Studies on cyanobacterial protein PipY shed light on structure, potential functions and vitamin B_6-dependent epilepsy

Lorena Tremiño[1], Alicia Forcada-Nadal[1,2], Asunción Contreras[2] and Vicente Rubio[1*]

[1]Instituto de Biomedicina de Valencia (IBV-CSIC), and CIBER de Enfermedades Raras (CIBERER-ISCIII), Valencia, Spain; [2]Departamento de Fisiología, Genética y Microbiología, Universidad de Alicante, Alicante, Spain

*Send correspondence to

Dr. Vicente Rubio

Instituto de Biomedicina de Valencia

Jaime Roig 11, Valencia-46010, Spain

E-mail: rubio@ibv.csic.es

Phone: +34 963391772. Fax: +34 963690800

Abstract

The *Synechococcus elongatus COG0325* gene, *pipY*, functionally interacts with nitrogen regulatory gene *pipX*. As a first step towards molecular understanding of these interactions PipY was characterized. This 221-residue protein is monomeric and hosts pyridoxal phosphate (PLP), binding it with limited affinity and losing it upon D-cycloserine incubation. PipY crystal structures with and without PLP revealed a single-domain monomer folded as the TIM barrel of type III-fold PLP enzymes, with PLP highly exposed, fitting a PipY role in PLP homeostasis. The mobile PLP phosphate-anchoring C-terminal helix might act as trigger for PLP exchange. Exploiting the universality of *COG0325* functions, we used PipY in site-directed mutagenesis studies to shed light on disease causation by epilepsy-associated mutations in the human *COG0325* gene, *PROSC*.

Keywords. *Synechococcus elongatus* PCC7942, COG0325, pyridoxal phosphate proteins.

Abbreviations. CD, circular dichroism; DCS, D-cycloserine; DTT, dithiothreitol; ODC, ornithine decarboxylase; PDB, Protein Data Bank; PEG, polyethylene glycol; PLP, pyridoxal 5'-phosphate; r.m.s.d., root-mean-square deviation

Cyanobacteria, the chloroplasts ancestors [1], are very ancient, widespread and abundant [2] photosynthetic bacteria that play key roles in global oxygen, CO_2 and nitrogen fluxes [3]. Nitrogen incorporation is highly regulated in cyanobacteria [4]. The small protein PipX [5,6] is a key player, as highlighted by transcriptomic studies in *Synechococcus elongatus* PCC7942 (from now on *S. elongatus*) [7]. PipX regulation, mediated by the carbon/energy/nitrogen signalling protein PII and the gene expression regulator NtcA [5,8,9], also involves newly discovered partners as PlmA [10].

Very recently, the synteny in cyanobacteria of the mutually adjacent *pipX* and *pipY* genes led to the identification in *S.elongatus* of functional interactions between these two genes [11]. *pipY* is a member of the highly widespread [12] and intriguing *COG0325* gene family [11] that includes *Escherichia coli yggS*, a gene that when inactivated causes a phenotype of pyridoxine toxicity and pleiotropic changes related to amino acid-related metabolism [12-14]. Inactivation of *pipY* in *S.elongatus* [11] also enhanced pyridoxine toxicity and led to increased sensitivity to antibiotics that target pyridoxal phosphate (PLP) such as D-cycloserine (DCS) [15]; these two traits were influenced by various amino acids [11]. Pleiotropic changes were also observed in null mutants of the *COG0325* genes *ylmE* of *Bacillus subtilis* [16] and *PROSC* of humans [17,18]. The deficiency of the human gene was reported to cause vitamin B_6-dependent epilepsy [17,18]. The pleiotropy of the changes associated with *COG0325* mutations is best explained if the *COG0325* products, which are PLP proteins [19], rather than being enzymes play key roles in PLP homeostasis. These roles remain uncharacterized molecularly.

Striving towards delineation of the roles of PipY, we decided to characterize physically and structurally PipY from *S. elongatus*. Such characterization appeared pertinent, given the extensive *in vivo* functional data collected for its corresponding

gene [11]. Furthermore, there was only one published analysis of the structure of another COG0325 protein, YBL036C from yeast [19], an organism that is phylogenetically distant from cyanobacteria. The YBL036C structure had been determined as a part of a structural genomics effort, and the biological *in vivo* role of YBL036C had not been specifically explored. In addition, the structural knowledge of YBL036C did not include the corresponding structure without PLP. However, the changes associated with the presence or absence of PLP may be highly relevant for a role of COG0325 proteins in PLP delivery to other targets. Although crystal structures of three bacterial COG0325 proteins have been deposited in the Protein Data Bank (PDB) (Supplementary Table S1), none of these structures has been analysed, justifying the need for thorough structural analysis of a bacterial COG0325 protein in the absence and the presence of PLP. We carry out here such structural analysis on *S. elongatus* PipY. The detailed structural information gained in the present studies by X-ray crystallography of PLP-containing and PLP-free PipY makes of PipY a true structural paradigm for COG0325 proteins. In addition, we have characterized some PipY traits in solution, including its mass, its spectroscopic properties and both its ability to bind PLP and its sensitivity to the PLP-targeting antibiotic DCS. The toxicity of this antibiotic for *S. elongatus* was previously shown to be increased by *pipY* inactivation [11].

We also get insight here on the effects of two PROSC mutations, Pro87Leu and Arg241Gln, reported in patients with vitamin B_6-dependent epilepsy [17,18]. Puzzlingly, the clinical severity inferred for these mutations did not parallel the ability of the corresponding mutant *PROSC* gene forms to complement an *E.coli yggS*$^-$ mutant for its phenotype of increased sensitivity to pyridoxine toxicity [17]. To try to explain these puzzling observations, we have introduced into PipY the mutations corresponding to Pro87Leu and Arg241Gln of PROSC, studying relevant characteristics of the mutant

PipY forms. This approach has exploited the apparent universality of COG0325 functions (for which the recent studies on *S. elongatus* provided strong confirmation [11]) and the sequence conservation between *S. elongatus* PipY and human PROSC (Supplementary Fig. S1). In this way we provide valuable hints on the respective ability and lack of ability of the Pro87Leu and Arg241Gln PROSC mutants to override the mutant phenotype of the *yggS*$^{-}$ *E.coli* mutant strain [17], as well as on the possible mechanism of disease causation by the Arg241Gln PROSC mutation.

MATERIALS AND METHODS

Production of wild-type and mutant PipY

The *S.elongatus pipY* open reading frame (gene *Synpcc7942_2060*, http://genome.microbedb.jp/cyanobase/GCA_000012525.1/genes/Synpcc7942_2060), PCR-amplified from genomic DNA using Deep Vent DNA polymerase (from New England Biolabs) and the primers 5'-GTCAGAGGCACATATGGCCCAAATTGC-3' and 5'-GGCCAAGACTCGAGCGATCGCGGC-3', was digested with *Nde*I/*Xho*I and ligated into the similarly digested pET-22b+ plasmid (from Novagen). The resulting pET-22b(+)-2060 plasmid, isolated from *E.coli* DH5α, was used for site-directed mutagenesis using the Quickchange approach (Stratagene, La Jolla, CA) and the following forward and reverse mutagenic primers: for the Pro63Leu mutation, 5'ACGGATCTGCTGGATTTGACTTGG3' and 5'AAGTCAAATCCAGCAGATCCGTCAG3'; and for the Arg210Gln mutation, 5'GCAACTTGGATTCAAGTCGGAACC3' and 5'TTCCGACTTGAATCCAAGTTGCC3'. The presence of the correct sequences and of the desired mutations was verified by Sanger sequencing of the plasmids.

To produce PipY with a C-terminal -GluHis$_6$ tag, the pET-22b(+)-2060 plasmid was transformed into *E.coli* BL21(DE3) (from Novagen), growing the transformed cells

in 1-L liquid LB-ampicillin (0.1 mg/ml) at 37ºC, to ~0.8 OD^{600}, followed by 3-hour induction with 1 mM isopropyl β-D-1-thiogalactopyranoside. All subsequent steps were at 4ºC. After centrifugation, the cell pellet, suspended in 30 ml of 25 mM HEPES pH 7.5/0.4 M NaCl/1 mM dithiothreitol (DTT)/1 mM phenyl methyl sulphonyl fluoride/20 mM imidazole, was sonicated to disrupt the cells, the suspension was centrifuged, and PipY was purified from the supernatant by Ni-affinity chromatography and imidazole gradient elution (His-Trap HP 1-ml column in an Åkta-FPLC, both from GE Healthcare). The resulting preparation was concentrated to ~10 mg/ml (Bradford assay [20] using bovine serum albumin as standard) and placed in 20 mM HEPES pH 7.5/16 mM NaCl/1mM DTT (storage buffer) by centrifugal ultrafiltration (Amicon Ultra, from Millipore). The Pro63Leu and Arg210Gln mutants were purified in the same way.

PipY crystallization, X-ray diffraction and structure determination

Crystals were grown by vapour-diffusion (21°C) in 0.8 µl sitting drops made of 1/1 mixtures of protein solution and crystallization solution. The protein solution contained 12.6 mg/ml PipY in storage buffer supplemented with either 10 mM DCS (PipY-Apo) or 10 mM D-Ala (PipY-PLP). The crystallization solution was, for PipY-Apo, 0.1 M Tris-HCl pH 8.5/0.2 M $MgCl_2$/25% w/v polyethylene glycol (PEG) 3350; and for PipY-PLP, 0.1 M Hepes pH 7.5/0.2 M calcium acetate/18% w/v PEG 8000. Crystals were flash-frozen in liquid nitrogen using as cryobuffers the crystallization solutions (enriched with 5% extra PEG 3350 for PipY-Apo). Synchrotron (Table 1) X-ray diffraction was at -170°C, using Pilatus3 6M detectors. Datasets at 1.90/1.93 Å resolution (PipY-Apo/PipY-PLP) were processed with XDS [21], scaled with AIMLESS [22], (Table 1) and phased by molecular replacement with Phaser [23], using for the PipY-Apo dataset a poly-alanine model of YggS from *E.coli* (PDB entry 1W8G)

lacking residues 1-20; and utilizing for PipY-PLP the A subunit of PipY-Apo once it was traced and refined. Refmac5 [24] was used for optimizing the positioning of the two PipY chains present in the asymmetric units of both crystals (Table 1), using rigid body refinement. It was also used for model refinement, alternating iteratively cycles of restrained refinement and of manual model building with Coot [25]. This last program was used for incorporating PLP (taken from the PDB) to the PipY-PLP model. Isotropic B factors and Translation/Libration/Screw (TLS) were used in the last refinement steps [26], choosing the TLS groups with the TLSMD server (http://skuld.bmsc.washington.edu/~tlsmd) [27]. All diffraction data were used throughout the refinement processes, except for the randomly selected 5% of the data used for R_{free} calculation. The stereochemistry of the model was checked and improved with PDB_REDO [28]. Analysis with Rampage [29] of the geometry of the main chain torsion angles revealed excellent values for both models. Graphical representations of the structures were generated with PyMOL (Pymol Molecular Graphics System, Version 1.6, Schrödinger, LLC, www.pymol.org).

Other techniques

Size exclusion chromatography was performed on a Superdex 200 10/300 column fitted on an ÄKTA FPLC (both from GE Healthcare), at 0.5 ml/min flow rate, using a solution of 10 mM Na phosphate pH 7.0/0.4 M NaCl/5 mM $MgCl_2$/5 mM EDTA/0.2 mM DTT, and monitoring the absorbance of the effluent at 280 nm. The mass of PipY was estimated from a semi-logarithmic plot prepared with protein standards of known masses.

Spectrophotometric determinations were carried out with a NanoDrop 1000 spectrophotometer (Thermo Scientific) on 2 µl samples. All reagents and PipY were in

storage buffer. Optical absorption values are given for 1-cm light path. Incubations of PipY with PLP or DCS were performed for 15 min in 50 µl solutions at 25ºC, with controls lacking PipY carried in parallel and subtracted. For incubation with DCS, PipY had been incubated previously with 0.34 mM PLP and then freed from unbound PLP by centrifugal gel filtration [30]. Plots of the absorbance at 425 nm of PipY versus PLP or DCS concentrations were fitted to hyperbolae using GraphPadPrism (GraphPad Software, San Diego, CA).

CD spectra in the far-UV region (195–250 nm) were obtained at 21 °C with a Jasco J-810 spectropolarimeter in 0.2 ml of 50 mM Na phosphate pH 7.5 and 0.25 mg/ml protein in a 0.1 cm-path cell. Each spectrum was the average of five scans.

The thermal stabilities of wild-type and mutant PipY proteins were investigated by the thermofluor approach [31] as previously reported [32], using 0.25 mg/ml PipY in a solution of 22 mM Hepes pH 7.5 with the indicated additions.

RESULTS

The *pipY* protein product

PipY was produced in *E. coli* from an expression plasmid and was obtained in pure form (Fig. 1A, inset). Similarly to the *E. coli* and yeast *COG0325* gene products, YggS and YBL036C, [12,13,19], PipY was monomeric, judged from its behaviour in size exclusion chromatography assays (Fig. 1A). As expected for a protein forming a protonated Schiff base with PLP [33], pure PipY was yellow due to a broad optical absorption peak having its maximum absorption at ~425 nm (Fig. 1B, black continuous line), a longer wavelength than the maximum for free PLP (Fig. 1B, dotted line). PipY incubation with excess PLP, followed by removal of the PLP excess by centrifugal gel filtration [30] increased the size of the 425-nm peak (Fig. 1B, dashed line), as if the pure

protein was incompletely saturated with PLP, perhaps because of partial dissociation of PLP during protein isolation and storage. In fact, prolonged PipY storage (several months) decreased the height of the 425-nm peak (Fig. 1B, grey line), as expected for PLP dissociation with time. Titration with PLP of this aged protein restored maximal absorbance with a K_D value for PLP of ~30 μM (Fig. 1C). As is typical for protein-bound PLP [34], the 425-nm peak disappeared with the addition of DCS (Fig. 1D), while absorbance at shorter wavelength (around 335 nm) increased. Titration of the decrease in absorbance at 425 nm versus DCS concentrations (Figs. 1D and 1E) yielded a K_D for DCS in the mM range, two to three orders of magnitude higher than K_D values for PLP-containing proteins that bind DCS as an analogue of an amino acid substrate (the case for alanine racemase [15]). This low affinity for DCS excludes specific binding of this antibiotic to PipY, in agreement with the reported lack of alanine racemase activity of PipY [11] and of the *COG0325* gene products of other species [13].

Overall characteristics of crystalline PipY

PipY crystal structures were determined at ~1.9 Å resolution with and without PLP (PipY-PLP and PipY-Apo, respectively; Table 1). No electron density that could correspond to D-Ala or DCS was observed in the crystals, despite the fact that these compounds were present at 10 mM concentrations in the respective crystallization solutions (see Materials and Methods). Each crystal had two protein chains in the asymmetric unit. These chains were related by twofold non-crystallographic symmetry. The small contact surfaces (~5% of the exposed surface) and the predictions of the PISA server (www.ebi.ac.uk/pdbe/pisa/) did not support a homodimeric nature of PipY in solution, agreeing with the size exclusion chromatography results that showed that PipY is a monomer (Fig. 1A).

Except for the initial and final 2-3 residues and the C-terminal His_6-tag, the entire polypeptide chains were visible in the crystal structures. In addition, one PLP molecule was found binding per chain in the PipY-PLP crystal. Both monomers in each asymmetric unit were essentially identical (r.m.s.d. <0.6 Å, Supplementary Table S1), with somewhat lower identity when PLP-containing and PLP-lacking PipY molecules were compared (r.m.s.d. 0.62-0.98 Å, Supplementary Table S1). There was high similarity with the reported structure [19] of the yeast *COG0325*-encoded protein YBL036C or with the structures of another three *COG0325* products from other bacteria that are deposited in the PDB (Supplementary Table S1; r.m.s.d. 1.29-1.75 Å; and Supplementary Fig S2A).

The structure of PipY

PipY has (Fig. 2A) the characteristic type-III fold of PLP-dependent enzymes [19,35], a modified TIM barrel [36] having an extra N-terminal α-helix preceding the first of the eight β/α repeating units and lacking the helix of the last repeat. Unlike the paradigms of this fold, alanine racemase, ornithine decarboxylase and broad specificity amino acid racemase [37-39], all of them composed of two-domain subunits forming homodimers, PipY is a single-domain monomer consisting exclusively of the barrel domain (Supplementary Fig. S2A). This is also the case for the highly similar structures previously reported for YBL036C [19] or deposited in the PDB for another three bacterial COG0325 proteins (Supplementary Table S1). All these COG0325 protein structures have a one-turn imperfect α helical extension of the C-terminal β-strand (helix 9, Fig. 2A and Supplementary Fig. S2A). This extension binds the phosphate of PLP (see below). However, PipY differs from all the other *COG0325*-encoded protein products of known structure in the length of the first α-helix (Supplementary Fig. S2),

which in PipY spans less than three full turns, while in all these other proteins is much longer on both the N- and C-terminal ends of this helix, which spans up to 6 turns (Fig. 2B). Perhaps this helix, which is lacking in canonical TIM barrels, could be involved in interactions with other proteins, with specific requirements for COG0325 proteins of different species.

PLP binding

PLP sits towards the centre of the barrel, on the C-edge side of the parallel β-sheet (Fig. 2A). It is found in the same site and orientation as in yeast YBL036C and in the deposited structures of COG0325 proteins. This location is the same that is found in the PLP enzymes presenting the PLP type III fold (Supplementary Figs. S2A-C). However, in PipY and the other COG0325 proteins of known structure the PLP is much more accessible than in the type III fold PLP enzymes [37-39]. In these dimeric enzymes the cofactor is shielded by the two-domain architecture and by the other subunit, being accessible only through a tunnel (not shown). The exposed PLP of PipY and of other COG0325 proteins should facilitate PLP exchange with other molecular partners. The PLP forms a Schiff base with Lys26 (Fig. 2C; and Supplementary Fig. S2C). Its phosphate is anchored via six hydrogen bonds on helix 9 (residues Gly212 and Thr213) and on the β7-α8 turn (Ser195 and, indirectly through water, Ser196). The cofactor makes other non-covalent interactions with six of the eight β/α repeats of the modified TIM barrel. Only β/α repeats 4 and 5 do not contact the PLP. The pyridoxal ring is partially sandwiched between Leu70 from β3 and Met194 from the β7-α8 turn, lying through its border on Arg210 from β8, with which an NH..N hydrogen bond is made. The exposed part of the pyridoxal ring makes contacts through its C2-methyl group with Met152 from β6, hydrogen bonding its C3-OH with Asn47 from β2, and

contacting Val24 of β1 via its C4 carbaldehyde and C5 hydroxymethyl groups. Except for the indirect interaction with Ser196 (Fig. 2C), all PLP interactions found in PipY are present in the other COG0325 structures, either with identical or conservatively replaced residues (Supplementary Fig.S2A).

Changes associated with PLP binding in the context of a role of PipY in PLP homeostasis

Our determination of the structures of PLP-bound and PLP-free forms of PipY allowed characterization of the changes associated with PLP binding (Fig. 2D). These changes are not large, as shown by the modest r.m.s.d. increase when comparing PipY-PLP and PipY-Apo (Supplementary Table S1). The largest changes are experienced by helix 9 (Fig. 2D), which is displaced towards the PLP phosphate that anchors on it. This helix drags in its movement the residues following it, which are displaced by up to 4.5 Å. Helices 1 and 2 are also dragged towards the phosphate, while Ser195 and the β7-α8 loop are displaced by the phosphate, with concomitant small displacement of helix 8, which drags with it helix 7 and the β6-α7 loop, and even helix 6 and the β5-α6 loop.

The movement of helix 9 might be the best candidate for triggering the PLP exchange required for the proposed role of PipY in PLP homeostasis [11-14,17,18]. Due to its exposed nature helix 9 could interact with other macromolecular players, acting as a trigger for PLP exchange, since a change in the position of helix 9 could importantly weaken PLP binding to PipY, given the PLP-anchoring role of this helix.

Exploiting PipY for investigating the disease-causing mechanisms of PROSC mutations

Recently, vitamin B_6-dependent epilepsy was reported in association with null mutations in the human *COG0325* gene *PROSC* [17,18]. However, some patients carried the missense PROSC mutations Pro87Leu and/or Arg241Gln, two mutations for which a loss-of-function effect was at best hypothetical [18]. In addition, there was no correspondence between the degree of clinical severity attributed to each of these mutations on the basis of the observations on the patients [18] and the ability of each mutant form to complement the increased pyridoxine toxicity phenotype of the *yggS*$^-$ *E.coli* mutant [17]. Since the PROSC Pro87 and Arg241 residues are conserved in *S.elongatus* PipY (respective PipY counterparts, Pro63 and Arg210, Supplementary Fig. S1), we investigated the effects of the corresponding PipY mutations on recombinantly produced PipY as an approximation to the effects of PROSC mutations. Both mutant proteins could be purified (Fig. 3A, bottom panel) and were monomeric (determined by size exclusion chromatography, data not shown). The Pro63Leu mutation did not negatively affect any of the parameters assessed: yield (Fig. 3A, top panel), thermal stability (Fig. 3B), proper fold of the pure protein (estimated by CD, Fig. 3C) and PLP content (assessed by the absorption spectrum, Fig. 3D). The latter was to be expected, since Pro63 sits on the opposite face of the TIM barrel than PLP (Fig. 2A) and, therefore, the mutations of this residue should not directly hamper PLP binding. In contrast, the Arg210Gln mutant, also obtained in pure (Fig. 3A, bottom panel) and well-folded form (Fig. 3C) was produced in quite low yield (Fig. 3A, top panel), exhibited decreased thermal stability (nearly 5ºC decrease, Fig. 3B) and had little if any PLP, judged from its low optical absorption in the 425 nm region (Fig. 3D). This last observation fits the fact that the mutation affects the invariant arginine shown by the structure to run parallel to the buried border of the pyridoxal ring, with which it makes an NH..N hydrogen bond (Fig. 2A,C). Our findings explain the results of

complementation assays for the Pro87Leu and Arg241Gln PROSC mutants [17] and the disease-causality of the Arg241Gln mutation (see Discussion)

DISCUSSION

The complementation of *yggS*$^{-}$ *E.coli* by *COG0325* genes of other species [12,13,17] and the phenotypic similarities of this *E.coli* mutant and of the *S.elongatus pipY*$^{-}$ mutant [11] are strong indications that the function of *COG0325* genes is conserved across phyla. The alleviation by PLP or by vitamin B_6 supplementation of the alterations observed in these mutants [11-14], best illustrated by the vitamin B_6 dependence of the epilepsy due to PROSC deficiency [17,18], indicates that this conserved function is related to PLP, in line with the fact that COG0325 products are PLP-containing proteins. Yet, previous work [11-13] has not favoured an enzymatic role of COG0325 proteins. In our own work, we have failed to generate crystals of PipY bound to D-Ala or to DCS despite extensive crystallization attempts with these compounds. These two ligands have been found binding in crystals of the structurally similar alanine racemase [40]. Furthermore, as exemplified in the PipY structure, COG0325 proteins lack key constant and obligatory features of type III-fold PLP enzymes [35]. These missing traits include the dimeric character, the presence of a second domain and the existence of an invariant substrate-binding arginine (Arg129 of Ala racemase, Supplementary Figs. S2A,C) and of a negative charge that stabilizes this arginine [37-41]. Thus, COG0325 proteins are believed to be involved in PLP homeostasis rather than being catalysts [11-14,17,18].

A PLP-homeostatic role would be instrumental to prevent the toxicity of free B_6-vitamers [13,42, and see 17] and would imply that COG0325 proteins deliver PLP to the apoenzymes of PLP-dependent enzymes. Such role fits: i) the observations of

increased toxicity of pyridoxine with *yggS*$^-$ *E. coli* [12,13] and of pyridoxine and PLP-targeting antibiotics with *pipY*$^-$ *S. elongatus* [11]; ii) the interactions of these traits with the supply of amino acids to these bacterial mutants [11-13]; iii) the observation of pleiotropic amino acid metabolism changes in *COG0325* mutants [12-14,16-18]; iv) the synthetic lethality caused in *S. elongatus* by the *pipY*/*cysK* double mutation [11]; and v) particularly, the finding of amino acid levels in the cerebrospinal fluid of a patient with PROSC deficiency that are highly suggestive of multiple transaminase deficiency [18]. In line with a homeostatic role of PipY, in the structures of PipY (present data) and of other COG0325 proteins ([19], and see PDB files given in Supplementary Table S1) PLP is exposed as if it was poised for transfer. Furthermore, the relatively large changes in the orientation of the outwards-protruding helix 9 depending on the presence or absence of PLP, taken together with the key role of this helix in PLP anchoring, renders this helix an obvious candidate for being a trigger that could be operated by partners to weaken the interactions of COG0325 proteins with PLP, promoting PLP transfer.

The alternative possibility that COG0325 proteins, although not catalytic by themselves, could be subunits of hypothetical heterodimeric PLP enzymes in which the COG0325 subunit would act as a common PLP-containing module appears unlikely. The pleiotropic metabolic effects of the null *COG0325* gene mutations demand multiple alternative partners for the COG0325 protein. It would be unlikely not to detect any of these partners in yeast-two-hybrid assays using PipY as the bait, but, in fact, no partner was detected [11]. Furthermore, the increased amino acid levels in the cerebrospinal fluid of a PROSC-deficient patient [18] fitted multiple transaminase deficiency. However, transaminases do not utilize for binding PLP the modified TIM barrel fold used by COG0325 proteins [35]. Similarly, a role of COG0325 in the regulation of a specific PLP enzyme, as exemplified for ornithine decarboxylase with its

enzyme/antizyme/antizyme inhibitor system [43], cannot be excluded, but, again, appears unlikely, largely because it would not fit the multi-target pleiotropy reflected in the metabolic phenotypes of COG0325 mutants.

Strong support for a PLP homeostatic role of COG0325 proteins was provided by observations made in vitamin B_6-dependent epilepsy associated to PROSC mutations [17,18]. By using PipY as a model of PROSC, we attempted here to shed light on why the PROSC Pro87Leu mutant complemented and the Arg241Gln mutant did not complement the pyridoxine toxicity phenotype of an *E.coli yggS*$^-$ strain [17]. The abundant production in *E.coli* of PLP-carrying Pro63Leu PipY could explain the complementation observed with the PROSC Pro87Leu mutant [17]. In contrast, lack of complementation would be expected for the poorly produced and PLP-lacking Arg210Gln mutant of PipY, as observed for the corresponding human PROSC mutant Arg241Gln [17], in line with the prior observation that there was no complementation by YggS when it carried a mutation that made it unable to bind PLP [13]. Concerning disease-causation and clinical severity, the lack of bound PLP and the decreased stability and yield of PipY carrying the Arg210Gln mutation clearly support a disease-causing role of the corresponding Arg241Gln PROSC mutation, in agreement with the finding of this last mutation in patients with vitamin B_6-dependent epilepsy. However, the PROSC deficiency due to the Arg241Gln mutation may be partial if the mutant can still bind PLP at increased concentrations of this cofactor, what could explain the relatively mild clinical severity associated with this mutation [18]. In contrast to the Arg210Gln PipY mutant, the PipY Pro63Leu mutant that mirrors the PROSC Pro87Leu mutant appeared normal. Thus studies using the PROSC Pro87Leu mutant itself expressed in human cells might be necessary to test experimentally the disease-causation and clinical severity attributed to this mutation [18]. In any case, it is

interesting that three of the six missense PROSC mutations observed thus far in patients with vitamin B_6-dependent epilepsy replace proline by leucine or the reverse (Supplementary Fig. S2A) [17,18]. Given the structural uniqueness of proline as a protein amino acid, it is conceivable that these changes could lead to increased PROSC misfolding, raising the possibility of using pharmacological chaperones to try to alleviate the deleterious effects of these missense mutations.

PROTEIN DATABANK FILES

Protein DataBank accession codes 5NM8 and 5NLC

ACKNOWLEDGEMENTS

We thank MA Castells (UA), JL Llacer, S. Ventas (IBV-CSIC) and N. Gougeard (CIBERER-IBV) for technical support, A. García and M. Orzáez (CIPF-Valencia) for CD facilities, and the IBV-CSIC crystallographic facility for crystal growth. Supported by grants from the Valencian Government (PrometeoII/2014/029) and Spanish Governments (BFU2014-58229-P to V.R.; BFU2012-33364 and BFU2015-66360-P to A.C.; FPI contract to LT) and to EC FP7/2007-2013 BioStruct-X (grant agreement N°283570, proposal 7687) for synchrotron access. We are grateful to Diamond (Oxfordshire, UK) and Alba (Barcelona, Spain) synchrotrons for access and for staff support. The authors declare no conflict of interest.

REFERENCES

1. Margulis L, Schwartz KV (1982) Five kingdoms. An illustrated guide to the phyla of life on earth. WH Freeman & Company, San Francisco.
2. Garcia-Pichel F, Belnap J, Neuer S, Schanz F (2003) Estimates of global cyanobacterial biomass and its distribution. Algol Stud 109:213–217
3. Hamilton TL, Bryant DA, Macalady JL (2016) The role of biology in planetary evolution: cyanobacterial primary production in low-oxygen Proterozoic oceans. Environ Microbiol 18:325-340.
4. Luque I, Forchhammer K (2008) Nitrogen assimilation and C/N balance sensing. In The Cyanobacteria: Molecular Biology, Genomics and Evolution (Herrero A and Flores E, eds.). Caister Academic Press, Poole, UK .
5. Llácer JL, Espinosa J, Castells MA, Contreras A, Forchhammer K, Rubio V (2010) Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. Proc Natl Acad Sci U S A 107:15397-15402.
6. Forcada-Nadal A, Palomino-Schätzlein M, Neira JL, Pineda-Lucena A, Rubio V. (2017) The PipX protein, when not bound to its targets, has its signaling C-terminal helix in a flexed conformation. Biochemistry, in press.
7. Espinosa J, Rodríguez-Mateos F, Salinas P, Lanza VF, Dixon R, de la Cruz F, Contreras A. (2014) PipX, the coactivator of NtcA, is a global regulator in cyanobacteria. Proc Natl Acad Sci U S A 111:E2423-E2430.
8. Burillo S, Luque I, Fuentes I, Contreras A (2004) Interactions between the nitrogen signal transduction protein PII and N-acetyl glutamate kinase in organisms that perform oxygenic photosynthesis. J Bacteriol 186:3346-3354.

9. Espinosa J, Forchhammer K, Burillo S, Contreras A (2006) Interaction network in cyanobacterial nitrogen regulation: PipX, a protein that interacts in a 2-oxoglutarate dependent manner with PII and NtcA. Mol Microbiol 61:457-469.

10. Labella JI, Obrebska A, Espinosa J, Salinas P, Forcada-Nadal A, Tremiño L, Rubio V, Contreras A. (2016) Expanding the cyanobacterial nitrogen regulatory network: The GntR-like regulator PlmA interacts with the PII-PipX complex. Front Microbiol 7:1677.

11. Labella JI, Cantos R, Espinosa J, Forcada-Nadal A, Rubio V, Contreras A (2017) PipY, a member of the conserved COG0325 family of PLP-binding proteins, expands the cyanobacterial nitrogen regulatory network. Front Microbiol, *in press*.

12. Prunetti L, El Yacoubi B, Schiavon CR, Kirkpatrick E, Huang L, Bailly M, ElBadawi-Sidhu M, Harrison K, Gregory JF, Fiehn O, Hanson AD, de Crécy-Lagard V (2016) Evidence that COG0325 proteins are involved in PLP homeostasis. Microbiology 162:694-706.

13. Ito T, Iimori J, Takayama S, Moriyama A, Yamauchi A, Hemmi H, Yoshimura T (2013) Conserved pyridoxal protein that regulates Ile and Val metabolism. J Bacteriol 195:5439-5449.

14. Ito T, Yamauchi A, Hemmi H, Yoshimura T. (2016) Ophthalmic acid accumulation in an Escherichia coli mutant lacking the conserved pyridoxal 5'-phosphate-binding protein YggS. J Biosci Bioeng 122:689-693.

15. Neuhaus FC (1967) D-Cycloserine and O-carbamyl-D-serine. In Antibiotics I. Mechanism of action (Gottlieb D and Shaw PD, eds.) Springer Verlag Berlin Heidelberg, pp 40-83.

16. Kolodkin-Gal I, Romero D, Cao S, Clardy J, Kolter R, Losick R. (2010) D-Amino acids trigger biofilm disassembly. Science 328:627-629.

17. Darin N, Reid E, Prunetti L, Samuelsson L, Husain RA, Wilson M, El Yacoubi B, Footitt E, Chong WK, Wilson LC, Prunty H, Pope S, Heales S, Lascelles K, Champion M, Wassmer E, Veggiotti P, de Crécy-Lagard V, Mills PB, Clayton PT. (2016) Mutations in PROSC disrupt cellular pyridoxal phosphate homeostasis and cause vitamin-B_6-dependent epilepsy. Am J Hum Genet 99:1325-1337.

18. Plecko B, Zweier M, Begemann A, Mathis D, Schmitt B, Striano P, Baethmann M, Vari MS, Beccaria F, Zara F, Crowther LM, Joset P, Sticht H, Papuc SM, Rauch A (2017) Confirmation of mutations in PROSC as a novel cause of vitamin B_6-dependent epilepsy. J Med Genet pii: jmedgenet-2017-104521.

19. Eswaramoorthy S, Gerchman S, Graziano V, Kycia H, Studier FW, Swaminathan S. (2003) Structure of a yeast hypothetical protein selected by a structural genomics approach. Acta Crystallogr D Biol Crystallogr 59:127-135.

20. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. Anal Biochem 72:248–254.

21. Kabsch W (2010) XDS. Acta Crystallogr D Biol Crystallogr 66:125-132.

22. Evans PR and Murshudov GN (2013) How good are my data and what is the resolution? Acta Cryst D Biol Crystallogr 69:1204–1214.

23. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. (2007) Phaser crystallographic software. J Appl Crystallogr 40:658-674.

24. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr 67:355-367.

25. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486–501.

26. Winn MD, Murshudov GN, Papiz MZ (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. Methods Enzymol 374:300–321.

27. Painter J, Merritt EA (2006) TLSMD web server for the generation of multi-group TLS models. J Appl Crystallogr 39:109–111.

28. Joosten RP, Salzemann J, Bloch V, Stockinger H, Berglund AC, Blanchet C, Bongcam-Rudloff E, Combet C, Da Costa AL, Deleage G, Diarena M, Fabbretti R, Fettahi G, Flegel V, Gisel A, Kasam V, Kervinen T, Korpelainen E, Mattila K, Pagni M, Reichstadt M, Breton V, Tickle IJ, Vriend G (2010) PDB_REDO: automated re-refinement of X-ray structure models in the PDB. J Appl Crystallogr 42:376-384.

29. Lovell SC, Davis IW, Arendall WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Structure validation by Cα geometry: φ, ψ and Cβ deviation. Proteins 50:437–450.

30. Penefsky HS (1977) Reversible binding of Pi by beef heart mitochondrial adenosine triphosphatase. J Biol Chem 252:2891-2899.

31. Vedadi M, Niesen FH, Allali-Hassani A, Fedorov OY, Finerty PJ Jr, Wasney GA, Yeung R, Arrowsmith C, Ball LJ, Berglund H, Hui R, Marsden BD, Nordlund P, Sundstrom M, Weigelt J, Edwards AM (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. Proc Natl Acad Sci U S A 103:15835-15840.

32. Sancho-Vaello E, Marco-Marín C, Gougeard N, Fernández-Murga L, Rüfenacht V, Mustedanagic M, Rubio V, Häberle J (2016) Understanding N-acetyl-L-glutamate synthase deficiency: Mutational spectrum, impact of clinical mutations on enzyme functionality, and structural considerations. Hum Mutat 37:679-694.

33. Fasella P (1967) Pyridoxal phosphate. Annu Rev Biochem 36:185-210.

34. Cook SP, Galve-Roperh I, Martínez del Pozo A, Rodríguez-Crespo I (2002) Direct calcium binding results in activation of brain serine racemase. J Biol Chem 277:27782-27792.

35. Percudani R, Peracchi A. (2003) A genomic overview of pyridoxal-phosphate-dependent enzymes. EMBO Rep 4:850-854.

36. Banner DW, Bloomer AC, Petsko GA, Phillips DC , Pogson CI, Wilson IA, Corran PH, Furth AJ, Milman JD, Offord RE, Priddle JD, Waley SG. (1975) Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5Å resolution: using amino acid sequence data. Nature 255:609-614.

37. Shaw JP, Petsko GA, Ringe D (1997) Determination of the structure of alanine racemase from *Bacillus stearothermophilus* at 1.9-Å resolution. Biochemistry 36:1329-1342.

38. Kern AD, Oliveira MA, Coffino P, Hackert ML (1999) Structure of mammalian ornithine decarboxylase at 1.6 Å resolution: stereochemical implications of PLP-dependent amino acid decarboxylases. Structure 7:567-581.

39. Espaillat A, Carrasco-López C, Bernardo-García N, Pietrosemoli N, Otero LH, Álvarez L, de Pedro MA, Pazos F, Davis BM, Waldor MK, Hermoso JA, Cava F (2014) Structural basis for the broad specificity of a new family of amino-acid racemases. Acta Crystallogr D Biol Crystallogr 70:79-90.

40. Wu D, Hu T, Zhang L, Chen J, Du J, Ding J, Jiang H, Shen X. (2008) Residues Asp164 and Glu165 at the substrate entryway function potently in substrate orientation of alanine racemase from *E. coli*: Enzymatic characterization with crystal structure analysis. Protein Sci 17:1066-1076

41. Watanabe A, Yoshimura T, Mikami B, Hayashi H, Kagamiyama H, Esaki N (2002) Reaction mechanism of alanine racemase from *Bacillus stearothermophilus*: X-ray

crystallographic studies of the enzyme bound with N-(5'-phosphopyridoxyl) alanine. J Biol Chem 277:19166-19172.

42. Ghatge MS, Contestabile R, di Salvo ML, Desai JV, Gandhi AK, Camara CM, Florio R, González IN, Parroni A, Schirch V, Safo MK (2012) Pyridoxal 5'-phosphate is a slow tight binding inhibitor of *E. coli* pyridoxal kinase. PLoS One 7:e41680.

43. Wu HY, Chen SF, Hsieh JY, Chou F, Wang YH, Lin WT, Lee PY, Yu YJ, Lin LY, Lin TS, Lin CL, Liu GY, Tzeng SR, Hung HC, Chan NL (2015) Structural basis of antizyme-mediated regulation of polyamine homeostasis. Proc Natl Acad Sci U S A 112:11229-11234.

44. Marina A, Alzari PM, Bravo J, Uriarte M, Barcelona B, Fita I, Rubio V (1999) Carbamate kinase: New structural machinery for making carbamoul phosphate, the common precursor of pyrimidines and arginine. Protein Sci 8: 934-940.

45. Gil F, Ramón-Maiques S, Marina A, Fita I, Rubio V. (1999) N-Acetyl-L-glutamate kinase from *Escherichia coli*: cloning of the gene, purification and crystallization of the recombinant enzyme and preliminary X-ray analysis of the free and ligand-bound forms. Acta Cristallogr D Biol Crystallogr 55:1350-1352.

46. Llácer JL, Contreras A, Forchhammer K, Marco-Marín C, Gil-Ortiz F, Maldonado R, Fita I, Rubio V (2007) The crystal structure of the complex of PII and acetylglutamate kinase reveals how PII controls the storage of nitrogen as arginine. Proc Natl Acad Sci U S A 104:17644-17649.

Table 1. X-Ray diffraction data and refinement statistics of PipY models

Crystal	PipY-PLP (PDB 5NM8)	PipY-Apo (PDB 5NLC)
Data collection		
Synchrotron/Beamline	Diamond/ I03	ALBA/ BL13 - XALOC
Wavelength (Å)	0.9763	0.9792
Space group	$P2_12_12_1$	$P2_1$
Unit cell a, b, c (Å)	46.6, 98.9, 103.7	67.6, 41.8, 77.8
α , β, γ, (°)	90, 90, 90	90, 108.2, 90
Solvent (%)	50.1	37.7
Matthews coefficient	2.43	1.97
Resolution range[a] (Å)	71.60-1.93 (1.97-1.93)	73.89-1.90 (1.94-1.90)
Reflections[a], total/unique	469,259 / 37,132 (26,106 /2,409)	150,200 / 32,801 (7,964 / 2,072)
Completeness[a] (%)	99.9 (98.7)	99.7 (99.2)
Multiplicity[a]	12.6 (10.8)	4.6 (3.8)
I/σ[a]	22.3 (4.5)	20.6 (4.0)
$CC_{1/2}$[a] (%)	99.9 (87.3)	99.8 (93.5)
Wilson B-factor ($Å^2$)	30.1	21.0
Rpim[a,b] (%)	2.0 (19.4)	2.3 (17.0)
Refinement		
Resolution[a] (Å)	28.39-1.93 (1.97-1.93)	44.61-1.90 (1.94-1.90)
Reflections, work/test	35,275 / 1,857	31,161 / 1,640
R_{factor}[c]/R_{free}[d] (%)	19.4 / 23.0	15.7 / 20.5
r.m.s.d. from ideal value		
Bond length (Å)	0.0096	0.0180
Bond angle (°)	1.4116	1.8560
Number of:		
Polypeptide chains	2	2
Protein atoms	3,322	3,358
Water molecules	162	375
PLP	2	0
Average B-factor ($Å^2$)		
Protein	46.0	26.3
H_2O	44.8	34.7
PLP	49.6	-
Ramachandran plot[e] (%)		
Favoured	98.8	97.9
Allowed	1.2	2.1
Outliers	0.0	0.0

[a]Values in parentheses are the data for the highest-resolution shell.

[b]$R_{pim} = \Sigma_{hkl}[1/(N-1)]^{1/2}\Sigma_i|I_{hkl,i} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\langle I_{hkl}\rangle$.

[c]$R_{factor} = \Sigma_{hkl}||F_{obs}|-|F_{calc}||/\Sigma_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

[d]R_{free} is the same as R_{factor} but calculated for the 5% data omitted from the refinement.

[e]Calculated with RAMPAGE [29]

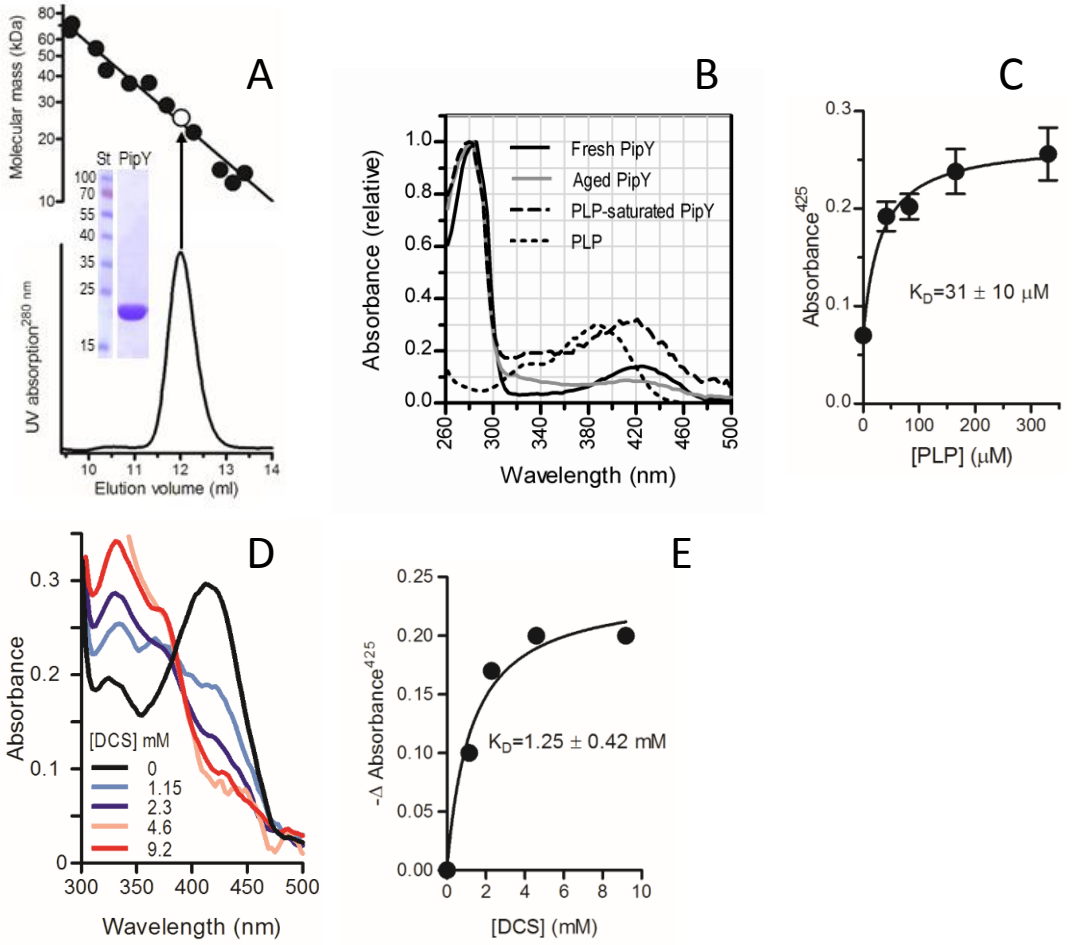


Fig. 1. PipY properties. **(A)** Size exclusion chromatography of pure PipY with elution monitored as UV absorption (bottom) and plotted semilogarithmically (top) together with the elution positions of molecular mass standards (closed circles). The open circle, corresponding to PipY, has been placed at the intersection of the PipY elution position and its sequence-deduced mass. The protein standards used and their masses (in parenthesis, in kDa) were: *Enterococcus faecalis* carbamate kinase (71.3) [44], bovine serum albumin (66.4), *E. coli* N-acetyl-L-glutamate kinase (54.3) [45], chicken ovalbumin (42.7), *S. elongatus* PII (37.2) [46], *Pseudomonas aeruginosa* PII (36.9) (our own unpublished data), bovine erythrocyte carbonic anhydrase (29), soybean trypsin inhibitor (21.5), cow s lactalbumin (14.2), pancreatic ribonuclease A (13.7) and horse heart cytochrome C (12.3). *Inset*, SDS-PAGE of pure PipY (St, protein standards of the indicated masses). **(B)** Optical absorption spectrum of PipY, freshly prepared (black continuous line), after aging for several months (grey continuous line), or following a 15-min incubation of 2 mg/ml PipY with 0.34 mM PLP, followed by excess PLP removal by centrifugal gel filtration [30] (broken line, large trace). Spectra are normalized relative to the absorption at 280 nm. For comparison, the spectrum of PLP, normalized to a maximum absorbance of 0.3, is shown (dotted line). **(C)** Influence of added PLP on the height of the protein-bound PLP peak of 0.66 mg/ml of aged PipY. **(D, E)** Influence of variable DCS concentrations on the absorption spectrum of PipY (0.91 mg/ml, incubated with 0.34 mM PLP and then freed from unbound PLP by centrifugal gel filtration) in the 300-500 nm range (D) or at 435 nm (E), fitting the data to a hyperbola.

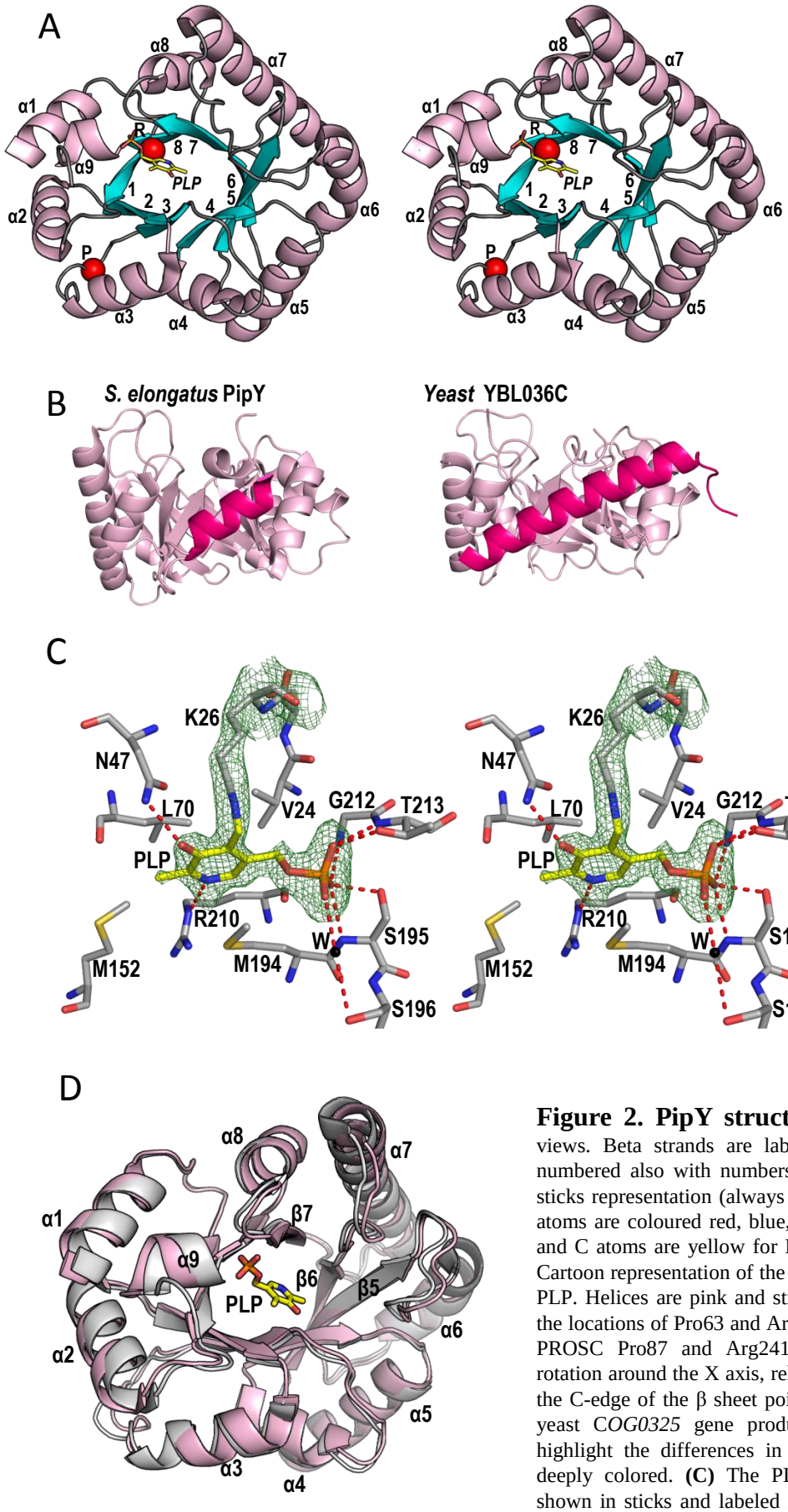


Figure 2. PipY structure. Panels A and C are stereo views. Beta strands are labeled with numbers. Helices are numbered also with numbers preceded by an α. When using sticks representation (always the case for PLP), O, N, S and P atoms are coloured red, blue, yellow and orange, respectively, and C atoms are yellow for PLP and grey for the protein. **(A)** Cartoon representation of the structure of PipY complexed with PLP. Helices are pink and strands are blue. Red spheres mark the locations of Pro63 and Arg210, which correspond to human PROSC Pro87 and Arg241. **(B)** Cartoon side views (90º rotation around the X axis, relative to the view in panel A, with the C-edge of the β sheet pointing to the top) of PipY and the yeast C*OG0325* gene product YBL036C (PDB 1CT5), to highlight the differences in the length of helix 1, which is deeply colored. **(C)** The PLP binding site. Side-chains are shown in sticks and labeled in single letter notation. The grid around the PLP-Lys26 adduct shows the Fo-Fc omit electron density map for this adduct at 2.5 σ. Broken red lines represent hydrogen bonds of PLP with the protein (**D)** PipY movements upon PLP binding shown by superimposing PipY-Apo (pink) and PipY-PLP (gray) in cartoon representation. Selected secondary structure elements are labeled.

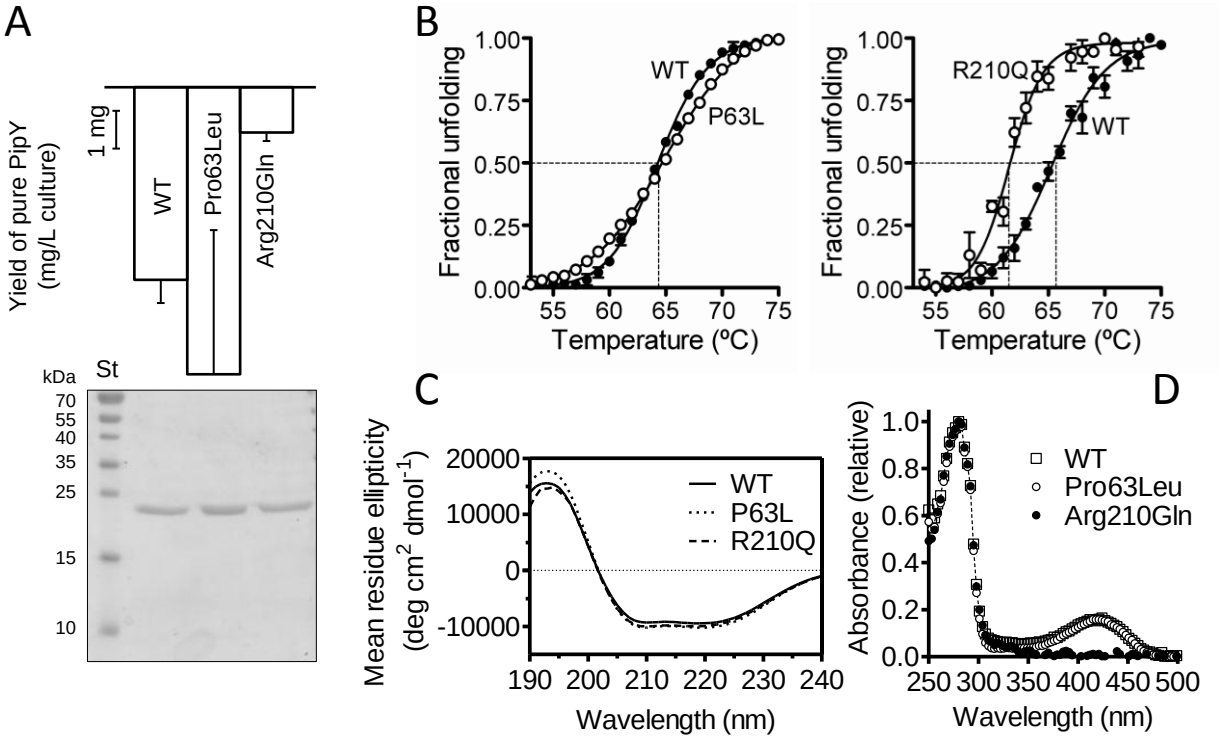


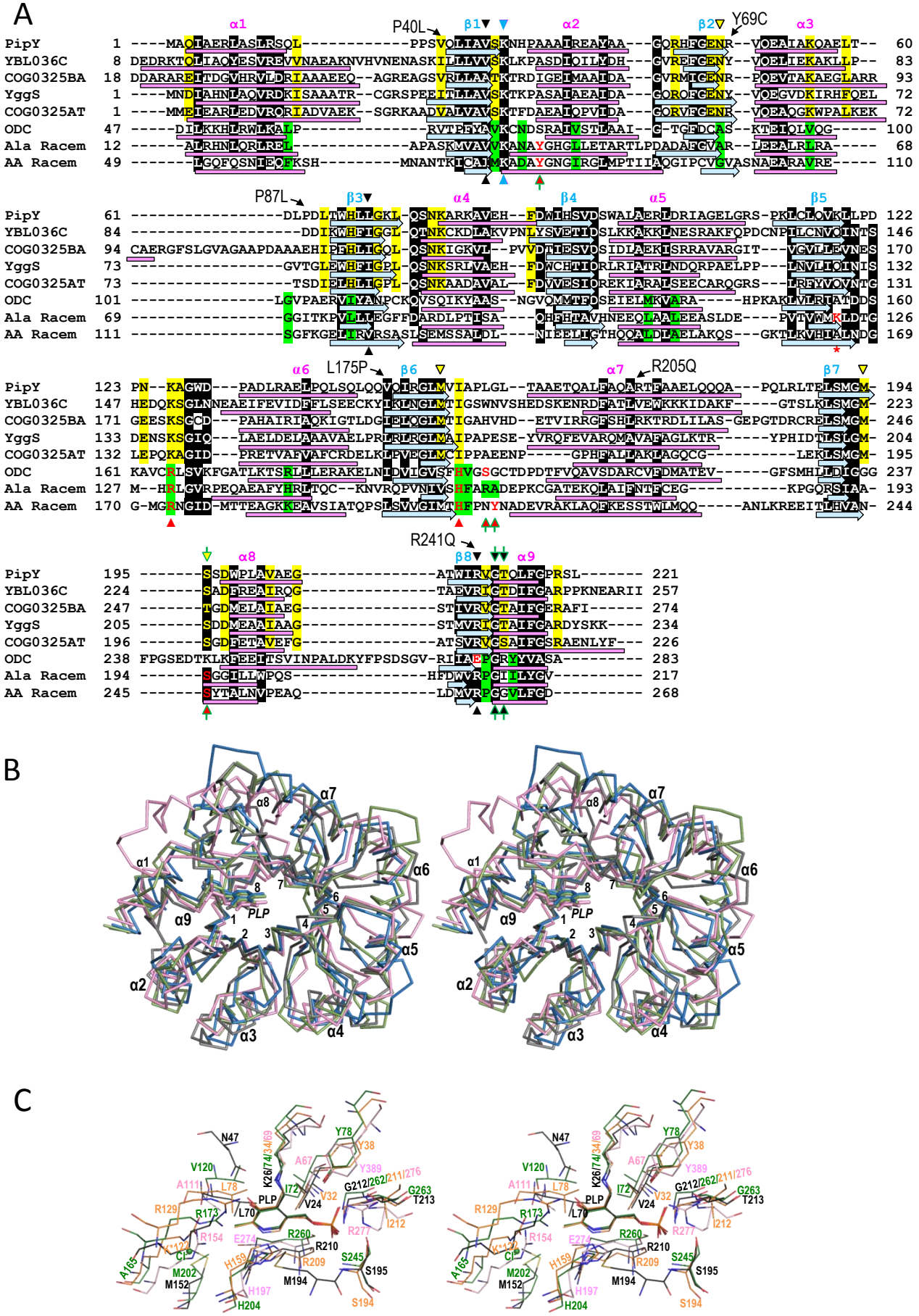
Figure 3. Comparisons between wild-type and Pro63Leu and Arg210Gln mutants of PipY. **(A)** Yield (top; mean SE; n ≥ 2) and purity (bottom; coomassie-stained SDS-PAGE, 15% gel) of the final protein preparations. **(B)** Thermal stability of the wild-type (closed circles) and the two mutants (open circles), estimated in thermofluor assays (see Materials and Methods). Buffer, pH and protein concentrations were the same for both panels, but the solution contained, in addition, 0.5 M NaCl and 20 mM imidazole (left panel) or 16 mM NaCl and 1mM DTT (right panel). The broken lines mark for each protein form the temperature at which the fluorescence change is 50% of the maximum. **(C)** CD spectra to exclude gross misfolding of the mutants. **(D)** Optical absorption spectrum to show that the Arg210Gln mutant lacks the peak of bound PLP. Spectra are the average of ≥ 2 preparations. They are normalized relative to the absorption at 280 nm.

```
PROSC     MWRAGSMSAELGVGCALRAVNERVQQAVARRPRDLPAIQPRLVAVSKTKPADMVIEAYGH 60
PipY      ----------------MAQIAERLASL----RSQL-PPSVQLIAVSKNHPAAAIREAYAA 39
                           :  : **: .       :*   . :*:****.:**  : ***.
                         α3         L     β3
PROSC     GQRTFGENYVQELLEKASNPKILSLCPEIKWHFIGHLQKQNVNKLMAVPNLFMLETVDSV 120
PipY      GQRHFGENRVQEAIAKQA--ELTDL-PDLTWHLLGKLQSNKARKAVE--HFDWIHSVDSW 94
          *** **** *** : * :  ::. * *::.**::*:**.::..* :   ::  :.:***

PROSC     KLADKVNSSWQRKGSPERLKVMVQINTSGEESKHGLPPSETIAIVEHINAKCPNLEFVGL 180
PipY      ALAERLDRIAGELGRSPKLC--LQVKLLPDPNKAGWDPADLRAELPQL-SQLQQVQIRGL 151
           **::::    . *   :*   :*::   : .* *  *::  * : :: ::  :::: **

PROSC     MTIGSFGHDLSQGPNPDFQLLLSLREELCKKLN---IPADQVELSMGMSADFQHAVEVGS 237
PipY      MVIAPLGLTAAET-----QALFAQARTFAAELQQQAPQLRLTELSMGMSSDWPLAVAEGA 206
          *.*. :*   ::      * *::  . :. :*:        .*******:*:  **  *:
          β8 Q    α9
PROSC     TNVRIGSTIFGERDYSKKPTPDKCAADVKAPLEVAQEH 275
PipY      TWIRVGTQLFGPRSL----------------------- 221
          * :*:*: :** *.
```

Supplementary Figure S1. Clustal ω (www.ebi.ac.uk) sequence alignment of *S. elongatus* PipY and human PROSC, indicating above the piled sequences the epilepsy-associated mutations Pro87Leu and Arg241Gln [17,18] that have been introduced here into the corresponding positions of PipY. Secondary structure elements surrrounding the mutated residue are indicated in PipY by double underlining (strands) or ondulating underlining (helices), identifying these elements above the sequences. Below the sequences, asterisks mark invariance, and double and single dots indicate higher and lower strengths of residue conservation in both structures.



Supplementary Figure S2

Supplementary Figure S2. Sequence and structural comparisons of PipY with other proteins. **(A)** Structure-aided sequence alignment of the five *COG0325* gene products of known structure (*S. elongatus* PipY and the products of yeast, *Bifidobacterium adolescentis*, *E. coli* and *Agrobacterium tumefaciens*, abbreviated PipY, YBL036C, COG0325BA, YggS and COG0325AT, with PDB file identifiers 5NM8, 1CT5, 3CPG, 1W8G and 3R79, respectively) and of the TIM barrel domains of mouse ornithine decarboxylase (ODC; PDB file 7ODC), *E. coli* alanine racemase (Ala Racem, PDB 2RJG) and *Vibrio cholerae* broad specificity amino acid racemase (AA Racem; PDB 4BEU). Residues conserved or conservatively replaced in at least 5 of the 7 aligned sequences are highlighted by black shadowing over white lettering. When the conservation is restricted to the 5 COG0325 products or to the 3 enzymes, the residues are highlighted yellow and green, respectively. Horizontal arrows (in light blue) and rectangles (pinkish) under each sequence mark β strands and α helices, respectively, with labelling of these secondary structure elements at the top of the alignment (blue and pink labels). The triangles at the top and bottom signal residues that make contacts with PLP in the COG0325 products and the enzymes, respectively, being blue for the Shiff base-making lysine, arrow-shaped an green-lined for phosphate-interacting residues, black if the interaction is conserved in all the aligned sequences, or yellow or red if conserved only in COG0325 products or the enzymes, respectively. In addition, yellow and red lettering is used in some occasions to confirm the interaction of a given residue with PLP when the interaction is not fully conserved. The red asterisk denotes the lysine that is carboxylated in Ala racemase. The curved small arrows and the adjacent mutations (shown in single letter code) above the PipY sequence indicate the corresponding single amino acid changes found in vitamin B6-dependent epilepsy [17, 18]. **(B)** Stereo view of the superposition of the structures (backbone representation) of PipY-PLP (grey) and the barrel domain of broad spectrum amino acid racemase (green; PDB entry 4BEU), alanine racemase (blue, 2RJG) and ornithine decarboxylase (pink, 7ODC). The bound PLP molecules are shown in sticks representation and are coloured as the corresponding protein chain. Beta strands are labeled with numbers while helices are labelled with an α followed by a number. **(C)** Detailed stereo view (sticks representation) of the superimposed PLP binding sites of PipY-PLP (black) and of the barrel domains of broad spectrum amino acid racemase (green; PDB 4BEU), alanine racemase (orange, 2RJG) and ornithine decarboxylase (pink, 7ODC). O, N, S and P atoms are coloured red, blue, yellow and orange, respectively. Residue labels are in single letter amino acid code and are in the same color as the corresponding structures. The pink asterisk denotes that lysine 122 is carboxylated in alanine racemase [41]. The green sphere is the Cl^- anion in broad spectrum amino acid racemase [39]. Tyr389 of ODC belongs to the other domain of the same subunit of this enzyme. Not shown in the figure because they obliterate the view of other residues are Tyr299 and Tyr255 from the second domain of the other subunit of amino acid racemase and alanine racemase, respectively. These residues extend from the front of the figure towards the O atom bound to the C3 atom of the pyridoxal ring, forming a hydrogen bond with it via their phenolic O atoms.

Supplementary Table S1. **Degree of similarity among the structures of PipY monomers studied here, and with the structures of other *COG0325* gene products, and of TIM barrel domains of amino acid racemases and ornithine decarboxylase.**

	Root mean square deviation for the superimposition of the indicated number of Cα atoms[a] (Å / number Cα atoms)			
	PipY-Apo A	**PipY-Apo B**	**PipY-PLP A**	**PipY-PLP B**
PipY-Apo B	0.59/215	-	-	-
PipY-PLP A	0.90/213	0.62/215	-	-
PipY-PLP B	0.98/210	0.70/212	0.39/212	-
COG0325 yeast YBL036C (1CT5)	-	-	1.63/195	1.61/194
COG0325 *E. coli* YggS (1W8G)	-	-	1.50/205	1.52/205
COG0325 *B. adolescentis* (3CPG)	1.75/202	1.69/202	-	-
COG0325 *A. tumefaciens* (3R79 A)	-	-	1.29/199	1.42/196
***Alanine racemase* (2RJG)**	-	-	2.39/191	-
***Broad spectrum AA racemase* (4BEU)**	-	-	-	2.31/200
***Ornithine decarboxylase* (7ODC)**	-	-	2.60/186	-

[a]The two chains in the same asymmetric unit are distinguished by the letters A and B. *B. adolescentis, Bifidobacterium adolescentis; A. tumefaciens, Agrobacterium tumefaciens*; AA, amino acids. The alanine racemase, broad spectrum amino acid racemase and ornithine decarboxylase are from *E. coli*, *Vibrio cholerae* and mouse, respectively. Protein DataBank accession numbers are given between parentheses. Except in the case of the protein from *B. adolescentis*, which contained no PLP and thus was compared with the monomers of PipY-Apo, all other proteins contained PLP and have been compared with PipY-PLP. In the cases of racemases and ornithine decarboxylase only the result of the best superimposition of PipY with one subunit of the indicated enzymes is shown.