simulated annealing σ_A-weighted F_o–F_c omit maps contoured at 3.0 σ. (C) Stereographic view (relaxed) of a superposition of the active sites of MmP5CDH complexed with L-glutamate (salmon carbons, black dashes, PDB 3V9K), glutarate (green carbons, green dashes), and succinate (gold carbons, gold dashes). The protein for the L-glutamate complex is shown in white. The red dashes represent the inferred nucleophilic attack distances (2.3 Å for GSA and glutarate semialdehyde; 2.6 Å for succinate semialdehyde). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

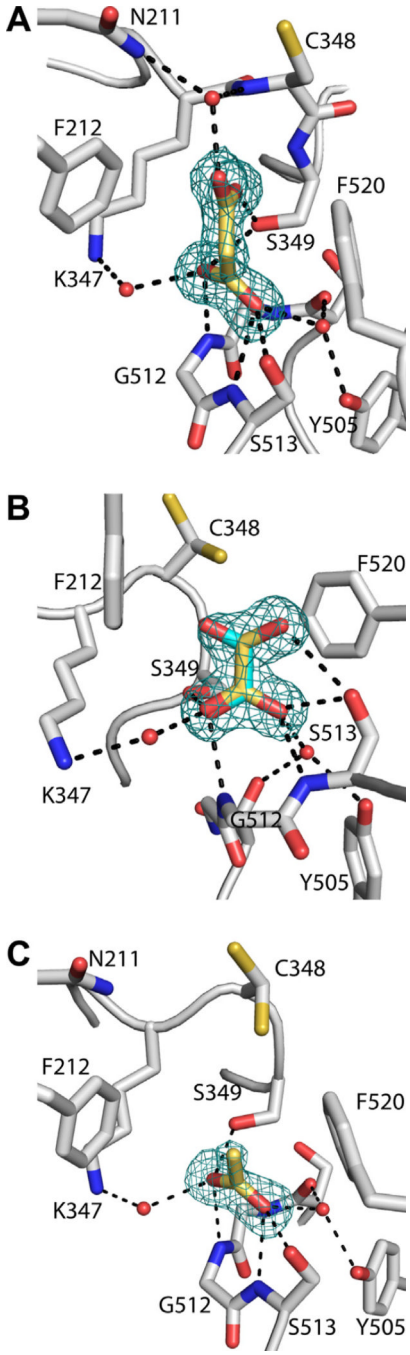

Fig. 5.

The binding of 3- and 2-carbon ligands to MmP5CDH. Electron density and interactions for (A) malonate, (B) glyoxylate, and (C) acetate. The cages represent simulated annealing σ_A-weighted F_o-F_c omit maps contoured at 3.0 σ. In panel B, conformations A and B of glyoxylate are colored gold and cyan, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

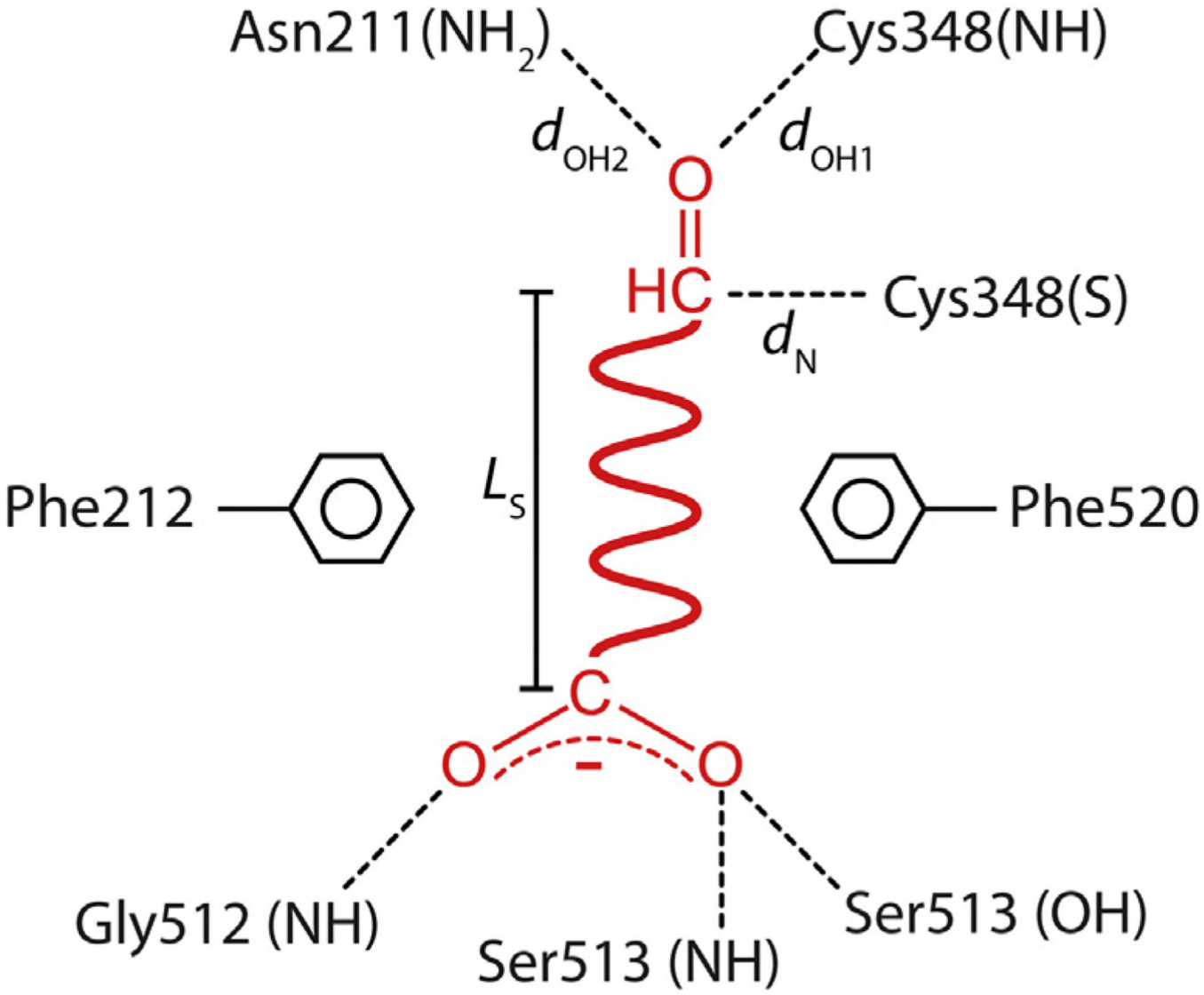

Fig. 6.
Cartoon representation of the essential elements of semialdehyde recognition by MmP5CDH.

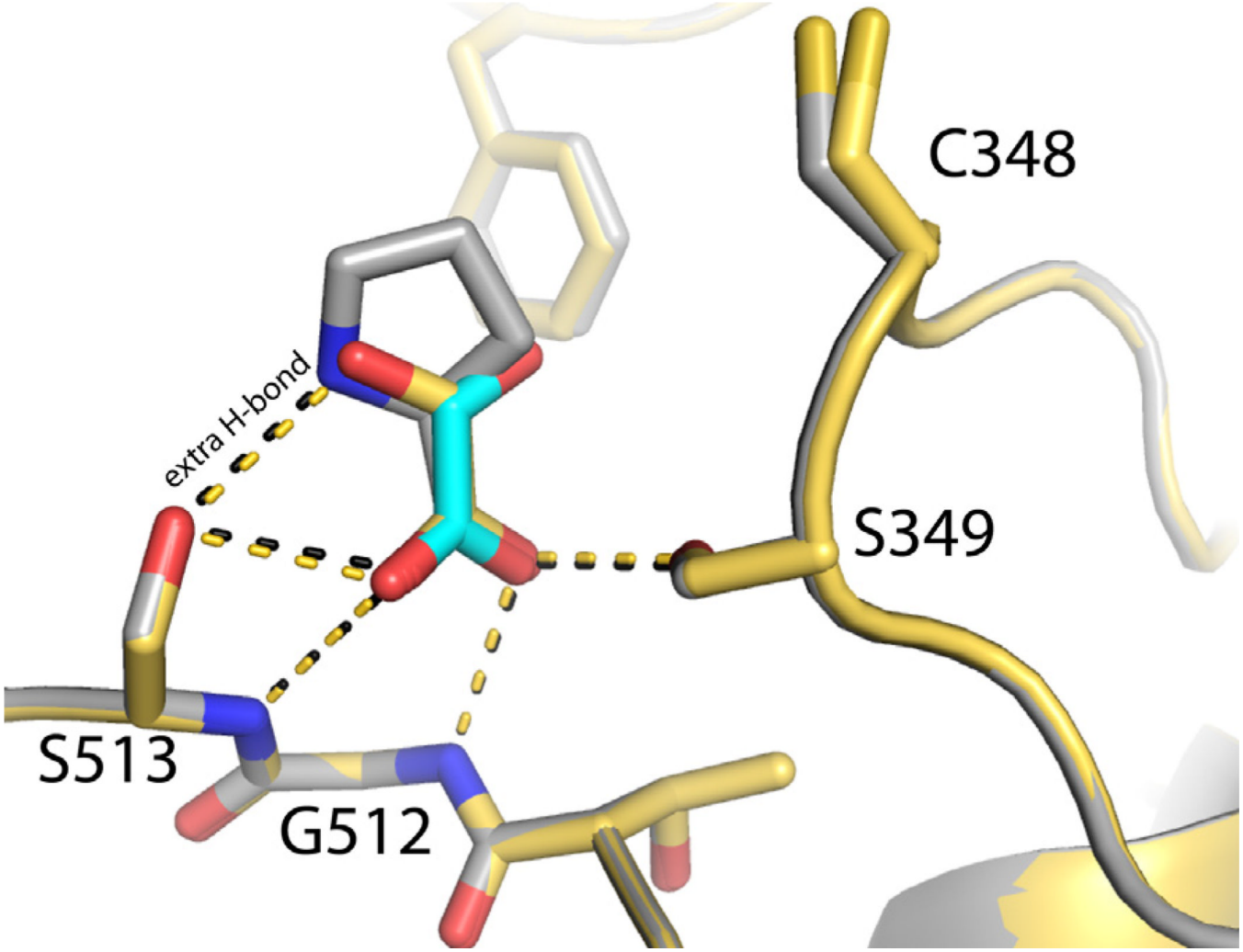

Fig. 7.
Comparison of the binding of L-proline and glyoxylate to MmP5CDH. The L-proline complex is colored gray with black hydrogen bonds. The glyoxylate complex is colored gold with gold hydrogen bonds. Conformations A and B of glyoxylate are colored gold and cyan, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1

Data collection and refinement statistics[a].

	Acetate	Glyoxylate	Malonate	Succinate	Glutarate
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 84.9	a = 84.9	a = 84.9	a = 84.9	a = 84.9
	b = 93.9	b = 94.2	b = 94.0	b = 94.1	b = 94.0
	c = 132.1	c = 132.1	c = 132.1	c = 132.3	c = 132.2
Wavelength	1.541	1.541	1.541	1.541	1.541
Resolution (Å)	19.64–1.67	19.67–1.67	19.65–1.67	19.68–1.67	19.66–1.81
	(1.76–1.67)	(1.76–1.67)	(1.76–1.67)	(1.76–1.67)	(1.91–1.81)
No. of observations	529,175	541,047	524,924	551,588	459,704
Unique reflections	119,816	120,763	120,594	118,798	94,657
R_{merge}(I)[b]	0.067 (0.297)	0.068 (0.259)	0.076 (0.436)	0.062 (0.363)	0.092 (0.676)
R_{meas}(I)[b]	0.076 (0.350)	0.077 (0.305)	0.086 (0.520)	0.070 (0.422)	0.116 (0.789)
R_{pim}(I)[b]	0.035 (0.181)	0.035 (0.157)	0.040 (0.278)	0.031 (0.209)	0.052 (0.397)
Mean $I/\sigma(I)$	14.1 (3.9)	14.8 (4.7)	13.3 (2.8)	16.0 (3.6)	11.3 (2.2)
Completeness (%)	98.2 (89.0)	98.7 (91.9)	98.5 (91.2)	97.0 (87.9)	98.3 (90.9)
Multiplicity	4.4 (3.3)	4.5 (3.5)	4.4 (3.2)	4.6 (3.5)	4.9 (3.7)
No. of protein atoms	8189	8185	8204	8208	8203
No. of waters	660	742	748	799	711
No. of PEG fragments	2	2	2	2	2
No. of ligands	4	3	2	2	2
R_{cryst}	0.168	0.159	0.170	0.164	0.166
R_{free}[c]	0.194	0.183	0.198	0.191	0.203
R.m.s. deviations†					
Bond lengths (Å)	0.006	0.006	0.006	0.006	0.006
Bond angles (°)	1.041	1.054	1.032	1.037	1.029
Ramachandran plot[d] (%)					
Favored	98.1	98.1	98.3	98.0	98.2
Allowed	1.9	1.9	1.7	2.0	1.8

	Acetate	Glyoxylate	Malonate	Succinate	Glutarate
Outliers	0	0	0	0	0
Average B (Å^2)					
Protein	14.4	12.5	14.6	15.2	16.7
Ligand	17.8	18.6	18.7	18.1	24.3
Water	20.6	19.5	21.2	22.8	22.3
PEG	24.1	21.5	24.2	27.5	29.0
PDB Code	4LGZ	4LH0	4LH1	4LH2	4LH3

[a] Values for the outer resolution shell of data are given in parenthesis.

[b] Definitions of R_{merge}, R_{meas}, and R_{pim} can be found in Weiss [37].

[c] Common 5% test set.

[d] The Ramachandran plot was generated with RAMPAGE [38].

Table 2

Kinetic parameters for the inhibition of MmP5CDH[a].

	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (s^{-1} M^{-1})	K_i (mM)
Glyoxylate	0.160 ± 0.002	26 ± 2	6200 ± 500	0.27 ± 0.02
Succinate	0.150 ± 0.002	31 ± 2	4800 ± 300	58 ± 4
Glutarate	0.150 ± 0.004	40 ± 4	3800 ± 400	30 ± 3
L-glutamate	0.150 ± 0.004	28 ± 3	5400 ± 600	12 ± 1

[a]Enzyme activity was measured at 20 °C with succinate semialdehyde as the variable substrate and the NAD^+ concentration fixed at 1 mM.

Table 3

Inferred geometrical parameters for active and inactive semialdehydes.

	L_S [a] (Å)	d_N [b] (Å)	d_{OH1} [c] (Å)	d_{OH2} [c] (Å)	Relative k_{cat}/K_m [d]
GSA	5.0	2.3	2.7	3.1	1.0
Glutarate semialdehyde	5.0	2.3	2.6	3.1	0.5
Succinate semialdehyde	3.3	2.6	2.8	3.3	0.2
Malonate semialdehyde	2.5	3.2	3.6	5.4	0
Glyoxylate	1.5	4.4	5.0	5.7	0

[a] Semialdehyde chain length, defined as the distance between the two terminal C atoms of the substrate in the E–S complex (Fig. 6).

[b] Nucleophilic attack distance, defined as the distance between the C atom of the substrate aldehyde group and the S atom of Cys348 (Fig. 6).

[c] Oxyanion hole hydrogen bond distances. d_{OH1} is the distance between the O atom of the substrate aldehyde group and the N atom of Cys348. d_{OH2} is distance between the O atom of the substrate aldehyde group and the side chain N atom of Asn211 (Fig. 6).

[d] Data from reference [21] expressed relative to P5C/GSA.