Published in final edited form as:
J Mol Biol. 2013 September 9; 425(17): 3106–3120. doi:10.1016/j.jmb.2013.05.027.

Structural Determinants of Oligomerization of !^{1}-Pyrroline-5-Carboxylate Dehydrogenase: Identification of a Hexamerization Hot Spot

Min Luo[1], **Ranjan K. Singh**[1], and **John J. Tanner**[1,2,*]

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

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

Abstract

The aldehyde dehydrogenase (ALDH) superfamily member !^{1}-pyrroline-5-carboxylate dehydrogenase (P5CDH) catalyzes the NAD^+-dependent oxidation of glutamate semialdehyde to glutamate, which is the final step of proline catabolism. Defects in P5CDH activity lead to the metabolic disorder type II hyperprolinemia, P5CDH is essential for virulence of the fungal pathogen *Cryptococcus neoformans*, and bacterial P5CDHs have been targeted for vaccine development. Although the enzyme oligomeric state is known to be important for ALDH function, the oligomerization of P5CDH has remained relatively unstudied. Here we determine the oligomeric states and quaternary structures of four bacterial P5CDHs using a combination of small-angle X-ray scattering, X-ray crystallography, and dynamic light scattering. The P5CDHs from *Thermus thermophilus* and *Deinococcus radiodurans* form trimer-of-dimers hexamers in solution, which is the first observation of a hexameric ALDH in solution. In contrast, two *Bacillus* P5CDHs form dimers in solution but do not assemble into a higher order oligomer. Site-directed mutagenesis was used to identify a hexamerization hot spot that is centered on an arginine residue in the NAD^+-binding domain. Mutation of this critical Arg residue to Ala in either of the hexameric enzymes prevents hexamer formation in solution. Paradoxically, the dimeric Arg-to-Ala *T. thermophilus* mutant enzyme packs as a hexamer in the crystal state, which illustrates the challenges associated with predicting the biological assembly in solution from crystal structures. The observation of different oligomeric states among P5CDHs suggests potential differences in cooperativity and protein-protein interactions.

Keywords

proline catabolism; aldehyde dehydrogenase; small-angle X-ray scattering; X-ray crystallography

*Corresponding author Department of Chemistry, University of Missouri-Columbia, Columbia, MO 65211, USA. Phone: (573) 884-1280. Fax: 573-882-2754. tannerjj@missouri.edu..

Accession numbers
Coordinates and structure factor amplitudes for TtP5CDHR100A have been deposited in the PDB under the accession number 4K57.

Introduction

The enzyme !^{1}-pyrroline-5-carboxylate (P5C) dehydrogenase (P5CDH, EC 1.5.1.12) catalyzes the NAD^{+}-dependent oxidation of glutamate semialdehyde to glutamate, which is the final step of proline catabolism (Fig. 1).[1] In humans, P5CDH also catalyzes the final step of hydroxyproline catabolism, which is the conversion of 4-hydroxyglutamate semialdehyde to 4-erythro-hydroxy-L-glutamate. The enzyme is widely distributed in eukaryotes and bacteria. In the former organisms, P5CDH is localized to the mitochondrial matrix. In some bacteria, mostly Gram-negative bacteria, P5CDH is combined with the first enzyme of proline catabolism, proline dehydrogenase (PRODH), into the bifunctional enzyme proline utilization A (PutA).[2,3] However, in Gram-positive bacteria, PRODH and P5CDH are separate monofunctional enzymes encoded by distinct genes. Monofunctional P5CDH is the subject of this research.

Defects in P5CDH underlie the autosomal recessive disorder type II hyperprolinemia.[4-8] Certain missense and frameshift mutations in the gene encoding P5CDH abrogate enzyme function, resulting in elevated levels of P5C and proline in plasma, urine, and cerebrospinal fluid.[9] Type II hyperprolinemia is causally linked to neurologic manifestations, such as increased incidence of seizures and intellectual and developmental disabilities,[10] although exactly how the enzyme deficiency contributes to these conditions is unclear. Possible mechanisms involve the role of proline as a neurotransmitter,[11-14] oxidative stress,[9] and mitochondrial dysfunction.[15]

P5CDHs from microorganisms have been gaining attention as well. A recent study of proline catabolism genetics in the fungal pathogen *Cryptococcus neoformans* showed that P5CDH is required for optimal production the major cryptococcal virulence factors.[16] This work also showed that a mutant strain of *C. neoformans* in which P5CDH was disabled is avirulent in a mouse model of infection. Recent immunological studies have shown that P5CDH is associated with the outer surface of two important bacterial pathogens, *Streptococcus pneumoniae* and *Staphylococcus aureus*.[17] Although the secretion mechanism and nature of the physical interaction that anchors the enzyme to the surface are unknown, P5CDH appears to be a new member of an emerging class of anchorless surface proteins found in Gram-positive bacteria that are thought to be potential virulence factors.[18] Furthermore, *S. pneumoniae* and *S. aureus* P5CDH are immunogenic and have been proposed as components of vaccines against *S. aureus*, a bacterial pathogen that causes significant morbidity, mortality, and healthcare costs worldwide.[17] Three-dimensional structural studies of P5CDHs should aid efforts to design inhibitors of fungal P5CDHs and new vaccines based on bacterial P5CDHs.

P5CDH is a member of the vast aldehyde dehydrogenase (ALDH) superfamily and is known as ALDH4A1. Like other ALDHs, P5CDH exhibits a 3-domain fold, as exemplified by *Thermus thermophilus* P5CDH, the first P5CDH to be structurally characterized (Fig. 2a).[19] The catalytic domain provides the essential Cys nucleophile that attacks the C atom of the substrate aldehyde group. The $NAD(P)^{+}$-binding domain exhibits the Rossmann fold. The oligomerization domain is a β substructure that protrudes from the NAD^{+}-binding domain and consists of a β-hairpin and the final β-strand of the polypeptide.

Oligomerization is an important aspect of the ALDH structure-function paradigm. All ALDHs form a domain-swapped dimer in which the oligomerization domain of one protomer engages the catalytic domain of the other protomer (Fig. 2b). In some ALDHs, such as ALDH1 and ALDH2, two of these dimers assemble into a dimer-of-dimer homotetramer.[20] Oligomerization appears to be important for ALDH function and stability.[20] For example, the inactive form of ALDH2 found in 40% of the East Asian

population has a glutamate to lysine substitution in an oligomerization interface.[21] Also, the tetrameric assembly of ALDH protomers appears to be essential for the half-of-the-sites reactivity exhibited by ALDH1A1 and ALDH2.[22]

In contrast to other ALDHs, relatively little is known about the oligomeric states and quaternary structures of P5CDHs. The oligomeric state in solution has been determined only for human P5CDH, which forms the classic ALDH homodimer, but does not assemble into a higher order oligomer.[23] To address this knowledge gap, we report an analysis of the oligomeric states and quaternary structures of the bacterial P5CDHs from *Thermus thermophilus* (TtP5CDH), *Deinococcus radiodurans* (DrP5CDH), *Bacillus halodurans* (BhP5CDH), and *Bacillus licheniformis* (BlP5CDH).

RESULTS

TtP5CDH and DrP5CDH form hexamers in solution

TtP5CDH and DrP5CDH were analyzed with small-angle X-ray scattering (SAXS), which can define solution conformation and assembly states in combination with crystallography.[24,25] The SAXS curve for TtP5CDH is shown in Fig. 3a. The Guinier plot exhibits good linearity and yields radius of gyration (R_g) of 43.4 ± 0.3 Å (Table 1). Calculations of the pair distribution function (P(r)) suggest R_g of 43.0 – 43.2 Å for maximum particle dimension (D_{max}) of 120 - 125 Å (Fig. S1). The P(r) function has one maximum centered at r = 60 Å (Fig. S1). Similar results were obtained for DrP5CDH, suggesting that the two enzymes adopt the same oligomeric state in solution (Fig. 3c). The Guinier R_g for DrP5CDH is 43.1 ± 0.1 Å. The R_g and D_{max} estimates from P(r) calculations are 43.6 Å and 120 - 125 Å, respectively. The P(r) function for DrP5CDH is nearly identical to that of TtP5CDH (Fig. S1).

The crystal structure of TtP5CDH was used to help determine the quaternary structures of TrP5CDH and DrP5CDH. The crystal structure of TtP5CDH was reported by Inagaki and coworkers in 2006.[19] The space group is *H*3 with one classic ALDH dimer (Fig. 2b) in the asymmetric unit. The R_g of the dimer is only 30 Å, which suggests that TtP5CDH assembles into a higher order oligomer in solution. Furthermore, the SAXS curve calculated from the dimer deviates substantially from the experimental curves (Figs. 3a, 3c). The *H*3 crystal lattice was inspected to identify a higher order assembly that is consistent with the SAXS data. As described previously,[19] application of the crystallographic 3-fold rotation to the asymmetric unit generates a trimer-of-dimers hexamer (Fig. 2c). The hexamer has R_g of 42.6 Å, which agrees with the experimental R_g of 43 Å. The theoretical SAXS curve calculated from the hexamer exhibits good agreement with the experimental ones for both TtP5CDH (Fig. 3a) and DrP5CDH (Fig. 3c). Ensembles containing both the hexamer and dimer models were also considered using the minimal ensemble search method.[26] Slightly better fits to the experimental profiles are obtained with ensembles consisting of 94 % hexamer and 6% dimer for TtP5CDH (Fig. 3a), and 87 % hexamer and 13 % dimer for DrP5CDH (Fig. 3C). These calculations suggest that the hexamer is the predominant oligomer in solution. Moreover, the *ab initio* SAXS envelope matches the size and shape of the hexamer (Fig. 3b). It is concluded that TtP5CDH and DrP5CDH exist primarily as a trimer-of-dimers hexamer in solution. This is the first observation of a hexameric ALDH in solution.

The hexamer oligomeric state was confirmed using dynamic light scattering (DLS) (Table 1 and Table S1) and the SAXS volume of correlation (V_c).[27] DLS data indicate a hydrodynamic radius (R_H) of 7.2 nm for TtP5CDH, which implies a molecular weight (M) of 351 kDa. The latter value is within 3 % of the predicted M of 342 kDa for a hexamer. Similarly, the estimated R_H and M of DrP5CDH are 7.3 nm and 361 kDa. The V_c values of

TtP5CDH and DrP5CDH are 1294 Å^2 and 1295 Å^2, respectively (Table 1). The corresponding M values are 314 kDa for TtP5CDH and 316 kDa for DrP5CDH (Table 1), which are within 8 % of the M of the hexamer. The DLS and V_c results confirm that TtP5CDH and DrP5CDH exist primarily as hexamers in solution.

BhP5CDH and BlP5CDH are dimeric in solution

The SAXS profile of BhP5CDH is profoundly different from those of TtP5CDH and DrP5CDH (Fig. 4a). In particular, the pronounced valley at q = 0.083 Å^{-1} and peak at q = 0.102 Å^{-1}, which are characteristic of hexameric P5CDHs, are absent in the SAXS profile for BhP5CDH. The R_g of BhP5CDH from Guinier analysis is 31.3 ± 0.1 Å, while the real space R_g from calculations of P(r) is 31.8 – 31. 9 Å for D_{max} = 95 – 105 Å. These values are much smaller than the R_g of 43 Å of hexameric P5CDHs, suggesting that BhP5CDH does not form a hexamer. Furthermore, the P(r) has a maximum at r = 35 Å, whereas the P(r) for hexameric P5CDH has a maximum at 60 Å (Fig. S1). The R_g calculated from a P5CDH dimer is 30 Å, implying that BhP5CDH forms the classic ALDH dimer but does not assemble into higher order oligomers. M estimated from SAXS V_c is 98 kDa (Table 1), which is suggestive of the dimer (114 kDa) and certainly eliminates tetramers (228 kDa) and hexamers (342 kDa) from consideration. Also, DLS data are consistent with BhP5CDH forming a dimer (R_H = 4.4 nm, M = 106 kDa, Table S1).

The hypothesis that BhP5CDH forms a dimer in solution was tested using crystallographic data. The structure of BhP5CDH was determined by the New York Structural Genomics Research Consortium. The enzyme crystallizes in space group $C2$ with three molecules in the asymmetric unit (PDB code 3QAN, unpublished). Two of the molecules form the classic ALDH dimer, while the crystallographic 2-fold rotation generates the other half of the dimer for the third molecule. Analysis of the crystal lattice with PDBePISA[28] indicates that the dimer is the most probable assembly in solution. Inspection of crystal packing with COOT confirmed that the hexamer is absent. The SAXS profile calculated from the BhP5CDH dimer agrees well with the experimental profile (Fig. 4a), and the SAXS reconstruction exhibits good agreement with the dimer (Fig. 4b). It is concluded that BhP5CDH is dimeric in solution.

The oligomeric state of BlP5CDH (75 % identical to BhP5CDH) was determined using DLS and analysis of crystal packing. The R_H is 4.6 nm, which corresponds to M = 119 kDa, consistent with a dimer (Table S1). The crystal structure of BlP5CDH was also determined by the New York Structural Genomics Research Consortium. The enzyme crystallizes in space group $P2_1$ with eight molecules in the asymmetric unit. As deposited, dimers are not evident in the asymmetric unit. However, application of crystallographic symmetry allows a different choice of asymmetric unit that contains four dimers. Analysis of the BlP5CDH crystal lattice with PDBePISA indicates that the dimer is the most probable assembly in solution. It is concluded that BlP5CDH is also dimeric in solution.

Site-directed mutagenesis rationale

Site-directed mutagenesis (to Ala) was used to identify residues important for hexamerization. TtP5CDH was chosen as the model for this study because high resolution crystal structures are available (e.g., PDB 2BHQ). The major interface between dimers in the hexamer is formed by a helix from the NAD$^+$-binding domain (α3, residues 96-111) and the oligomerization domain (red in Fig. 2a). These structural elements form a symmetric interface lining the inside surface of the tunnel that surrounds the 3-fold axis (red in Fig. 5a). This interface buries 1100 Å^2 of surface area. For comparison, the interfacial area of the domain-swapped dimer interface is 2900 Å^2. A smaller dimer-dimer contact surface (600

Å^2) is located on the outside of the hexamer (blue in Fig. 5b). The α14 helix of the catalytic domain (residues 454-463) and α3 are prominent in this interface (Fig. 2a).

The α3 helix was targeted for site-directed mutagenesis because it participates in both dimer-dimer interfaces (Fig. 5c). Specifically, three positively charged residues that form dimer-dimer electrostatic interactions were individually mutated to Ala: Arg100, Lys104, and Arg111. Arg100 is unique in that it is the only residue in the protein that interacts with two protomers outside of its own dimer (3.2 Å cutoff). These dimer-dimer interactions include ion pairs with Asp166, Glu168, and Glu458, and a hydrogen bond with Tyr154 (Fig. 5c). Arg100 also forms an intermolecular stacking interaction with Arg461. The other two residues targeted for mutagenesis, Lys104 and Arg111, form dimer-dimer ion pairs with the carboxyl-terminus of the polypeptide chain (Fig. 5c).

Arg153 of the oligomerization domain was also mutated to Ala. This residue is interesting because it is next to the two-fold axis of the major dimer-dimer interface, and thus its guanidinium group stacks in parallel with that of the symmetry-related related Arg153 (Fig. 5c).

Steady-state kinetic measurements

The kinetic constants for the native and mutant enzymes were estimated using P5C as the variable substrate and NAD^+ fixed at 1 mM in order to assess whether mutation of the hexamer interface causes any gross change in enzyme activity (Table S2, Fig. S2). The catalytic efficiencies (k_{cat}/K_m) of the TtP5CDH mutant enzymes are within 20 % of that of the native enzyme, indicating that these particular mutations do not have an obvious, substantial effect on activity.

Hexamerization hot spot

SAXS analysis of TtP5CDHR100A clearly shows that mutation of Arg100 to Ala is sufficient to disrupt the hexamer. The SAXS curve for TtP5CDHR100A lacks the valley and peak features that are diagnostic of the hexamer (Fig. 6a). The R_g from Guinier analysis is 32.1 ± 0.2 Å, which is close to the value of 30 Å calculated from the dimer and substantially smaller than that of the hexamer (43 Å). Furthermore, the SAXS curve calculated from the dimer exhibits excellent agreement with the experimental one (Fig. 6a), and the SAXS envelope matches the dimer (Fig. 6b). Also, the R_H estimated from DLS is 4.7 nm, which is substantially smaller than the R_H of 7.2 for the TtP5CDH hexamer (Table S1). The corresponding M from DLS is 123 kDa, which is similar to value of 114 kDa expected for the dimer. M estimated from SAXS V_c is 100 kDa (Table 1), which is also consistent with the dimer. It is concluded that TtP5CDHR100A exists in solution primarily as a dimer.

The analogous mutation was generated for DrP5CDH (DrP5CDHR102A). The SAXS profile of DrP5CDHR102A is likewise indicative of a dimer (Fig. 6a). The R_g is 32.6 ± 0.1 Å, and the SAXS curve calculated from the dimer agrees well with the experimental curve. The SAXS envelope exhibits good agreement with the dimer (Fig. 6b). Also, the M of DrP5CDHR102A estimated from DLS is 126 kDa, which is similar to the value of 114 kDa expected for a dimer (Table S1). M estimated from SAXS V_c is 111 kDa (Table 1), which is also consistent with the dimer. These data show that DrP5CDHR102A likewise exists primarily as a dimer in solution.

In contrast, mutation of Arg111, Lys104, or Arg153 individually to Ala does not disrupt the hexamer. The SAXS curve for TtP5CDHR111A clearly exhibits the valley and peak features that are diagnostic of the hexamer (Fig. 6a). The R_g from Guinier analysis is 46.1 ± 0.2 Å, which is consistent with a hexamer. Furthermore, the profile calculated from the hexamer matches the experimental profile (Fig. 6a), and the shape reconstruction resembles the

hexamer (Fig. 6b). Minimum ensemble search calculations also suggest that the hexamer is the predominant (88%) oligomer in solution (Fig. 6a). The *M* value estimated from SAXS V_c is 319 kDa (Table 1), which is just 7 % lower than *M* of the hexamer. TtP5CDHR111A, TtP5CDHK104A, and TtP5CDHR153A were analyzed with DLS. The hydrodynamic radii for these proteins span the range 6.8 – 7.2 nm, implying *M* = 304 - 337 kDa, which is consistent with the hexamer being the major species in solution (Table S1). These results suggest that TtP5CDHR111A, TtP5CDHK104A, and TtP5CDHR153A exist primarily as hexamers in solution.

Crystallization of dimeric TtP5CDHs

TtP5CDHR100A was targeted for crystallization to better understand why mutation of Arg100 to Ala disrupts hexamerization. Surprisingly, the mutant enzyme crystallized in the same *H*3 lattice as TtP5CDH (Table 2), and thus the trimer-of-dimers hexamer is evident in crystalline TtP5CDHR100A despite loss of the numerous interactions formed by Arg100.

Electron density maps indicated that the Arg100Ala mutation causes a cascade of conformational changes in the vicinity of residue 100 (Fig. 7a). Glu168 rotates into the void created by removal of the Arg100 side chain, while Tyr154 moves into the space vacated by Glu168 (Fig. 7b). Also, electron density for the side chains of Arg153 and Glu458 is weak and diffuse, implying disorder.

Formation of the *H*3 lattice by TtP5CDHR100A prompted the generation of double and triple mutants designed to induce a new crystal form that is devoid of the hexamer (Table S3).

This goal was achieved with the triple mutant R100A/K104A/R111A, which crystallizes in space group *P*1. A moderate resolution diffraction data set was obtained (Table 2), which was sufficient to determine the arrangement of dimers in the crystal lattice. The calculations show that the triclinic cell contains 4 dimers (Fig. 8). Analysis of protein-protein interfaces using manual inspection and PDBePISA shows that the hexamer is not present in the lattice.

Discussion

Global sequence identity apparently is not a good predictor of the oligomeric state of P5CDHs. We previously showed using analytical ultracentrifugation that human P5CDH is a dimer in solution. The human enzyme has 30% sequence identity to the bacterial enzymes studied here (Table S4). The *Bacillus* enzymes are also dimeric despite having 50% identity to the hexameric enzymes TtP5CDH and DrP5CDH. Thus, prediction of the oligomeric state of P5CDH from sequence requires a more careful examination of both sequence and structure. This conclusion is consistent with the hot spot theory of protein-protein interaction, which posits that a few critical residues in the interface account for most of the binding energy.[29,30]

We used structure-guided alanine scanning mutagenesis to identify a hexamerization hot spot for bacterial P5CDHs. The hot spot is centered on Arg100, which is consistent with early work showing that protein-protein interaction hot spots tend to be enriched in arginine, surrounded by less important residues.[29] Arg100 is unique in that it is the only residue that interacts with two other protomers of the hexamer (Fig. 5c). This key residue forms dimer-dimer interactions with five residues: Tyr154, Asp166, Glu168, Glu458, and Arg461. Arg100 is also present in DrP5CDH but is substituted by Ala/Asn in the *Bacillus* enzymes and Gln in human P5CDH (Fig. S3), consistent with it being essential for hexamerization. Arg461 is also found in DrP5CDH, while Asp166 is conservatively substituted with Glu. On the other hand, Arg461 and Asp166 are not conserved in the three dimeric P5CDHs. Also, a

clear trend is not evident for Tyr154 or Glu168. This analysis suggests that the triplet of Arg100, Asp/Glu166, and Arg461 is a hexamerization hot spot, which could be used to predict the oligomeric states of P5CDHs. For example, *S. aureus* P5CDH, which has been proposed as a vaccine reagent, is predicted to be dimeric since it has Glu in place of Arg100.

The prediction of stable oligomeric species in solution from crystal structures is an important area of research. PDBePISA is used widely for this purpose and has been adopted by the PDB for predicting the biological assembly of deposited structures (remark 350 of PDB entries). TtP5CDHR100A is interesting in this regard because it forms a hexamer in the crystal yet is predominately dimeric in solution under the conditions used for SAXS and DLS (1-5 mg/mL). Analysis of the TtP5CDHR100A *H*3 lattice with PDBePISA indicates that the hexamer is the most probable assembly in solution, which contradicts the SAXS data. However, the PDBePISA results do suggest that the R100A hexamer is less stable than the native one. For example, considering only the protein component and omitting solvent, the mutation causes 5 % and 7 % decreases in the surface area buried in the major and minor hexamer interfaces, respectively. Also, the complexation significance score of the major hexamer interface decreases from 1.0 for TtP5CDH (maximum possible value) to only 0.2 for TtP5CDHR100A. Thus, the program was able to discern a difference in the interfaces of the two crystal structures. Our results suggest that the prediction of the solution oligomeric state from crystal packing remains challenging in some cases.

Consideration of dynamic self-association equilibrium provides plausible explanations for the discrepancy in the oligomeric state of TtP5CDHR100A in the aqueous and solid states. The SAXS data show that the dimer-hexamer equilibrium of TtP5CDHR100A lies far on the dimer side under conditions of 1-10 mg/mL protein in Tris/NaCl buffers at pH 7.5. It is possible that the conditions used for crystallization shift the equilibrium to the hexamer, enabling crystal growth. However, DLS performed on TtP5CDHR100A under solution conditions similar to those of the initial crystallization drop yielded R_H of 4.3 nm and M of 101 kDa (Table S1). Thus, the combination of low pH and high concentration of 2-methyl-2,4-pentanediol does not dramatically shift the equilibrium toward the hexamer. Another explanation is that the *H*3 crystal form may have sufficiently favorable free energy to allow crystal growth even at very low hexamer concentration. As hexamers join the growing crystal, the dimer-hexamer equilibrium shifts toward the hexamer according to Le Chatelier's principle, sustaining crystal growth.

The role of oligomerization in P5CDH function is unknown, but possibilities include cooperativity and substrate channeling. Positive cooperatively in cofactor binding and half-of-the-sites reactivity are exhibited by some tetrameric ALDHs.[22,31] Analogous kinetic studies have not been performed for P5CDHs, but the observation of a higher order oligomer for TtP5CDH and DrP5CDH suggests that such experiments are worth pursuing. Oligomerization of P5CDH could also be connected with substrate channeling.[32,33] The potential for substrate channeling between monofunctional PRODH and P5CDH is suggested by the observation that the two enzymes are combined into a single polypeptide chain (PutA) in some bacteria. Eisenberg's group first proposed that such fused proteins could be used to predict protein-protein interactions (Rosetta Stone hypothesis of protein-protein interactions).[34] Kinetic studies have indeed shown that PutAs exhibit substrate channeling.[35,36] Thus, it is possible that monofunctional PRODH and P5CDH physically interact and engage in intermolecular substrate channeling. If so, it seems likely that the oligomeric states and quaternary structures of the interacting proteins will play an important role in dictating PRODH-P5CDH associations. This idea also remains to be tested.

Experimental Procedures

Expression and purification of TtP5CDH

The gene encoding P5CDH from *Thermus thermophilus* HB27 (NCBI Reference Sequence YP_005182.1, 516 residues) was cloned from genomic DNA into pKA8H using *NdeI* and *BamHI* sites by the University of Missouri DNA Core. The expressed protein includes an N-terminal His_8 tag and Tobacco Etch Virus Protease (TEVP) cleavage site. Treatment with TEVP results in Gly-His followed by the TtP5CDH polypeptide.

TtP5CDH was expressed in *Escherichia coli* BL21(DE3)pLysS. A 10 mL overnight culture was used to inoculate 1 L of LB media. The cells were grown at 37 °C at 250 rpm and induced with 0.5 mM IPTG at OD_{600} = 0.6 with an induction temperature of 22 °C at 200 rpm for 18 hours. Cells were harvested by centrifugation, resuspended in 50 mM HEPES, 300 mM NaCl, 10 mM imidazole, and 5% glycerol at pH 8.0, and flash-frozen in liquid nitrogen. The frozen cells were thawed with protease inhibitors (AEBSF, TPCK, E64, Pepstatin, Leupeptin) and broken by sonication. Cell debris and unbroken cells were separated by centrifugation at 16,500 rpm for one hour in a SS34 rotor. The supernatant was applied to a His-Trap HP column charged with Ni^{2+} (GE Healthcare) equilibrated with 20 mM HEPES, 300 mM NaCl, and 5% glycerol at pH 8.0. TtP5CDH was eluted with equilibration buffer containing 300 mM imidazole. Fractions were pooled, and TEVP, THP, and 20X TEV buffer (1 M Tris-HCl, 10 mM EDTA, pH 8.0) were added so that the final solution contained 5 mg of TEVP per 30 mg of protein in 50 mM Tris–HCl, 0.5 mM THP, and 0.5 mM EDTA, pH 8.0. The sample was incubated at 30 °C for 2 hours and dialyzed overnight at 4 °C before injection onto the His-Trap HP column. The tag-free protein was collected at 30 mM imidazole and dialyzed in 50 mM Tris-HCl, 0.5 mM THP, 5% glycerol, 0.5 mM EDTA, pH 8.0 in preparation for anion exchange chromatography (HiTrap Q). The protein was eluted with a linear 0 - 1 M NaCl gradient. Size exclusion chromatography (SEC) (Superdex 200, 25 mL) was used as the final step of purification.

Expression and purification of DrP5CDH

A gene encoding DrP5CDH (NCBI Reference Sequence NP_294537, 523 residues) with codons optimized for expression in *E. coli* was synthesized (Genscript) and subcloned into pKA8H using *NdeI* and *BamHI* sites. The expressed protein includes an N-terminal His_8 tag and TEVP site. Cleavage with TEVP produces Gly-His followed by the DrP5CDH polypeptide.

DrP5CDH was expressed in BL21(DE3)pLysS (induction at OD_{600} = 0.8 with 0.5 mM IPTG for 5 hours at 22 °C). The cells were collected by centrifugation, resuspended in 50 mM Tris, 100 mM NaCl, 10 mM imidazole, 5% glycerol, at pH 7.5, and frozen at −80 °C.

The frozen cells were thawed at 4 °C in the presence of protease inhibitors (0.1 mM TPCK, 0.05 mM AEBSF, 0.1 μM Pepstatin, 0.01 mM Leupeptin, 5μM E-64) and broken using sonication. The mixture was centrifuged at 16500 rpm in an SS34 rotor for 1 hour at 4°C, filtered through a 0.45 μm filter (Millipore) and loaded on a HisTrap HP column (5 mL) that had been charged with $NiCl_2$ and equilibrated in 50 mM Tris, 300 mM NaCl, 10 mM imidazole, and 5% glycerol at pH 7.5. Washing steps were performed using the loading buffer supplemented with 10 mM imidazole followed by 30 mM imidazole. The protein was eluted with 300 mM imidazole. The histidine tag was removed by incubating the protein with 0.2 mg/ml TEVP for 1 hour at 28°C followed by dialysis at 4 °C against 50 mM Tris, 50 mM NaCl, and 5% glycerol at pH 7.5. The mixture was applied to the HisTrap HP column to separate the cleaved protein, which appeared in the flow-through, from the tag and TEVP. The cleaved protein was dialyzed overnight at 4 °C into 50 mM Tris, 0.5 mM EDTA, 0.5 mM DTT, and 5% glycerol at pH 7.8 in preparation for anion exchange

chromatography (HiTrap Q). The sample was loaded onto the column using a buffer of 50 mM Tris and 5% glycerol at pH 7.8, and a linear NaCl gradient was applied. DrP5CDH eluted at 280 - 340 mM NaCl. The protein concentration was estimated using the bicinchoninic acid method (Pierce kit) with bovine serum albumin as the standard.

Expression and purification of BhP5CDH and BlP5CDH

Expression constructs for BhP5CDH (NCBI Reference Sequence NP_243603.1, 515 residues) and BlP5CDH (NCBI Reference Sequence YP_077616.1, 516 residues) were obtained from the New York Structural Genomics Research Consortium. The expressed enzymes have C-terminal His_8-tags and a TEVP cleavage site. TEVP cleavage produces the enzyme polypeptide followed by AENLYFQ. The enzymes were expressed and purified using the protocols described for TtP5CDH.

Site-directed mutagenesis

Site-directed mutants of TtP5CDH and DrP5CDH were generated using the QuikChange II site-directed mutagenesis kit (Agilent) using the primers listed in Table S3. The mutations were confirmed with sequencing performed by the University of Missouri DNA core. The mutant enzymes were purified as described above for the native enzymes.

SAXS

SAXS experiments were performed at beamline 12.3.1 of the Advanced Light Source via the mail-in program.[37,38] Prior to data collection, all protein samples were subjected to SEC using a Superdex 200 column. The column buffer was typically 50 mM Tris, 5% glycerol, 0.5 mM THP, and 50 mM NaCl at pH 7.5. In some cases (e.g., DrP5CDH), the SEC fractions were pooled, concentrated to ~12 mg/mL, and dialyzed at 4 °C for 24 hours against 50 mM Tris, 50 mM NaCl, 0.5 mM EDTA, 0.5 mM THP, and 5% glycerol at pH 7.8. For each protein, scattering intensities were measured at three nominal protein concentrations (1 - 10 mg/mL). For each protein concentration, exposure times of 0.5, 1.0, 3.0, and 6.0 sec were used. Scattering curves collected from the protein samples were corrected for background scattering using intensity data collected from the SEC effluent or dialysis buffer.

The SAXS data were analyzed as follows. A composite scattering curve for each sample was generated with PRIMUS[39] by scaling and merging the high q region from one of the longer time exposures with the low q region from a shorter time exposure. The scattering curves were multiplied with a concentration factor and overlaid on each other to check for concentration dependent variation of the profile. No substantial concentration effects were observed for any of the samples. PRIMUS was also used to perform Guinier analysis. FoXS was used to calculate theoretical scattering profiles from atomic models.[40] FoXS was also used to perform minimum ensemble calculations.[26] GNOM was used to calculate pair distribution functions.[41] MOLEMAN was used to calculate R_g from atomic coordinates.[42] The SASTBX server[43] was used for shape reconstruction calculations.

Estimation of molecular weight from SAXS volume of correlation

The molecular weight was estimated from the volume of correlation (V_c) as described recently by Rambo and Tainer.[27] Briefly, V_c is a new SAXS invariant defined as the ratio of the zero angle scattering intensity, $I(0)$, to the total scattered intensity. The latter quantity is equal to the integral $\int qI(q)\mathrm{d}q$ performed over the entire range of the scattering data. V_c thus has units of Å^2. Rambo and Tainer showed that V_c is independent of solute concentration and the aforementioned integral converges for both folded-compact and unfolded-flexible particles. For proteins, they also demonstrated that the molecular weight (M) in units of Da

can be estimated from a single SAXS curve by the relationship, $M = V_c^2 R_G^{-1}/0.1231$. An analogous relationship was provided for RNA.

V_c was calculated as follows. First, $I(0)$ and R_G were estimated from Guinier analysis using PRIMUS. These values were used to extrapolate the experimental SAXS curves to $q = 0$ using the equation, $I(q) = I(0) \exp(-q^2 R_G^2/3)$. Extrapolation to $q = 0$ is needed for accurate calculation of the total scattered intensity. The extrapolated region consisted of 18 – 19 points with q spacing of 0.000610 $Å^{-1}$, which matches the q spacing of the experimental data. The area under the curve of $qI(q)$ versus q was calculated using the Polygon Area utility of Origin 9 software, and V_c was calculated as the ratio of $I(0)$ to the area.

Crystallization of TtP5CDHR100A

Crystals of TtP5CDHR100A were grown in sitting drops at room temperature using the protocol for TtP5CDH.[44] Briefly, screening using commercially available kits (Hampton Research) yielded positive results in several conditions, which led to either the twinned form described earlier[44] or the more desired *H*3 form that was used for structure determination. Space group *H*3 crystals of TtP5CDHR100A were grown using sitting drops formed by mixing 1 μL of the protein stock solution (7 mg/mL protein in 50 mM Tris-HCl at pH 7.5, 100 mM NaCl, 5 % (v/v) glycerol, 0.5 mM THP, and 0.5 mM EDTA) and 1 μL of the reservoir containing 45 % 2-methyl-2,4-pentanediol and 0.05 M sodium citrate buffer at pH 5.2. The reservoir was used as the cryoprotectant. The space group is *H*3 with unit cell dimensions of $a = 173$ Å and $c = 279$ Å and two protein molecules in the asymmetric unit.

Crystallization of TtP5CDHR100A/K104A/R111A

Triclinic crystals of the triple mutant TtP5CDHR100A/K104A/R111A were grown at 295 K with the sitting-drop method of vapor diffusion. Initial conditions were identified using commercially available crystal screens (Hampton Research). TtP5CDHR100A/K104A/R111A was crystallized using a reservoir of 0.2 M $MgCl_2$, 0.1 M Tris-HCl pH 8.5, and 30% w/v PEG 4000. The protein stock solution contained 5 mg/mL TtP5CDH triple mutant in the buffer of 50 mM Tris, 100 mM NaCl, 0.5 mM THP, 0.5 mM EDTA, and 5% glycerol at pH 7.8. Crystals were cryoprotected with 0.2 M $MgCl_2$, 0.1 M Tris-HCl at pH 8.5, 30% w/v PEG 4000, and 28% PEG 200. The space group is *P*1 with unit cell dimensions a = 65.1 Å, b = 102.7 Å, c = 160.6 Å, α = 86.3°, β=87.5°, and γ = 79.4°. The asymmetric unit includes 8 protein molecules (4 dimers), which implies 48 % solvent and V_M of 2.35 $Å^3$/Da.

X-ray diffraction data collection, phasing, and refinement

X-ray diffraction data from crystals of TtP5CDHR100A in space group *H*3 were collected at beamline 4.2.2 of Advanced Light Source. The 1.54 Å resolution data set used for refinement consisted of 360 frames with an oscillation width of 0.5° per image, detector distance 110 mm, and exposure time 2 s/image. The data were processed with XDS[45] and SCALA[46] via CCP4i.[47] Refinement using PHENIX[48] commenced from the coordinates of TtP5CDH with Arg100 truncated to Ala. COOT[49] was used for model building. Data collection and refinement statistics are listed in Table 2.

X-ray diffraction data from crystals of the triple mutant TtP5CDHR100A/K104A/R111A were collected on an in-house Rigaku rotating anode generator coupled to an R-AXIS IV++ detector. Two data sets were recorded and merged. The first one consisted of 360 frames collected with an oscillation width of 0.5°, detector distance of 270 mm, and exposure time of 7 minutes per frame. The second set consisted of 360 frames collected with an oscillation width of 0.5°, detector distance of 200 mm, and exposure time of 5.2 minutes per frame. The data were integrated with MOSFLM[50] and merged to 2.42 Å resolution with SCALA. The arrangement of the dimers in the asymmetric unit was determined using molecular

replacement as implemented in PHASER[51] with a search model derived from a dimer of TtP5CDHR100A with Lys104 and Arg111 truncated to Ala. A clear solution with 4 dimers in the unit cell and log-likelihood gain of 31444 was obtained. Rigid body refinement yielded $R_{work} = R_{free} = 0.290$ for all data to 2.42 Å resolution.

DLS

These experiments were performed on a Protein Solutions DynaPro 99 Molecular Sizing Instrument (Wyatt Technology) at 20 °C with a wavelength of 836.3 nm and scattering angle of 90°. Protein concentrations were in the range 0.5 - 2 mg/ml in 50 mM Tris buffer at pH 7.5. Prior to DLS, each sample was centrifuged for 10 min at 13,000 × g at 4°C and passed through a 0.22 μm Millipore filter (Whatman). The data were collected with acquisition time of 10 s with at least 18 acquisitions. The DLS data were analyzed by the program DYNAMICS v.5.26.38 by performing regularization fit using the regularization algorithm on the measured autocorrelation functions.

Steady-state kinetics

The P5CDH activity of TtP5CDH and TtP5CDH mutant enzymes were measured at 20 °C by monitoring NADH production at 340 nm as described previously.[23] The final assay mixture (1 mL) contained 6 πg/ml P5CDH (0.1 πM), 1 mM NAD^+, and various concentrations of P5C in 0.1 M potassium phosphate buffer at pH 7.5. The pH of the P5C stock solution was adjusted to 7.5 before adding to the reaction mixture. Kinetic constants (Table S2) were estimated by fitting the initial rate data to the Michaelis-Menten equation using Origin 9.0 (Fig. S2).

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. Tommi White for purifying *T. thermophilus* genomic DNA and Dr. Mingyi Zhou for cloning the TtP5CDH gene. We thank Prof. Steven Almo and the New York Structural Genomics Research Consortium for providing the BhP5CDH and BlP5CDH clones, and Prof. Donald Becker for providing P5C. We thank Kevin Dyer of the SIBYLS Mail-In SAXS Program for collecting the SAXS data. X-ray scattering and diffraction technologies and their applications to the determination of macromolecular shapes and conformations at the SIBYLS beamline at the Advanced Light Source, Lawrence Berkeley National Laboratory, are supported in part by the DOE program Integrated Diffraction Analysis Technologies (IDAT) under Contract Number DE-AC02-05CH11231 with the U.S. Department of Energy. We thank Drs. Jay Nix for help with diffraction data collection and processing. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Abbreviations used

PRODH	proline dehydrogenase
P5C	!^{1}-pyrroline-5-carboxylate
P5CDH	!^{1}-pyrroline-5-carboxylate dehydrogenase
PutA	proline utilization A
ALDH	aldehyde dehydrogenase
TtP5CDH	!^{1}-pyrroline-5-carboxylate dehydrogenase from *Thermus thermophilus*

DrP5CDH	!^{1}-pyrroline-5-carboxylate dehydrogenase from *Deinococcus radiodurans*
BhP5CDH	!^{1}-pyrroline-5-carboxylate dehydrogenase from *Bacillus halodurans*
BlP5CDH	!^{1}-pyrroline-5-carboxylate dehydrogenase from *Bacillus licheniformis*
SAXS	small-angle X-ray scattering
DLS	dynamic light scattering
TEVP	Tobacco Etch Virus protease
SEC	size exclusion chromatography

References

1. Phang JM. The regulatory functions of proline and pyrroline-5-carboxylic acid. Curr. Top. Cell. Reg. 1985; 25:92–132.
2. Tanner JJ. Structural biology of proline catabolism. Amino Acids. 2008; 35:719–30. [PubMed: 18369526]
3. Singh RK, Tanner JJ. Unique Structural Features and Sequence Motifs of Proline Utilization A (PutA). Front. Biosci. 2012; 17:556–568.
4. Efron ML. Familial Hyperprolinemia. Report Of A Second Case, Associated With Congenital Renal Malformations, Hereditary Hematuria And Mild Mental Retardation, With Demonstration Of An Enzyme Defect. N. Engl. J. Med. 1965; 272:1243–54. [PubMed: 14290545]
5. Baumgartner MR, Rabier D, Nassogne MC, Dufier JL, Padovani JP, Kamoun P, Valle D, Saudubray JM. Delta1-pyrroline-5-carboxylate synthase deficiency: neurodegeneration, cataracts and connective tissue manifestations combined with hyperammonaemia and reduced ornithine, citrulline, arginine and proline. Eur. J. Pediatr. 2005; 164:31–6. [PubMed: 15517380]
6. Geraghty MT, Vaughn D, Nicholson AJ, Lin WW, Jimenez-Sanchez G, Obie C, Flynn MP, Valle D, Hu CA. Mutations in the Delta1-pyrroline 5-carboxylate dehydrogenase gene cause type II hyperprolinemia. Hum. Mol. Genet. 1998; 7:1411–5. [PubMed: 9700195]
7. Scriver, CR.; Sly, WS.; Childs, B.; Beaudet, AL.; Valle, D.; Kinzler, KW.; Vogelstein, B., editors. The Metabolic and Molecular Bases of Inherited Disease. 8th edit.. McGraw-Hill; New York: 2001.
8. Valle D, Goodman SI, Applegarth DA, Shih VE, Phang JM. Type II hyperprolinemia. Delta1-pyrroline-5-carboxylic acid dehydrogenase deficiency in cultured skin fibroblasts and circulating lymphocytes. J. Clin. Invest. 1976; 58:598–603. [PubMed: 956388]
9. Wyse AT, Netto CA. Behavioral and neurochemical effects of proline. Metab. Brain Dis. 2011; 26:159–72. [PubMed: 21643764]
10. Phang, JM.; Hu, CA.; Valle, D. Disorders of proline and hydroxyproline metabolism.. In: Scriver, CR.; Beaudet, AL.; Sly, WS.; Valle, D., editors. Metabolic and molecular basis of inherited disease. McGraw Hill; New York: 2001. p. 1821-1838.
11. Gogos JA, Santha M, Takacs Z, Beck KD, Luine V, Lucas LR, Nadler JV, Karayiorgou M. The gene encoding proline dehydrogenase modulates sensorimotor gating in mice. Nat. Genet. 1999; 21:434–9. [PubMed: 10192398]
12. Felix D, Kunzle H. The role of proline in nervous transmission. Adv. Biochem. Psychopharmacol. 1976; 15:165–73. [PubMed: 15411]
13. Takemoto Y, Semba R. Immunohistochemical evidence for the localization of neurons containing the putative transmitter L-proline in rat brain. Brain Res. 2006; 1073-1074:311–5. [PubMed: 16458270]
14. Fremeau RT Jr. Caron MG, Blakely RD. Molecular cloning and expression of a high affinity L-proline transporter expressed in putative glutamatergic pathways of rat brain. Neuron. 1992; 8:915–26. [PubMed: 1350201]
15. He F, DiMario PJ. Drosophila delta-1-pyrroline-5-carboxylate dehydrogenase (P5CDh) is required for proline breakdown and mitochondrial integrity-Establishing a fly model for human type II hyperprolinemia. Mitochondrion. 2011; 11:397–404. [PubMed: 21168532]

16. Lee IR, Lui EY, Chow EW, Arras SD, Morrow CA, Fraser JA. Reactive Oxygen Species Homeostasis and Virulence of the Fungal Pathogen Cryptococcus neoformans Requires an Intact Proline Catabolism Pathway. Genetics. 2013 doi: 10.1534/genetics.113.150326.
17. Lijek RS, Luque SL, Liu Q, Parker D, Bae T, Weiser JN. Protection from the acquisition of Staphylococcus aureus nasal carriage by cross-reactive antibody to a pneumococcal dehydrogenase. Proc. Natl. Acad. Sci. USA. 2012; 109:13823–8. [PubMed: 22869727]
18. Chhatwal GS. Anchorless adhesins and invasins of Gram-positive bacteria: a new class of virulence factors. Trends Microbiol. 2002; 10:205–8. [PubMed: 11973142]
19. Inagaki E, Ohshima N, Takahashi H, Kuroishi C, Yokoyama S, Tahirov TH. Crystal structure of Thermus thermophilus Delta1-pyrroline-5-carboxylate dehydrogenase. J. Mol. Biol. 2006; 362:490–501. [PubMed: 16934832]
20. Rodriguez-Zavala JS, Weiner H. Structural aspects of aldehyde dehydrogenase that influence dimer-tetramer formation. Biochemistry. 2002; 41:8229–37. [PubMed: 12081471]
21. Larson HN, Weiner H, Hurley TD. Disruption of the coenzyme binding site and dimer interface revealed in the crystal structure of mitochondrial aldehyde dehydrogenase “Asian” variant. J. Biol. Chem. 2005; 280:30550–6. [PubMed: 15983043]
22. Belem YS, Pablo PJ, Jose RZ. New insights into the half-of-the-sites reactivity of human aldehyde dehydrogenase 1A1. Proteins. 2013
23. Srivastava D, Singh RK, Moxley MA, Henzl MT, Becker DF, Tanner JJ. The Three-Dimensional Structural Basis of Type II Hyperprolinemia. J. Mol. Biol. 2012; 420:176–189. [PubMed: 22516612]
24. Putnam CD, Hammel M, Hura GL, Tainer JA. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q. Rev. Biophys. 2007; 40:191–285. [PubMed: 18078545]
25. Rambo RP, Tainer JA. Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. Curr. Opin. Struct. Biol. 2010; 20:128–37. [PubMed: 20097063]
26. Pelikan M, Hura GL, Hammel M. Structure and flexibility within proteins as identified through small angle X-ray scattering. Gen. Physiol. Biophys. 2009; 28:174–89. [PubMed: 19592714]
27. Rambo RP, Tainer JA. Accurate assessment of mass, models and resolution by small-angle scattering. Nature. 2013; 496:477–81. [PubMed: 23619693]
28. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 2007; 372:774–97. [PubMed: 17681537]
29. Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. J. Mol. Biol. 1998; 280:1–9. [PubMed: 9653027]
30. Moreira IS, Fernandes PA, Ramos MJ. Hot spots--a review of the protein-protein interface determinant amino-acid residues. Proteins. 2007; 68:803–12. [PubMed: 17546660]
31. Wei B, Weiner H. Making an Oriental equivalent of the yeast cytosolic aldehyde dehydrogenase as well as making one with positive cooperativity in coenzyme binding by mutations of glutamate 492 and arginine 480. Chemico-biological interactions. 2001; 130-132:173–9. [PubMed: 11306041]
32. Arentson BW, Sanyal N, Becker DF. Substrate channeling in proline metabolism. Front. Biosci. 2012; 17:375–88.
33. Anderson KS. Fundamental mechanisms of substrate channeling. Methods Enzymol. 1999; 308:111–45. [PubMed: 10507003]
34. Marcotte EM, Pellegrini M, Ng HL, Rice DW, Yeates TO, Eisenberg D. Detecting protein function and protein-protein interactions from genome sequences. Science. 1999; 285:751–3. [PubMed: 10427000]
35. Srivastava D, Schuermann JP, White TA, Krishnan N, Sanyal N, Hura GL, Tan A, Henzl MT, Becker DF, Tanner JJ. Crystal structure of the bifunctional proline utilization A flavoenzyme from Bradyrhizobium japonicum. Proc. Natl. Acad. Sci. USA. 2010; 107:2878–83. [PubMed: 20133651]

36. Surber MW, Maloy S. The PutA protein of Salmonella typhimurium catalyzes the two steps of proline degradation via a leaky channel. Arch. Biochem. Biophys. 1998; 354:281–287. [PubMed: 9637737]
37. Hura GL, Menon AL, Hammel M, Rambo RP, Poole FL 2nd, Tsutakawa SE, Jenney FE Jr. Classen S, Frankel KA, Hopkins RC, Yang SJ, Scott JW, Dillard BD, Adams MW, Tainer JA. Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). Nat. Methods. 2009; 6:606–612. [PubMed: 19620974]
38. Classen S, Hura GL, Holton JM, Rambo RP, Rodic I, McGuire PJ, Dyer K, Hammel M, Meigs G, Frankel KA, Tainer JA. Implementation and performance of SIBYLS: a dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the Advanced Light Source. J. Appl. Crystallogr. 2013; 46:1–13. [PubMed: 23396808]
39. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J. Appl. Crystallogr. 2003; 36:1277–1282.
40. Schneidman-Duhovny D, Hammel M, Sali A. FoXS: a web server for rapid computation and fitting of SAXS profiles. Nucleic Acids Res. 2010; 38:W540–4. [PubMed: 20507903]
41. Svergun D. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Crystallogr. 1992; 25:495–503.
42. Kleywegt GJ. Validation of protein models from Calpha coordinates alone. J. Mol. Biol. 1997; 273:371–6. [PubMed: 9344745]
43. Liu H, Hexemer A, Zwart PH. The Small Angle Scattering ToolBox (SASTBX): an open-source software for biomolecular small-angle scattering. J. Appl. Crystallogr. 2012; 45:587–593.
44. Inagaki E, Takahashi H, Kuroishi C, Tahirov TH. Crystallization and avoiding the problem of hemihedral twinning in crystals of Delta1-pyrroline-5-carboxylate dehydrogenase from Thermus thermophilus. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 2005; 61:609–11.
45. Kabsch W. XDS. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:125–32. [PubMed: 20124692]
46. Evans P. Scaling and assessment of data quality. Acta Cryst. 2006; D62:72–82.
47. Potterton E, Briggs P, Turkenburg M, Dodson E. A graphical user interface to the CCP4 program suite. Acta Cryst. 2003; D59:1131–7.
48. Adams PD, Gopal K, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Pai RK, Read RJ, Romo TD, Sacchettini JC, Sauter NK, Storoni LC, Terwilliger TC. Recent developments in the PHENIX software for automated crystallographic structure determination. J. Synchrotron Rad. 2004; 11:53–5.
49. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Cryst. 2004; D60:2126–32.
50. Leslie AG. The integration of macromolecular diffraction data. Acta Cryst. 2006; D62:48–57.
51. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J. Appl. Crystallogr. 2007; 40:658–674. [PubMed: 19461840]
52. Weiss M. Global indicators of X-ray data quality. J. Appl. Cryst. 2001; 34:130–135.
53. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 2010; D66:12–21.

- We determined the oligomeric states and quaternary structures of four P5CDHs.
- We report the first observation of a hexameric aldehyde dehydrogenase in solution.
- Global sequence identity is not predictive of the oligomeric state of P5CDH.
- Ala scanning reveals a hexamerization hot spot centered on an Arg residue.
- Predicting the solution oligomeric state from the crystal can be challenging.

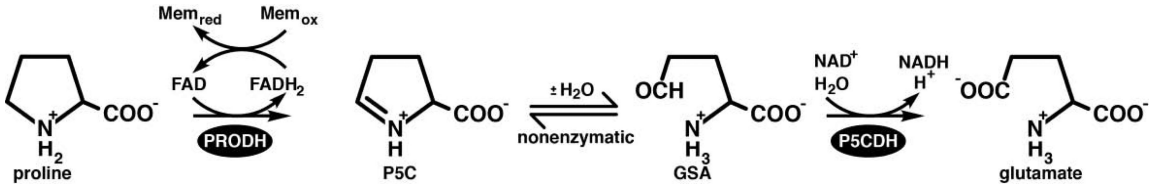

Fig. 1.
The reactions of proline catabolism.

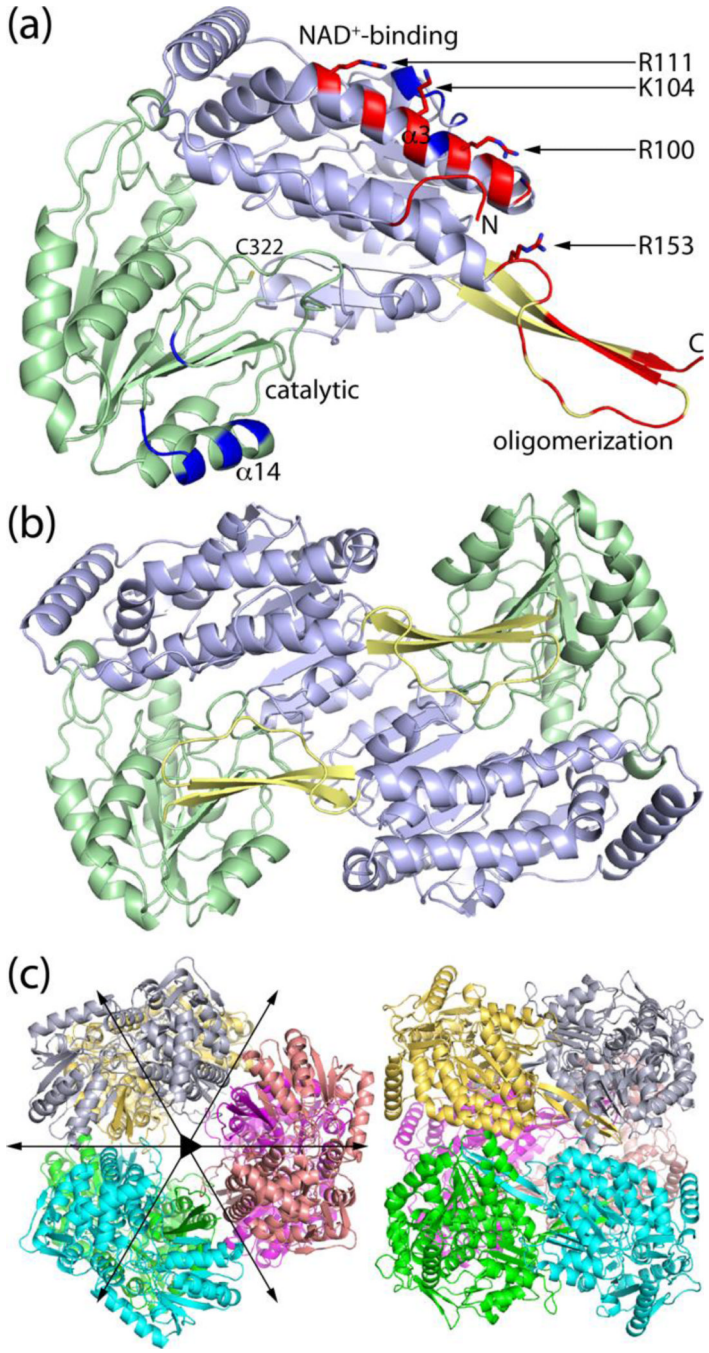

Fig. 2.

Structure of TtP5CDH (PDB code 2BHQ). (A) The protomer is shown with the catalytic domain in green, NAD^+-binding domain in light blue, and oligomerization domain in yellow. Red patches indicate residues that form the major dimer-dimer interface of hexametric P5CDHs. Dark blue denotes residues involved in the minor dimer-dimer interface of hexametric P5CDHs. (B) Structure of the classic ALDH dimer. (C) Two views of the TtP5CDH hexamer deduced from SAXS and X-ray crystallography.

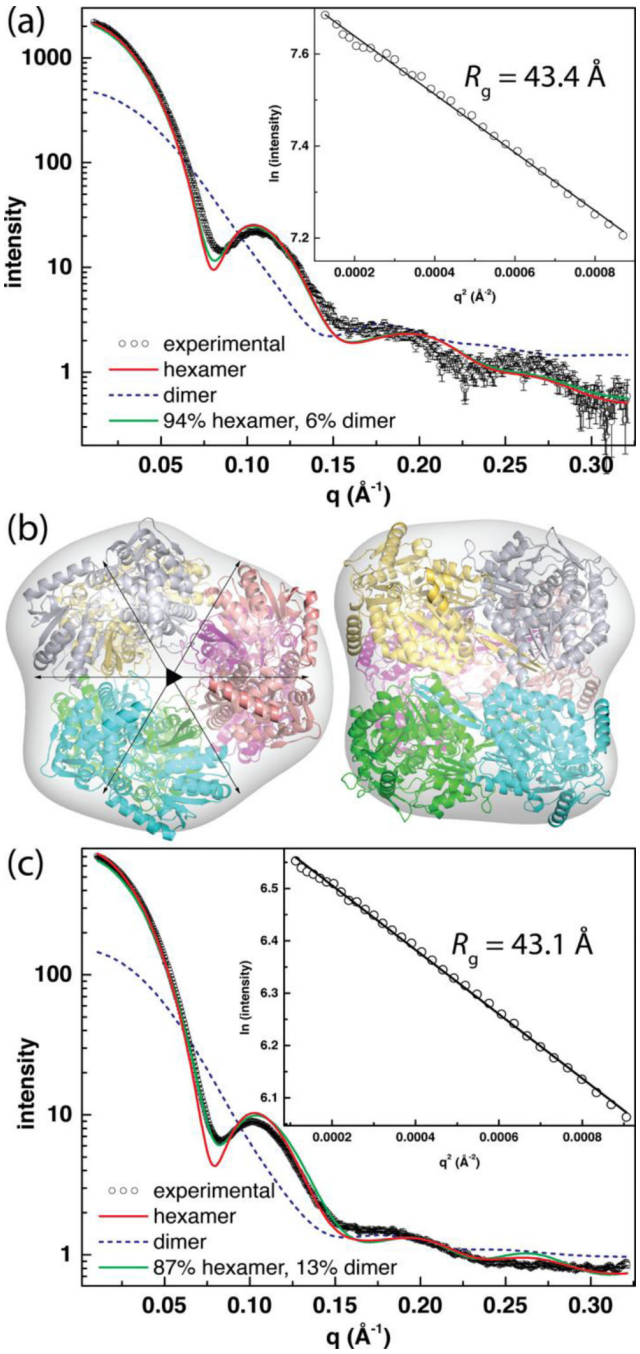

Fig. 3.

SAXS analyses of TtP5CDH and DrP5CDH. (A) Experimental and calculated SAXS curves for TtP5CDH. The inset shows a Guinier plot spanning the range of qR_g from 0.489 to 1.28. The linear fit of the Guinier plot has R^2 of 0.996. (B) Superposition of the TtP5CDH SAXS shape reconstruction and the hexamer generated from crystallographic symmetry. The SAXS envelope was calculated using the SASTBX server. Two orthogonal views are shown. (C) Experimental and calculated SAXS curves for DrP5CDH. The inset shows a Guinier plot spanning the range of qR_g from 0.458 to 1.30 (R^2 = 0.999).

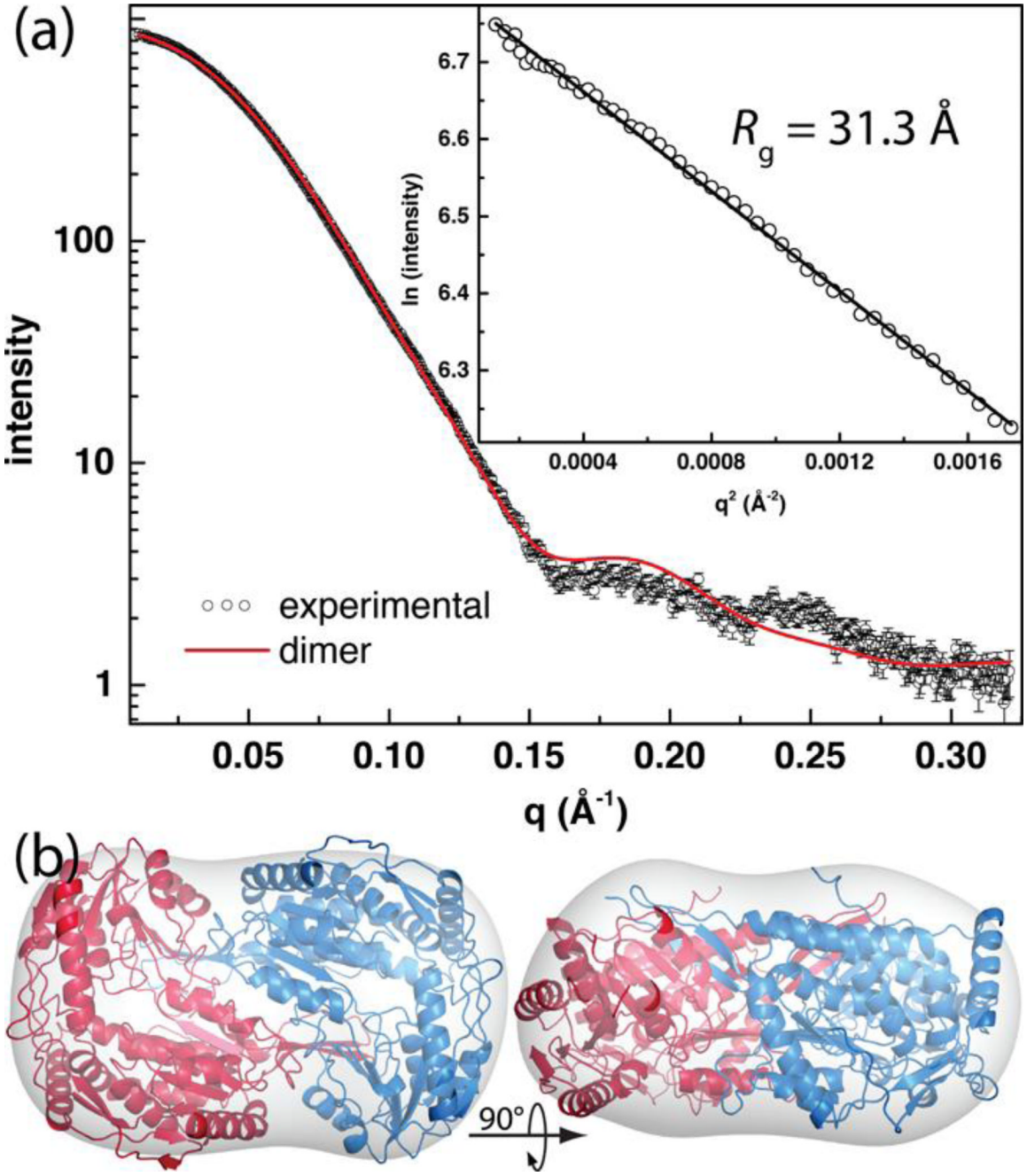

Fig. 4.

SAXS analysis of BhP5CDH. (A) Experimental and calculated SAXS curves. The inset shows a Guinier plot spanning the qR_g range of 0.352 – 1.30 (R^2 = 0.998). (B) Superposition of the SAXS shape reconstruction and the dimer.

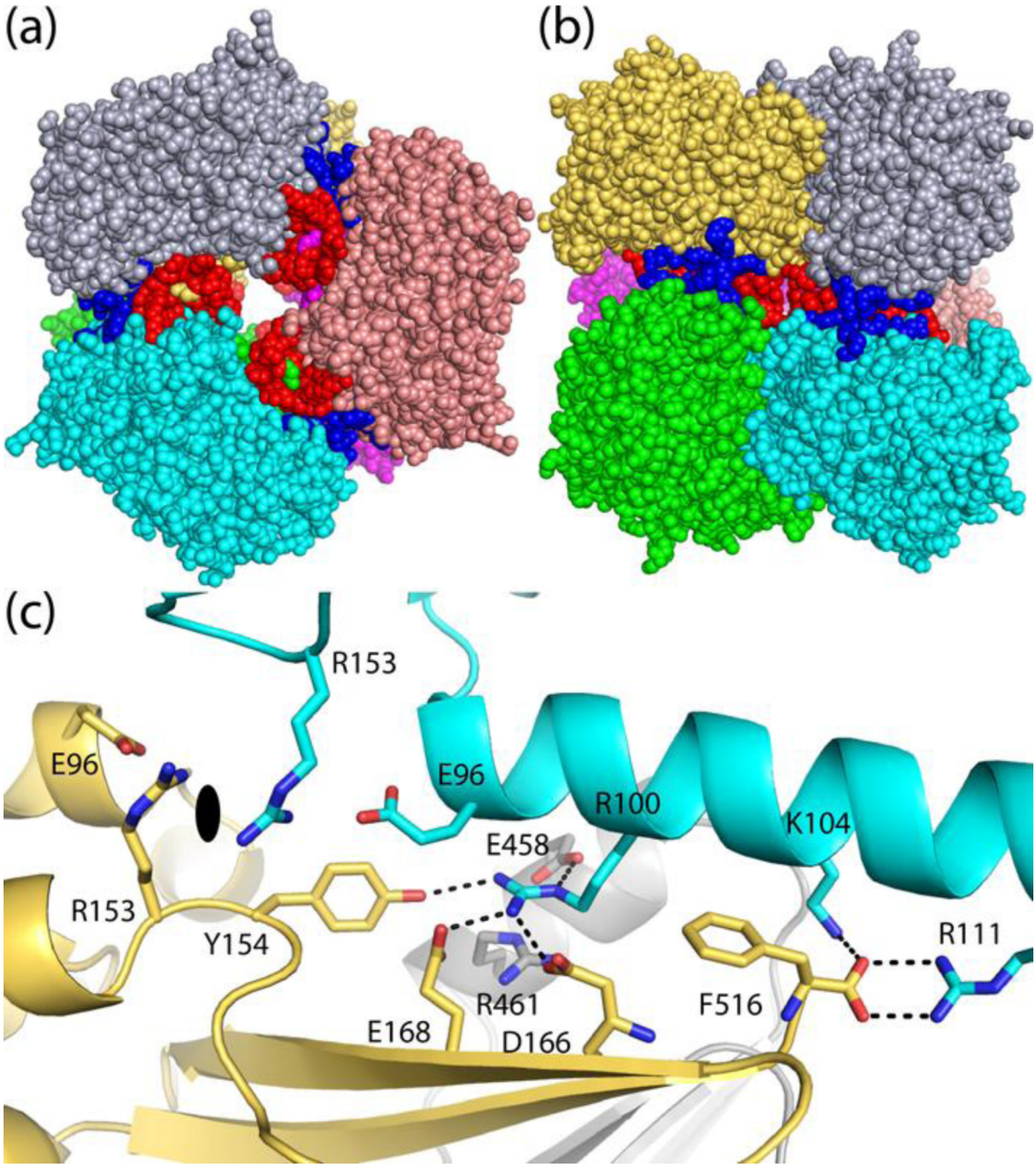

Fig. 5.

Dimer-dimer interfaces within the P5CDH hexamer. (A and B) Orthogonal views of the hexamer with the A-B dimer colored cyan-green, C-D dimer in gray-yellow, and E-F dimer in brown-magenta. Red and blue denote residues in the major and minor dimer-dimer interfaces, respectively. (C) Dimer-dimer interactions formed by the residues targeted for mutagenesis (Arg100, Lys104, Arg111, and Arg153). As in panels A and B, protomer A is colored cyan, and dimer C-D is colored gray-yellow. The oval denotes the approximate location of one of the 2-fold axes of the hexamer.

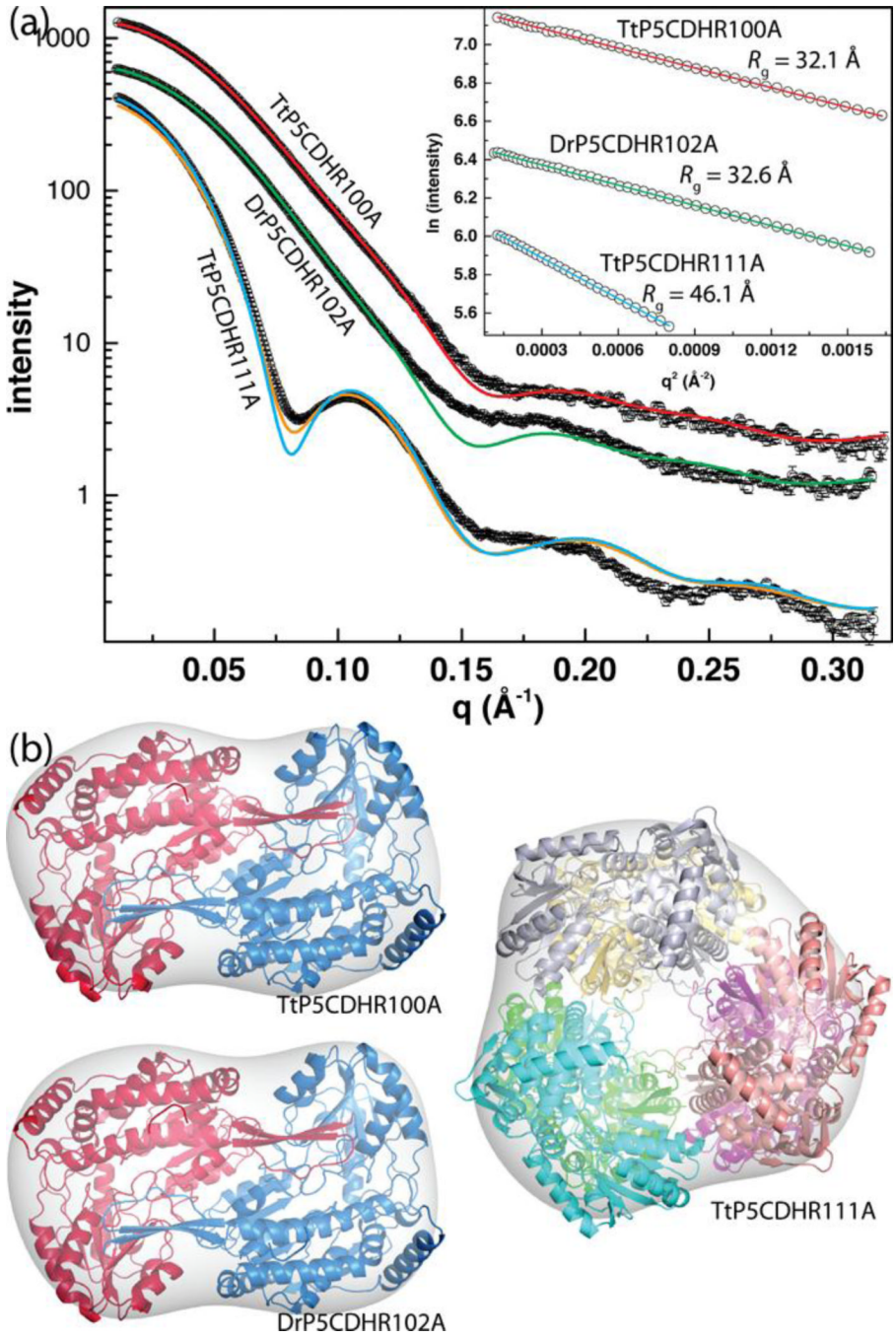

Fig 6.

SAXS analysis of P5CDH mutant enzymes. (A) Experimental and theoretical SAXS curves. Arbitrary offsets were applied to the curves for clarity. The theoretical SAXS curves for TtP5CDHR100A (red) and DrP5CDHR102A (green) were calculated from a TtP5CDH dimer. The theoretical curve in cyan for TtP5CDHR111A was calculated from a TtP5CDH hexamer, while the curve in orange was calculated from an ensemble consisting of 88% hexamer and 12 % dimer. The inset shows Guinier plots spanning the qR_g ranges of 0.361-1.30 for TtP5CDHR100A (R^2 = 0.9992), 0.347-1.30 for DrP5CDHR102A (R^2 =0.9996), and 0.519-1.31 for TtP5CDHR111A (R^2 = 0.9997). (B) Superpositions of the

SAXS shape reconstructions and oligomer models for TtP5CDHR100A, DrP5CDHR102A, and TtP5CDHR111A.

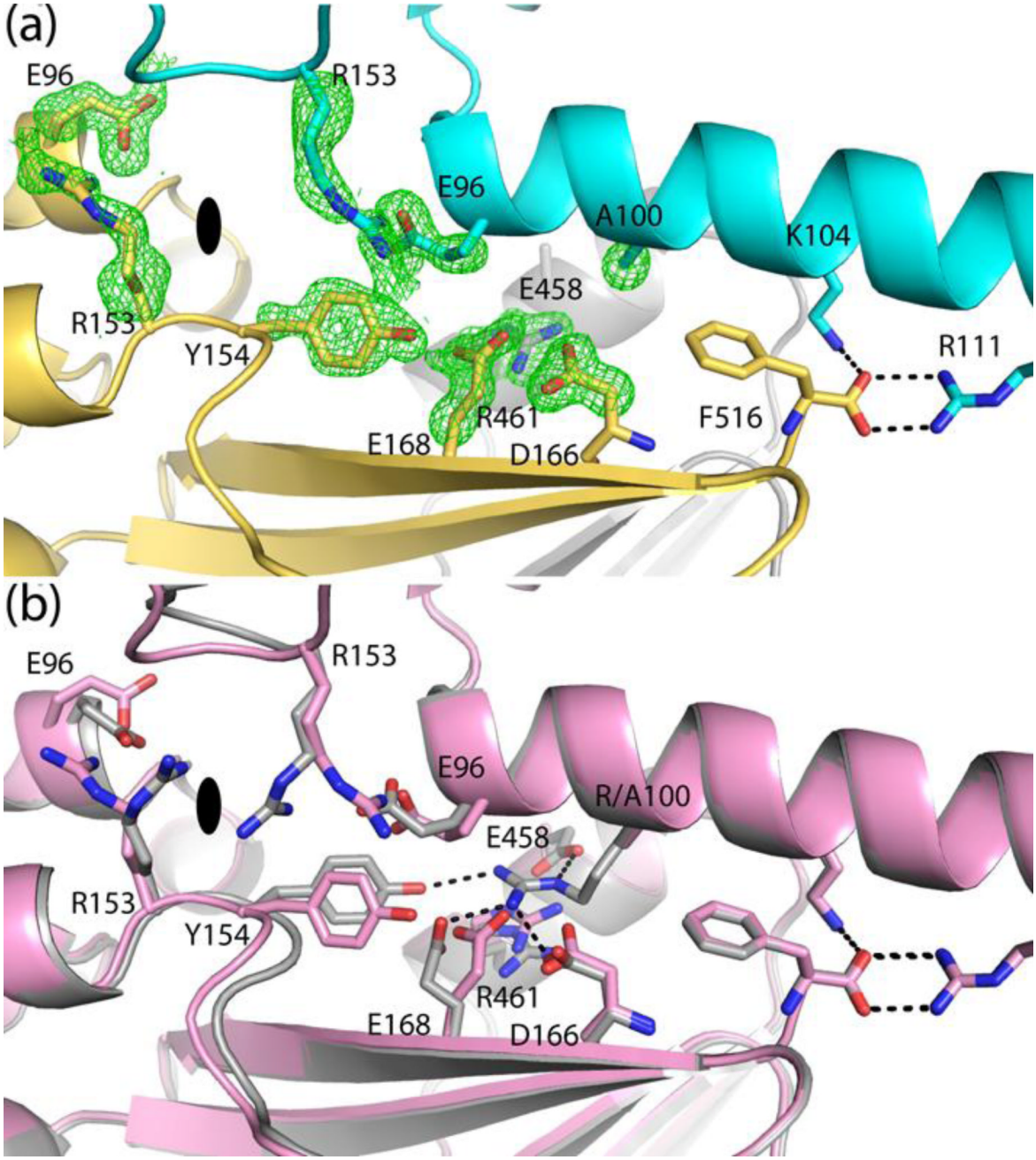

Fig. 7.

Hexamer interface of TtP5CDHR100A. (A) Electron density for residues of TtP5CDHR100A that change conformation. The cage represents a simulated annealing σ_A-weighted F_o - F_c omit map (3σ). The orientation and coloring scheme are identical to those of Fig. 5c. (B) Comparison of the hexamer interfaces of TtP5CDH (gray) and TtP5CDHR100A (pink).

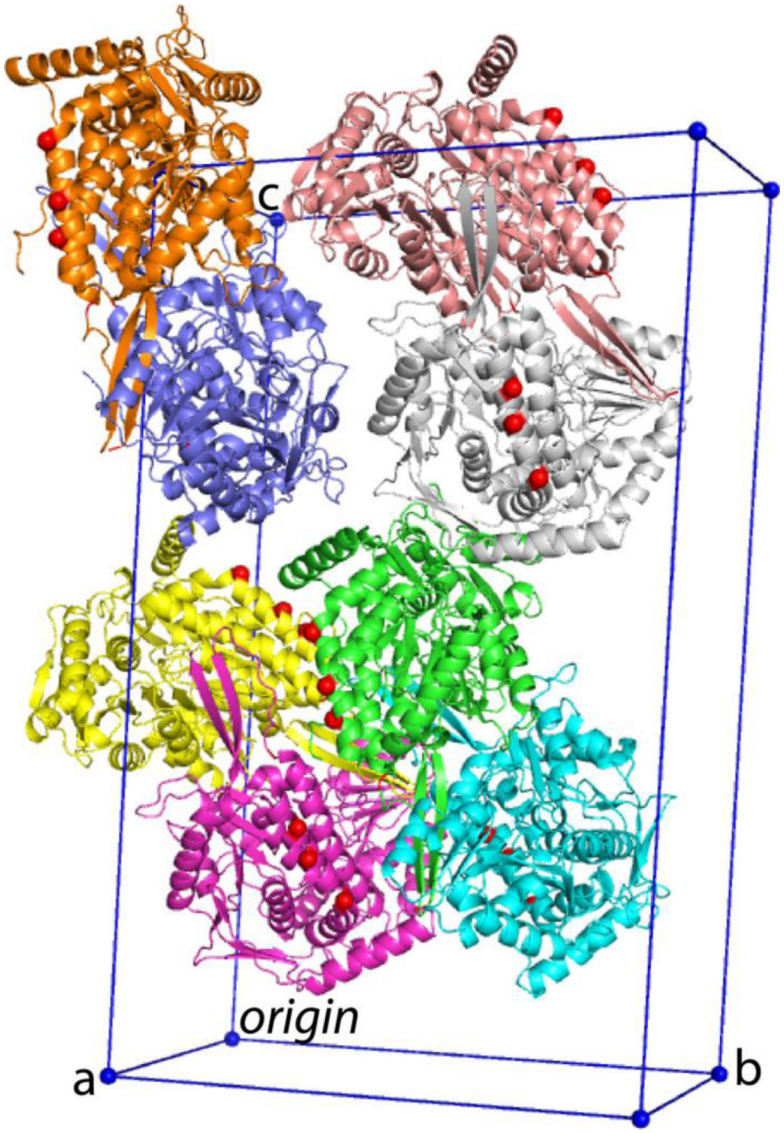

Fig. 8.
Triclinic unit cell of TtP5CDHR100A/K104A/R111A. The red spheres denote the locations of the mutated residues, Ala100, Ala104, and Ala111.

Table 1

Parameters derived from SAXS and DLS

	Guinier R_G (Å)	Real space R_G (Å)	V_c (Å^2)	SAXS M (kDa)[a]	DLS M (kDa)[b]	Oligomeric state
TtP5CDH	43	43	1294	314	351	Hexamer
TtP5CDHR100A	32	33	628	100	123	Dimer
TtP5CDHK104A	ND[c]	ND[c]	ND[c]	ND[c]	337	Hexamer
TtP5CDHR111A	46	44	1346	319	314	Hexamer
TtP5CDHR153A	ND[c]	ND[c]	ND[c]	ND[c]	304	Hexamer
TtP5CDHR100A/K104A/R111A	ND[c]	ND[c]	ND[c]	ND[c]	120	Dimer
DrP5CDH	43	44	1295	316	361	Hexamer
DrP5CDHR102A	33	33	666	111	126	Dimer
BhP5CDH	31	32	616	98	106	Dimer
BlP5CDH	ND[c]	ND[c]	ND[c]	ND[c]	119	Dimer

[a] The molecular weight estimated from the volume of correlation as $M = V_C^2 R_G^{-1}/0.1231$.

[b] The molecular weight estimated from dynamic light scattering.

[c] Not determined.

Table 2

X-ray diffraction data collection and refinement statistics[a]

	TtP5CDH R100A	TtP5CDH R100A/K104A/R111A
Wavelength (Å)	1.000	1.542
Space group	*H*3	*P*1
Unit cell parameters (Å, °)	*a* = 102.7, *c* = 279.5	*a* = 65.1, *b* = 102.7 *c* = 160.6, α = 86.3, β = 87.5, γ = 79.4
Resolution (Å)	47.3 - 1.54 (1.62 - 1.54)	83.2-2.42 (2.55-2.42)
Total observations	835877	287864
Unique reflections	158507	114331
Multiplicity	5.3 (2.8)	2.5 (1.8)
R_{merge} [b]	0.045 (0.461)	0.038 (0.110)
R_{meas} [b]	0.050 (0.549)	0.047 (0.156)
R_{pim} [b]	0.021(0.291)	0.028 (0.110)
$\langle I/\sigma(I)\rangle$	20.7 (2.3)	15.4 (5.2)
Completeness (%)	96.8 (78.5)	73.5 (64.4)
Predicted oligomeric state[c]	hexamer	dimer
R_{work}	0.168 (0.282)	
R_{free} [d]	0.185 (0.291)	
Number of atoms	8608	
Protein residues	1032	
Water molecules	623	
RMSD bond lengths (Å)	0.006	
RMSD bond angles (°)	1.04	
Ramachandran plot[e]		
Favored (%)	98.64	
Outliers (%)	0	
MolProbity score (percentile)	99	
Average *B*-factors		
Protein (Å^2)	22.5	
Water (Å^2)	29.7	
Coordinate error (Å)[f]	0.13	
PDB code	4K57	

[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.[52]

[c] Oligomeric state predicted from crystal packing using PDBePISA.

[d] 5% random test set.

[e] The Ramachandran plot was generated with MolProbity.[53]

[f] Maximum - likelihood based coordinate error from PHENIX.