

## Unique structural features and sequence motifs of proline utilization A (PutA)

Ranjan K. Singh<sup>1</sup>, John J. Tanner<sup>1</sup>

<sup>1</sup>*Departments of Chemistry and Biochemistry, University of Missouri-Columbia, Columbia, MO 65211, USA*

### TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Comparison of a bifunctional PutA with monofunctional PRODH and P5CDH
  - 3.1. Structure of a minimalist PutA
  - 3.2. The PutA PRODH barrel has an extra helix involved in FAD binding and substrate channeling
  - 3.3. An extra helix at the C-terminus of PutA plugs the substrate-channeling cavity
  - 3.4. BjPutA has domains not found in the monofunctional enzymes
4. The conserved C-terminal motif of branch 1 PutAs
5. Beyond the minimalist PutA
  - 5.1. The DNA-binding domain of trifunctional PutA
  - 5.2. The C-terminal domain of unknown function
6. Summary
7. Acknowledgements
8. References

## 1. ABSTRACT

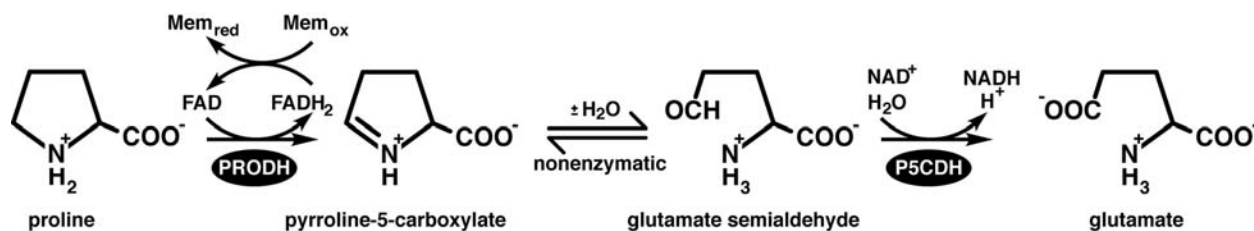
Proline utilization A proteins (PutAs) are bifunctional enzymes that catalyze the oxidation of proline to glutamate using spatially separated proline dehydrogenase and pyrroline-5-carboxylate dehydrogenase active sites. Here we use the crystal structure of the minimalist PutA from *Bradyrhizobium japonicum* (BjPutA) along with sequence analysis to identify unique structural features of PutAs. This analysis shows that PutAs have secondary structural elements and domains not found in the related monofunctional enzymes. Some of these extra features are predicted to be important for substrate channeling in BjPutA. Multiple sequence alignment analysis shows that some PutAs have a 17-residue conserved motif in the C-terminal 20-30 residues of the polypeptide chain. The BjPutA structure shows that this motif helps seal the internal substrate-channeling cavity from the bulk medium. Finally, it is shown that some PutAs have a 100-200 residue domain of unknown function in the C-terminus that is not found in minimalist PutAs. Remote homology detection suggests that this domain is homologous to the oligomerization beta-hairpin and Rossmann fold domain of BjPutA.

## 2. INTRODUCTION

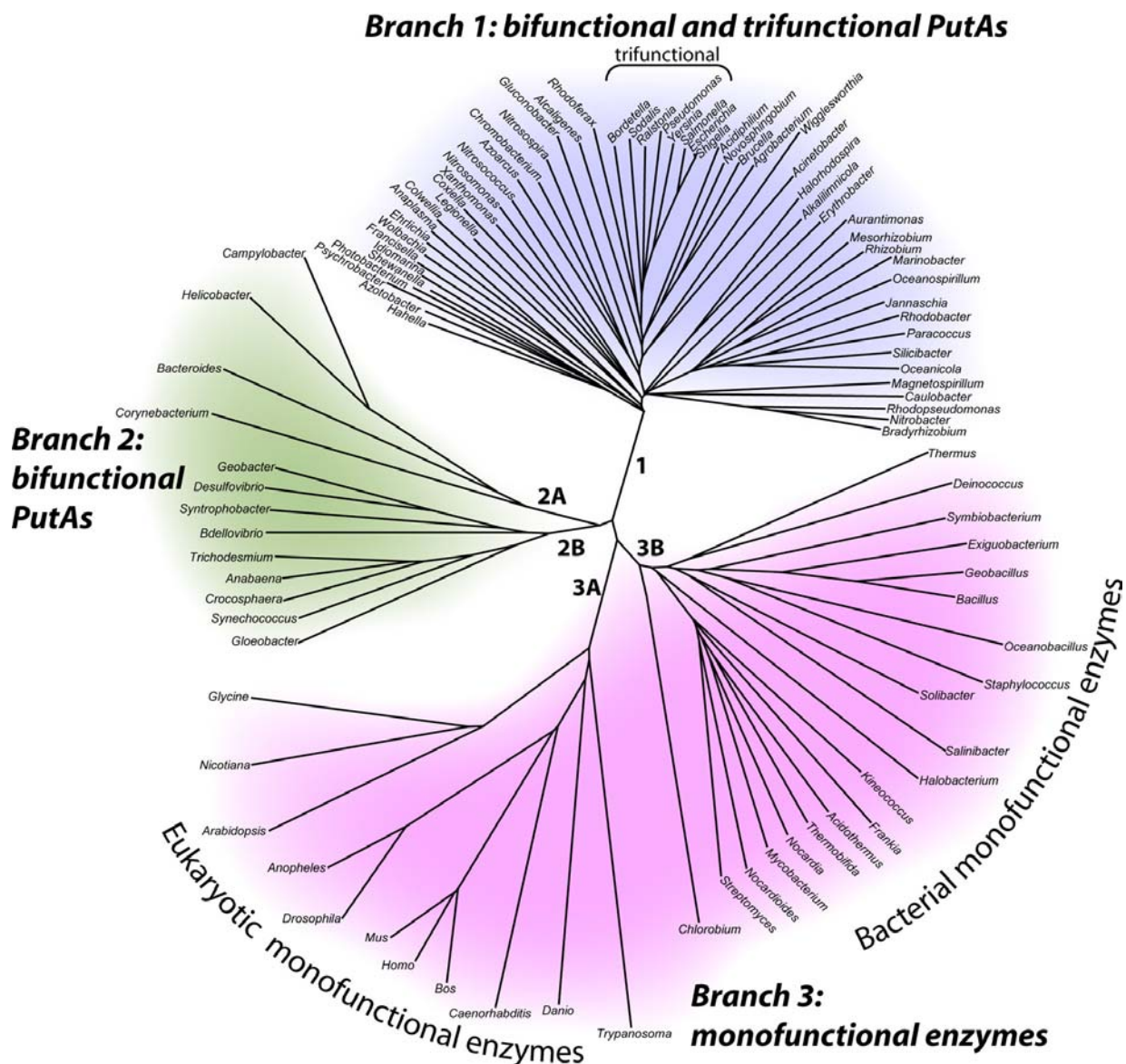
The oxidation of proline to glutamate, i.e. proline catabolism, is catalyzed by two enzymes, proline dehydrogenase (PRODH) and pyrroline-5-carboxylate (P5C) dehydrogenase (P5CDH) (Figure 1). The former catalyzes the oxidation of L-proline to P5C with concomitant reduction of an enzyme-bound FAD cofactor. The latter enzyme catalyzes the oxidation of L-glutamate semi-aldehyde (GSA) to L-glutamate using NAD<sup>+</sup> as the electron acceptor. Note that the product of the PRODH reaction is not the substrate for P5CDH. Instead, the two reactions are coupled by the hydrolysis of P5C, which has traditionally been thought of as a nonenzymatic process. As noted by Phang over two decades ago, researchers typically refer to P5C and GSA interchangeably, since the two species are not distinguished in most experiments (1). Phang's observation remains valid today.

One interesting feature of proline catabolism is that PRODH and P5CDH are combined into a single polypeptide in some organisms (Figure 2). The combined enzymes are known as proline utilization A (PutA) and were discovered by Roth's group in the late 1970s during

## Unique features of puta



**Figure 1.** The reactions of proline catabolism.



**Figure 2.** Phylogenetic tree representing the organization of proline catabolic enzymes in bacteria and eukaryotes. PutAs are found in branches 1 and 2. Monofunctional PRODH and P5CDH enzymes are found in branch 3. A cluster of trifunctional PutAs is indicated.

their studies of proline utilization in *Salmonella typhimurium* (2). Analysis of genome sequence data suggests that PutAs are limited to Gram-negative bacteria (Figure 2, branches 1, 2), whereas PRODH and P5CDH are

separate enzymes encoded by distinct genes in Gram-positive bacteria (branch 3B) (3). In eukaryotes, PRODH and P5CDH are also separate enzymes and are localized to mitochondria (branch 3A). Human PRODH is a p53-

## Unique features of puta

induced tumor suppressor protein localized to the inner mitochondrial membrane and is often referred to as POX to emphasize its role as a superoxide-generating oxidase (4-12). Human P5CDH (ALDH4 (13)) is also induced by p53 (14) and is located in the mitochondrial matrix. ALDH4 has been characterized biochemically, including elucidation of the oligomeric state in solution (dimer) and kinetic mechanism (15, 16).

The PutA part of the PutA/PRODH/P5CDH family tree has two branches (3, 17). Branch 1 primarily consists of PutAs from alpha-, beta-, and gamma-proteobacteria. Branch 2 includes PutAs from delta- and epsilon-proteobacteria as well as cyanobacteria. The PutAs in branch 1 have chain lengths from 999 to almost 1400 residues, and the pairwise sequence identities are greater than 38 %. The polypeptide length for branch 2 PutAs ranges from around 980 to almost 1300 residues, and the pairwise sequence identity range can be as low as 23 %. Thus, branch 2 PutAs appear to be a more diverse group than branch 1 PutAs. Between branches 1 and 2, the pairwise sequence identities are typically less than 30 %. Nevertheless, the residues in the PRODH and P5CDH active sites are highly conserved, indicating that the three-dimensional structures of the catalytic domains are conserved by PutAs. Whether the three-dimensional arrangement of the catalytic and other domains is likewise conserved remains to be determined.

PutAs are further classified as bifunctional or trifunctional. Bifunctional PutAs exhibit only PRODH and P5CDH catalytic activities, have polypeptide chain lengths in the range of ~980 residues to over 1300 residues, and are found in both PutA branches. Bifunctional PutAs from *Bradyrhizobium japonicum* (BjPutA, (18-20)) and *Helicobacter* species (21-23) have been studied. Trifunctional PutAs constitute a subset of branch 1 PutAs and are distinguished by the presence of a DNA-binding domain (a ribbon-helix-helix domain) in the first ~50 residues of the polypeptide chain. The polypeptide chain length of trifunctional PutAs are in the range of ~1270-1361. In addition to functioning as dual PRODH/P5CDH enzymes, trifunctional PutAs have a third function of repressing transcription of the *put* regulon, which contains the genes encoding PutA and the proline transporter PutP, when proline levels are low (24-27). High levels of proline in the bacterium's environment cause PutA to disengage from the *put* control region thus activating transcription of *putA* and *putP*. Thus, trifunctional PutAs are remarkable proteins that link transcription and metabolism in response to an environmental cue (proline level). Trifunctional PutAs from *S. typhimurium* (25, 26, 28-31) and *Escherichia coli* (EcPutA) (24, 32-46) have been studied. PutA from *E. coli* is the most studied trifunctional PutA and is considered to be the archetypal trifunctional PutA.

The observation that enzymes catalyzing successive reactions in a metabolic pathway are combined into a single polypeptide chain as in PutA has intriguing implications. First, the covalent linking of the two active sites may allow the transfer of the reaction

product of one enzyme to the next without equilibrating with the bulk medium. Substrate channeling is the term used for such kinetic mechanisms, and Arentson *et al.* provide a review of substrate channeling in proline metabolism in this issue (47). Two limiting channeling mechanisms are possible: direct transfer and proximity. In the former, the intermediate moves through an internal cavity or tunnel connecting the two active sites without leaving the confines of the protein. Proximity refers to a spectrum of cases in which the reaction product dissociates from the first enzyme but encounters a locally high concentration of the second enzyme and thus does not truly equilibrate with the bulk medium. There are potential advantages of substrate channeling, including decreased transit time between enzymes, protection of labile intermediates, and isolation of intermediates from competing enzymatic reactions (48-50). The latter point is relevant for proline metabolism, because P5C/GSA is common to proline catabolism, proline biosynthesis, and arginine biosynthesis. Another implication of fused enzymes was recognized by Eisenberg's group in 1999 and concerns the prediction of protein-protein interactions (51). The basic idea is that the observation that two proteins are separate in some organisms and fused in others implies that the two separate proteins form a functional association. Thus, protein fusion (or gene fusion) provides the 'Rosetta stone' for identifying potential protein-protein interactions. In proline catabolism, for example, the Rosetta Stone hypothesis suggests the possibility that monofunctional PRODH and P5CDH interact and engage in intermolecular substrate channeling.

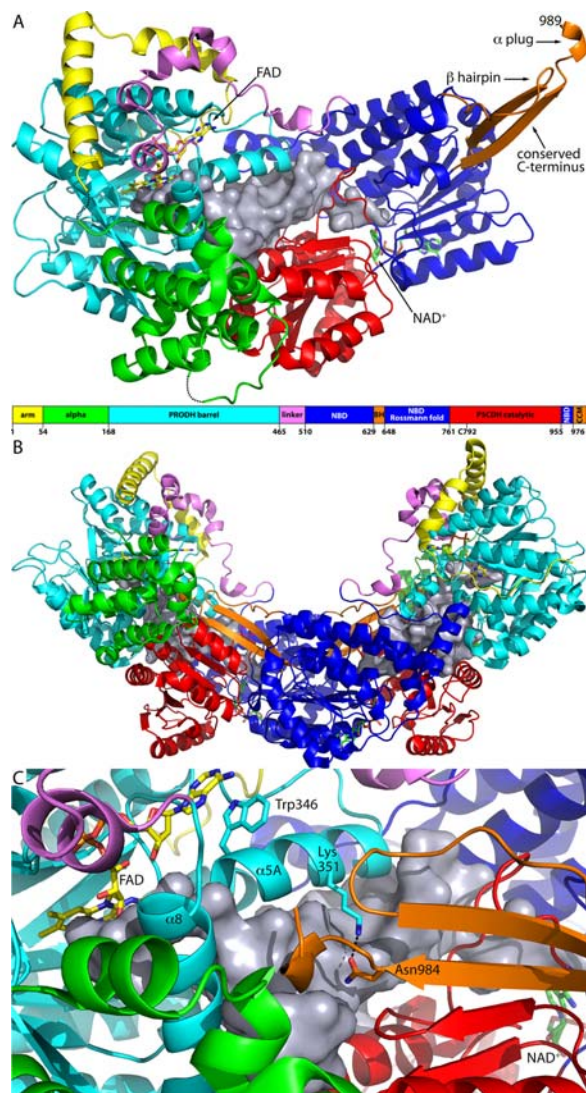
Three-dimensional structural studies have contributed to our understanding of proline catabolic proteins. Crystal structures have been solved for the monofunctional PRODH and P5CDH enzymes from *Thermus thermophilus* (TtPRODH (17, 52, 53), TtP5CDH (54, 55)), the DNA-binding domains of two trifunctional PutAs (56-58), a PRODH domain construct of EcPutA (EcPutA86-630, (46, 59-62)), and full-length BjPutA (63). Here, we use the BjPutA structure as a platform for identifying the unique features that distinguish PutAs from their monofunctional relatives and to gain insight into the structure and function of the C-terminal domains of PutAs.

## 3. COMPARISON OF A BIFUNCTIONAL PUTA WITH MONOFUNCTIONAL PRODH AND P5CDH

### 3.1. Structure of a minimalist PutA

PutA from *Bradyrhizobium japonicum* (BjPutA) is one of the simplest of bifunctional PutAs. At 999 residues in length, it is the shortest branch 1 PutA known and is thus considered to be a minimalist PutA. We recently reported the crystal structure of BjPutA (PDB code 3haz), which is the first structure of a full-length PutA. The structure shows that PutAs are more than simply a fusion of two catalytic domains. The protomer comprises seven domains: arm, alpha, PRODH barrel, linker, NAD<sup>+</sup>-binding, P5CDH catalytic, and oligomerization domain (Figure 3A). The PRODH active site is located in a

## Unique features of puta



**Figure 3.** Structure of BjPutA. (A) Protomer structure with the domains colored according to the domain diagram. The silver surface represents the substrate-channeling cavity. FAD and NAD<sup>+</sup> are represented as yellow and green sticks, respectively. Abbreviations used in the domain diagram are as follows: NBD, NAD<sup>+</sup>-binding domain; BH, beta-hairpin; CCM, conserved C-terminal motif. (B) The domain-swapped dimer of BjPutA. The domains are colored according to the domain diagram in panel A. The silver surfaces represent the substrate-channeling cavities of the two protomers. (C) Close-up view of the oligomerization domain covering the cavity of the other protomer. This figure and others were prepared with PyMol (72).

distorted (beta-alpha)<sub>8</sub> barrel. The barrel structure is very similar to that of EcPutA (59). The P5CDH active site is located in the interface between the NAD<sup>+</sup>-binding domain and the P5CDH catalytic domain. The two active sites are separated by 41 Å and connected by a large, irregularly shaped internal cavity (silver surface in Figure 3A). This cavity most likely functions in substrate channeling, as described in this issue (47).

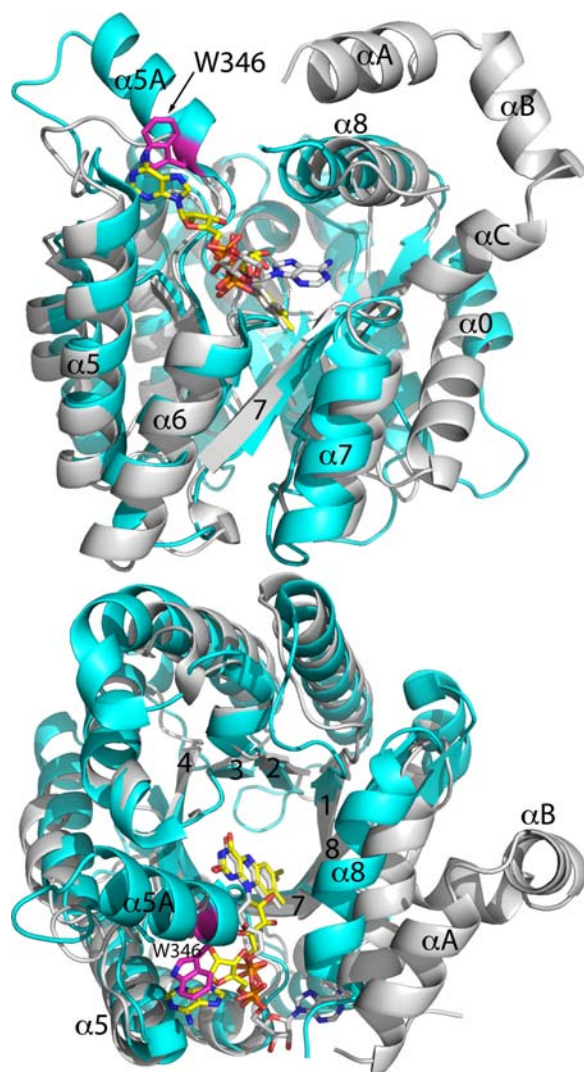
Oligomerization is essential for substrate channeling in BjPutA. The enzyme forms a U-shaped dimer (Figure 3B), and two of these dimers assemble into a ring-shaped dimer-of-dimers tetramer (see reference (63)). The dimer shown in Figure 3B is the relevant oligomeric state for understanding substrate channeling, so we will describe it in detail here. Dimerization is mediated by an oligomerization domain (orange in Figure 3) that protrudes from the NAD<sup>+</sup>-binding domain. The oligomerization domain consists of two elements that are far apart in sequence, a beta-hairpin formed by residues 629-647 and a beta-strand followed by a short helix at the C-terminus of polypeptide chain (residues 976-989). The C-terminal strand of the oligomerization domain of one protomer forms main chain hydrogen bonds with the beta-sheet of the P5CDH catalytic domain of the other protomer to form a large, twisted intermolecular beta-sheet (Figures 3B and 3C). This type of oligomerization is an example of domain swapping (64). Both the bipartite oligomerization domain and the domain-swapped dimer are conserved by aldehyde dehydrogenases. In BjPutA, the oligomerization domain not only mediates dimerization but also covers the cavity of the other protomer (Figure 3C). Without the lid provided by the oligomerization domain, the cavity would be open to the bulk medium. Thus dimerization appears to be essential for formation of the substrate channeling cavity. The sealed, internal cavity is consistent with direct transfer rather than proximity as the predominant mechanism of substrate channeling in BjPutA.

### 3.2. The PutA PRODH barrel has an extra helix involved in FAD binding and substrate channeling

Both the PutA PRODH domain and the monofunctional enzyme TtPRODH exhibit the distorted (beta-alpha)<sub>8</sub> barrel that is characteristic of the PRODH/PutA family. As shown in Figure 4, the two folds are nearly identical; the RMSD between the two proteins is 2.0 Å for 256 aligned residues. In both structures, the FAD is bound at the C-terminal ends of the strands of the barrel. The fold is unusual in that the final helix, alpha8, sits atop the barrel rather than alongside beta8 as in the classical triosephosphate isomerase barrel. Helix 8 contains residues that are critical for substrate recognition, including a conserved Arg-Arg motif that has been shown to bind the substrate carboxylate group (60). Both PutA and the monofunctional PRODH exhibit this distortion of the classical (beta-alpha)<sub>8</sub> barrel fold.

Despite the similarities in overall fold and amino acid sequence (28 % identity), there is an important difference between monofunctional PRODH and PutA. The PutA PRODH barrel has an extra helix (alpha5a) inserted between beta5 and alpha5 (Figure 4). This additional secondary structural element is also present in EcPutA86-630, suggesting that it is conserved in branch 1 PutAs. Helix 5a contains a conserved tryptophan residue (Trp346 in BjPutA, Trp438 in EcPutA) that stacks against the FAD adenine (Figures 3C and 4), and this interaction presumably contributes to the different FAD conformations in PutA and TtPRODH, as described previously (17). The BjPutA structure reveals a new function for helix 5a. The helix forms a large section of the wall of the internal





**Figure 4.** Comparison of the monofunctional enzyme TtPRODH (white) and the PRODH barrel of BjPutA (cyan). Strands of the barrel are labeled 1-8 using the standard convention for  $(\beta\alpha)_8$  barrels. The extra helix of the PutA barrel is denoted alpha5a. Trp346 of alpha5a is colored magenta. The three helices of TtPRODH that precede the barrel are labeled alphaA, alphaB, and alphaC. Two orthogonal views are shown.

substrate-channeling cavity (Figure 3C), and its absence would leave a large hole to the bulk medium. Thus, helix 5a seems to be essential for channeling in PutA.

Finally, we note that TtPRODH has three additional helices that precede the start of the barrel (alphaA, alphaB, alphaC in Figure 4). These helices are located in the vicinity of the alpha domain of BjPutA, but they do not superimpose well with any of the helices of the alpha domain.

### 3.3. An extra helix at the C-terminus of PutA plugs the substrate-channeling cavity

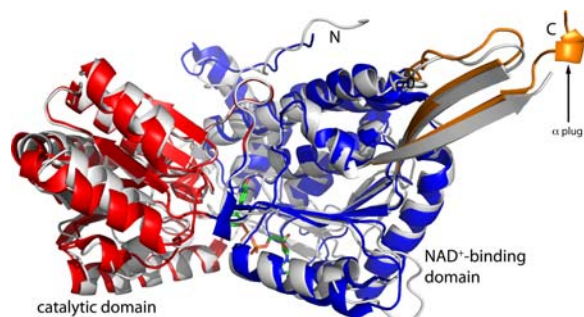
The structure of the monofunctional enzyme TtP5CDH is very similar to that of the P5CDH half of

BjPutA (Figure 5). The sequence identity of TtP5CDH to PutA is about 38 %, implying substantial structural similarity. Indeed, the RMSD between the two proteins is 1.4 Å for 463 residues. Both enzymes exhibit the characteristic aldehyde dehydrogenase fold (65), which consists of three domains, NAD<sup>+</sup>-binding, catalytic, and oligomerization. (The oligomerization is also referred to as the bridging domain (65) and beta-flap (63)). The NAD<sup>+</sup>-binding domain is structurally contiguous but actually consists of three separate sections of the polypeptide chain, as indicated in the domain diagram in Figure 3A. Sections 1 and 2 are separated in primary structure by the beta-hairpin of the oligomerization domain, while sections 2 and 3 are separated by the P5CDH catalytic domain. Section 2 exhibits a variation of the Rossmann dinucleotide binding fold. The classical Rossmann fold consists of two beta-alpha-beta-alpha-beta motifs that form a 6-stranded parallel beta sheet with relative strand order 321456 (66). The Rossmann domain of aldehyde dehydrogenase lacks the final strand and helix and is thus referred to as a non-classical Rossmann fold domain. The catalytic domains of TtP5CDH and BjPutA are also quite similar, as are the details of the P5CDH active sites (63).

Closer inspection of BjPutA and TtP5CDH, however, reveals an important difference at the C-terminus (Figure 5). Structure-based alignment of BjPutA and TtP5CDH shows that the BjPutA chain extends 14 residues past the final residue of TtP5CDH (Phe516). Some of these extra residues were resolved in the BjPutA structure, and they form a turn of helix that plugs a hole in the cavity wall (Figure 3C). Without this plug, there would be a significant hole in the cavity leading to the bulk medium. Interestingly, the extra residues of BjPutA are incompatible with the observed oligomeric state of TtP5CDH. TtP5CDH forms a hexamer, which can be thought of as a trimer of domain-swapped dimers. If the TtP5CDH chain were longer as in BjPutA, the C-terminus would clash with another dimer of the hexamer.

### 3.4. BjPutA has domains not found in the monofunctional enzymes

PutAs also have extra domains not found in the monofunctional enzymes, and the BjPutA structure exhibits three of them: arm, alpha domain, and linker (Figure 3A). The alpha-helical arm at the N-terminus (yellow in Figure 3A) wraps around the PRODH barrel and sits below the linker. The arm domain is observed in both BjPutA and EcPutA86-630, indicating that it is conserved by branch 1 PutAs. The arm connects to the alpha domain, which is a globular domain consisting of 6 helices (green in Figure 3A). The alpha domain contacts both the PRODH and P5CDH domains, as well as the oligomerization domain of the other protomer of the domain-swapped dimer (Figures 3B and 3C). Its strategic location at the confluence of three domains suggests that it is critical for properly orienting the two active sites for channeling and for formation of the internal cavity. We note that the alpha domain was disordered in the structure of EcPutA86-630, indicating that contacts with the P5CDH half of the enzyme are required for proper folding. As noted above, the alphaA, alphaB, and alphaC helices of TtPRODH are in the same general



**Figure 5.** Comparison of the monofunctional enzyme TtP5CDH (white) and the P5CDH half of BjPutA. BjPutA is colored according to the legend in Figure 3A. NAD<sup>+</sup> is drawn in green sticks.

location as the alpha domain of BjPutA. Whether these helices play a similar role as the PutA alpha domain is unknown.

The polypeptide that links the PRODH barrel to the NAD<sup>+</sup>-binding domain is also unique to PutA. The linker is not simply a flexible tether that keeps the two catalytic domains in close proximity, but rather has a well-defined structure that appears to be essential for maintaining the tertiary and quaternary structure of the enzyme. The linker (residues 465-509, violet in Figure 3) joins the C-terminus of helix 8 of the PRODH barrel to the N-terminus of the NAD<sup>+</sup>-binding domain. The 45-residue linker consists of 5 short helical segments that form a meandering U-turn, effectively redirecting the chain toward the P5CDH domain. We note that a nearly identical meandering U-turn is also found in the structure of EcPutA86-630, suggesting that the linker structure is conserved by branch 1 PutAs. Because of its wide, curved path, the linker traverses 100 Å, although the two residues it connects are separated by only 30 Å. The linker forms extensive interactions with other domains, which likely are essential for maintaining its three-dimensional structure. The majority of these interactions are with the PRODH domain (2000 Å<sup>2</sup> of inter-domain buried surface area) and the arm (1300 Å<sup>2</sup>). The large contact area with the arm reflects the fact that the linker sits atop the arm, essentially tracking its curved path around the barrel. In fact, it is possible that the primary role of the arm is to help stabilize the conformation of the linker. In summary, the linker is a structural element that interacts with disparate parts of the polypeptide chain via noncovalent interactions, and as a result is important for properly orienting the two active sites and creating the substrate channeling cavity.

#### 4. THE CONSERVED C-TERMINAL MOTIF OF BRANCH 1 PUTAS

The second element of the oligomerization domain contains a conserved sequence motif, which to our knowledge has not been described previously. Multiple sequence alignment (MSA) analysis reveals a conserved stretch of 17-residues located in the C-terminal 20-30 residues of branch 1 PutAs (Figure 6). The motif is found in both bifunctional and trifunctional branch 1 PutAs, but it

does not appear to be present in branch 2 PutAs and monofunctional P5CDHs. Based on the sequences analyzed, the consensus motif is Exxxxv[N or D]t[T or A]AaGGnaxL, where upper case denotes identity, lowercase denotes presence in over half of the sequences, and x denotes no significant conservation.

The BjPutA structure provides the three-dimensional context of the conserved motif. The first 12 residues of the motif were resolved in the structure. The motif begins at the N-terminus of the final strand of the enzyme and extends beyond the helical plug. The identically conserved Glu at position 1 forms an intersubunit hydrogen bond to the backbone of Lys965. The Asn at position 7, which is Asp in some sequences, forms an intersubunit hydrogen bond with Lys351 (Figure 3C). Lys351 is highly conserved in branch 1 PutAs (Arg in some sequences) and is part of helix 5a of the PRODH barrel. As described above, helix 5A is not found in monofunctional PRODHs. Positions 9-11 form the turn of helix that plugs a hole in the cavity wall. This analysis suggests that the C-terminal motif is important for formation of the substrate-channeling cavity.

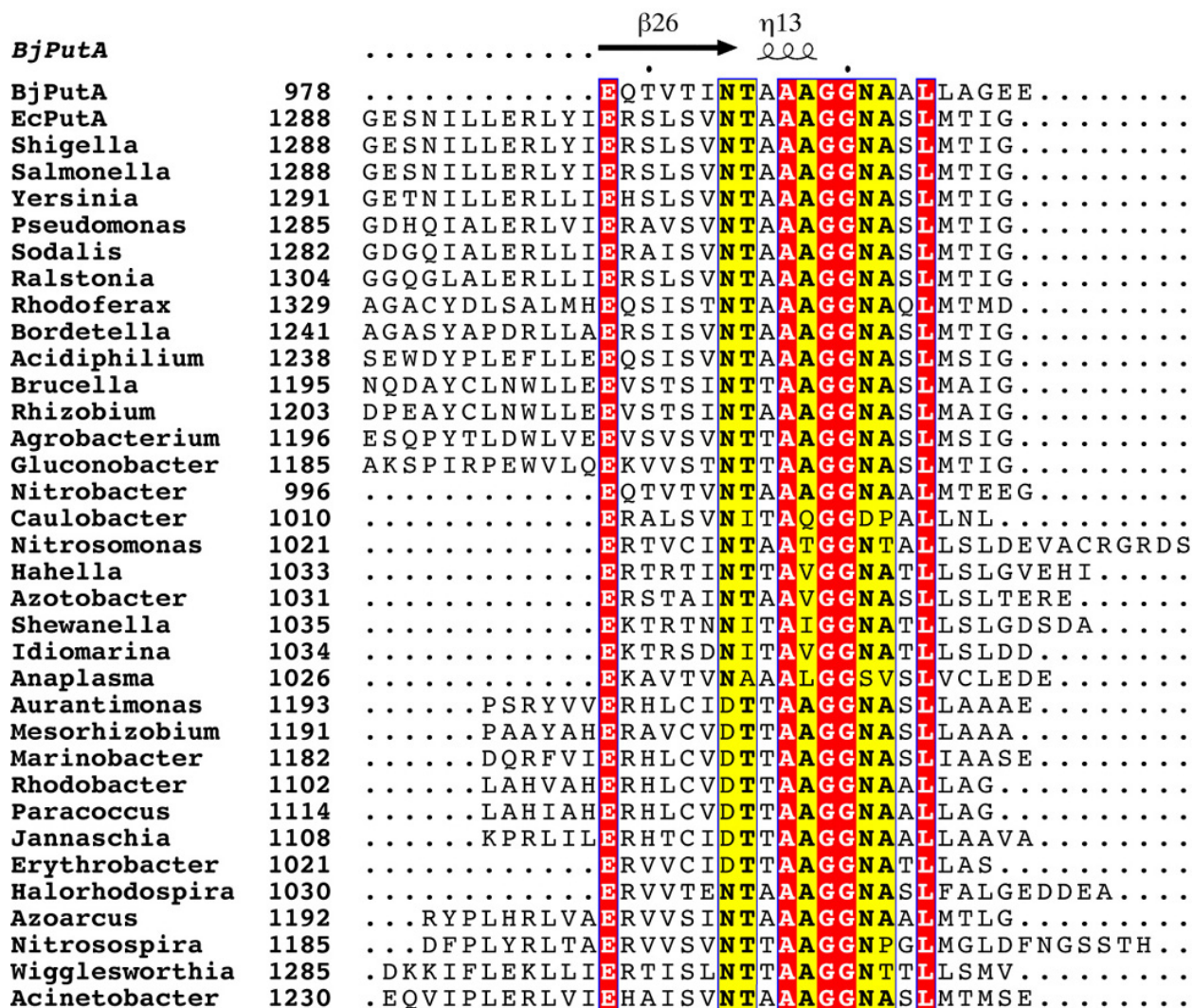
#### 5. BEYOND THE MINIMALIST PUTA

Many PutAs have chain lengths that are much longer than that of the 999-residue BjPutA, so it is natural to ask how these extra residues beyond the minimalist PutA are incorporated into the polypeptide chain. MSAs provide information about which domains are shared among PutAs and the locations of the extra domains of long PutAs. We will focus here on branch 1 PutAs, because the sequence identity is high within this group and the resulting trends are obvious. Although we have analyzed many branch 1 PutAs using MSAs, we present an alignment of just three due to space limitations. The results presented here are valid for the larger group. Figure 7 shows an MSA of BjPutA with a representative long bifunctional branch 1 PutA (*Azoarcus* PutA) and the archetypal trifunctional PutA, EcPutA. The alignment exhibits a long region of high identity (48 - 59 % pairwise) corresponding to BjPutA residues 5-974. This region corresponds to the arm, alpha, PRODH, linker, NAD<sup>+</sup>-binding, and P5CDH catalytic domains of BjPutA. Note that there is a gap in the long PutAs corresponding to the beta-hairpin of the oligomerization domain (Figure 7, beta11). The second region of high identity corresponds to the conserved C-terminal motif. These results suggest that all the domains described here for BjPutA are also present in the long PutAs except for the beta-hairpin, which appears to be abbreviated or absent. The MSA also indicates an extra domain at the N-terminus of trifunctional PutAs and another domain immediately preceding the conserved C-terminal motif in long branch 1 PutAs. The N-terminal domain is the ribbon-helix-helix DNA-binding domain. The function of the extra C-terminal domain is unknown.

##### 5.1. The DNA-binding domain of trifunctional PutA

The DNA-binding domains of EcPutA and *Pseudomonas putida* PutA have been extensively characterized using X-ray crystallography, NMR, and an





**Figure 6.** Section of an MSA of branch 1 PutAs showing the conserved motif at the C-terminus. The trifunctional PutAs are EcPutA through *Acidiphilium*, plus *Wigglesworthia*. The other PutAs are bifunctional. The secondary structure elements are from the BjPutA structure (PDB code 3haz). This figure and others were prepared with ClustalW2 (73) and ESPrnt (74).

array of biophysical and biochemical techniques (56-58, 67). The PutA DNA-binding domain has the ribbon-helix-helix (RHH) fold, which identifies PutA as a member of a large superfamily of transcription factors that includes Arc, MetJ, CopG, and NikR, among others. Schreiter and Drennan have written an excellent review on the RHH superfamily (68). The DNA-binding domain connects to the arm domain via a ~35-residue polypeptide (Figure 7). No structural information is available for these residues, thus it is not known whether this polypeptide is a flexible tether or has a well-defined three-dimensional structure. The sequence identity of the linker is relatively low (results not shown), which perhaps argues in favor of a flexible tether. Furthermore, analysis of the sequences of several trifunctional PutAs using the Disopred server suggests a high probability of disorder within residues 50-80 (69).

## 5.2. The C-terminal domain of unknown function

MSAs indicate that trifunctional PutAs have a

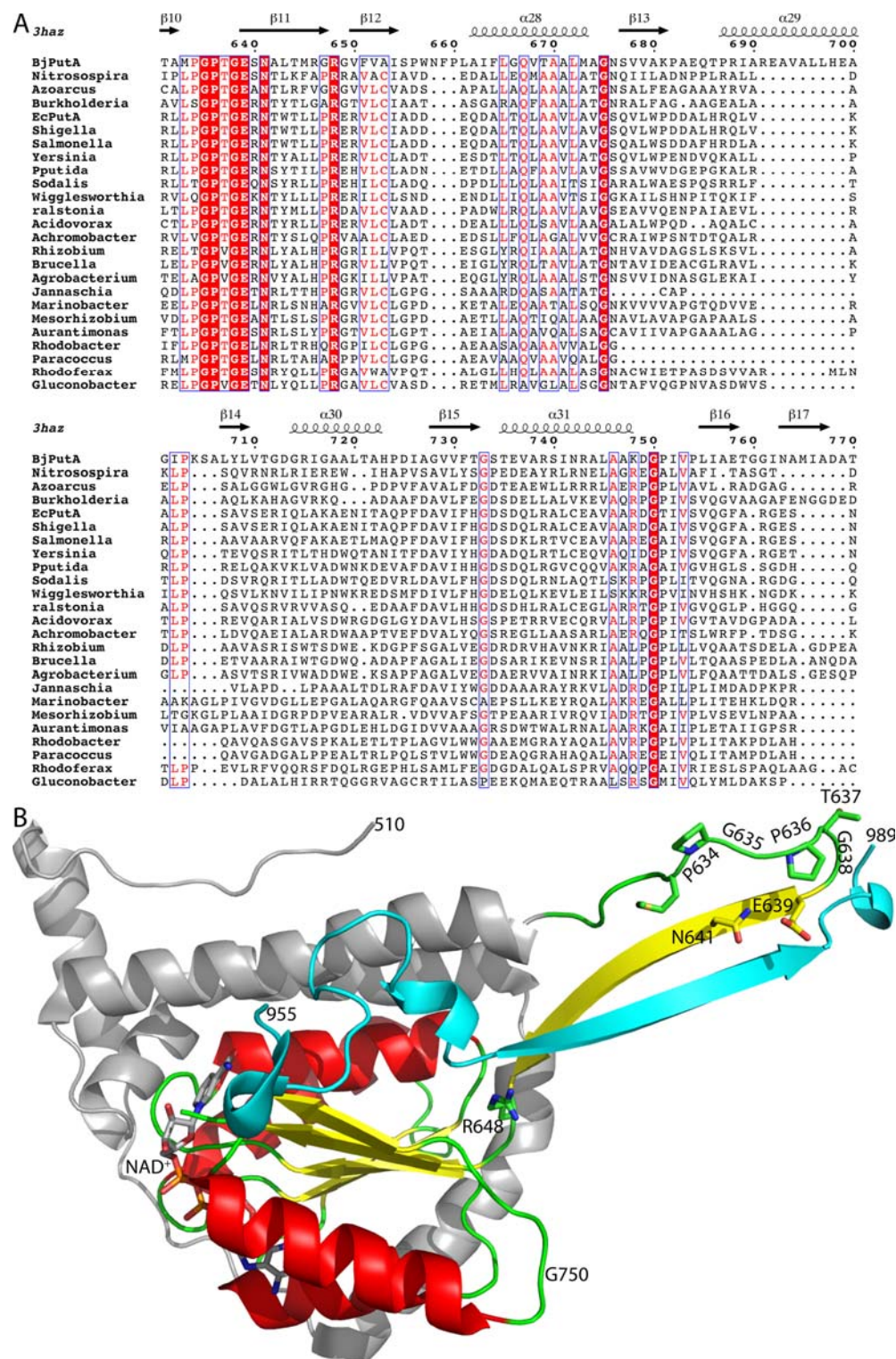
~200-residue domain near the C-terminus (Figure 7, EcPutA). Bifunctional PutAs of branch 1 that are longer than about 1100 residues also have this domain; the PutA from *Azoarcus* sp. BH72 is one example (Figure 7, Az). Interestingly, long bifunctional PutAs of branch 2 also have an extra domain in the C-terminus, but it is not clear whether it is related to the C-terminal domain of branch 1 PutAs. Thus, we will restrict our discussion here to the conserved C-terminal domain of branch 1 PutAs, which we denote by CTD. Using BjPutA as a reference, the CTD is inserted between the third section of the NAD<sup>+</sup>-binding domain and the conserved C-terminal motif of the oligomerization domain. The size of the CTD ranges from about 130 residues to 220 residues.

Although the function of the CTD is unknown, amino acid sequence analysis provides some intriguing ideas to test. Analysis of over 2 dozen branch 1 PutA CTDs using the remote homology detection algorithm

**Figure 7.** MSA of three branch 1 PutAs: BjPutA (Bj, GenBank BAC52526.1), *Azoarcus* sp. BH72 PutA (Az, GenBank CAL96369.1), and EcPutA (Ec, GenBank AAB59985.1). BjPutA is a minimalist PutA. Az is a long bifunctional PutA. Ec is a trifunctional PutA. The secondary structure elements above the sequence blocks are from the BjPutA structure (PDB code 3haz). The secondary structure elements for the N-terminal ribbon-helix-helix domain are from a structure of the EcPutA DNA-binding domain (PDB code 2gpe). Symbols below the sequence blocks denote the following: green squares, substrate-channeling cavity; triangles, proline binding site; hexagons, FAD binding site; diamonds, NAD<sup>+</sup> binding site; ovals, GSA binding site; stars, catalytic Cys of the PSCDH catalytic domain and the Glu that is predicted to assist in hydrolysis of the thioacylenzyme.







**Figure 8.** Homology of the CTD to the BjPutA beta-hairpin and Rossmann fold domain. (A) MSA of BjPutA with the CTDs of several branch 1 PutAs. (B) The structure of the NAD<sup>+</sup>-binding and oligomerization domains of BjPutA. The core region of homology with the CTD (residues 629-760) is colored according to secondary structure, with alpha helices in red, beta strands in yellow, and loops in green. Conserved residues of the CTD are indicated. Residues 510-628 are colored silver. Residues 955-989 are colored cyan.

## Unique features of puta

HHSearch (70, 71) suggests that the CTD is homologous to the beta-hairpin and Rossmann fold regions of aldehyde dehydrogenases. For example, using the CTD of EcPutA as the query, HHSearch returned 33 aldehyde dehydrogenases with probability scores ranging from 99.1 % to 99.9 %. The top match was BjPutA with a probability score of 99.9 and E-value of 7.7E-26. The alignment shows that the CTD of EcPutA is 25 % identical to residues 551-761 of BjPutA. In the other cases tested, the BjPutA beta-hairpin/Rossmann domain was also identified as the closest homolog, and the probability score was in the range 99.7-100.0 %. These results are indicative of meaningful homology.

The homology is evident in an MSA of BjPutA with the CTDs of several branch 1 PutAs (Figure 8A). The core region of homology corresponds to the beta-hairpin of the oligomerization domain and the Rossmann fold domain of BjPutA. The structure of this region is highlighted in green, yellow, and red in Figure 8B. The highest similarity is found in the beta-hairpin, which contains the conserved motif IpGptGExN. This result is significant, because multiple sequence alignments show that long branch 1 PutAs have a gap corresponding to the beta-hairpin of BjPutA (Figure 7, beta11). Thus, it appears that the beta-hairpin of the oligomerization domain has been shifted to the CTD in long PutAs.

In summary, remote homology detection suggests that the CTD includes a beta-hairpin that is homologous oligomerization domain of aldehyde dehydrogenases plus a Rossmann fold domain. Thus, long branch 1 PutAs are predicted to have two Rossmann fold domains, one that has high identity (~50 %) to the Rossmann fold domain of BjPutA and a second one in the CTD that has lower identity (14 - 33 %).

## 6. SUMMARY

Comparison of the PutA structures to those of the monofunctional enzymes is useful for thinking about the possibility of protein-protein interactions predicted by the Rosetta Stone hypothesis. Several unique features of PutAs are absent in the monofunctional enzymes, including helix alpha5a, the arm domain, the alpha domain, and the linker domain. These components appear to be important for orienting the catalytic domains of the PutA protomer so that the two active sites face each other and for sealing the substrate-channeling cavity from the bulk medium. The absence of these PutA-specific structural features in the monofunctional enzymes, at least those from branch 3B, perhaps argues against the formation of an efficient PRODH-P5CDH channeling complex. However, kinetic measurements of substrate channeling and biophysical measurements of protein-protein association are still needed to test the Rosetta Stone hypothesis for monofunctional proline catabolic enzymes.

The CTD is the only PutA domain that has not been structurally characterized. Remote homology detection suggests the tantalizing hypothesis that the CTD of branch 1 PutAs contains a beta-hairpin like the one in

the oligomerization domain of aldehyde dehydrogenases and a Rossmann fold domain. Interestingly, remote homology detection did not find meaningful homology for the C-terminal domains of branch 2 enzymes. Thus, the C-terminal domains of branch 1 and 2 enzymes may differ in structure. These results raise several intriguing questions. Does the CTD bind NAD<sup>+</sup> or is it a pseudo-NAD<sup>+</sup>-binding domain that primarily plays a structural role? We note that the oxidation of GSA to Glu requires only one equivalent of NAD<sup>+</sup> and thus only one functional NAD<sup>+</sup>-binding domain is expected. Furthermore, the CTD is missing the NFP motif found in BjPutA residues Asn658-Phe659-Pro660 (Figure 8A). This motif is highly conserved by P5CDHs (including the first Rossmann domain of all PutAs), and the Asn residue is thought to help anchor the substrate in the oxyanion hole by forming a hydrogen bond with the epsilon O atom of GSA (54). The absence of this critical residue argues against the CTD participating in catalysis. If the CTD plays a structural role, does the predicted beta-hairpin interact with the C-terminal strand and participate in domain-swapped dimerization? And, does it help cover the substrate-channeling cavity as in BjPutA? New biochemical and structural studies on PutAs designed to answer these questions represent an exciting next phase of research in proline catabolism.

## 7. ACKNOWLEDGEMENTS

We thank Dr. Donald Becker for insightful discussions. This was supported in part by the National Institutes of Health Grants GM065546 and GM061068.

## 8. REFERENCES

1. J. M. Phang: The regulatory functions of proline and pyrroline-5-carboxylic acid. *Curr. Top. Cell. Reg.*, 25, 92-132 (1985)
2. B. Ratzkin and J. Roth: Cluster of genes controlling proline degradation in *Salmonella typhimurium*. *J. Bacteriol.*, 133(2), 744-54 (1978)
3. J. J. Tanner: Structural biology of proline catabolism. *Amino Acids*, 35(4), 719-30 (2008)
4. S. J. Downing, J. M. Phang, E. M. Kowaloff, D. Valle and R. J. Smith: Proline oxidase in cultured mammalian cells. *J. Cell. Physiol.*, 91(3), 369-76 (1977)
5. E. M. Kowaloff, J. M. Phang, A. S. Granger and S. J. Downing: Regulation of proline oxidase activity by lactate. *Proc. Natl. Acad. Sci. U S A*, 74(12), 5368-71 (1977)
6. S. P. Donald, X. Y. Sun, C. A. Hu, J. Yu, J. M. Mei, D. Valle and J. M. Phang: Proline oxidase, encoded by p53-induced gene-6, catalyzes the generation of proline-dependent reactive oxygen species. *Cancer Res.*, 61(5), 1810-5 (2001)
7. C. A. Hu, S. P. Donald, J. Yu, W. W. Lin, Z. Liu, G. Steel, C. Obie, D. Valle and J. M. Phang: Overexpression of proline oxidase induces proline-dependent and

## Unique features of puta

mitochondria-mediated apoptosis. *Mol. Cell. Biochem.* (2006)

8. Y. Liu, G. L. Borchert, A. Surazynski, C. A. Hu and J. M. Phang: Proline oxidase activates both intrinsic and extrinsic pathways for apoptosis: the role of ROS/superoxides, NFAT and MEK/ERK signaling. *Oncogene*, 25(41), 5640-7 (2006)

9. J. Pandhare, S. K. Cooper and J. M. Phang: Proline oxidase, a proapoptotic gene, is induced by troglitazone: evidence for both peroxisome proliferator-activated receptor gamma-dependent and -independent mechanisms. *J. Biol. Chem.*, 281(4), 2044-52 (2006)

10. S. K. Cooper, J. Pandhare, S. P. Donald and J. M. Phang: A novel function for hydroxyproline oxidase in apoptosis through generation of reactive oxygen species. *J. Biol. Chem.*, 283(16), 10485-92 (2008)

11. J. M. Phang, S. P. Donald, J. Pandhare and Y. Liu: The metabolism of proline, a stress substrate, modulates carcinogenic pathways. *Amino Acids*, 35(4), 681-90 (2008)

12. J. M. Phang, J. Pandhare and Y. Liu: The metabolism of proline as microenvironmental stress substrate. *J Nutr*, 138(10), 2008S-2015S (2008)

13. A. Yoshida, A. Rzhetsky, L. C. Hsu and C. Chang: Human aldehyde dehydrogenase gene family. *Eur. J. Biochem.*, 251(3), 549-57 (1998)

14. K. A. Yoon, Y. Nakamura and H. Arakawa: Identification of ALDH4 as a p53-inducible gene and its protective role in cellular stresses. *J. Hum. Genet.*, 49(3), 134-40 (2004)

15. C. M. Forte-McRobbie and R. Pietruszko: Purification and characterization of human liver "high Km" aldehyde dehydrogenase and its identification as glutamic gamma-semialdehyde dehydrogenase. *J. Biol. Chem.*, 261(5), 2154-63 (1986)

16. C. Forte-McRobbie and R. Pietruszko: Human glutamic-gamma-semialdehyde dehydrogenase. Kinetic mechanism. *Biochem. J.*, 261(3), 935-43 (1989)

17. T. A. White, N. Krishnan, D. F. Becker and J. J. Tanner: Structure and kinetics of monofunctional proline dehydrogenase from *Thermus thermophilus*. *J. Biol. Chem.*, 282(19), 14316-27 (2007)

18. N. Krishnan and D. F. Becker: Characterization of a bifunctional PutA homologue from *Bradyrhizobium japonicum* and identification of an active site residue that modulates proline reduction of the flavin adenine dinucleotide cofactor. *Biochemistry*, 44(25), 9130-9 (2005)

19. J. P. Schuermann, T. A. White, D. Srivastava, D. B. Karr and J. J. Tanner: Three crystal forms of the bifunctional enzyme proline utilization A (PutA) from

*Bradyrhizobium japonicum*. *Acta Cryst.*, F64(Pt 10), 949-53 (2008)

20. P. F. Straub, P. H. Reynolds, S. Althomsons, V. Mett, Y. Zhu, G. Shearer and D. H. Kohl: Isolation, DNA sequence analysis, and mutagenesis of a proline dehydrogenase gene (putA) from *Bradyrhizobium japonicum*. *Appl. Environ. Microbiol.*, 62(1), 221-9 (1996)

21. N. Krishnan and D. F. Becker: Oxygen Reactivity of PutA from *Helicobacter* Species and Proline-Linked Oxidative Stress. *J. Bacteriol.*, 188(4), 1227-35 (2006)

22. N. Krishnan, A. R. Doster, G. E. Duhamel and D. F. Becker: Characterization of a *Helicobacter hepaticus* putA mutant strain in host colonization and oxidative stress. *Infect. Immun.*, 76(7), 3037-44 (2008)

23. K. Nakajima, S. Inatsu, T. Mizote, Y. Nagata, K. Aoyama, Y. Fukuda and K. Nagata: Possible involvement of put A gene in *Helicobacter pylori* colonization in the stomach and motility. *Biomed. Res.*, 29(1), 9-18 (2008)

24. E. D. Brown and J. M. Wood: Redesignated purification yields a fully functional PutA protein dimer from *Escherichia coli*. *J. Biol. Chem.*, 267(18), 13086-92 (1992)

25. R. Menzel and J. Roth: Regulation of genes for Proline Utilization in *Salmonella typhimurium*: Autogenous Repression by the putA gene Product. *J. Mol. Biol.*, 148, 21-44 (1981)

26. P. Ostrovsky de Spicer, K. O'Brien and S. Maloy: Regulation of proline utilization in *Salmonella typhimurium*: a membrane-associated dehydrogenase binds DNA *in vitro*. *J Bacteriol*, 173(1), 211-9 (1991)

27. S. Vilchez, M. Manzanera and J. L. Ramos: Control of expression of divergent *Pseudomonas putida* put promoters for proline catabolism. *Appl. Environ. Microbiol.*, 66(12), 5221-5 (2000)

28. R. Menzel and J. Roth: Enzymatic properties of the purified putA protein from *Salmonella typhimurium*. *J. Biol. Chem.*, 256(18), 9762-6 (1981)

29. P. Ostrovsky de Spicer and S. Maloy: PutA protein, a membrane-associated flavin dehydrogenase, acts as a redox-dependent transcriptional regulator. *Proc. Natl. Acad. Sci. USA*, 90(9), 4295-8 (1993)

30. M. W. Surber and S. Maloy: The PutA protein of *Salmonella typhimurium* catalyzes the two steps of proline degradation via a leaky channel. *Arch. Biochem. Biophys.*, 354(2), 281-287 (1998)

31. M. W. Surber and S. Maloy: Regulation of Flavin Dehydrogenase Compartmentalization: Requirements for PutA-Membrane Association in *Salmonella typhimurium*. *Biochim. Biophys. Acta*, 1421, 5-18 (1999)



## Unique features of puta

32. J. M. Wood and D. Zadworny: Amplification of the *put* Genes and Identification of the *put* Gene Products in *Escherichia coli* K12. *Can. J. Biochem.*, 58, 787-796 (1980)
33. J. M. Wood: Genetics of L-proline Utilization in *Escherichia coli*. *J. Bacteriol.*, 146, 895-901 (1981)
34. J. L. Abrahamson, L. G. Baker, J. T. Stephenson and J. M. Wood: Proline dehydrogenase from *Escherichia coli* K12. Properties of the membrane-associated enzyme. *Eur. J. Biochem.*, 134(1), 77-82 (1983)
35. S. Graham, J. T. Stephenson and J. M. Wood: Proline Dehydrogenase from *Escherichia coli* K12, Reconstitution of Functional Membrane Association. *J. Biol. Chem.*, 259(4), 2656-2661 (1984)
36. J. M. Wood: Membrane association of proline dehydrogenase in *Escherichia coli* is redox dependent. *Proc. Natl. Acad. Sci. USA*, 84(2), 373-377 (1987)
37. E. D. Brown and J. M. Wood: Conformational change and membrane association of the PutA protein are coincident with reduction of its FAD cofactor by proline. *J. Biol. Chem.*, 268(12), 8972-9 (1993)
38. M. Ling, S. W. Allen and J. M. Wood: Sequence analysis identifies the proline dehydrogenase and delta 1-pyrroline-5-carboxylate dehydrogenase domains of the multifunctional *Escherichia coli* PutA protein. *J. Mol. Biol.*, 243(5), 950-6 (1994)
39. D. F. Becker and E. A. Thomas: Redox properties of the PutA protein from *Escherichia coli* and the influence of the flavin redox state on PutA-DNA interactions. *Biochemistry*, 40(15), 4714-21 (2001)
40. M. P. Vinod, P. Bellur and D. F. Becker: Electrochemical and functional characterization of the proline dehydrogenase domain of the PutA flavoprotein from *Escherichia coli*. *Biochemistry*, 41, 6525-6532 (2002)
41. W. Zhu, Y. Gincherman, P. Docherty, C. D. Spilling and D. F. Becker: Effects of proline analog binding on the spectroscopic and redox properties of PutA. *Arch. Biochem. Biophys.*, 408(1), 131-6 (2002)
42. W. Zhu and D. F. Becker: Flavin redox state triggers conformational changes in the PutA protein from *Escherichia coli*. *Biochemistry*, 42(18), 5469-77 (2003)
43. B. A. Baban, M. P. Vinod, J. J. Tanner and D. F. Becker: Probing a hydrogen bond pair and the FAD redox properties in the proline dehydrogenase domain of *Escherichia coli* PutA. *Biochim. Biophys. Acta*, 1701(1-2), 49-59 (2004)
44. W. Zhang, Y. Zhou and D. F. Becker: Regulation of PutA-membrane associations by flavin adenine dinucleotide reduction. *Biochemistry*, 43(41), 13165-74 (2004)
45. W. Zhu and D. F. Becker: Exploring the proline-dependent conformational change in the multifunctional PutA flavoprotein by tryptophan fluorescence spectroscopy. *Biochemistry*, 44(37), 12297-306 (2005)
46. W. Zhang, M. Zhang, W. Zhu, Y. Zhou, S. Wanduragala, D. Rewinkel, J. J. Tanner and D. F. Becker: Redox-induced changes in flavin structure and roles of flavin N(5) and the ribityl 2'-OH group in regulating PutA--membrane binding. *Biochemistry*, 46(2), 483-91 (2007)
47. B. W. Arentson, N. Sanyal and D. F. Becker: Substrate channeling in proline metabolism. *Front. Biosci.* (2011)
48. K. S. Anderson: Fundamental mechanisms of substrate channeling. *Methods Enzymol.*, 308, 111-45 (1999)
49. E. W. Miles, S. Rhee and D. R. Davies: The molecular basis of substrate channeling. *J. Biol. Chem.*, 274(18), 12193-6 (1999)
50. X. Huang, H. M. Holden and F. M. Raushel: Channeling of substrates and intermediates in enzyme-catalyzed reactions. *Annu. Rev. Biochem.*, 70, 149-80 (2001)
51. E. M. Marcotte, M. Pellegrini, H. L. Ng, D. W. Rice, T. O. Yeates and D. Eisenberg: Detecting protein function and protein-protein interactions from genome sequences. *Science*, 285(5428), 751-3 (1999)
52. T. A. White and J. J. Tanner: Cloning, purification and crystallization of *Thermus thermophilus* proline dehydrogenase. *Acta Cryst.*, F61(Pt 8), 737-739 (2005)
53. T. A. White, W. H. Johnson, Jr., C. P. Whitman and J. J. Tanner: Structural basis for the inactivation of *Thermus thermophilus* proline dehydrogenase by N-propargylglycine. *Biochemistry*, 47(20), 5573-80 (2008)
54. E. Inagaki, N. Ohshima, H. Takahashi, C. Kuroishi, S. Yokoyama and T. H. Tahirov: Crystal structure of *Thermus thermophilus* Delta1-pyrroline-5-carboxylate dehydrogenase. *J. Mol. Biol.*, 362(3), 490-501 (2006)
55. E. Inagaki, N. Ohshima, K. Sakamoto, N. D. Babayeva, H. Kato, S. Yokoyama and T. H. Tahirov: New insights into the binding mode of coenzymes: structure of *Thermus thermophilus* [Delta]1-pyrroline-5-carboxylate dehydrogenase complexed with NADP<sup>+</sup>. *Acta Cryst.*, F63, 462-465 (2007)
56. J. D. Larson, J. L. Jenkins, J. P. Schuermann, Y. Zhou, D. F. Becker and J. J. Tanner: Crystal structures of the DNA-binding domain of *Escherichia coli* proline utilization A flavoprotein and analysis of the role of Lys9 in DNA recognition. *Protein Sci.*, 15, 1-12 (2006)
57. Y. Zhou, J. D. Larson, C. A. Bottoms, E. C. Arturo, M. T. Henzl, J. L. Jenkins, J. C. Nix, D. F. Becker and J. J. Tanner: Structural basis of the transcriptional regulation of

## Unique features of puta

the proline utilization regulon by multifunctional PutA. *J. Mol. Biol.*, 381(1), 174-88 (2008)

58. S. Halouska, Y. Zhou, D. F. Becker and R. Powers: Solution structure of the *Pseudomonas putida* protein PpPutA45 and its DNA complex. *Proteins*, 75(1), 12-27 (2009)

59. Y. H. Lee, S. Nadaraia, D. Gu, D. F. Becker and J. J. Tanner: Structure of the proline dehydrogenase domain of the multifunctional PutA flavoprotein. *Nat. Struct. Biol.*, 10(2), 109-114 (2003)

60. M. Zhang, T. A. White, J. P. Schuermann, B. A. Baban, D. F. Becker and J. J. Tanner: Structures of the *Escherichia coli* PutA proline dehydrogenase domain in complex with competitive inhibitors. *Biochemistry*, 43(39), 12539-48 (2004)

61. E. L. Ostrander, J. D. Larson, J. P. Schuermann and J. J. Tanner: A conserved active site tyrosine residue of proline dehydrogenase helps enforce the preference for proline over hydroxyproline as the substrate. *Biochemistry*, 48(5), 951-9 (2009)

62. D. Srivastava, W. Zhu, W. H. Johnson, Jr., C. P. Whitman, D. F. Becker and J. J. Tanner: The structure of the proline utilization A proline dehydrogenase domain inactivated by N-propargylglycine provides insight into conformational changes induced by substrate binding and flavin reduction. *Biochemistry*, 49(3), 560-569 (2010)

63. D. Srivastava, J. P. Schuermann, T. A. White, N. Krishnan, N. Sanyal, G. L. Hura, A. Tan, M. T. Henzl, D. F. Becker and J. J. Tanner: Crystal structure of the bifunctional proline utilization A flavoenzyme from *Bradyrhizobium japonicum*. *Proc. Natl. Acad. Sci. U. S. A.*, 107(7), 2878-83 (2010)

64. M. J. Bennett, M. P. Schlunegger and D. Eisenberg: 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci.*, 4(12), 2455-68 (1995)

65. Z. J. Liu, Y. J. Sun, J. Rose, Y. J. Chung, C. D. Hsiao, W. R. Chang, I. Kuo, J. Perozich, R. Lindahl, J. Hempel and B. C. Wang: The first structure of an aldehyde dehydrogenase reveals novel interactions between NAD and the Rossmann fold. *Nat. Struct. Biol.*, 4(4), 317-26 (1997)

66. C. A. Bottoms, P. E. Smith and J. J. Tanner: A structurally conserved water molecule in Rossmann dinucleotide-binding domains. *Protein Sci.*, 11(9), 2125-37 (2002)

67. D. Gu, Y. Zhou, V. Kallhoff, B. Baban, J. J. Tanner and D. F. Becker: Identification and characterization of the DNA-binding domain of the multifunctional PutA flavoenzyme. *J. Biol. Chem.*, 279(30), 31171-6 (2004)

68. E. R. Schreiter and C. L. Drennan: Ribbon-helix-helix transcription factors: variations on a theme. *Nat. Rev. Microbiol.*, 5(9), 710-20 (2007)

69. H. Watanabe and J. W. Hastings: Specificities and properties of three reduced pyridine nucleotide-flavin mononucleotide reductases coupling to bacterial luciferase. *Mol. Cell. Biochem.*, 44, 181-187 (1982)

70. J. Soding: Protein homology detection by HMM-HMM comparison. *Bioinformatics*, 21(7), 951-60 (2005)

71. K. Arnold, L. Bordoli, J. Kopp and T. Schwede: The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, 22(2), 195-201 (2006)

72. W. L. DeLano: The PyMOL User's Manual. DeLano Scientific, Palo Alto, CA, USA (2002)

73. R. Chenna, H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins and J. D. Thompson: Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.*, 31(13), 3497-500 (2003)

74. P. Gouet, E. Courcelle, D. I. Stuart and F. Metoz: ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics*, 15(4), 305-8 (1999)

**Abbreviations:** PutA: proline utilization A, NAD<sup>+</sup>: nicotinamide adenine dinucleotide, PRODH: proline dehydrogenase, P5C: pyrroline-5-carboxylate, P5CDH: pyrroline-5-carboxylate dehydrogenase, FAD: flavin adenine dinucleotide, GSA: glutamate semialdehyde, BjPutA: PutA from *Bradyrhizobium japonicum*, EcPutA: PutA from *Escherichia coli*, TtPRODH: monofunctional PRODH from *Thermus thermophilus*, TtP5CDH: monofunctional P5CDH from *Thermus thermophilus*, MSA: multiple sequence alignment, RHH: ribbon-helix-helix, CTD: C-terminal domain of long branch 1 PutAs

**Key Words:** Proline Utilization A, PutA, Substrate Channeling, Proline Catabolism, Proline Metabolism, Proline Dehydrogenase, Pyrroline-5-Carboxylate Dehydrogenase, Pyrroline-5-Carboxylate, Glutamate Semialdehyde, Domain Repeat, Aldehyde Dehydrogenase, Flavoenzyme, Remote Homology Detection, Review

**Send correspondence to:** John J. Tanner, Department of Chemistry, University of Missouri-Columbia, Columbia, MO 65211, USA. Tel: 573-884-1280, Fax: 573-882-2754, E-mail: tannerjj@missouri.edu

<http://www.bioscience.org/current/vol17.htm>