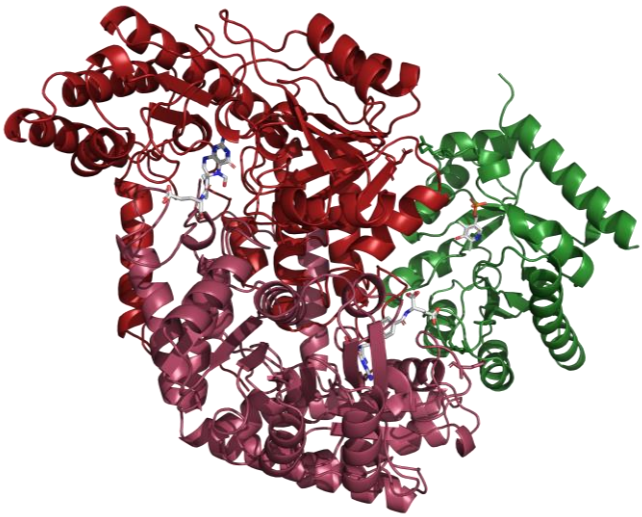
SAPIENZA
UNIVERSITÀ DI ROMA

Ph.D. Course in Biochemistry

Department of Biochemical Sciences "A. Rossi Fanelli"

XXXIII Cycle

Regulation of vitamin B_6 metabolism in *Escherichia coli*



Tutor:

Prof. Roberto Contestabile

Supervisor:

Prof. Martino Luigi Di Salvo

Coordinator:

Prof. Stefano Gianni

Ph.D. Student

Federico D'Alessio

Abstract

The catalytically active form of vitamin B_6, pyridoxal 5'-phosphate (PLP), acts as a coenzyme in a variety of different enzymatic reactions. PLP is essential for the normal growth and development of all the organisms, but only plants and microorganisms can synthetise it *de novo*, while the other organisms must recycle it from nutrients.

The PLP is a very reactive molecule thanks to its aldehyde group and its intracellular concentration must be kept low to avoid undesired toxic reaction in the organism. The regulation of the free PLP concentration in cells occurs through a variety of mechanism, such as its dephosphorylation to PL or, as recently discovered, through the binding to PLP Binding Protein. In *Escherichia coli* an important candidate for this role has recently been discovered: YggS is a non-catalytic protein able to bind PLP. It belongs to the COG0325 family, a class of protein sharing structure homology with PLP-dependent enzymes, such as alanine racemase and some decarboxylases that have the same Fold-type III.

The metabolism of PLP in *E. coli* has been studied for years. In this work, our group have deepened the regulation of the PLP metabolism, studying and elaborating the state of the art, and crossing the available literature data with those produced in our lab about the regulation of the expression of the genes involved in PLP homeostasis and focusing the analysis on the most important genes. Our studies have also analysed *in vivo* the phenotypes linked to the genes involved in PLP homeostasis when *E. coli* is grown in different media and in presence of vitamers, in order to better understand the role of the analysed genes. The study has also considered the expression of the genes involved in PLP homeostasis in presence of vitamin B_6 vitamers and during different growth phases.

Finally, given the importance of the gene *yggS* in *E. coli*, the attention was focused on the characterization of the protein YggS, whose role has not yet been discovered. This protein is hypothesised to be both a PLP binding protein and a carrier of this important cofactor to the apo PLP-dependent enzymes. Our studies were focused on YggS capacity to bind PLP and, using mutant forms of this protein, we have verified and studied the transfer mechanism of PLP to the apo PLP-dependent enzymes.

Index

1. Introduction

1.1 Pyridoxal-5'-phosphate: an overview of the biological active form of the vitamin B_6

Vitamin B_6 is a water-soluble compound represented by six vitamers that are interconvertible thanks to different enzymes. The six vitamers of the vitamin B_6 are: pyridoxine (PN), pyridoxamine (PM), pyridoxal (PL) and their phosphorylated forms on their 5' carbon: pyridoxine-5'-phosphate (PNP), pyridoxamine-5'-phosphate (PMP) and the bioactive form pyridoxal-5'-phosphate (PLP) (fig 1.1). The vitamers share the same structure of the 2-methyl-3-hydroxypyridine and differ for the group on the C4 position.

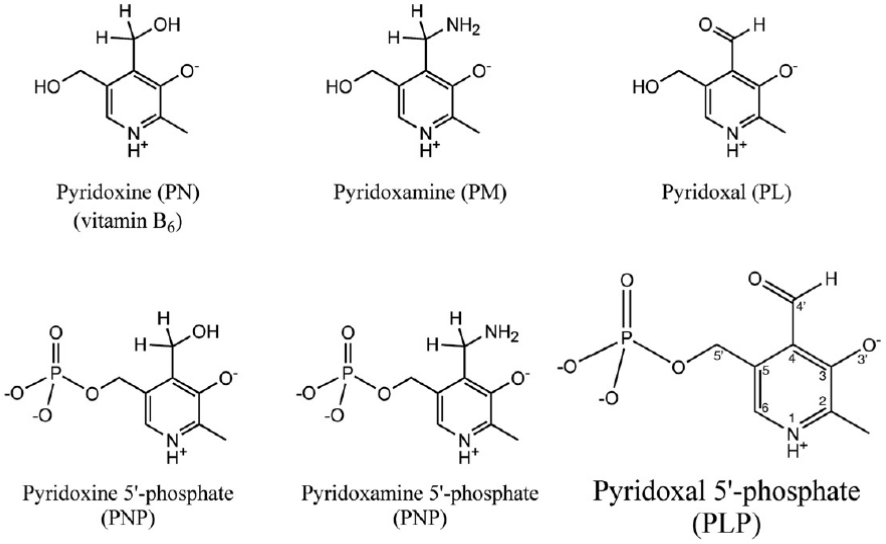


Figure 1.1: Structures of the six B_6 vitamers. Carbon atom numbering is shown on the PLP structure (Di Salvo et al, 2011)

The biologically active vitamer, pyridoxal 5'-phoshate (since here called PLP), acts as a cofactor for more than 150 enzymes (Percudani and Peracchi 2003), and it is essential for normal growth and development of all the organisms. In fact, as enzymatic cofactor, PLP is involved in a lot of essential

pathways such as: the synthesis, degradation and transformation of biogenic amines, polyamines and amino acids, transsulfuration, supply of one carbon units, synthesis of tetrapyrrolic molecules, synthesis and degradation of the neurotransmitters (Mills et al. 2005). As a free (not bound to enzyme) molecule, PLP has other important roles, such as chaperon for correct enzymes folding (Cellini et al. 2014), modulator for transcription factors in bacteria (Huq et al. 2007; Tramonti et al. 2018), and as a virulence factor for pathogenic microorganism such as *Helicobacter pylori*, *Mycobacterium tuberculosis* and *Actinobacillus pleuropneumoniae* (Dick et al. 2010; Grubman et al. 2010; Xie et al. 2017).

PLP is characterized by an aldehyde group on the C4 carbon and, as a small aldehyde, is a very reactive molecule, so in cells there must be mechanisms for its homeostasis because its free form must be kept in low concentration in order to avoid undesired reaction with other reactive biological molecules (i.e. the amino and thiol groups).

1.2 Vitamin B_6 metabolism

Vitamin B_6 represented by PLP and its vitamers is essential for all the organisms. Despite the considerable importance of this molecule, only plants and microorganism can synthetize *de novo* this vitamin. They do this using two different and mutually exclusive biosynthetic pathways: the deoxyxylulose 5-phosphate (DXP)-dependent pathway and the DXP-independent pathway (fig 1.2). All other organisms must assimilate the vitamers from nutrients and interconvert them, depending on their needing, using a so-called salvage pathway (Fitzpatrick, Moccand, and Roux 2010).

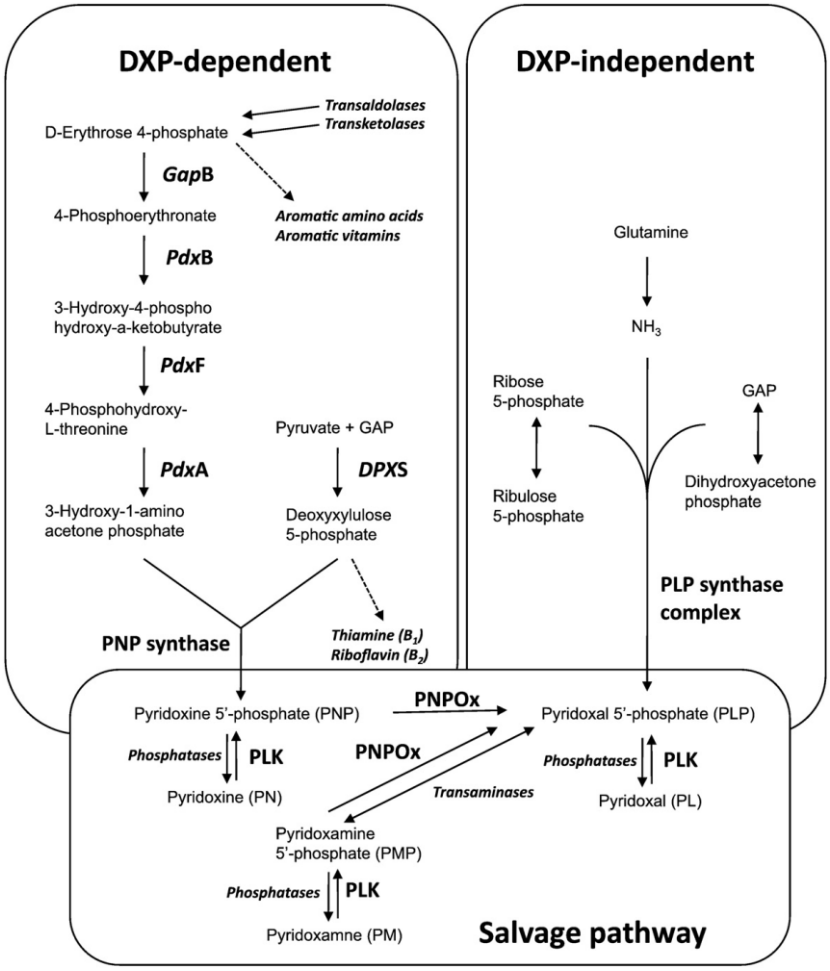


Fig 1.2: Vitamin B6 biosynthetic pathways: De novo DXP-dependent pathway (present in some eubacteria): GapB, D-erythrose 4-phosphate dehydrogenase; PdxB, erythronate-4-phosphate dehydrogenase; PdxF/SerC, phosphoserine aminotransferase; PdxA, 4-hydroxythreonine-4-phosphate dehydrogenase; DPXS, 1-deoxy-D-xylulose-5-phosphate synthase; PNP synthase, from PdxJ gene. De novo DXP-independent pathway (present in other eubacteria, fungi, plants and Archaea): PLP synthase complex: synthase domain from Pdx1 gene; glutaminase domain from Pdx2 gene. Salvage pathway (present in all organisms, including mammals): PLK, pyridoxal kinase from PdxK gene and pyridoxal kinase 2 from PdxY gene; PNPOx, pyridoxine 5′-phosphate oxidase from PdxH gene. (Di Salvo et al, 2011)

1.2.1 Vitamin B_6 *de novo* biosynthesis

As already mentioned above, the unique living beings able to biosynthesize vitamin B_6 are plants and microorganisms. Two different and independent vitamin B_6 *de novo* biosynthetic pathways are known: a DXP-dependent and a DXP-independent pathway, depending on the usage of 1-deoxy-D-xylulose-5-phosphate as a reaction intermediate (Fitzpatrick et al. 2007).

The DXP-dependent *de novo* route was the first to be discovered and studied in *Escherichia coli* and for a long time was believed to be ubiquitous but, actually, it is used only from some eubacteria (Fitzpatrick et al. 2007; Di Salvo, Contestabile, and Safo 2011). This biosynthetic pathway is composed by two convergent branches (Fig 1.2). The first branch starts with the conversion of D-erythrose 4-phosphate into 4-phosphoerythronate, while the second one starts from the condensation of pyruvate and glyceraldehyde 3-phosphate to produce 1-deoxy-D-xylulose-5-phosphate. The two branches converge on the reaction catalysed by pyridoxine 5'-phosphate synthase (PdxJ) to form pyridoxine 5'-phosphate (Laber et al. 1999). The reaction steps used in these two branches are catalysed by five enzymes encoded by the genes *gapB, pdxB, pdxF, pdxA,* and *dxpS*. Being PLP the bioactive form of the vitamin B_6, PNP generated in the pathway described above must be oxidised to PLP by PNP oxidase (PNPOx) encoded by the gene *pdxH* (Fitzpatrick et al. 2007).

The so-called DXP-independent *de novo* route was discovered by chance in Fungi, and after that it was clear that it is more widely diffused than the DXP-dependent route, being found in Archaea, most Eubacteria and plants (Ehrenshaft et al. 1999)(Mittenhuber 2001). This pathway use glutamine, ribose or ribulose 5-phosphate and either glyceraldehyde 3-phosphate or

dihydroxyacetone phosphate to produce directly PLP thanks to the catalytic action of the PLP synthase complex encoded by *pdx1* and *pdx2* genes (Zein et al. 2006, Raschle, Amrhein, and Fitzpatrick 2005).

1.2.2 The vitamin B_6 salvage pathway

Some organisms have developed a pathway that bypasses the necessity to produce PLP recycling the vitamers that they absorb from the environment as a nutrient or from the degraded PLP dependent enzymes. This alternative route to PLP production is known as "salvage pathway" and it is made by few enzymes able to interconvert the vitamin B_6 vitamers into PLP and vice versa (fig 1.2). The key enzymes of the salvage pathway are: the ATP-dependent pyridoxal kinase (PDXK) that is able to convert each vitamer into their relative phosphorylated form, and the pyridoxine 5'-phosphate oxidase (PNPOx) which oxidases pyridoxine 5'-phosphate into PLP using FMN as a cofactor. These two enzymes are linked also in the regulation of the PLP level in the cell binding this molecule and being regulated themselves by product inhibition. The binding of the PLP on these two proteins contributes to its homeostasis and regulates its production trough different mechanisms that will be discussed in the following paragraphs.

Moreover, in the PLP salvage pathway are present other enzymes such as the phosphorylases that are required to dephosphorylate PLP and other phosphorylated vitamers in order to deliver them through the membranes to the other cells. For example, in human the intestine only absorbs non phosphorylated vitamers (Darin et al. 2016) and therefore, as studied in rats, the vitamers are first dephosphorylated by the intestinal phosphatases (Middleton 1990). Recently PL reductase, which converts PN to PL, has been

proposed as an important component of the salvage pathway (Ito and Downs, 2020).

1.3 Vitamin B_6 in humans

Vitamin B_6 in humans is absorbed from food and is produced from the intestinal microflora. The major nutritional rich foods for vitamin B_6 are: fish, beef liver and other organ meats, potatoes and starchy vegetables, and also fruits. In animal-derived food the vitamin B_6 is mainly present as PLP associated with glycogen phosphorylase and as PMP, while in vegetables it is present as PN or PN-5'-β-D-glucoside (McCormick 1989).

The metabolism of the vitamin B_6 in human involves different organs and enzymes as shown in the pathway (Fig 1.3) from dietary assumption, recycle and catabolism of excess PLP (Darin et al. 2016). After their dietary assumption, the phosphorylated vitamers cannot cross the membrane of the intestine (Darin et al. 2016), so some phosphatases such as the tissue-non-specific alkaline phosphatase (TNSALP) hydrolase them in order to allow their transport across the membrane (Said 2004). Once in the intestine PM and PM can be converted before their transport to the liver (Albersen et al. 2013) or, once in the liver in PLP. In the liver PM, PL and PN are first phosphorylated from the pyridoxal kinase and then PMP and PNP are oxidised by the PNPOx to form PLP that, once bound to the albumin (Bohney, Fonda, and Feldhoff 1992), is exported to other cells. The liver is also the organ where the excess PLP is dephosphorylated to PL and then oxidised by an aldehyde dehydrogenase or an aldehyde oxidase (AOX) (Stanulović et al. 1976) to pyridoxic acid to be excreted in urine.

In order to reach the brain, the PLP must dissociate from albumin and be converted in PL by a tissue non-specific alkaline phosphatase. As PL it can

pass the blood brain barrier (BBB) probably by facilitated diffusion and it is "trapped" in the brain cells as PLP by the action of the pyridoxal kinase (Spector and Johanson 2007).

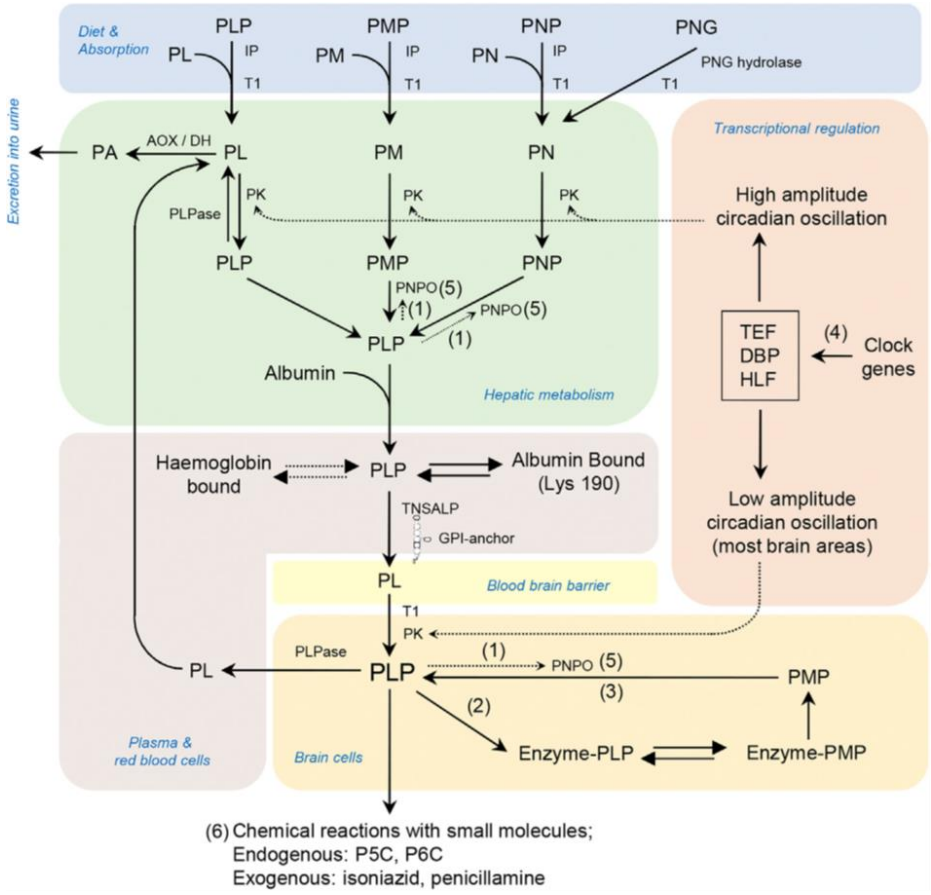


Fig 1.3: **Enzymes and transporters involved in mammalian CNS PLP Synthesis and homeostasis and known human genetic vitamin-B_6 dependent epilepsies**

Pyridoxal 5'-phosphate (PLP); pyridoxamine 5'-phosphate (PMP); pyridoxal (PL); pyridoxine (PN); pyridoxine 5'-phosphate (PNP); pyridoxine-50-b-D-glucoside (PNG); intestinal phosphatases (IP); transporter (identity unknown; T1); pyridoxal kinase (PK); pyridox(am)ine 5'-phosphate oxidase (PNPO); tissue non-specific alkaline phosphatase (TNSALP); pyridoxal phosphatase (PLPase); aldehyde oxidase (Mo cofactor)/b-NAD dehydrogenase (AOX/DH). (1) PNPO is controlled by feedback inhibition from PLP. (2) PLP functions as a co-factor, forming Schiff bases with the ε-amino group of lysine residues of proteins. (3) PLP can be formed by recycling the cofactor from degraded enzymes (salvage pathway). (4) PLP levels are maintained, in part, by circadian-clockcontrolled transcription factors with PAR bZip transcription factors (DBP, HLF, and TEF) targeting PK. (5) PNPO mutations cause a B_6-dependent epilepsy disorder. (6) Disorders resulting in accumulation of L-D,1-pyrroline-5-carboxylic acid (P5C) and D,1-piperideine-6-carboxylic acid (P6C) (hyperprolinaemia type II and pyridoxine-dependent epilepsy due to mutations in ALDH4A1 and ALDH7A1, respectively) cause decreases in bioavailable PLP as do reactions with exogenous small molecules. (Darin et al, 2016)

Once PLP is in the cell, its free form must be kept under a certain concentration because of its reactiveness. In order to limit unwanted reactions and toxicity effects given by the free form of PLP, the cells have developed a very efficient protein whose function is not well defined, but that is able to bind PLP. In human the protein that must control the concentration of PLP is the PLP Binding Protein (PLPBP), a protein that does not have any catalytic activity (Darin et al. 2016).

Due to its importance in the homeostasis of the PLP, a deep focus on this protein will be discussed in a dedicated paragraph.

Imbalances in PLP homeostasis can cause severe diseases, PLP deficiency in fact is implicated in several neurological and non-neurological disorders. The vitamin assumed with diet is commonly sufficient since most of the food contains vitamin B_6. The most of the PLP imbalance are done by inborn error on the recycle pathway genes or by the assumption of PLP inactivating drugs (Di Salvo et al. 2011). The effect of the PLP concentration imbalance will be discussed in the following paragraphs to better describe the linkage between the low and the high concentration with the most important and studied associated pathologies.

1.4 PLP related diseases in humans

In humans, an insufficient concentration or an excess of PLP can be associated with both mild and severe different pathologies. PLP deficiency due to drug assumption may produce symptoms such as unconsciousness, seizures, sleeplessness, headache, restlessness, agitation, tremors, and hallucination, while the deficiency linked with genetic mutation is implicated in severe pathologies, including neonatal epileptic encephalopathy, seizures, autism, Down syndrome, schizophrenia, autoimmune polyglandular disease,

Parkinson's, Alzheimer's, epilepsy, attention deficit hyperactivity disorders and learning disability (Di Salvo et al. 2011).

Some pathologies linked with the low concentration of PLP, such as: symptoms of the premenstrual syndrome (Wyatt et al. 1999), nausea and vomiting during pregnancy (Ebrahimi, Maltepe, and Einarson 2010), carpal tunnel syndrome (Ryan-Harshman and Aldoori 2007) can easily be solved with the assumption of this molecule. More severe pathologies linked with low level of PLP are some neurological diseases and there are also evidences that a low concentration of PLP can contribute to the onset of diabetes and cancer.

Excess of PLP can be toxic too, in fact in humans the effects of large doses (more than 200mg/day) have been observed (Windebank 1985) mostly at the level of peripheral nervous system as sensory and motor neuropathy leading to numbness in hands and feet that usually finish when supplementation is interrupted (Foca 1985). Another symptom of an excess of PLP can be the seizure susceptibility by diminishing GABA levels through several mechanisms (di Salvo, Safo, and Contestabile 2012).

1.4.1 PLP deficiency related diseases

The essential role of PLP in the metabolism of neurotransmitters including GABA, dopamine, epinephrine, serotonin, histamine and D-serine is the main cause of pathological phenotypes when inborn errors lead to a deficiency in the PLP concentration (Hatch et al. 2016).

PLP deficiency occurs through different mechanisms, an example is the PNPOx deficiency due to inborn defect of the enzyme that affects the PLP synthesis and recycle causing a pathology called neonatal epileptic

encephalopathy. The patients affected by this pathology present a particular electroencephalogram pattern and some of them can be treated with PLP but not with PN (Wang and Kuo 2007).

Hypophosphatasia is a pathology that affects the PLP concentration in the brain and it is due to a molecular defect on the gene coding for the tissue non-specific alkaline phosphatase (TNSALP). Phosphorylated PLP cannot cross the blood brain barrier and its deficiency in the brain impairs the synthesis of the neurotransmitters causing seizures (Ware et al. 2014).

Symptoms similar to the patients with PNPOx deficiency have been observed also in people with PLP binding protein (PLPBP) deficiency because the intracellular transport of PLP were damaged (Darin et al. 2016).

PLP deficiency is given also by the accumulation of metabolites that inactivate PLP causing pyridoxine dependent epilepsy (PDE). These cases are characterized by deficiency of ALDH7A1 (α-aminoadipic semialdehyde dehydrogenase) or ALDH4A1 (P5C dehydrogenase) (Wilson et al. 2019).

All the deficiency described above, when in their most severe cases (neonatal presentations) can have other consequences including anemia, lactic acidosis and hypoglycemia (Wilson et al. 2019).

1.4.2 PLP role in cancer

Despite until the early 1980s the shortage of PLP was considered a strategy against cancer due to the importance of the PLP in the proliferation of the cells (Galluzzi et al. 2013), recent studies have evidenced the opposite relation between vitamin B_6 and cancer. In fact, PLP deficiency is linked to an increase of tumor insurgence, in particular in the gastrointestinal tract (Gylling et al. 2017; Kayashima et al. 2011) and lung (Zuo et al. 2019). It is

clear that PLP has a multifaced role in cancer being both an antioxidant and an essential micronutrient for cell proliferation (Contestabile et al. 2020).

PLP is not only a cofactor for several enzymatic reactions, but is also an efficient ROS scavenger, as the other vitamers, thanks to the presence of both hydroxyl (-OH) and amine (-NH2) substituents on the pyrimidine ring which can directly react with the peroxy radicals (Contestabile et al. 2020; Kannan and Jain 2004).

It has been proposed that in conditions of metabolic stress or in the case of shortage of necessary nutrients, some cellular processes such as DNA acetylation/methylation, synthesis of DNA precursors and ROS production can be altered causing DNA damage, which can drive cells toward cancer (Turgeon, Perry, and Poulogiannis 2018). Although many studies converge towards a protective role of vitamin B_6 in cancer, the molecular mechanisms are not completely understood. PLP, as antioxidant molecule, could play an important role in mediating the cross talk between metabolism and DNA damage, in fact PLP acts against the formation of advanced glycation end products (AGEs), that are genotoxic compounds associated with senescence and diabetes (Booth et al. 1997).

PLP does not have only an antioxidant role in cancer, in fact it plays a crucial role as a cofactor of the serine hydroxy methyltransferase (SHMT) whose folate dependent reaction is the main source of one carbon unit in the metabolism, playing an important role in the synthesis of the thymidylate (Fig 1.4)

An eventual PLP deficiency can cause a decrease in activity of the SHMT causing a misincorporation of uracil in the DNA (Giardina et al. 2018; MacFarlane et al. 2011; Paone et al. 2014). SHMT is not the only PLP dependent enzyme whose malfunction reflect its effect on the DNA, in fact also glycine decarboxylase (that also depend on folate) is essential for the

synthesis of purines and so, for DNA metabolism (Contestabile et al. 2020; Fleming and Copp 1998).

Given the implications that vitamin B_6 have in DNA metabolism, it is not surprising that low levels of this vitamin are associated with cancer.

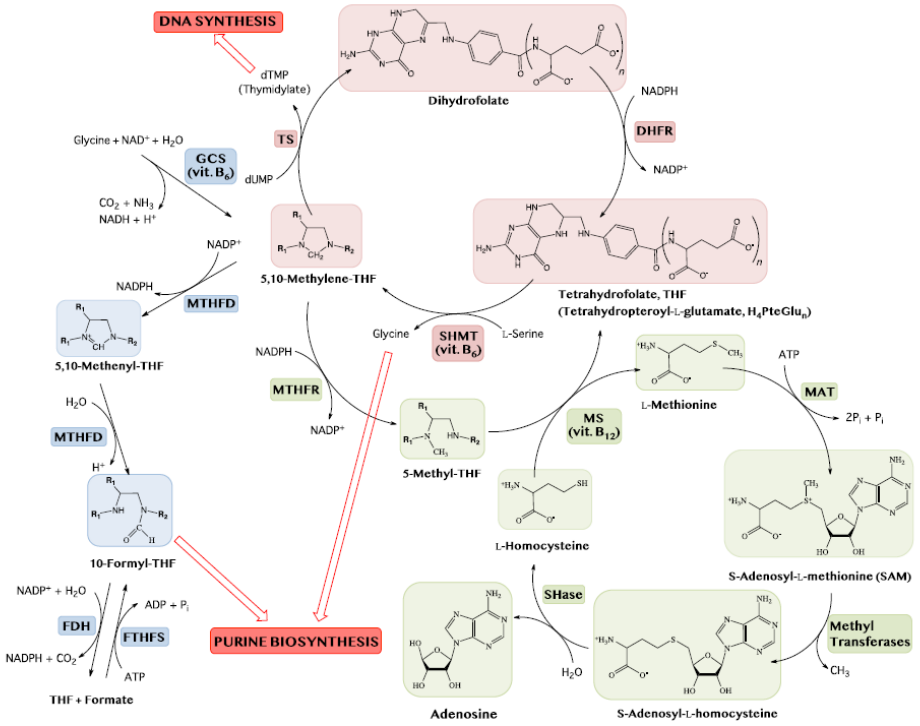


Fig 1.4: Schematic of B_9 metabolism comprising the thymidylate cycle (red diagram), the methionine cycle (green diagram) and the purine biosynthesis pathway (blue diagram). The enzymes involved are: dihydrofolate reductase (DHFR); thymidylate synthase (TS); serine hydroxymethyltransferase (SHMT); methylenetetrahydrofolate reductase (MTHFR); methionine synthase (MS); methionine adenosyltransferases (MAT); S-adenosylhomocysteinase (SHase); glycine cleavage system (GCS); methylenetetrahydrofolate dehydrogenase (MTHFD); 10-formyltetrahydrofolate dehydrogenase (FDH); formyltetrahydrofolate synthetase (FTHFS). (Contestabile et al, 2020)

1.4.3 PLP role in diabetes

This relationship between PLP deficiency and diabetes has been previously described (Rubí 2012), but the molecular and cellular mechanisms underlying this relationship have not yet been completely understood.

PLP deficiency may impact on diabetes in different ways; as reported by Kotake and collaborators (Kotake et al. 1975) the metabolites produced in the altered tryptophan degradation pathway can interfere with insulin activity, causing insulin resistance. Moreover, due to the role of PLP in controlling the expression of genes involved in adipogenesis and in the metabolic pathway of homocysteine, a deficiency of this vitamin could cause insulin resistance (Liu et al. 2016; Moreno-Navarrete et al. 2016).

Studies performed by Rubì and collaborators reported that PLP deficiency may play a role in type I diabetes onset, probably because of the function of this vitamer as a cofactor for two enzymes both present in the pancreatic islet: glutamic acid decarboxylase (GAD65), which synthesizes γ–aminobutyric acid (GABA), that, as part of the GABA cycle it could represent a reserve energy source for the cells based on glutamate; and L-amino acid decarboxylase (L-AADC), a key enzyme in dopamine synthesis (Fig 1.5) (Rubí 2012). Dopamine and serotonin produced by L-AADC could be involved in the regulation of insulin synthesis and secretion, thus, lower levels of dopamine can favour the onset of the type I diabetes (Rorsman et al. 1995). Since GAD65, having the amino acids 250-273 sequence similar to the amino acids 28-50 sequence of the coxsackievirus B protein P2-C, can act as an autoantigen in pancreatic cells (Atkinson et al. 1994). It was proposed that, as a result of this molecular mimicry, cross-reactive T-cell proliferation will occur, leading to autoimmune destruction of the B cells (Rorsman et al. 1995). Also, the key point for the regulation of GAD65 is the change from

apo- to holo-enzyme. In fact, PLP binding domain surrounds the auto-antigenic region described above, thus, a PLP-deficiency results in higher levels of apo-GAD65 exposing the antigenic domain and not able to synthesize GABA (Fig 1.5) (Rubí 2012).

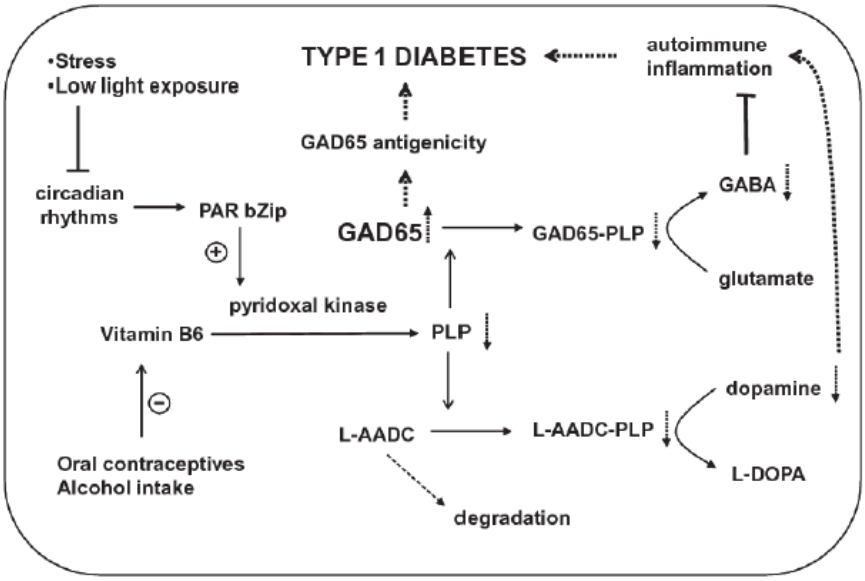


Fig 1.5: **Hypotetical model for PLP deficiency and the onset of type I diabetes.** B_6 vitamers are converted into PLP by pyridoxal kinase, whose expression is promoted by the transcription factor PAR bZip. Alteration in circadian cycle, alcohol consumption and use of contraceptives decrease PLP synthesis. GAD65 devoid of PLP does not synthesize GABA. Alteration of L-AADC function lowers dopamine levels. GAD65, glutamic acid decarboxylase; PLP, pyridoxal 5'-phosphate; PAR bZip, transcription factor belonging to proline and acidic amino acid-rich basic leucine zipper transcription factor family; L-AADC, L-amino acid decarboxylase; L-DOPA, levodopa. (Rubì, 2012)

1.5 PLP homeostasis: its importance and the involved proteins

PLP is a very reactive molecule thanks to the aldehyde group on the C4' position that can easily forms aldimines with primary and secondary amines (Di Salvo et al. 2011). Its concentration as a free compound must be kept low to avoid undesired reactions and toxic effect on the organisms as observed in human with diseases like sensory neuropathy (Critcher 2015). Also in the microorganisms free PLP levels must be kept low in order to avoid unwanted inhibition as observed in *Candida utilis*, where PLP can inhibit the enzyme 6-phosphogluconic dehydrogenase binding the lysine residues in the active site through a Schiff base (Rippa, Signorini, and Pontremoli 1967).

Given the dangerousness of the PLP unregulated concentrations, different mechanisms are adopted in the cells to control the level of this molecule and in eukaryotic the concentration is maintained as low as 1µM still providing enough PLP to saturate the PLP-dependent enzymes.

To manage the PLP level to the correct concentration, the most established method is the dephosphorylation to PL by phosphatase. Another mechanism is the conversion of the PL into pyridoxic acid catalysed by the aldehyde oxidase or the NAD-dependent dehydrogenase enzymes (Stanulović et al. 1976). In mammals, the pyridoxal derives from the phosphatase activity of a specific PLP-phosphatase (Gao and Fonda 1994) while in microorganism there should be specific phosphatases, in fact recently it has been discovered a specific PLP-phosphatase in *E. coli* encoded by the gene *ybhA* (Sugimoto et al. 2017) that reduces the PLP toxicity.

An important player in PLP homeostasis is the PLPBP and its orthologous that was first identified in *E. coli* as a member of the COG0325 family (Prunetti et al. 2016).

It is interesting that also the PLP producing enzymes (PDXK and PNPOx) can be regulated binding this molecule tightly (Musayev et al. 2003; Yang and Schirch 2000) contributing to its homeostasis. About pyridoxal kinase (PDXK) in *E. coli*, it was observed that PLP can bind the protein in the active site on the residue Lys229 (fig 1.6), forming an inhibited complex that can be partially reactivated transferring the bound PLP to apo-B_6 enzymes, that was suggested to be an important mechanism of activation of the PLP-dependent enzymes (Ghatge et al. 2012).

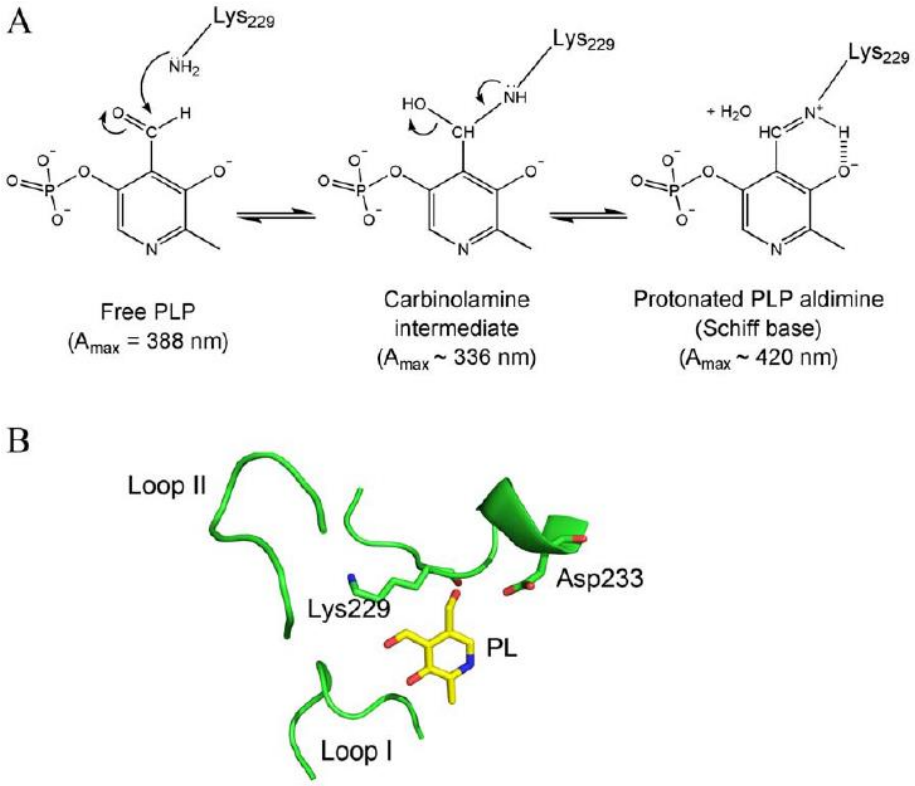


Fig 1.6: **Mechanism of reaction between PLP and the active site K229**
A) A scheme showing the structures of the carbinolamine intermediate and the enolimine form of the protonated PLP aldimine.
B) Active site structure of the binary complex of ePL kinase and PL showing the position of K229. (Ghatge et al., 2012)

PNPOx is also able to bind PLP on second site different from the catalytic one how is been proved by the experiments of Safo and colleagues (Safo et al. 2001). The PLP binds tightly to the second site, in fact experiments lead by Yang and collaborators resulted in the observation that PLP still bind the second site on PNPOx after gel filtration chromatography (Yang and Schirch 2000). The second binding site has not yet been correctly identified but is surely different from the catalytic site, as demonstrated by the experiments performed by Barile and collaborators using fluorometrics measurements, in which they measured the dissociation constants of PLP for *E. coli* WT PNPOx and a quadruple catalytic site-mutant (unable to bind PLP at the active site) (Fig 1.7) obtaining comparable values of the dissociation constants. Furthermore, PLP titration experiments on concentrated solutions of the quadruple mutated PNPOx gave a different binding stoichiometry compared to the WT enzyme (Barile et al, 2019).

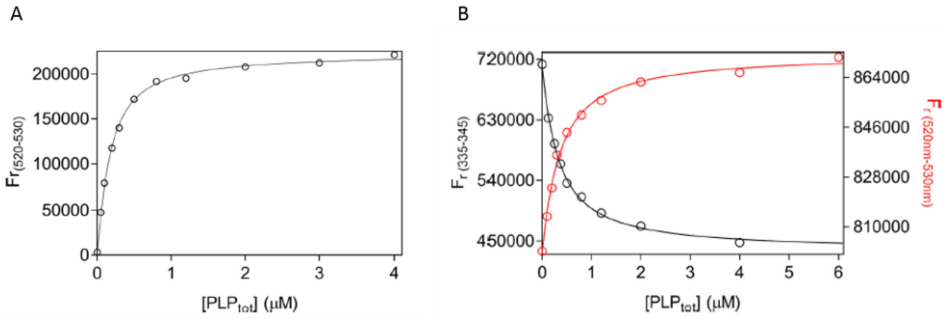


Fig 1.7: PLP binding equilibrium constant analysis

A) PLP equilibrium curve for WT PNPOx. K_d calculated is 150±0.4 nM;
B) PLP equilibrium curve for the apo form of WT PNPOx (in black) and of the quadruple mutant of PNPOx (in red). K_d calculated are 280±0.1 nM and 450±0.2 nM, respectively. (Barile et al, 2019)

In presence of high concentration of PLP, human PNPOx activity is reduced but not totally inhibited showing a residual catalytic capacity contrarly to

what has been observed in the *E. coli* PNPOx that is completely inhibited. The mixed-type inhibition mechanism of the two enzymes is similar, suggesting that the second binding site could be the same on both proteins (Barile et al. 2020).

1.5.1 PLP Binding Protein

Unbalanced levels of free PLP can be toxic in cells and so far, no protein has been assigned specifically to its homeostasis. A good candidate for this important role is the PLP Binding Protein (PLPBP), also called YggS in *E. coli*, PROSC in human and PipY in *S. elongatus* (Darin et al. 2016; Labella et al. 2017; Prunetti et al. 2016).

YggS was identified in *E. coli* as a member of the COG0325 family (Eswaramoorthy et al. 2003; Ito et al. 2013). This protein family shares with such decarboxylases and the alanine racemases a protein folding typical of some PLP-dependent enzymes called Fold-type III (Percudani and Peracchi 2003) and characterized by the $(\alpha/\beta)_8$ barrel structure. The yeast hypothetical protein YBL036C of the family COG0325 have been crystallised in a $(\alpha/\beta)_8$ barrel structure and, also if it was not predicted to be a PLP binding protein, it was covalently bound to a molecule of PLP (Eswaramoorthy et al. 2003; Percudani and Peracchi 2003). These structures have an α-helical extension of the C-terminal β-strand binding the phosphate of PLP, which could act as a trigger for PLP exchange (Labella et al. 2020) (fig 1.8).

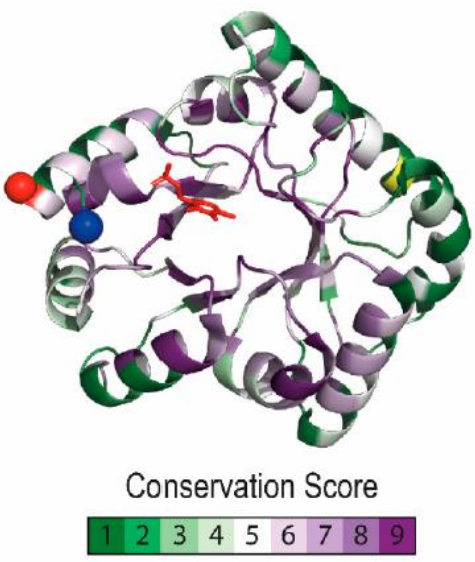


Fig 1.8: Pyridoxal phosphate binding protein (PLPBP) residue conservation and distinctive structural features.
Conservation scores were calculated using CONSURF with COG0325 alignment and a phylogenetic tree from the EGGNOG database as query. The color code illustrates lowest to highest conservation in a green to purple gradient, with yellow for non-informative residues, plotted over the Synpcc7942_2060 (PDB 5NM8) structure. The N-terminal and C-terminal ends are indicated with a red and blue sphere, respectively. The PLP molecule is colored in red. (Labella et al, 2020)

This protein family is highly conserved with proteins ranging between a molecular mass 24-30 KDa, present in almost all the kingdom of life including mammals, bacteria, yeast and plants (Fig 1.9). A similar degree of conservation suggests that these proteins play a crucial role for the cellular function (Ito et al. 2013).

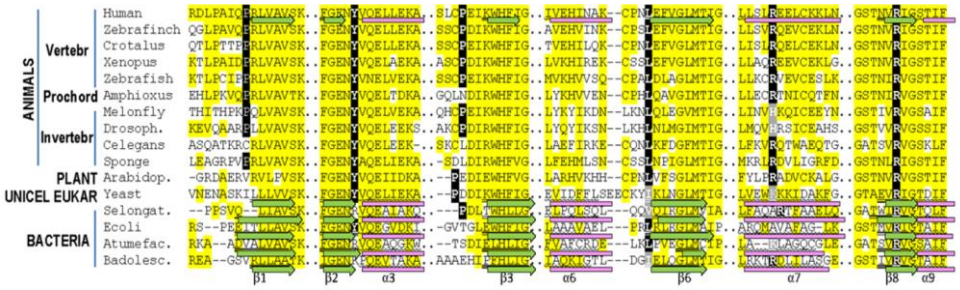


Fig 1.9: **Alignment of short regions of the PLPHP sequence (in single letter amino acid code)**

The individual species and the taxonomic group are given to the left. The Clustal *ω* alignment has been manually modified to allow the stacking of the same secondary structure elements revealed in crystal. (Tremino et al, 2018)

Generally, these proteins are monomeric as revealed by size exclusion chromatography analysis of a purified PipY from *S. elongatus*. Human protein PLPBP expressed in *E. coli* by Tremino and collaborators has been purified and analysed using the size exclusion chromatography revealing the presence of a second pick corresponding to the dimer. A new preparation of the protein in presence of β-mercaptoethanol reduced the formation of the dimeric fraction, suggesting the presence of an equilibrium where the dimer is formed by the covalent bond between two monomers (Tremiño, Forcada-Nadal, and Rubio 2018).

Proteins present in the COG0325 family binds PLP and are structurally similar to other Fold-type III PLP dependent enzymes, however, they don’t show activity against amino acids and the main substrates of the PLP dependent enzymes (Ito et al. 2013).

A PLP homeostasis role was hypothesized for PLP binding proteins because a knockout of the genes coding for these proteins led to different phenotypes such as: PN toxicity and an imbalance in the amino/ketoacids pools in *E. coli*,

sensitivity to antibiotics that target essential PLP holoenzymes in *S. elongatus* and, in human, mutation on the gene coding for PLPBP cause vitamin B_6 dependent epilepsy (Darin et al. 2016; Labella et al. 2020; Plecko et al. 2017; Tremiño et al. 2018).

Deletion of *yggS* gene in *E. coli* leads to a pyridoxine sensitiveness as observed by Prunetti and collaborators that can be complemented by an *in trans* expression of *Zea mays* COG0325 family protein (Fig 1.10). The recovery of the ΔyggS phenotype using a homologous protein was studied independently by Ito and Darin with the respective collaborators. The amino acids relative concentrations in *E. coli* are modified by *yggS* knockout and this phenotype can be reverted, along with the phenomenon of the PN toxicity, through complementation with other PLP Binding Protein encoding genes as human (PLPBP), yeast (YBL036C) and *Bacillus subtilis* (*ylmE*) variants suggesting a conserved role for these proteins (Darin et al. 2016; Ito et al. 2013). PN toxicity was associated to a toxic accumulation of PNP due to the loss of activity of PNPOx in ΔyggS background, but purified PNPOx from ΔyggS *E. coli* have shown full activity (Prunetti et al. 2016).

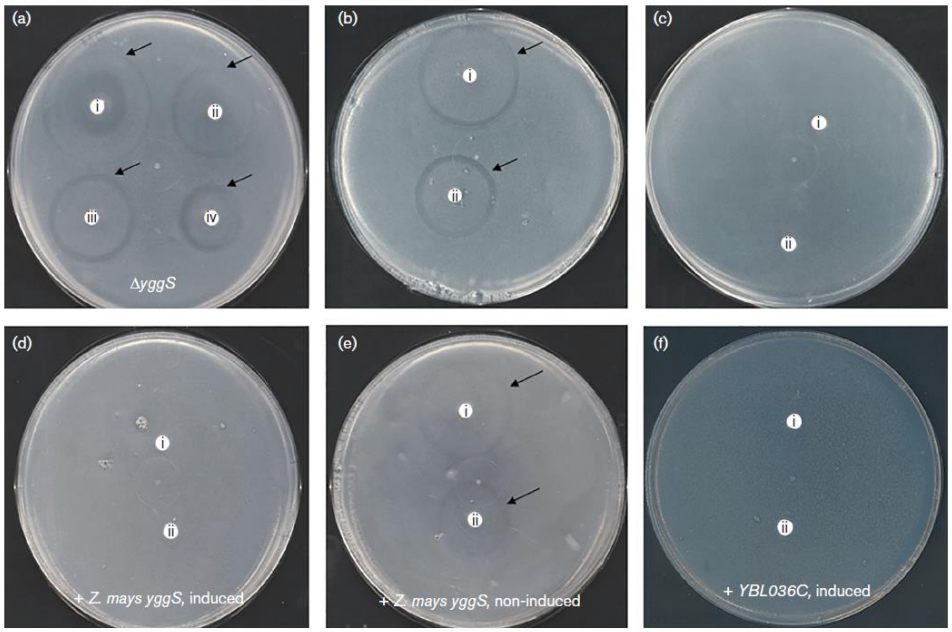


Fig 1.10: Toxic effect of PN on the E. coli ΔyggS strain and the universality of YggS function.
(a) DyggS (VDC6594) lawn and drops of PN at concentrations of 5.9 mM (i), 590 mM (ii), 295 mM (iii) or 59 mM (iv);
PN at concentrations of 5.9 mM (i) or 590 mM (ii) was used in experiments (b–f). (b) ΔyggS pBAD18 lawn in presence of 0.2 % arabinose; (c) ΔyggS in pBAD24yggSEc (pBY291.3) lawn in presence of 0.2 % arabinose; (d) ΔyggS transformed with pBAD18 LOC100191932 (pBY298.7) lawn in presence of 0.2 % arabinose; (e) ΔyggS transformed with pBAD18 LOC100191932 in absence of 0.2 % arabinose; (f) ΔyggS transformed with pBEY YBL036C (pBEY329.12) in the presence of 50 ng aTet ml21.
The arrows indicate the presence of the ring of toxicity. (Prunetti et al, 2016)

Although in several organism PLPBP is not essential for life (Ito et al. 2013; Labella et al. 2017; Prunetti et al. 2016), in humans a court of people has been reported thus far with vitamin B6 dependent epilepsy associated with PLPBP mutation (Darin et al. 2016; Plecko et al. 2017). The mutations found in these patients are both truncating and splicing mutation that lead to PLPBP

loss of function with early in life nervous system pathology as epilepsy with developmental delay if not treated with vitamin B_6 supplementation (Tremiño et al. 2018).

Given its diffusion between the living beings and the conservation of PLPBP it is clear that this protein, whose exactly function has not been yet unveiled, has a crucial role in the organisms and given the phenotype when it is mutated, PLPBP is probably directly involved in PLP metabolism.

It was hypothesised the role of PLPBP as a "sponge" for the homeostasis of the PLP in order to have a reservoir, but this role must be confirmed.

Recent studies of Vu and collaborators, have suggested a role of YggS in *Salmonella enterica* as an enhancer of the PLP-PMP cycling on PNPOx limiting the PLP inhibition effect on the oxidase (Vu, Ito, and Downs 2020).

1.6 Aim of the thesis

Pyridoxal 5'-phosphate (PLP) is the bioactive form of vitamin B_6 and is involved in more than 150 enzymatic reactions. This cofactor is involved in several metabolic pathways, such as the synthesis of neurotransmitters, amino acid metabolism, glucose homeostasis, and heme metabolism. In different organisms, PLP has also biological functions other than catalysis, such as those of reactive oxygen species (ROS) scavenger in plants and transcriptional regulator in Eubacteria.

This work aims to deepening our understanding of vitamin B_6 metabolism in *Escherichia coli* and involves the study of the physiologic effect of deletions of several genes related to PLP metabolism, and the effect of vitamin B_6 vitamers on *E. coli* growth.

PLP-binding protein encoded by the gene *yggS*. The effort of these studies has permitted to explore which is its physiological role of YggS in order to assign a specific function to this protein. The hypothesis is that YggS could be a PLP homeostatic protein that acts as a "sponge" to reduce free PLP concentration, and could also act as a carrier protein, able to transfer PLP to the apo-PLP-dependent enzymes that need it as a cofactor. To this aim, we studied the transfer mechanism and the possible interaction of YggS with the PLP-dependent enzyme serine hydroxymethyltranferase.

2. *Materials and Methods*

Materials

2.1 Materials, bacterial strains, plasmids and reagents

All the chemicals used in this study were purchased from Sigma-Aldrich (Saint Luis, MO, USA) and Carlo Erba (Milan, Italy) except for the chromatographic gels, infact DEAE-sepharose and Phenil-sepharose were purchased from GE Healthcare (Milwaukee, WI, USA) and NiNTA-agarose from Quiagen.

The oligonucleotides used for in situ mutagenesis were synthetized by Metabion international AG (Planegg, Germany) and the DNA sequencing were operated by Microsinth AG (Balgach, Switzerland).

Growth media and supplements

Media utilised for bacterial cultures were previously sterilised using relative temperature and pressure of 120°C and 1 atm for 20 minutes in autoclave.
Bacteria were grown both in liquid and solid medium. Selective liquid medium was only added with the appropriate antibiotic when the temperature decreased under 55°C. Selective solid medium was added of 15g/L of agarose before sterilization and then supplemented with the necessary antibiotic when temperature reached 55°C.
Luria Bertani (LB) growth medium is a complex medium containing all the necessary nutrients for an optimal growth of *E. coli*. LB medium used in our experiments was composed of: 10g/L Tryptone, 5g/L Yeast extract, 5g/L NaCl.
Minimal medium M9 is a defined medium used to grow cells under nutritional control and it was composed of: 0.6% anhydrous Na2HPO4, 0.3%

KH2PO4, 0.05% NaCl, 0.1% NH4Cl, 0.4% glucose, 2 mM MgSO4, 0.1 mM CaCl2.

Growth media were supplemented with the following antibiotic doses: 40µg/mL Kanamycin or 100µg/mL Ampicillin.

Bacterial strains

- *Escherichia coli* DH5α strain was used for transformation and selection of mutants;
- *Escherichia coli* BL21(DE3) was used for the expression of WT YggS because this gene is already present in the genome and eventual recombination does not affect the correct protein expression. This DE3 strain contain a lysogen of the bacteriophage λDE3 under control of the promoter LacUV5 so a T7 RNA polymerase gene can be expressed on request;
- *Escherichia coli* HMS174(DE3) is a non-recombinant strain, in this study was used to express all the mutant forms of YggS. As the previous this is a DE3 *E. coli* strain so, it contains the same T7 RNA polymerase gene;
- *Escherichia coli* GS1993 is a ΔGlyA strain that in this study was used for the expression of *e*SHMT both WT and mutants after its transformation with the plasmid pBS*glyA*;
- *Escherichia coli* Rosetta (DE3)/pET19*h*PLPP.
- *Escherichia coli* GC10/pBAD24*yggS*.
- The following strains of *Escherichia coli* were bought from the Keio collection and used for the growth curves:
    - *Escherichia coli* BW25113
    - *Escherichia coli* BW25113 Δ*yggS*

- *Escherichia coli* BW25113 Δ*pdxH*
- *Escherichia coli* BW25113 Δ*pdxJ*
- *Escherichia coli* BW25113 Δ*pdxK*
- *Escherichia coli* BW25113 Δ*pdxY*

Plasmids

- pET28a(+)/*yggS* WT and mutant forms;
- pBS*glyA* WT and mutant forms;
- pET19/*h*PLPP;
- pBAD24/*yggS* WT
- pBAD24/*yggS* K38A
- pBAD24/*yggS* S132A
- pBAD24/*yggS* R229A

Table 2.1 Oligonucleotides for molecular cloning

YggS	
yggS_NcoI_for	GGCCATGGACGATATTGCGCATAACCTG
yggS_XhoI_rev	GGCTCGAGTTTTTTAGAGTAATCACGCGCACC
YggS N-His_up	CGGCATATGAACGATATTGCGCATAACCTG
YggSN-His_down	GCGAATTCTTATTTTTTAGAGTAATCACGCGCAC
K36A_for	CAGTCAGTgcaACAAAACCTGCGAGCG
K36A_rev	GGTTTTGTtgcACTGACTGCAAGCAGC
K36/38A_for	TGCAACAGCGCCTGCGAGCGCCATC
K36/38A_rev	CTCGCAGGCGCTGTTGCACTGACTG
K89A_for	CAGTCTAATGCGAGCCGCCTGGTGGCA
K89A_rev	CAGGCGGCTCGCATTAGACTGCAACGG
E134A_for	AGTGATgcaAACAGTAAGTCCGG
E134A_rev	CTTACTGTTtgcATCACTAATGTTAA
K137A_for	AACAGTgcgTCCGGGATTCAAC
K137A_rev	CCCGGAcgcACTGTTTTCATC
K233/234A_for	GCGTGATTACTCTGCAGCACTCGAGCACCACC
K233/234A _rev	CGCACTAATGAGACGTCGTGAGCTCGTGGTGG
ΔK233/234A_for	TAAGAATTCGAGCTCCGTCGACAAGCTTGC
ΔK233/234A _rev	CGGAGCTCGAATTCTTAGAGTAATCACGCGC
eSHMT	
D89A_for	TCGGCGCTGCCTACGCTAACGTCCAGCC
D89A_rev	GTTAGCGTAGGCAGCGCCGAACAGTTC
K354A_for	ACGATCCGGCGAGCCCGTTTGTGACCTC
K354A_rev	AAACGGGCTCGCCGGATCGTTCGGTACG

Table 2.2 Oligonucleotides for RT-qPCR

QpdxB_for	ATTGGTTTTTCCGCTGCACC
QpdxB_rev	CGCGCTGGACTAACTCATCC
QecopdxH_for	ATGGTGGTCGCTACCGTGG
QecopdxH_rev	CTATCACGCGGGCGGCTATG
QpdxJ_for	TGACCATATCGCTACGCTGC
QpdxJ_rev	CCAGGCAGCAAAAATGTGGC
QpdxK_for	TGCCGTGCCTGCTATCAAAC
QpdxK_rev	AGGATGGTCTTTGCGTAGCG
QecopdxY_for	AATGGACTGGCTGCGTGATG
QecopdxY_rev	GCACATGAAACTCTGCGACACC
QyggS_for	CATACCATCGACCGTTTGCG
QyggS_rev	GCCATTTGGCGTGCAACTTC
rRNA16S_for	GACGTTACCCGCAGAAGAAGC
rRNA16S_rev	GTGGACTACCAGGGTATCTAATCC

Methods

2.2 Molecular cloning of WT *yggS*

The coding sequence of the *yggS* gene from *E. coli* was amplified by PCR using primers yggS_NcoI_for and yggS_XhoI_rev (Table 2.1) and pBAD24*yggS* as template. The amplified ~700 bp segment was inserted into a pET28b(+) vector between *Nco*I and *Xho*I restriction sites. In this way the *yggS* gene contains the HisTAG coding region at the C-terminal terminus. This new construct, named pET28-YggS, was used to transform *E. coli* BL21(DE3) competent cells for protein expression. The nucleotide sequence of the insert was determined for both strands and no differences were detected with respect to the sequence reported in data banks. In order to produce an YggS protein form with the HisTAG coding region at the N-terminal terminus, the same steps were performed. Specifically, the *yggS* gene were amplified using primers YggS N-His_up and YggS N-His_down and the same pBAD24*yggS* as template. The resulting fragment was inserted into pET28b(+) vector between *Nde*I and *EcoR*I restriction sites. Summarizing, two versions of YggS were produced where the His-TAG is present at the N- and C-terminal, respectively.

2.3 Site-directed mutagenesis

YggS

The *yggS* gene inserted into the pET28-YggS plasmid was used as template in site-directed mutagenesis reactions carried out using two complementary oligonucleotides containing the mutation as primers (K36A_for and K36A_rev, K36/38A_for and K36/38A_rev, K89A_for and K89A_rev, E134A_for and E134A_rev, K137A_for and K137A_rev, K233/234A_for

and K233/234A _rev, ΔK233/234A_for and ΔK233/234A _rev). Sequence of the constructs was verified by DNA sequencing. Expression and purification of YggS mutant forms was performed as described for wild-type YggS.

eSHMT

The *glyA* gene inserted into the pBS/*glyA* plasmid was used as template in site-directed mutagenesis reactions carried out using two complementary oligonucleotides containing the mutation as primers (D89A_for and D89A_rev, K354A_for and K354A_rev).

2.4 RNA analysis through RT-qPCR

RNA was isolated from three independent cultures (three biological replicates) grown up 1 OD_{600} in M9 minimum medium added with 0.1 mM PL, 0.1 mM PN and 0.5% casaminoacids (CAA), using the NucleoSpin RNA kit from Macherey and Nagel. RNA concentration and quality were evaluated by measuring the OD at 260 nm and the ratio 260/280 nm, respectively, and by electrophoresis on 1.2% agarose gels. RT-qPCR reactions were performed in two steps. Reverse transcription of DNase-treated RNAs (0.5 μg) was carried out using the Maxima® first strand cDNA synthesis kit (Thermo Scientific) with the random primers provided in the kit. Real Time PCR was performed on a Real Time PCR Instrument (Mx3000P QPCR system, Agilent technologies) with a two-step reaction using Maxima® SYBR Green qPCR Master mix (Thermo Scientific) and the oligonucleotides reported in Table 2.2. The relative expression of each target gene was determined by the Pfaffl method (Pfaffl, 2001) using the rRNA 16S gene as normalizer. The fold induction resulting from the different samples was reported as averaged

values, and the error (as confidence interval) was measured as standard deviation. The statistical significance of the obtained values was measured using the P-value, calculated using the Student's t-test.

2.5 Expression and purification of the proteins

YggS

An overnight culture (50 mL) of *E. coli* BL21(DE3) cells transformed with plasmid pET28-YggS was used to inoculate 4 L of Luria–Bertani broth containing kanamycin (40 $mg \cdot L^{-1}$). Bacteria were allowed to grow for approximately 5 h at 37 °C (until their OD_{600} reached to ~ 0.6), then the growing temperature was lowered to 28 °C and the expression of YggS induced with 0.2 mM isopropyl thio-β-D-galactopyranoside (IPTG). Bacteria were harvested after 18 h and suspended in 150 mL of 20 mM potassium-phosphate buffer, pH 7.3, containing 5 mM EDTA, PLP and 2 mg/ml lysozyme. Cell lysis was carried out by sonication on ice (3-min in short 20-s pulses with 20-s intervals). Lysate was centrifuged at 12000 g for 30 min and the pellet was discarded. The supernatant was precipitated by addition of ammonium sulphate to 30% saturation, centrifuged, and the resulted pellet was discarded. Ammonium sulfate was added to the resulting supernatant so to reach the 70% saturation concentration. The sample was centrifuged at 12000 g for 30 min and the pellet suspended in the 20 mM potassium-phosphate buffer, pH 7.3 to dilute the ammonium sulfate to a final concentration less than 10%. The sample was loaded onto a 10 ml Ni-NTA Superflow (QIAGEN) column, previously equilibrated with 20 mM potassium phosphate buffer, pH 7.3 containing 10 mM imidazole. The column was washed with 50 mL of the same buffer, 50 mL of the same

buffer containing 20 mM imidazole, and eluted with a linear 20 to 300 mM imidazole gradient (the buffer containing imidazole was adjusted to pH 7.3 with HCl). Fractions containing the expected protein (molecular mass ~26.000 Da) were pooled and dialyzed overnight against 50 mM Na-HEPES buffer pH 7.6, containing 5 mM β-mercaptoethanol. The protein was purified to homogeneity, as judged by SDS-PAGE analysis and the purification yield was about 35 mg YggS per litre of bacterial culture. Recombinant YggS was stored at 4 °C. Protein subunit concentration was calculated using a theoretical extinction coefficient at 280 nm of 16960 M^{-1} cm^{-1} (calculated with the Expasy ProtParam tool). PLP content was calculated after the addition of 0.2M NaOH, measuring the absorbance at 388 nm.

eSHMT

An overnight culture (50 mL) of *E. coli* GS1993 pBS*glyA* cells was used to inoculate 4 L of Luria–Bertani broth containing ampicillin (100 $mg \cdot L^{-1}$). Bacteria were allowed to grow for approximately 18h at 37 °C, then they were harvested and suspended in 120 mL of 30 mM TRIS hydrochloride buffer, pH 7.6, containing 20 mM EDTA, PLP and 2 mg/ml lysozyme. Cell lysis was carried out by sonication on ice (3-min in short 20-s pulses with 20-s intervals). Lysate was centrifuged at 12000 g for 30 min and the pellet was discarded. The supernatant was precipitated by addition of ammonium sulfate to 50% saturation, centrifuged, and the precipitate was discarded. Ammonium sulfate was added to the resulting supernatant so to reach the 75% saturation concentration. The sample was centrifuged at 12000 g for 30 min and the pellet suspended in the 20 mM potassium-phosphate buffer, pH 7.2, containing 100 μM EDTA and 1 mM β-mercaptoethanol. Then the sample was dialyzed overnight against the 20 mM potassium-phosphate buffer, pH 7.2, containing 100 μM EDTA and 1 mM β-mercaptoethanol. The

dialyzed sample was loaded onto a DEAE-Sephadex column which had been equilibrated with 20 mM potassium-phosphate, pH 7.2. The column was washed with the equilibrating buffer until the absorbance at 280 nm was below 0.2. The *e*SHMT enzyme was then eluted with a linear salt gradient (0 to 400 mM sodium chloride). The mixing chamber contained 300 ml 20 mM potassium phosphate, pH 7.2, and the reservoir contained 200 ml of 100 mM potassium phosphate with 400 mM sodium chloride, pH 6.4. The yellow fractions, showing the absorbance bands at 420 nm (corresponding at the PLP internal aldimine) were pooled, and the protein was precipitated by the addition of ammonium sulfate to 75% of saturation. After centrifugation, the precipitate was dissolved in 50 mM potassium-phosphate buffer, pH 7.2, and loaded onto a phenyl-sepharose column which had been equilibrated with the same buffer containing 40% ammonium sulfate. The protein was then eluted with a linear 40% to 0 ammonium sulfate gradient. The yellow fractions were pooled and ammonium sulfate was added to the protein sample to the 75% saturation concentration. Then the precipitate was dissolved in 20 mM potassium-phosphate buffer, pH 7.2, containing 200 μM DTT and 100 μM EDTA and dialysed over night against the same buffer. The purified protein was pure as determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Protein subunit concentration (~45.000 Da M.W.) was calculated using a theoretical extinction coefficient at 280 nm of 44884 M^{-1} cm^{-1} and 42790 (calculated with the Expasy ProtParam tool) for the holo and apo forms, respectively. PLP content was calculated after the addition of 0.2M NaOH, measuring the absorbance at 388 nm.

hPLPP

An overnight culture (50 mL) of *E. coli* Rosetta (DE3)/pET19*h*PLPP cells was used to inoculate 4 L of Luria–Bertani broth containing ampicillin (100

mg·L^{-1}). Bacteria were allowed to grow for approximately 5 h at 37 °C (until their OD_{600} reached to ~ 0.6), then the growing temperature was lowered to 28 °C and the expression of *h*PLPP induced with 0.2 mM isopropyl thio-β-D-galactopyranoside (IPTG). Bacteria were harvested after 18 h and suspended in 50 mM sodium-phosphate buffer, pH 8, containing 300 mM sodium chloride and 2 mg/ml lysozyme. Cell lysis was carried out by sonication on ice (3-min in short 20-s pulses with 20-s intervals). Lysate was centrifuged at 12000 g for 30 min and the pellet was discarded. The supernatant was loaded onto a 10 ml Ni-NTA Superflow (QIAGEN) column, previously equilibrated with 50 mM sodium-phosphate buffer, pH 8.0 containing 300 mM sodium chloride and 10 mM imidazole. The column was washed with 50 mL of the same buffer and the protein was eluted with a linear 10 to 300 mM imidazole gradient. Fractions containing the expected protein (molecular mass ~31.000 Da) were pooled and dialyzed overnight against 50 mM Na-HEPES buffer pH 7.2, containing 150 mM sodium chloride. The protein was purified to homogeneity, as judged by SDS-PAGE analysis. Protein subunit concentration was calculated using a theoretical extinction coefficient at 280 nm of 18450 M^{-1} cm^{-1} (calculated with the Expasy ProtParam tool).

2.6 Preparation of the apo-forms of the proteins

Both apo-YggS and apo-SHMT were prepared using the same protocol. In order to remove the molecules of PLP bound through a Schiff base to the protein YggS and SHMT, 40 mg of both the proteins have been treated separately with L-Cysteine. Thanks to the reactivity of the PLP with the thiol group of the cysteine, there is the formation of a compound called Thiazolidine (Fig. 2.1) that is more stable than the internal aldimine between Lysine-PLP but less affine so it dissociates from the protein.

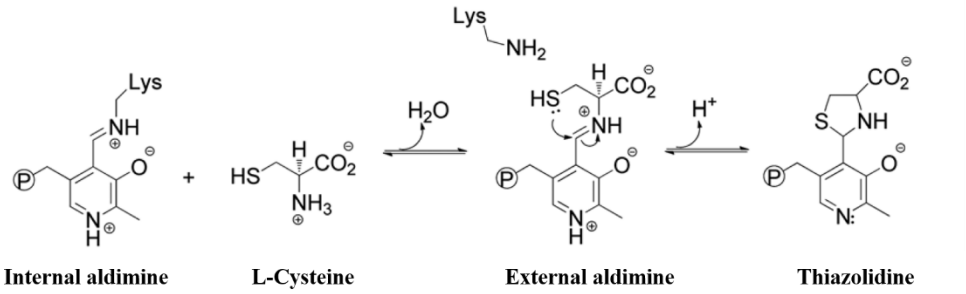


Figure 2.1 Reaction scheme showing the reaction between PLP bound to a lysine residue and L-cysteine. (Lowther et al, 2012)

An amount of 40 mg of purified YggS or SHMT was added to the buffer 20 mM potassium-phosphate (KPi), pH 7.3, containing 5 mM β-mercaptoethanol, 25% ammonium sulphate. The solution was added with 100 mM L-cysteine and kept under mixing until it became clear.

The protein solution was loaded in a chromatographic column containing 7 mL of Phenil Sephadex (GE Healtcare) and washed with the same buffer containing only 50 mM L-cysteine. A second wash of the chromatographic column was done with the same buffer without L-cysteine in order to separate the protein from the thiazolidine. The separation of the thiazolidine were monitored measuring the absorption spectrum at 333 nm, the major absorbance wavelength for thiazolidine.

Elution of the apo-protein happened using the buffer 20mM potassium-phosphate (KPi), pH 7.3, 5 mM β-mercaptoethanol and the fractions containing the protein were selected using spectrophotometric and SDS-PAGE analysis and collected. The obtained proteins were dialyzed overnight against 50 mM NaHEPES buffer pH 7.6 containing 5 mM β-mercaptoethanol.

2.7 Size Exclusion Chromatography

Gel filtration of YggS wild-type and mutant forms was performed on a Superdex 200 10/300 GL column (GE Healthcare, Little Chalfont, UK) at room temperature and at a flow rate of 0.5 ml/min in 50 mM Na-HEPES buffer pH 7.6, containing 150 mM NaCl and 0.5 mM DTT. Elution profiles were obtained from absorbance at 280 nm. Calibration was performed with aldolase (158 kDa) conalbumin (75 kDa) and ovalbumin (44 kDa), carbonic anhydrase (29 kDa), RNase (17.7 kDa). Protein samples (100 µl) were subjected to chromatography at concentration of different concentrations.

2.8 Spectroscopic measurements

Concerning YggS, the spectrophotometric measurements were carried out at 20 °C in 50 mM Na-HEPES buffer pH 7.6. UV-visible spectra were recorded with a Hewlett-Packard 8453 diode-array spectrophotometer. Far-UV (190-250 nm) CD spectra were measured with a Jasco 710 spectropolarimeter equipped with a DP 520 processor using 0.2 and 1 cm path length quartz cuvettes and results were expressed as ellipticity [θ]. Far UV measurements were carried out in 20 mM KPi buffer pH 7.6. Protein samples were heated from 30° C and 80°C, with a heating rate of 1°C per min controlled by a Jasco programmable Peltier element, monitoring the dichroic activity at 220 nm every 0.5 °C. Denaturation curves were fitted with the sigmoidal trace described in the **Eqn. 1**.

2.9 Determination of the dissociation constant of PLP from YggS

PLP binding equilibria were analyzed taking advantage of the protein intrinsic fluorescence quenching observed upon the binding event. The dissociation constant of the binding equilibrium was then calculated from saturation curves obtained measuring the protein fluorescence emission intensity as a function of increasing PLP concentration. The cofactor (from 0 to 2 μM and from 0 to 15 μM for wt and mutant forms, respectively) was added to apoenzyme samples (50 nM wt and 200nM mutants) at 25 °C in 50 mM Na-HEPES pH 7.6. Preliminary experiments showed that the binding equilibrium was established within the mixing time. Fluorescence emission spectra (300 - 450 nm; 5 nm emission slit) were recorded after mixing PLP and apoenzyme with a FluoroMax-3 Jobin Yvon Horiba spectrofluorimeter, with excitation wavelength set at 280 nm (3 nm excitation slit), with a 1 cm path length quartz cell. For data analysis, we have removed the contribution of PLP at the fluorescence spectrum of the protein samples. Data were analyzed according to a quadratic equation, **Eqn. 2.**

2.10 Kinetic studies

2.10.1 PLP transfer studies

PLP transfer studies were performed with two different methods because the mutation on lysine 354 of the eSHMT cause a loss of activity probably due to the interaction of this residue with the tetrahydrofolate.

The kinetics were studied using a Hewlett-Packard 8453 diode-array spectrophotometer and monitoring in both the assay the variation of absorbance at 340 nm over time.

PLP transfer kinetic assay with L-serine and THF

For this assay 20 µM apo-YggS and 20 µM PLP have been mixed mixed in a vial, then in a different vial have been mixed 70 µM apo-eSHMT and the wanted concentration of hPLPP. All the volumes in the vials were adjusted to have 200 µL of PLP transfer reaction and keep at 37 °C for all the experimental time.

The cuvette has been filled with 598 µM of: 50mM NaHEPES pH 7.6, 25mM L-Serine, 50 µM Tetrahydrofolate (THF), 0,5 µM Methyltetrahydrofolate dehydrogenase (Mtd), 250 µM $NADP^+$.

At different times, the absorbance variation was measured for one minute in order to have a baseline given by the oxidation of the THF; after the minute, 2 µL of the protein reaction mix that was prepared were added into the cuvette and after 3 minutes have been measured the initial velocity of the eSHMT. This enzyme catalyses the reduction of $NADP^+$ to NADPH in order to convert L-Serine in Glycine. The coupled reaction uses Mtd to convert

methyltetrahydrofolate and regenerates $NADP^+$ causing an absorbance increase at 340nm.

The initial velocity analysed in a graph over time of measurement, permit to study the kinetic of the reactivation of the apo-eSHMT.

PLP transfer kinetic assay with L-*allo*-threonine

For this assay 20 µM apo-YggS and 20 µM PLP have been mixed mixed in a vial and in a different vial have been mixed 20 µM apo-eSHMT. All the volumes in the vials were adjusted to have 200 µL of PLP transfer reaction and keep at 37 °C for all the experimental time.

The cuvette has been filled with 558 µL of:

50mM NaHEPES pH 7.6, 10 µL of NADH (8 mg/mL), 10 µL of Alcool dehydrogenase (ADH) (4 mg/mL), 20 µL of L-*allo*-threonine (0,5M).

At different times, the absorbance variation was measured for one minute in order to have a baseline; after the minute, 2 µL of the protein reaction mix that was prepared were added into the cuvette and after 3 minutes have been measured the initial velocity of the eSHMT. This enzyme has a threonine aldolase activity and through the oxidation of NADH to NAD^+ it converts threonine in ethanol and glycine. The coupled reaction uses ADH to convert ethanol in acetaldehyde and regenerates NAD^+ to NADH causing an absorbance decrease at 340nm.

The initial velocity analysed in a graph over time of measurement, permit to study the kinetic of the reactivation of the apo-eSHMT.

2.11 Data analysis

Eqn.1: $$\theta_{220} = \left(\Delta\theta \frac{T^n}{T_m^n + T^n}\right) + const$$

θ_{220}: measured ellipticity at 220 nm

$\Delta\theta$: maximum ellipticity change

T: temperature expressed in Celsius

T_m: apparent melting temperature

n: steepness of the sigmoid curve

Eqn.2:

$$F_{rel} = F_0 - \left(F_0 - F_{inf}\right) \times \frac{1 - \frac{[PLP] + [P] + K_d + [PLP]_x - \sqrt{(-K_d - [PLP] - [P] - [PLP]_x)^2 (-(4 * ([P] * [PLP]) + [P] * [PLP]_x)}}{2}}{[P]}$$

F_{rel}: relative fluorescence measured

$F_{0:}$ fluorescence in the absence of PLP

F_{inf}: fluorescence at infinite PLP concentration

[P]: total protein subunit concentration

[PLP]: residual cofactor concentration bound to the apo-protein

$[PLP]_x$: PLP concentrations used for the titration

K_d: dissociation constant of the equilibrium HOLO ⇆ APO + PLP

2.12 Software used for figures and graphs

All the figures and the graphs reported in this work (except for those taken from the literature and whose reference are reported in the legends) were made using the following software or online services:

- Schemes and enzymatic pathways were realized using Microsoft Office PowerPoint;
- Molecular structures were drawn using Chemdraw 2D;
- The graphs of the reported data were made using GraphPad Prism 7;
- The promoter sequence alignment reported in Figure 3.23 was made using MeMe Suite.

Results

3. Regulation of vitamin B_6 metabolism in *E. coli*

Given the importance of the pyridoxal 5' phosphate as a cofactor for the most different reactions in *E. coli*, and given the importance of its homeostasis and regulation due to the toxicity of its free form, in this work we decided to undertake several studies aimed to a better understanding of the mechanisms of regulation of the intracellular concentration of this molecule.

In the following chapters, we studied the different possible mechanism of regulation of PLP. Starting from the identification of *E. coli* proteins that bind PLP, and elucidating the structures of the genes coding for them. Then, we analyzed the transcriptional levels of the previously identified genes, and hypotesized a possible regulation system that involve the promoter sequences of these genes. As last, we have performed an analysis on the translational level of the genes, and the relative protein abundance in the cell.

3.1 PLP-binding proteins in *E. coli*

Almost 1.5% of all genes in most prokaryotic genomes encode PLP-dependent enzymes (Percudani and Peracchi 2003). They catalyse many different reactions, such as transamination, decarboxylation, racemization, side chain cleavage, β- and γ-elimination or replacement (John 1995). Using the B_6 database (Percudani and Peracchi 2003) (http://bioinformatics.unipr.it/B6db), a tool for the description and classification of vitamin B_6-dependent enzymatic activities and of the corresponding protein families, all PLP-binding proteins found in *Escherichia coli* MG1655 proteome (57 in total) were retrieved and summarized in Table 3.1. Twenty enzymes are involved in amino acid

biosynthesis, while 10 are responsible for the biosynthesis of other molecules, such as coenzymes, polysaccharides, or polyamines. Many are catabolic enzymes, and also a putative PLP carrier protein (YggS, see below) and two PLP-dependent transcriptional regulators (YdcR and YjiR, see below) were reported. Importantly 26% of these proteins are essential for *E. coli* life, being the related knock out strains unable to grow in minimal medium. This indicates how PLP is important for bacterial life.

Table 3.1. PLP-binding proteins of *E. coli.*

Cyan: enzymes involved in amino acid biosynthesis; **Yellow**: enzymes involved in the biosynthesis of molecules as cofactors, polysaccharides and polyamines; **Orange**: different catabolic enzymes; **Green**: enzymes involved in iron-sulphur cluster assembly; **Purple**: PLP-dependent transcriptional regulators; **Grey**: a putative PLP carrier protein; **White**: putative amino transferases.

Protein name	**Activity**	**UniProt**	**EC**	**Biological process**	**Growth in minimal medium**
IlvE	branched-chain-amino-acid aminotransferase	P0AB80	2.6.1.42	Branched-chain amino acid biosynthesis	No
IlvA	Threonine deaminase	P04968	4.3.1.19	Branched-chain amino acid biosynthesis	No
CysK	cysteine synthase A	P0ABK5	2.5.1.47	Cys biosynthesis	No
CysM	cysteine synthase B	P16703	2.5.1.47	Cys biosynthesis	Yes
TrpB	tryptophan synthase, β	P0A879	4.2.1.20	Trp biosynthesis	No
LysA	diaminopimelate decarboxylase	P00861	4.1.1.20	Lys biosynthesis	No

ArgD	N-acetylornithine aminotransferase / N-succinyldiaminopimelate aminotransferase	P18335	2.6.1.11	Arg and Lys biosynthesis	Yes
ThrC	threonine synthase	P00934	4.2.3.1	Thr biosynthesis	No
MetB	O-succinylhomoserine(thiol)-lyase	P00935	2.5.1.48	Cys and Met biosynthesis	No
MetC	cystathionine β-lyase / L-cysteine desulfhydrase / alanine racemase	P06721	4.4.1.28	Met biosynthesis	No
AlaC	glutamate—pyruvate aminotransferase	P77434	2.6.1.2	Ala biosynthesis	Yes
AlaA	glutamate—pyruvate aminotransferase	P0A959	2.6.1.2	Ala biosynthesis	Yes
TyrB	tyrosine aminotransferase	P04693	2.6.1.57	Asp, Tyr, Leu, Phe biosynthesis	Yes
HisC	histidinol-phosphate aminotransferase	P06986	2.6.1.9	His biosynthesis	No
MalY	cystathionine β-lyase	P23256	4.4.1.13	Met biosynthesis	Yes
AspC	aspartate aminotransferase	P00509	2.6.1.1	Phe biosynthesis	Yes
LtaE	L-threonine aldolase	P75823	4.2.1.48	Gly biosynthesis	Yes
Alr	alanine racemase 1	P0A6B4	5.1.1.1	D-Ala biosynthesis	Yes
DadX	alanine racemase 2	P29012	5.1.1.1	D-Ala biosynthesis	Yes
AvtA	valine—pyruvate aminotransferase	P09053	2.6.1.66	Ala and Val biosynthesis	Yes

SerC	phosphoserine/phosphohydroxythreonine aminotransferase	P23721	2.6.1.52	Ser, Lys and PLP biosynthesis	No
SpeA	arginine decarboxylase, biosynthetic	P21170	4.1.1.19	Polyamine biosynthesis	Yes
SpeC	ornithine decarboxylase	P21169	4.1.1.17	Spermidine and putrescine biosynthesis	Yes
SelA	selenocysteine synthase	P0A821	2.9.1.1	Selenocysteine biosynthesis	Yes
WecE	dTDP-4-dehydro-6-deoxy-D-glucose transaminase	P27833	2.6.1.33	Polysaccharide biosynthesis	Yes
ArnB	UDP-4-amino-4-deoxy-L-arabinose aminotransferase	P77690	2.6.1.87	Polysaccharide biosynthesis	Yes
PabC	aminodeoxychorismate lyase	P28305	4.1.3.38	Folate biosynthesis	Yes
BioF	8-amino-7-oxononanoate synthase	P12998	2.3.1.47	Biotin biosynthesis	No
BioA	adenosylmethionine-8-amino-7-oxononanoate aminotransferase	P12995	2.6.1.62	Biotin biosynthesis	No
HemL	glutamate-1-semialdehyde aminotransferase	P23893	5.4.3.8	Tetrapyrrole biosynthesis	No also in LB
TdcB	catabolic threonine dehydratase	P0AGF6	4.3.1.19	Ser and Thr catabolism	No
DsdA	D-serine ammonia-lyase	P00926	4.3.1.18	Ser and Thr catabolism	Yes

Kbl	2-amino-3-ketobutyrate coenzyme A ligase	P0AB77	2.3.1.29	Thr catabolism	Yes
AstC	succinylornithine transaminase	P77581	2.6.1.81	Arg and ornithine metabolism	Yes
GlyA	Serine hydroxymethyltransferase	P0A825	2.1.2.1	Gly, Ser and tetrahydrofolate metabolism	No
GadA	glutamate decarboxylase A	P69908	4.1.1.15	Glu metabolism	Yes
GadB	glutamate decarboxylase B	P69910	4.1.1.15	Glu metabolism	Yes
GcvP	glycine decarboxylase	P33195	1.4.4.2	Gly metabolism	Yes
TnaA	tryptophanase / L-cysteine desulfhydrase	P0A853	4.1.99.1	Trp metabolism	Yes
CadA	lysine decarboxylase 1	P0A9H3	4.1.1.18	Lys metabolism	Yes
LdcC	lysine decarboxylase 2	P52095	4.1.1.18	Lys metabolism	Yes
YgeX	2,3-diaminopropionate ammonia-lyase	P66899	4.3.1.15	D-Ser catabolism	Yes
DcyD	D-cysteine desulfhydrase	P76316	4.4.1.15	D-amino acid metabolism	Yes
GabT	4-aminobutyrate aminotransferase	P22256	2.6.1.19	Gamma-aminobutyrate metabolism	Yes
PuuE	4-aminobutyrate aminotransferase	P50457	2.6.1.19	Gamma-aminobutyrate metabolism	Yes
PatA	putrescine aminotransferase	P42588	2.6.1.82	Putrescine catabolism	Yes

YbdL	methionine-oxo-acid transaminase	P77806	2.6.1.88	L-kynurenine metabolism	Yes
EpmB	lysine 2,3-aminomutase	P39280	5.4.3.2	Post-translational modification	Yes
GlgP	Glycogen phosphorilase	P0AC86	2.4.1.1	Glycogen metabolism	Yes
MalP	maltodextrin phosphorylase	P00490	2.4.1.1	Glycogen metabolism	Yes
SufS	L-cysteine desulfurase	P77444	2.8.1.7	Iron-sulfur cluster assembly	Yes
CsdA	cysteine sulfinate desulfinase	Q46925	2.8.1.7	Iron-sulfur cluster assembly	Yes
YdcR	MocR-like transcription factor	P77730		Regulation of transcription	Yes
YjiR	MocR-like transcription factor	P39389		Regulation of transcription	Yes
YggS	PLP homeostasis protein	P67080		PLP binding	Yes
YhfX	Putative	P45550			Yes
YhfS	putative aminotransferase	P45545			Yes

3.2 *E. coli* genes involved in PLP homeostasis

In the *E. coli* genome, almost all genes encoding enzymes involved in PLP biosynthesis and salvage pathways are part of complex operons. Because bacterial operons often contain genes with related functions, it has been proposed that the inclusion of PLP biosynthetic genes in complex, multifunctional operons may act genetically to integrate PLP biosynthesis into several branches of intermediary metabolism (Lam and Winkler 1992).

In the following paragraphs all the operons and genes having something to do with PLP homeostasis are described. These genes are spaced one each other, spread out all over the *E. coli* chromosome. Only *pdxH* and *pdxY* are so close to each other to be co-transcribed (see below).

3.2.1 Genes involved in PLP biosynthesis

As stated in the Introduction, *E. coli* synthesizes PLP through the so-called deoxyxylulose-5-phosphate (DXP)-dependent pathway, using the primary metabolites erythrose-4-phosphate, glyceraldehyde-3-phoshpate, and pyruvate (Fig. 3.1).

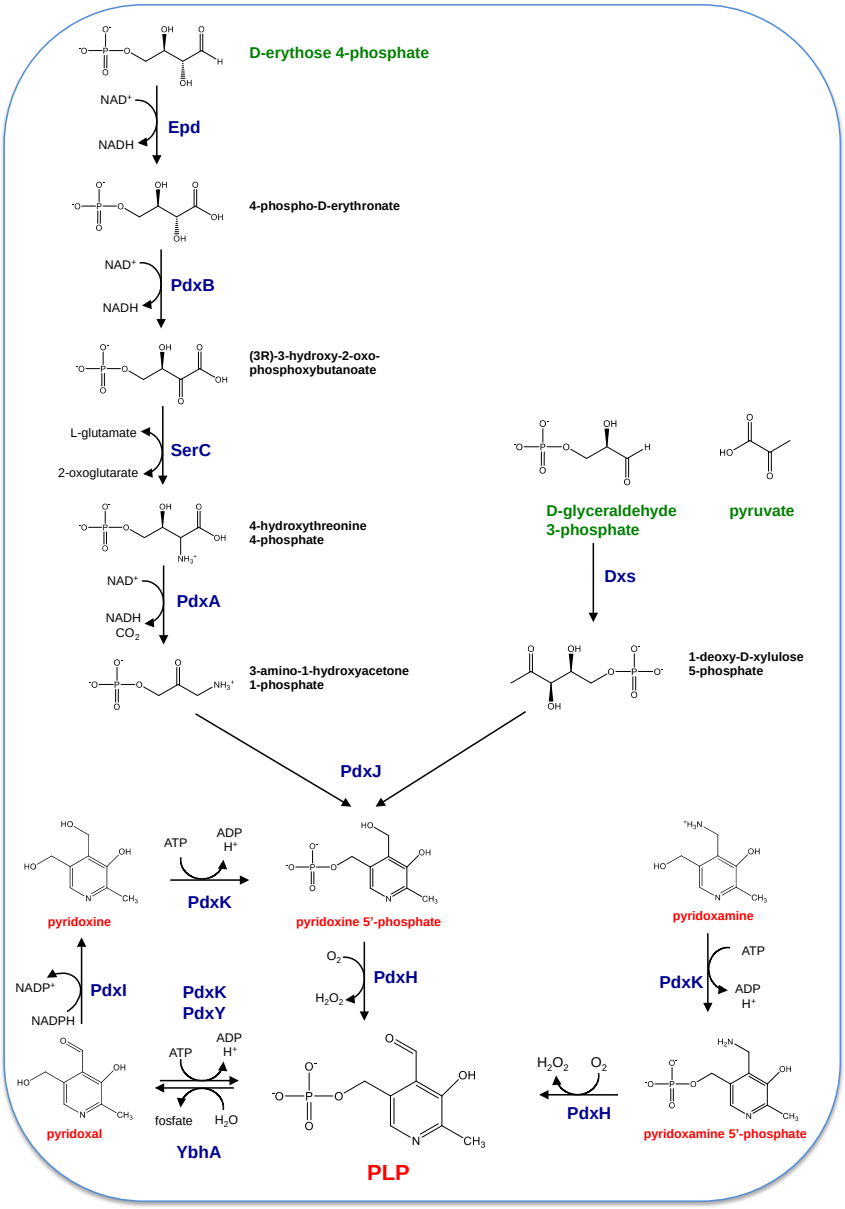


Figure 3.1: Scheme of PLP biosynthesis and salvage pathways. Primary metabolites, erythrose-4-phosphate, glyceraldehyde-3-phoshpate, and pyruvate, used as substrate are in green, whereas intermediate metabolites are in black. The enzymes involved are in blue.

epd

The *epd (or gapB)* gene located between 3,072,672 and 3,073,691 kb on the bacterial chromosome encodes D-erythrose-4-phosphate dehydrogenase (UniProt P0A9B6). This tetrameric enzyme catalyses the first reaction in the *de novo* PLP biosynthesis: the NAD-dependent oxidation of D-erythrose-4-phosphate (Fig. 3.2). This enzyme is homologous to glyceraldehyde 3-phosphate dehydrogenase (GAPDH, encoded by the *gapA* gene), the key enzyme of the glycolysis and gluconeogenesis pathways, with which it shares more than 40% amino acid sequence identity (Zhao et al. 1995). Nevertheless, Epd shows an efficient non-phosphorylating erythrose-4-phosphate dehydrogenase activity but a low phosphorylating glyceraldehyde-3-phosphate dehydrogenase activity, whereas GAPDH shows a highly efficient phosphorylating glyceraldehyde-3-phosphate dehydrogenase activity and a low phosphorylating erythrose-4-phosphate dehydrogenase activity (Boschi-Muller et al. 1997). However, this later low activity of the highly abundant *gapA*-encoded GAPDH can allow growth of a *epd* mutant strain (Yang et al. 1998b).

The *epd* gene is co-transcribed either with *pgk* and *fbaA* or only with *pgk.* (Fig. 3.2) (Bardey et al. 2005). *pgk* and *fbaA* are essential genes encoding the glycolytic enzymes phosphoglycerate kinase (UniProt P0A799) and fructose-bisphosphate aldolase (UniProt P0AB71), respectively. The expression of *epd* is actually disjointed from that of *pgk* and *fbaA*, being these latter genes also transcribed from 3 promoters located downstream *epd* (Bardey et al. 2005).

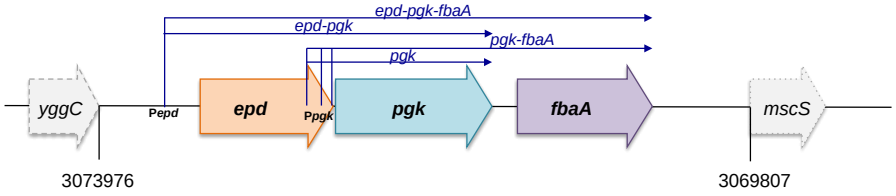


Fig 3.2: Schematic representations of the operon containing *epd*. Numbers are the genome positions in kb. The blue arrows represent mRNAs. The position of promoters is indicated.

pdxB

The *pdxB* gene (located between position 2,436,715 and 2,437,851 kb on the *E. coli* chromosome) encodes 4-phosphoerythronate dehydrogenase (UniProt P05459), which catalyses the second reaction of the PLP biosynthetic pathway, consisting in the NAD-dependent oxidation of 4-phosphoerythronate to 2-oxo-3-hydroxy-4-phosphobutanoate (Fig. 3.1). This protein is evolutionary homolog to D-3-phosphoglycerate dehydrogenase (Fig. 3.3), encoded by *serA*, which catalyzes the first step of the major serine biosynthesis pathway (Schoenlein et al. 1989).

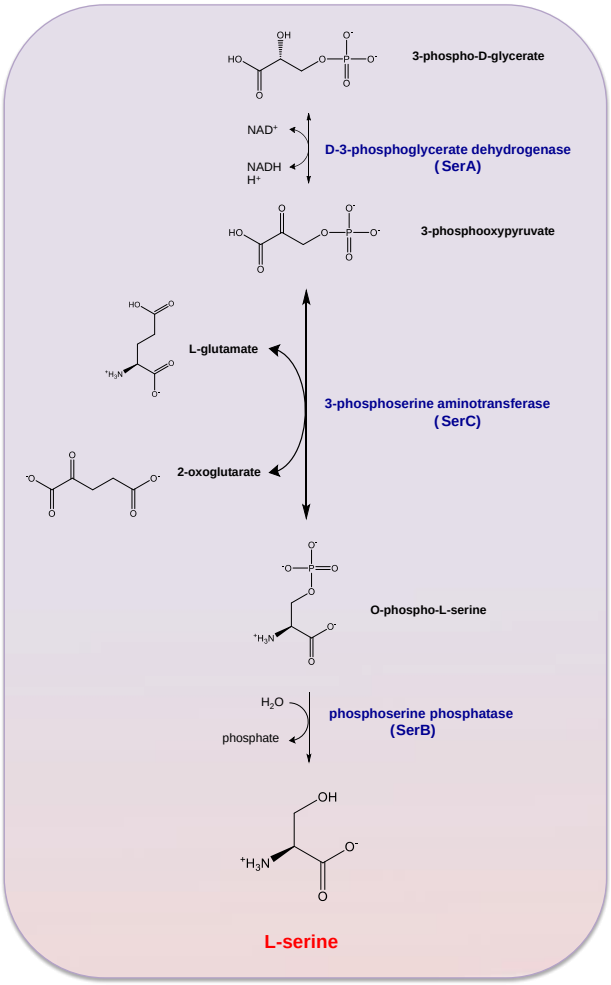


Fig 3.3: Scheme of the major serine biosynthesis pathway.

In *E. coli*, *pdxB* is the first gene of a multifunctional operon, which contains other three genes, *usg*, *truA* and *dedA* (Fig. 3.4). The *usg* gene encodes a putative semialdehyde dehydrogenase (UniProt P08390), as suggested by the sequence similarity to aspartate β-semialdehyde dehydrogenase (Henikoff and Wallace 1988). The *truA* gene product is tRNA pseudouridine synthase I (UniProt P07649), which catalyses the modification of specific uridine residues in the anticodon stem and loop of certain tRNA species. The resulting pseudouridine residues could improve the efficiency of translation

by stabilizing the structure of tRNAs and ensuring reading frame maintenance (Urbonavicius et al. 2001). The last gene of the *pdxB* operon is *dedA*, encoding an uncharacterized inner membrane protein, member of the DedA family.

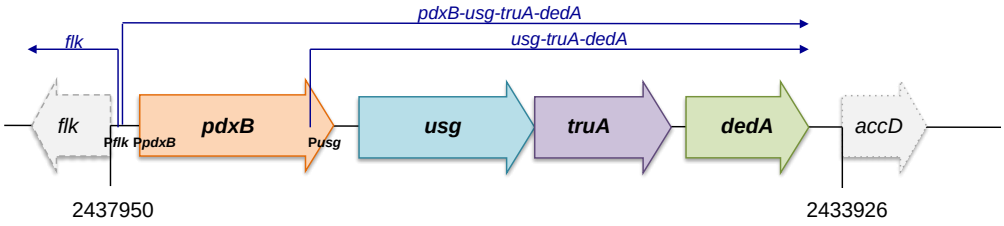


Fig 3.4: Schematic representations of the operon containing *pdxB*. Numbers are the genome positions in kb. The blue arrows represent mRNAs. The position of promoters is indicated

Of these four genes, only *pdxB* is absolutely required for growth of bacteria on minimal medium, being the *pdxB* mutant auxotrophic for PN (Arps and Winkler 1987a). Although these genes form an operon, their functions appear to be unrelated to each other. However, the close association of *pdxB* and *truA* is evolutionarily conserved among different enterobacterial species (Arps and Winkler 1987a).

Two different transcription units were identified: the first unit includes all four genes forming the operon and it is transcribed from the P_{pdxB} promoter, whereas the other unit contains the last three genes and is transcribed from an internal promoter (P_{usg}) located near the end of the *pdxB* coding region (Arps and Winkler 1987b). The P_{pdxB} promoter shares the same -10 region with an overlapping divergent promoter which controls the expression of the *flk* gene encoding a flagellar regulator (UniProt P15286) (Schoenlein et al. 1989;

Pease et al. 2002). However, any functional association between the *pdxB* and *flk* genes has yet to be found.

serC

The *serC* gene (located between position 957,653 and 958,741 kb on the *E. coli* chromosome) codes for the phosphoserine/phosphohydroxythreonine aminotransferase (UniProt P23721), an enzyme which is involved in both PLP and serine biosynthesis by using different substrates (Lam and Winkler 1990). In the PLP biosynthesis pathway, SerC adds an amino group to 2-oxo-3-hydroxy-4-phosphobutanoate (produced by PdxB) to form 3-hydroxy-4-phospho-L-threonine (Fig. 3.1); whereas in the serine pathway, SerC adds an amino group to 3-phosphohydroxypyruvate to form 3-phosphoserine (Fig. 3.3). In both transamination reactions the amino-group donor is L-glutamate, which is then converted into 2-oxoglutarate. The enzyme does not show any activity with non-phosphorylated substrates (Drewke et al. 1996). Interestingly, since SerC uses PLP as cofactor for its enzymatic activities, PLP is involved in its own biosynthesis (Drewke et al. 1996).

The *serC* gene forms an operon with the *aroA* gene, which encodes the enzyme 3-phosposhikimate 1-carboxyvinyltransferase (UniProt P0A6D3) (Duncan and Coggins 1986). This enzyme is involved in the chorismate pathway, which leads to the biosynthesis of aromatic amino acids. Mutants of *serC* and *aroA* are auxotrophs for serine/pyridoxine and aromatic amino acids, respectively; thus both genes are essential for the growth in minimal medium (Lam and Winkler 1990). The *serC-aroA* operon is transcribed from the P_{serC} promoter located upstream *serC* (Fig. 3.5). As consequence, the number of encoded enzymes may be simultaneously regulated at the transcriptional level. However, *serC* can also be transcribed as a monocistronic mRNA from the same promoter. In addition, a Rho-factor

independent transcription terminator located between *serC* and *aroA* is responsible for the transcriptional attenuation, which leads to an increased expression of *serC* with respect to *aroA* (Man et al. 1997).

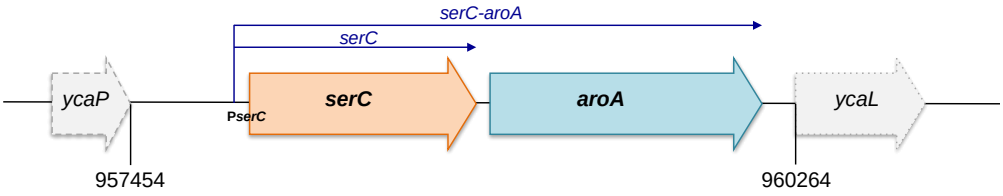


Fig 3.5: Schematic representations of the operon containing *serC*. Numbers are the genome positions in kb. The blue arrows represent mRNAs. The position of promoters is indicated

Interestingly, the arrangement of *serC* and *aroA* genes in a single operon does not seem to be accidental, since it can guarantee a coordinated regulation of the PLP, serine and chorismate biosynthetic pathways. In fact, it should be pointed out that these pathways are connected by three important features. First, the PLP and chorismate pathways share the utilization of erythrose 4-phosphate as starting substrate. Second, equimolar amounts of serine and chorismate are required for the biosynthesis of iron chelator enterochelin. Third, PLP, serine, and aromatic compounds are also linked in numerous steps and branches of intermediary metabolism. For example, chorismate is the precursor of tetrahydrofolate, which captures one-carbon units when serine is cleaved to glycine by the PLP-dependent enzyme serine hydroxymethyltransferase (Man et al. 1997).

pdxA

4-Hydroxythreonine-4-phosphate dehydrogenase (UniProt P19624), encoded by the *pdxA* gene (encompassing position 52,427 to 53,416 kb), catalyses the

fourth step in the PLP biosynthesis pathway (Fig. 3.1). PdxA is a NAD^+-dependent dehydrogenase responsible for the oxidation of 4-hydroxythreonine-4-phosphate to 3-amino-1-hydroxyacetone-1-phosphate (Banks and Cane 2004). A *pdxA* mutant is able to grow on minimal medium only if supplemented with PN or PL.

pdxA is part of a complex operon, which can give rise to different transcription units: *lptD* (from either impp1 or impp2 promoters), *lptD-surA-pdxA* (from impp3 promoter), *surA-pdxA-rmsA-apaGH*, *pdxA-rsmA-apaGH*, *pdxA-rmsA*, *rmsa-apaGH*, and *apaGH* (Roa et al. 1989) (Fig. 3.6).

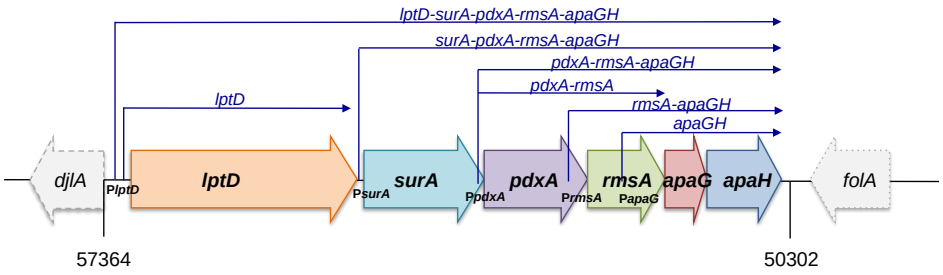


Fig 3.6: Schematic representations of the operon containing *pdxA*. Numbers are the genome positions in kb. The blue arrows represent mRNAs. The position of promoters is indicated

The first gene of the operon, *lptD* (lipopolysaccharide transport), encodes an essential protein of the outer membrane (UniProt P31554), required for the lipolysaccharide assembly (Wu et al. 2006). Considering that the decrease in the *lptD* expression causes the increase in the sensitivity to organic solvents (Abe et al. 2003) and *lptD* mutants show an increase in the membrane permeability (Braun and Silhavy 2002), LptD is important for both organic

solvent tolerance and outer membrane integrity. The *surA* gene, downstream of *lptD*, codes for a periplasmic peptidyl-prolyl isomerase (UniProt P0ABZ6), which is a chaperone required for folding and assembly of several outer membrane proteins, including OmpA, OmpF and LamB. It has been reported that *surA* mutants are unable to survive in stationary phase when pH is high (Lazar et al. 1998). Coherently, *surA* expression increases in cultures of high density as well as at low or high pH conditions. Another gene forming the operon is *rsmA* (or *ksgA*) whose gene product is a methyltransferase (UniProt P06992), which catalyses the methylation of two adjacent adenosines in the 16S rRNA of 30S ribosome. Interestingly, the KsgA protein binds specifically to its own mRNA and it may regulate its own translation (van Gemen et al. 1989). It has been shown that the mutation of *rsmA* induces resistance to kasugamycin, an inhibitor of translation initiation, without substantial defects of the growth. The second last gene of the operon is *apaG* encoding a DUF525 domain-containing protein (UniProt P62672), whereas the last gene is *apaH* coding for a diadenosine tetraphosphatase (Uniprot P05637). Both *apaG* and *apaH* belong to a group of genes involved in the mutagenesis activated by stress responses (Al Mamun et al. 2012). In particular, the ApaH protein hydrolyses diadenosine tetraphosphate (Ap_4), produced as a side product of aminoacyl-tRNA synthetases, as well as molecules related to Ap_4. Recently, it has been reported that ApaH removes caps containing Ap_4 and Gp_4 at the 5' end of RNA molecules, generating a diphosphorylated RNA product (Luciano et al. 2019).

Concerning the regulation of all these genes, the *lptD-surA-pdxA-rsmA-apaG-apaH* operon contains at least 8 promoters from which the different transcription units originate. In the P_{lptD} region, located upstream of *lptD* (Fig. 3.6), three different promoter sequences were identified: impp1, impp2

and impp3. The first two sequences are dependent from the σ^D factor and they are responsible for the constitutive expression of the *lptD* gene in equal measure (Ohtsu et al. 2004). On the contrary, the impp3 sequence is a σ^E dependent-promoter; it controls the transcription of polycistronic *lptD-surA-pdxA* mRNA. Specifically, σ^E factor is the sigma factor of RNA polymerase responding to heat shock or other environment stresses that cause an accumulation of misfolded polypeptides (Dartigalongue et al. 2001). P_{surA} is another promoter of the operon recognised by σ^E factor, it leads to the formation of *surA-pdxA-rmsA-apaGH* co-transcript. Interestingly, *surA* is always co-transcribed with *pdxA* and insertion mutants in *surA* cause PN auxotrophy (Pease et al. 2002). The transcription of *pdxA* also occurs from the P_{pdxA} promoter, which drives the formation of *pdxA-rsmA-apaGH* and *pdxA-rmsA* transcripts. Downstream of *pdxA*, other two promoters are present: P_{rmsA} promoter located at the end of *pdxA* gene (Roa et al. 1989) and a promoter within the *rmsA* gene, from which the transcription of both *apaG* and *apaH* genes starts (Blanchin-Roland et al. 1986).

dxs

The essential gene *dxs* (position 438,315- 440,177 kb, in the genome) encodes 1-deoxy-D-xylulose-5-phosphate synthase (UniProt P77488), which catalyses the thiamine diphosphate-dependent reaction of condensation between pyruvate and D-glyceraldehyde-3-phosphate (GAP) to yield DXP (Fig. 3.1). This metabolite is also an intermediate of the isoprenoid biosynthesis and of the thiamine biosynthesis pathways (Sprenger et al. 1997; Lois et al. 1998).

The transcript containing *dxs* is the *xseB-ispA-dxs-yajO* mRNA originated from the P_{xseB} promoter (Lois et al. 1998) (Fig. 3.7). The *xseB* gene product is the small subunit of Exonuclease VII (UniProt P0A8G9). The *ispA* gene is

linked to *dxs*, being also involved in the isoprenoid biosynthesis. It is an essential gene encoding the enzyme geranyl diphosphate/farnesyl diphosphate synthase (UniProt P22939), which catalyses two sequential reactions in the polyisoprenoid biosynthetic pathway (Fujisaki et al. 1986). The *yajO* gene can be transcribed also as single transcriptional units from its own promoter. It is functionally linked to *dxs*, because encodes a 1-deoxyxylulose-5-phosphate synthase (UniProt P77735), providing an alternative route to synthesize DXP from ribulose 5-phosphate, albeit less efficiently (Kirby et al. 2015).

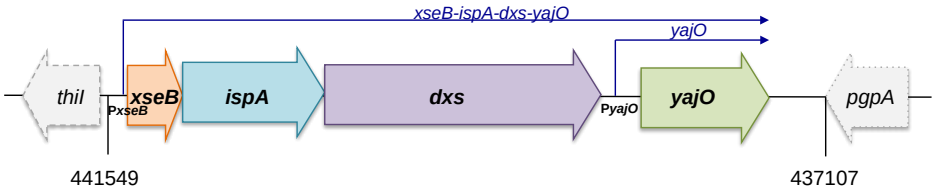


Fig 3.7: Schematic representations of the operon containing *dxs*. Numbers are the genome positions in kb. The blue arrows represent mRNAs. The position of promoters is indicated

pdxJ

Pyridoxine 5'-phosphate synthase (PdxJ, UniProt P0A794) is a homooctameric enzyme that catalyses the complex intramolecular condensation reaction between DXP and 3-amino-1-hydroxyacetone 1-phosphate to form PNP (Fig. 3.1) (Laber et al. 1999; Yeh et al. 2002).

The *pdxJ* gene located in the *E. coli* chromosome between 2,700,998 and 2,701,729 kb can be transcribed from two different promoters, which give

two different transcription units. The P_{rnc} promoter directs transcription of five genes (*rnc-era-recO-pdxJ-acpS)*, while transcription of the *pdxJ-acpS* mRNA initiates from P_{pdxJ} (Fig. 3.8) (Lam et al. 1992). The *rnc* and *era* genes are both important for the efficient ribosomal processing and maturation. The *rnc* gene encodes ribonuclease III (RNase III; UniProt P0A7Y0), an endonuclease specific for double-stranded RNA. The main substrates for RNase III in *E. coli* are the rRNA transcripts from the seven rRNA operons, leading to ribosomal RNA processing. However, this endonuclease is considered a global regulator because, by cleaving transcripts at either the 3′ or 5′ UTRs of genes and also within coding regions, plays also a role in posttranscriptional gene expression control. Moreover, RNase III is involved in the defence against viral infection (Court et al. 2013). Despite the important role that RNase III plays in rRNA processing, it is not essential for cell viability. In the absence of RNase III, the full-length ribosomal operon transcript is processed by other ribonucleases to generate 16S and 23S rRNAs, albeit more slowly (Deutscher 2009). In contrast, *era* is essential for viability. This encodes a GTPase (UniProt P06616) that interacts with RNase III at the level of the precursor rRNA, carrying out the maturation of the 30S ribosomal subunit. RecO (UniProt P0A7H3) has no role in ribosome biogenesis, but is instead involved in RecA-mediated homologous recombination (Matsunaga et al. 1996). The downstream *acpS* gene is essential for viability (Lam et al. 1992). It encodes the holo-acyl-carrier-protein synthase (AcpS; UniProt P24224), involved in lipid synthesis by transferring the 4-phosphopantetheine moiety of CoA to acyl-carrier protein activating this important intermediate of fatty acid biosynthesis (Flugel et al. 2000).

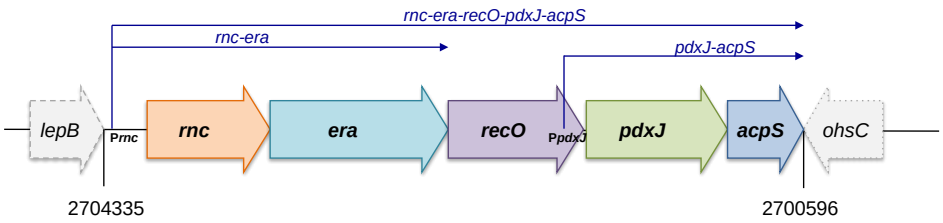


Fig 3.8: Schematic representations of the operon containing *pdxJ*. Numbers are the genome positions in kb. The blue arrows represent mRNAs. The position of promoters is indicated

Three major transcripts were found (Fig. 3.8). The abundant 1.9 kb *rnc-era* mRNA, and the less-abundant 3.7 kb *rnc-era-recO-pdxJ-acpS* are expressed from P_{rnc}, whereas the 1.3 kb *pdxJ-acpS* mRNA is transcribed from P_{pdxJ} (Matsunaga et al. 1996). Expression studies carried out using various genetic fusions showed that *pdxJ* is prevalently transcribed from P_{rnc}, but is translated 2.5-fold more efficiently from the 1.3 kb *pdxJ-acpS* mRNA than from the 3.7 kb *rnc-era-recO-pdxJ-acp*S transcript. The expression of the *rnc-era-recO* operon is negatively autoregulated at the level of mRNA stability by RNase III, which cleaves a stem-loop in its own 5′ UTR, causing the mRNA to become vulnerable to degradation by other ribonucleases. In contrast, no apparent effect of RNase III is observed on the P_{pdxJ}-dependent 1.3 kb *pdxJ-acpS* mRNA (Matsunaga et al. 1996).

pdxH and *pdxY*

The *pdxH* gene (located at position 1,717,351-1,718,007 kb) encodes the FMN-dependent enzyme (UniProt P0AFI7) catalysing the oxidation of PNP to PLP, as final step of PLP biosynthesis, and the oxidation of PMP to PLP in the PLP salvage pathway (Fig. 3.1). In this reaction, the final electron

acceptor is oxygen, which is reduced to H_2O_2. *pdxH* appears to be an essential gene for *E. coli* under both aerobic and anaerobic conditions, suggesting that, in the absence of oxygen, an unidentified compound acts as electron acceptor (Lam and Winkler 1992). *pdxH* mutants are PLP auxotrophs (Lam and Winkler 1992).

Three different transcription units were found: *pdxH-tyrS-pdxY, tyrS-pdxY*, and probably *pdxY* (Fig. 3.9). This is the sole case in which two genes involved in PLP biosynthesis and salvage pathway are co-transcribed. In fact, downstream of the *tyrS* essential gene, encoding the tyrosyl-tRNA synthetase (UniProt P0AGJ9), responsible for the attachment of L-tyrosine to tRNA(Tyr), there is the *pdxY* gene encoding pyridoxal kinase 2 (UniProt P77150; see next paragraph). Importantly, *tyrS* mutations are lethal. About 20% of *tyrS* transcripts originated from P_{pdxH}, whereas the remaining 80% originated from P_{tyrS}. This latter promoter is about four times stronger than P_{pdxH}, explaining the difference in steady-state amounts of the *pdxH-tyrS* and *tyrS* transcripts (Lam and Winkler 1992). An independent *pdxY* promoter has been identified by a genome-wide study using *gfp* transcriptional fusions (Zaslaver et al. 2006) (Fig. 3.9). If it is not from this promoter, *pdxY* transcription would originate only from the P_{pdxH} and P_{tyrS} promoters. Therefore, *tyrS* and *pdxY* would be co-transcribed. Blockage of transcription from P_{pdxH}, by an insertion in *pdxH*, reduced the *tyrS* transcript amount by about 20%, while reduced PL kinase activity by about 25%. This suggests a low level of coupling between the transcription of *pdxH* and that of *pdxY*. About 92% of *tyrS* transcripts terminate at a terminator located between *tyrS* and *pdxY*, and only about 8% read through into *pdxY*. Therefore, this terminator attenuates *pdxY* expression with respect to that of *tyrS*. Nevertheless, co-transcription of *pdxY* and *tyrS* may provide a point of

genetic integration that coordinates incorporation of amino acids into proteins with PLP coenzyme supply (Yang et al. 1998a).

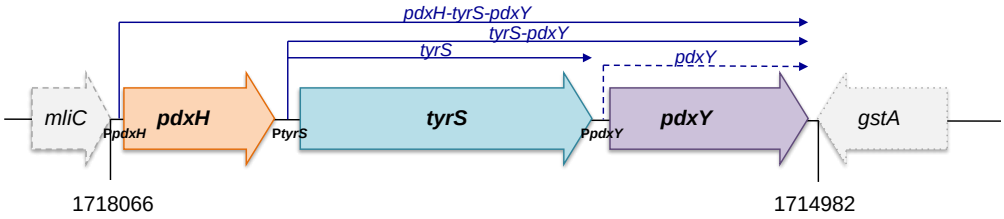


Fig 3.9: Schematic representations of the operon containing *pdxH* and *pdxY*. Numbers are the genome positions in kb. The blue arrows represent mRNAs. The position of promoters is indicated

The DXP-dependent pathway of PLP biosynthesis is typical of *E. coli* and γ-proteobacteria. In contrast, all archaea, fungi, plants and most bacteria rely on the action of the PLP synthase complex (coded by the *Pdx1* and *Pdx2* genes in plants, and by *pdxST* operon in bacteria), which directly produces PLP from glutamine, either ribose or ribulose 5- phosphate and either glyceraldehyde 3-phosphate or dihydroxyacetone phosphate (Fig. 3.10) (Fitzpatrick et al. 2007).

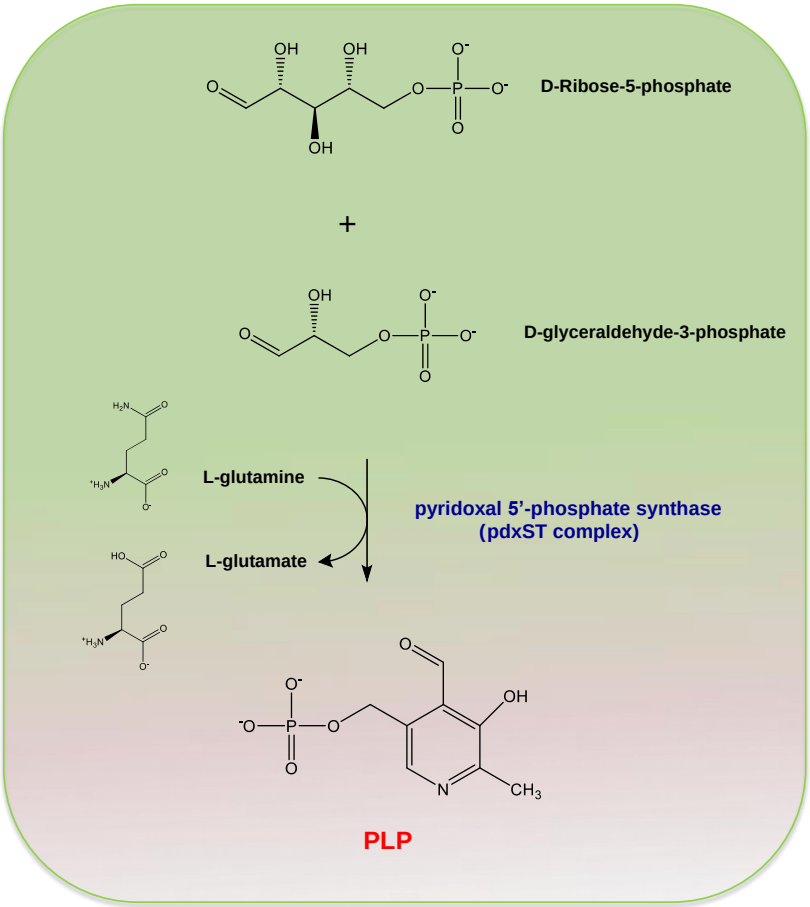


Fig 3.10: Scheme of the biosynthetic pathway for the PLP used by the pdxST complex.

The γ division of proteobacteria represents the most recent lineage of prokaryotic evolution. It has been suggested that, in the primordial environment where these bacteria lived, PLP was present so that the Pdx1-Pdx2 function was lost from their genome. However, when, following changes in the environment, PLP became scarcer and the need to produce it

arose, the γ-proteobacteria did not evolve a novel PLP biosynthesis pathway from scratch, but drew on the existing metabolism to create a pathway that produces PNP (Schoenlein et al. 1989). The salvage pathway evolved before this event (Mittenhuber 2001) and therefore PNP oxidase was already present to convert PNP in PLP. PdxJ is evolutionary distinct but highly similar to Pdx1 in structural and mechanistic terms, suggesting that it is an example of convergent evolution (Fitzpatrick et al. 2007). The function of Dxs is not only limited to PLP biosynthesis, but is also involved in isoprenoid biosynthesis (Lois et al. 1998). On the other hand, the first branch of PLP biosynthesis probably derives from the serine biosynthesis pathway by gene recruitment (Schoenlein et al. 1989). In fact, at a glance, the intermediates of the two routes are very similar (Figs. 3.6 and 3.7). Moreover, the two pathways share the same enzyme, SerC, and as mentioned before, *pdxB* and *serA* are homologous (Schoenlein et al. 1989). We want to point out that a strong evidence supporting this hypothesis derives from the observation that a Δ*pdxB* strain is still able to slowly grow at 30°C in the absence of vitamin B_6 supplementation (Lam and Winkler 1990), suggesting that PLP can be synthesized by different routes. In 2010 Kim et al, by overexpressing different *E. coli* genes, characterized a serendipitous pathway that leads to PLP synthesis in the absence of PdxB. This pathway uses an intermediate of serine biosynthesis, 3-phosphohydroxypyruvate, and three enzymes to bypass the reactions catalysed by PdxB and SerC in PLP biosynthesis. It consists in NudL, a putative CoA pyrophosphohydrolase, which hydrolyzes 3-phosphohydroxypyruvate to 3-hydroxypyruvate (3HP) that is spontaneously decarboxylated to produce glycolaldehyde; LtaE, a low-specificity threonine aldolase whose physiological function is unknown, that condensates glycolaldehyde and glycine to a mixture of 4-hydroxythreonine and 4-hydroxy-allo-threonine; and finally, a promiscuous homoserine kinase

(ThrB), which produces 4-phosphohydroxythreonine (Kim et al. 2010). Recently, a naturally selected Δ*pdxB* strain grown after 10 days in minimal medium containing glucose as sole carbon source was identified to contain a 4-step bypass pathway that restores PLP synthesis, producing levels of vitamers similar to those of the wild type strain. The bypass pathway consists of multiple phosphatases, which produce erythronate from the normal substrate 4-phosphoerytronate, 3-phosphoglycerate dehydrogenase (SerA) capable to oxidise erythronate to (3R)-3,4-dihydroxy-2-oxobutanoate, which is then converted to 4-hydroxythreonine by SerC, and lastly ThrB, the kinase producing 4-phosphohydroxythreonine (Kim et al. 2019). It is worth noticing that these serendipitous pathways tap into serine biosynthesis pathway.

3.2.2 Genes of salvage pathway.

pdxK

The *pdxK* gene located at position 2,536,386-2,537,237 kb is one of the genes involved in PLP metabolism transcribed as a single transcription unit (Fig. 3.11). It encodes pyridoxal kinase I (UniProt P40191), the enzyme that, by using ATP, phosphorylates PL, PN and PM (Fig. 3.1) (Yang et al. 1996). Therefore, in *E. coli* two kinases able to phosphorylate B_6 vitamers are present. This issue derived from an *in vivo* study done on a series of *E. coli pdxB* mutant strains unable to carry out the *de novo* synthesis of PLP (Yang et al. 1996). In particular, a *pdxB-pdxK* mutant could grow in minimal medium supplemented with PL, but not supplemented with PN, whereas a *pdxB-pdxY* mutant grew when either PL or PN were present. In contrast, growth of the *pdxB pdxK pdxY* triple mutant was observed only if precursors that bypass the need for *pdxB* were present, but not on PL, PN, or PM-containing media. These experiments clearly demonstrated that PdxK and

PdxY are the only physiologically significant PL, PN, and PM kinases in *E. coli*, with PdxY being specific for PL (Yang et al. 1996). The reason why two kinases are present in the *E. coli* cell is unknown. It is worth noticing that *in vitro* measurements of the activity showed that the purified PdxY is endowed with only 1% of the kinase activity measured with PdxK (di Salvo et al. 2004). For this reason, it has also been suggested that PdxY may play the role of kinase in another metabolic pathway, but it may also have enough PL kinase activity to support growth of *E. coli* in the absence of both *the novo* PLP synthesis and PdxK activity. Nevertheless, the crystal structure of PdxY, obtained using the protein resulting from the purification procedure and without the addition of exogenous ligands, showed that both active sites of the enzyme dimer were filled by very tightly bound ligands. In one active site mixture of PL and PLP were present, while in the other active site a covalently attached form of PL was visible. The presence of these tightly bound ligands made it unlikely that the protein is involved in a metabolic pathway different from the PLP salvage pathway (Safo et al. 2004).

The *pdxK* promoter was mapped in a genome-wide study that identified several transcription start sites in *E. coli* (Mendoza-Vargas et al. 2009), but it has not yet characterized in detail. Downstream of *pdxK* and transcribed in the opposite direction, there is *yfeK*, which codes for an uncharacterized protein (UniProt Q47702).

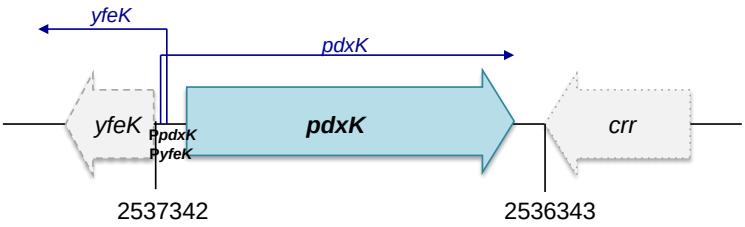


Fig 3.11: Schematic representations of the operon containing *pdxK*. Numbers are the genome positions in kb. The blue arrows represent mRNAs. The position of promoters is indicated

ybhA

The *ybhA* gene (located between 797,613 and 798,431kb) encodes a phosphatase (UniProt P21829) belonging to the superfamily of haloacid dehalogenase (HAD)-like hydrolases. YbhA in *vitro* has a phosphatase activity for PLP, erythrose-4-phosphate, fructose-1,6-bisphosphate, flavin mononucleotide, thiamine-pyrophosphate, glucose-6-phosphate and ribose-5-phosphate. Nevertheless, the purified enzyme showed the lowest Michaelis-Menten constant for PLP among the examined substrates (Kuznetsova et al. 2006). *In vivo* overexpression of YhbA attenuates the inhibitory effect on bacterial growth caused by the excess of PL in the medium. Conversely, overexpression of YbhA in cells grown in minimal medium leads to both reduced growth rate and decreased levels of intracellular PLP. This growth defect can be rescued by addition of PL in the growth medium (Sugimoto et al. 2018). Therefore, YbhA has been indicated as a specific phosphatase for PLP in *E. coli* (Fig. 3.1), by balancing the activity of the two pyridoxal kinases PdxK and PdxY. Although its preferred substrate, besides PLP, is fructose 1,6-bisphosphate (Kuznetsova et al. 2006), it does not function physiologically as fructose 1,6- bisphosphatase. This proved that the primary

role of YbhA is as a PLP phosphatase for the PLP homeostasis, not in gluconeogenesis process (Sugimoto et al. 2018). The role of YbhA was confirmed by the work of Kim et al 2019: they found mutations in the *ybhA* gene in 9 of the 11 naturally selected Δ*pdxB* strains capable to grow after 10 days in minimal medium, suggesting that in this background loss of a PLP phosphatase would be advantageous (Kim et al. 2019).

The gene *pgl*, encoding the enzyme of the oxidative pentose phosphate pathway 6-phosphogluconolactonase (UniProt P52697), is transcribed in the opposite direction. Therefore, *ybhA* constitute a single transcription unit (Fig. 3.12).

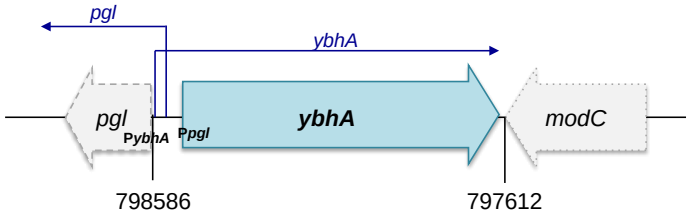


Fig 3.12: Schematic representations of the operon containing *ybhA*. Numbers are the genome positions in kb. The blue arrows represent mRNAs. The position of promoters is indicated

pdxI

Recently, another component has been added to the *E. coli* salvage pathway. It consists in the NADPH-dependent PL reductase, PdxI (UniProt P25906), which converts the PL coming from the environment to PN (Fig. 3.1). The enzyme, encoded by *pdxI*, also known as pyridoxine 4'-dehydrogenase belongs to the aldo keto reductase family (AKR-family) and was identified in a large-scale metabolomics screen for enzymatic activities of uncharacterized proteins (Sévin et al. 2017). The suggestion that PdxI is a PL reductase

derives from a recent paper from Ito and Downs (2020). They have observed that in a genetic KO background for *pdxI* and the other genes involved in PLP production (*pdxJ* and *pdxH*), external supplemented PL was not converted in PN as in the wild type strain, because the amount of PN inside the cell was significantly less than in the wild type (Ito and Downs 2020). Kinetic characterization suggested that PdxI is able to convert PL to PN, but it is not able to catalyse the reverse reaction and also that the reduction of PL to PN is favoured over the phosphorylation of PL to PLP, leading to a shortage of PL inside the cell. Therefore, following this suggestion in *E. coli* PLP is preferably synthesized using external PL by the pathway (PL→PN→PNP→PLP) rather than the pathway through the kinases PdxK and PdxY (PL→PLP).

The downstream gene *ydbD* encodes an uncharacterized protein probably involved in detoxification of methylglyoxal (UniProt P25907). The promoters of these two gene were only hypothesised and it is still unknown if *pdxI* gene is transcribed as a single transcription unit, or together with *ybdD* (Fig. 3.13). It is puzzling that a microarray analysis indicated that transcription of *pdxI* is strongly activated by NsrR (Filenko et al. 2007), but a subsequent ChIP-chip experiment indicated an NsrR-binding site on *ybdD* promoter (Partridge et al. 2009). The transcriptional regulator NsrR controls the expression of genes encoding enzymes that reduce or oxidize NO in response to NO endogenously generated or produced by host cells.

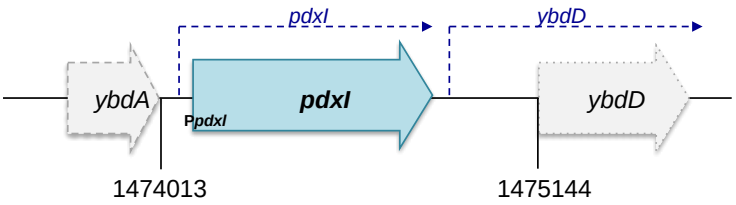


Fig 3.13: Schematic representations of the operon containing *pdxI*. Numbers are the genome positions in kb. The blue arrows represent mRNAs. The position of promoters is indicated

3.2.3 Other PLP- binding proteins

yggS

The *yggS* gene (located at 3,095,098-3,095,802 kb) encodes a PLP-binding protein involved in PLP homeostasis (UniProt P67080). It is widely distributed in bacteria and eukaryotes. The *yggS* mutant strain accumulates PNP, but not the others vitamers, and shows a toxicity ring when grown on solid medium containing excess exogenous PN (Prunetti et al. 2016). Moreover, with respect to the wild type strain, the knock out mutation causes an accumulation of certain metabolites, such as L-valine and 2-ketobutyrate, in the growth medium (Ito et al. 2013). Recently it has been proposed that the elevated levels of PNP caused by *yggS* mutation perturb the threonine biosynthesis and then the isoleucine/valine biosynthetic pathway, causing accumulation of metabolites like valine that prevents bacterial growth (Ito et al. 2019).

Physical clustering associations, resulting from a comparative genomic analysis, linked YggS to cell division and murein synthesis genes, as well as to PLP salvage and to the PLP-dependent enzyme GlyA (Prunetti et al. 2016). Nevertheless, in *E. coli yggS* may be the first gene of a polycistronic operon (*yggSTU-rdgB-hemW;* Fig. 3.14), although no correlation has been found between *yggS* and the downstream genes (Ito et al. 2013). Moreover, no experimental evidence of transcription of the *yggS* operon is present. The *yggT* gene encodes an integral membrane protein (UniProt P64564) probably involved in K^+ uptake. In fact, under hyperosmotic conditions, overexpression of *yggT* compensates for the growth defect of *E. coli* mutant cells lacking the major K^+ uptake systems (Ito et al. 2009). The *yggU* gene product is a small uncharacterized protein (UniProt P52060). RdgB, encoded by the downstream gene (*rdgB* or *yggV*) is a nucleoside triphosphate

pyrophosphatase (UniProt P52061), which hydrolyses deoxyinosine triphosphate and xanthosine triphosphate, mutagenic products of purine nucleotide deamination. RdgB may therefore remove misincorporated bases, such as xanthine or hypoxanthine, from the pool of DNA precursors (Bradshaw and Kuzminov 2003). Transcription of *rdgB*, together with the downstream gene *hemW*, is also mediated by the alternative sigma factor σ^{32}, which governs the immediate response to temperature stress (Nonaka et al. 2006). Moreover, under anaerobiosis, *rdgB* expression is repressed by FNR (fumarate and nitrate reduction regulatory protein), for which a binding site was identified upstream of the gene. The *hemW* (or *yggW*) gene codes for a heme chaperone (UniProt P52062). HemW is a radical SAM-family protein that appears to be able to transfer heme into enzymes of respiratory chains (Haskamp et al. 2018).

In the opposite direction with respect to the *yggS*-containing operon, there is *yggR* encoding a putative transporter annotated as Type II/IV secretion system family protein (UniProt P52052). Because the *yggS* and *yggR* start codons are only 17 bp apart, the two genes probably share a regulatory region in their respective promoters.

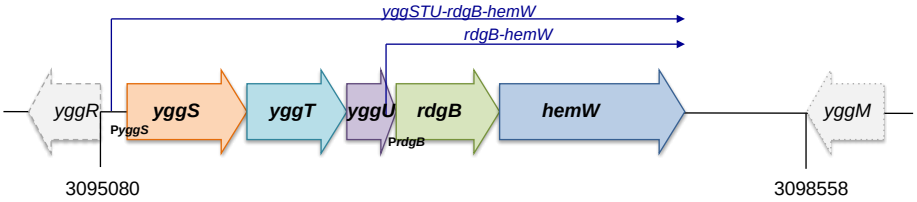


Fig 3.14: Schematic representations of the operon containing *yggS*. Numbers are the genome positions in kb. The blue arrows represent mRNAs. The position of promoters is indicated

MocR-like transcription factors

All bacteria, except *Chlamidiae*, possess the MocR-TFs, formed by the joining of a DNA binding domain and the protein scaffold of PLP-dependent enzymes (Bramucci et al. 2011). MocR-TFs control several different metabolic processes that involve vitamin B_6 and amino acids (Tramonti et al. 2018). *E. coli* has been reported to contain two MocR-TFs (Suvorova and Rodionov 2016), encoded by *ydcR* and *yjiR* genes (Table 3.1), although their function is still unknown. Unlike other well-studied MocR-TFs, i.e. *S. thiphymurium* PtsJ, *Bacillus clausii* and *Listeria monocytogenes* PdxR, etc, in which the regulatory gene is adjacent to the target gene (Tramonti et al. 2018), in the case of *E. coli* MocR-TFs no putative target genes involved in vitamin B_6 biosynthesis and salvage pathways located in close proximity could be identified.

The *ydcR* gene is located between 1,510,003 and 1,511,409 kb. Downstream and in the same direction, there is the *ydcSTUV-patD* operon (Fig. 3.15). The *ydcSTUV* genes encode a putative ABC transporter for putrescine transport, though the major putrescine transporter is PuuP. In particular, YdcS constitutes the periplasmic binding protein, YdcT is the ATP-binding component, which may provide driving force for putrescine transport, while YdcU and YdcV are the predicted inner membrane components of the ABC transporter. PatD, whose gene is transcribed from the *ydcS* promoter, encodes a γ-aminobutyraldehyde dehydrogenase, belonging to one of the two pathways of putrescine degradation. The *ydcSTUV-patD* operon is subject to the control of the stationary phase sigma factor σ^S, promoting transcription initiation from two σ^S-binding sites within the *ydcS* promoter. Moreover, expression is induced under nitrogen-limited growth from the transcriptional regulator Nac, whereas putrescine does not exert any control (Schneider et al.

2013). These observations suggest that the *ydcSTUV-patD* operon is induced in response to several stresses in *E. coli* and do not indicate any possible involvement of the YdcR MocR-TF.

The second gene encoding a MocR-TF is *yjiR*, located between 4,570,162 and 4,571,574 kb, in an unexplored region of the *E. coli* chromosome (Fig. 3.15B). In fact, the adjacent *yjiS* gene encodes an uncharacterized small protein (UniProt P39390).

In summary, considering all the operons containing genes involved in PLP homeostasis, it is possible to assume that the biosynthesis and salvage pathways are strictly associated with primary metabolism of *E. coli*. In fact, as reported in Table 3.2, they are transcribed with gene products involved in glycolysis, biosynthesis of amino acids, isoprenoid, fatty acid, RNA and DNA metabolism, and response to stress.

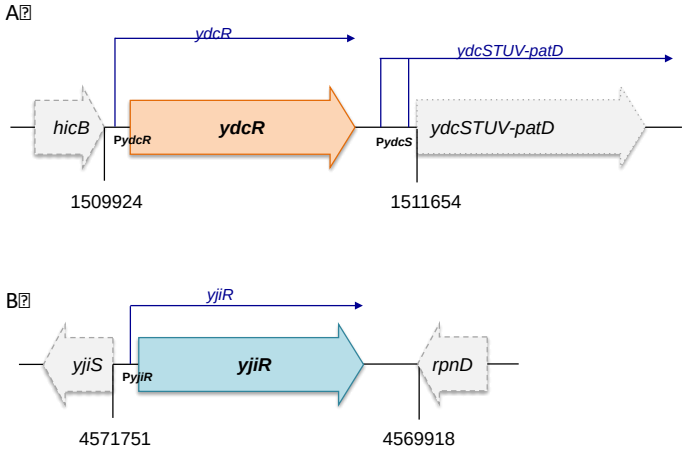


Fig 3.15: Schematic representations of the operon containing *ydcR* and *yjiR*. Numbers are the genome positions in kb. The blue arrows represent mRNAs. The position of promoters is indicated

Table 3.2: Biological process of genes associated with B_6 metabolism genes.
Essential means that the corresponding gene knock out strain does not grow in minimal medium without supplementation. The essential genes are green highlighted.

B_6 gene	**Gene in operon**	**Biological process**	**Essential**
epd	*pgK*	Glycolysis and gluconeogenesis	Yes
	fba	Glycolysis and gluconeogenesis	Yes
pdxB	*usg*	Amino acid biosynthesis	No
	truA	RNA modification	No
	dedA	n.r.	No
serC	*aroA*	Aromatic amino acid biosynthesis	Yes
pdxA	*lptD*	Membrane organization / response to stress	Yes
	surA	Protein folding	No
	rmsA	rRNA processing and modification	No
	apaG	Response to stress	No
	apaH	Response to stress/ RNA decapping	No
dxs	*xseB*	DNA catabolism	No
	ispA	Isoprenoid biosynthesis	Yes
	yajO	Isoprenoid biosynthesis	No
pdxJ	*rnc*	RNA processing	Yes
	era	Ribosome biogenesis	Yes
	recO	DNA repair/response to stress	No

	acpS	Fatty acid biosynthesis	Yes
pdxH	*tyrS*	Translation	Yes
yggS	*yggT*	Response to stress	No
	yggU	n.r.	No
	rdgB	Nucleotide metabolism	No
	hemW	Oxidation-reduction process	No

3.3. Regulation of PLP metabolism

3.3.1 State of the art on transcriptional regulation

Despite the importance of PLP biosynthesis in *E. coli*, the transcriptional regulation of the involved genes is poorly understood. It is known that three global regulators are involved in the regulation of *epd* and *serC* expression. Being *epd* homologous to GAPDH, it is not a surprise that Cra and CRP are involved in its transcriptional regulation. Indeed, the catabolite repressor-activator (Cra) protein controls the direction of carbon flux by activating genes encoding enzymes involved in oxidative and gluconeogenic carbon flow and by repressing genes concerned with the fermentative carbon flow (Ramseier 1996). On the other and, the cyclic AMP receptor protein (CRP) is an important transcription factor that regulates transcription initiation for a large number of genes. It has been identified a long time ago as a regulator of genes required for catabolism of sugars other than glucose (Kolb et al. 1993), but it was later understood that it also regulates transcription of many genes required for energy production, amino acid metabolism, nucleotide metabolism and ion transport systems (Zheng et al. 2004). In particular, the expression of *epd* is negatively affected by Cra, by binding at the -6/+12 region of the *epd* promoter, and is activated by the CRP-cAMP regulator, which binds at the -81/-60 region. Both regulators are also involved in the transcriptional regulation of *gapA* (Thouvenot et al. 2004) that can replace *epd*, in its absence. The *serC-aroA* operon is repressed by the CRP-cAMP complex and is activated by Lrp, through their binding at specific, conserved sequences of the promoter. The leucine-responsive global regulator Lrp regulates the expression of more than 40 genes in *E. coli*, including several operons involved in amino acid biosynthesis (Calvo and Matthews 1994). The combined effect of CRP and Lrp is to maximise the expression of the

serC-aroA operon in cells growing in glucose-containing minimal medium, when a rapid synthesis of vitamin and metabolites is beneficial to cell growth, and to attenuate it in amino acid-rich media, where serine and B_6 vitamers are present (Man et al. 1997).

It has been reported that the expression of the operons *pdxA-rsmA-apaGH* and *pdxA-rms* from P_{pdxA} is positively regulated by the growth rate through the Fis protein (Pease et al. 2002). Although the exact location of the binding site of this regulator on the P_{pdxA} promoter is unknown, the interaction with Fis was confirmed in a genome-wide analysis (Cho et al. 2008). Fis is a nucleoid-associated protein, whose deletion affects the transcription of 21% of *E. coli* genes. In the case of *pdxA*, it has been suggested that Fis is responsible for the positive growth rate regulation (Pease et al. 2002).

3.3.2 Transcriptional level analysis of the genes involved in the PLP metabolism

In this work, the transcriptional levels of the genes involved in the PLP metabolism (*pdxB*, *pdxJ*, *pdxH*, *pdxK*, *pdxY*, *yggS*, and *ybhA*) have been analysed during the exponential growth phase and during the stationary growth phase in different growth conditions. The mRNA was extracted from cultures of *Escherichia coli* BW25113 grown in minimum medium M9 and in M9 added with 0.1 mM PN, 0.1 mM PL and 0.5% casamino acids (CAA) as previously described in the "Methods" section. The expression of each gene observed in BW25113 grown in M9 minimal medium during the exponential phase was keep as reference.

The primers were designed to amplify only the first part of the gene (±250bp) and the analysis were performed using a qRT-PCR (BIO-RAD CFX Connect

real time) following the protocol for the preparation and the analysis described in the “Methods”.

pdxB

The expression of *pdxB* is slightly influenced by the presence of PN and PL in the exponential phase as it can be seen in the figure 3.16. On the opposite, the stationary phase is characterised by a strong decrease of the expression respect to the exponential phase, and this decrease is partially recovered by the addition of PN and, even better, with PL. Addition of CAA to the medium causes the suppression of the gene expression in the exponential phase and leave unaltered the level during the stationary phase. As in the experiments with the vitamers, the stationary phase has a lower expression than the exponential phase.

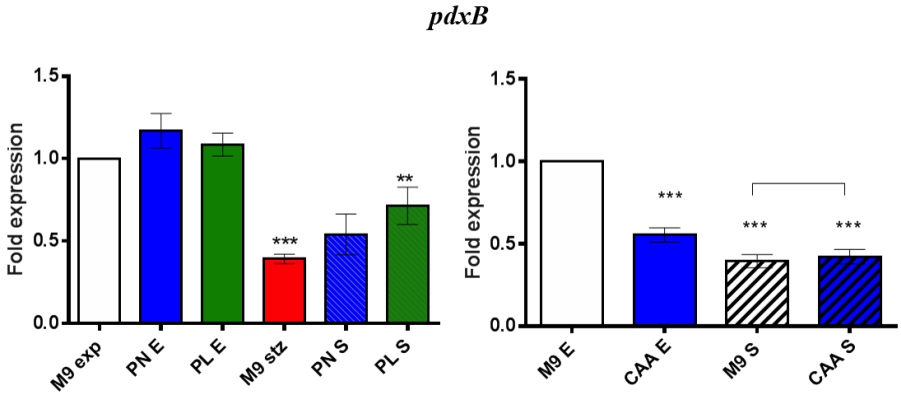


Fig 3.16: Trascriptional level analysis of the gene *pdxB*. Exp or E: indicate the exponential phase; stz or S: indicate the stationary phase. PN: pyridoxine, PL: pyridoxal, CAA: casaminoacids

pdxJ

The expression of this gene, as it can be seen in figure 3.17, doesn't seem influenced by the presence of vitamers PN and PL during the exponential phase with unchanged expression level. During the stationary phase, in *E. coli*, the expression of *pdxJ* have a small decrease that doesn't change in presence of PL, but that recover the same level of the exponential phase when PN is added to the medium. About the addition of CAA to the medium, it doesn't seem to have any effect either during the exponential nor the stationary phase. The only significant difference is between the exponential and the stationary phase in which it can be noted a decrease of *pdxJ* expression.

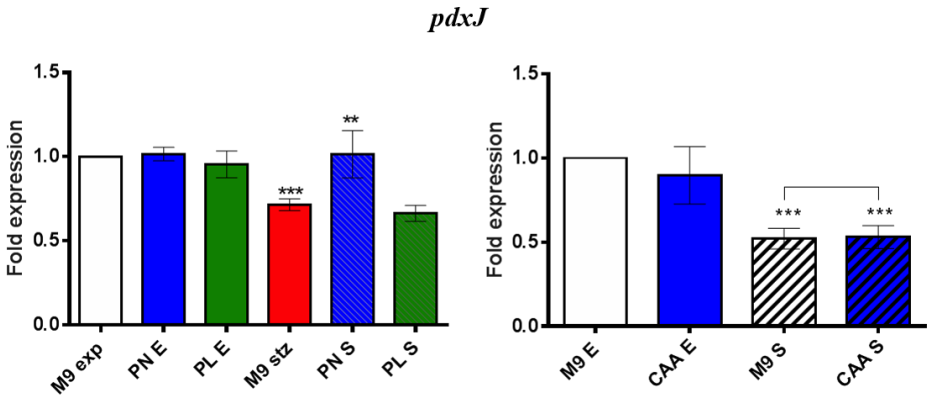


Fig 3.17: Trascriptional level analysis of the gene *pdxJ*. Exp or E: indicate the exponential phase; stz or S: indicate the stationary phase. PN: pyridoxine, PL: pyridoxal, CAA: casaminoacids

pdxH

Figure 3.18 shows that the expression level in presence of PN and PL doesn't change significantly during the exponential phase but decreases by about half during the stationary phase. The presence of casamino acids influences the expression of *pdxH* leaving unalterate the amount of mRNA during the exponential phase, but almost doubling its amount of messenger RNA during the stationary phase restoring a level similar to that of the exponential phase.

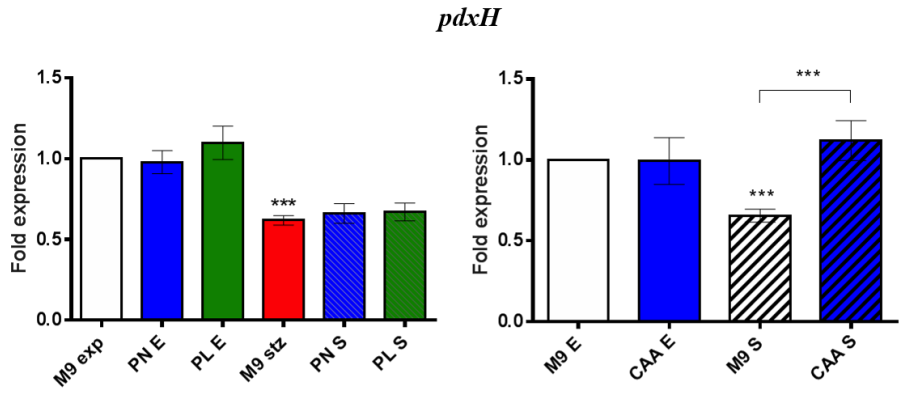


Fig 3.18: Trascriptional level analysis of the gene *pdxH*. Exp or E: indicate the exponential phase; stz or S: indicate the stationary phase. PN: pyridoxine, PL: pyridoxal, CAA: casaminoacids

pdxK

As shown in figure 3.19, the effect of PN and PL doesn't stimulate the expression of the pyridoxal kinase either during the exponential phase or the stationary phase. During the stationary phase, the analysis reveals a general minor expression of *pdxK* respect to the exponential phase also in presence of both the vitamers and a slight inhibition of the gene expression given by the PL during the stationary phase.

In presence of CAA, it is possible to notice the general suppression of the gene *pdxK* both in the exponential and in the stationary phase in which the effect of CAA is stronger.

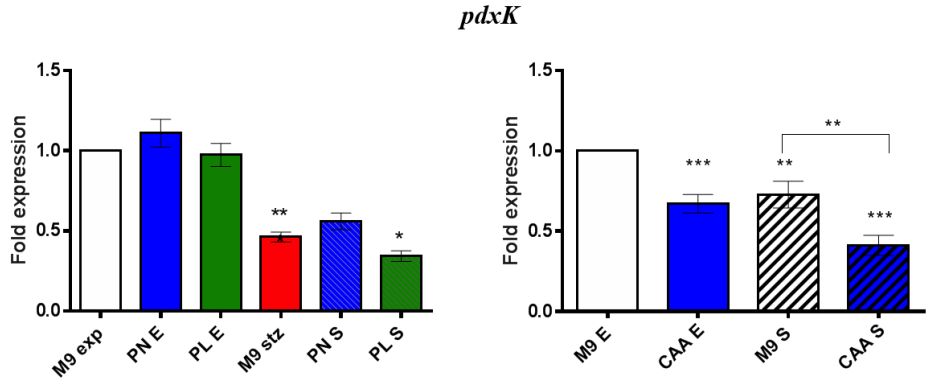


Fig 3.19: Trascriptional level analysis of the gene *pdxB*. Exp or E: indicate the exponential phase; stz or S: indicate the stationary phase. PN: pyridoxine, PL: pyridoxal, CAA: casaminoacids

pdxY

As it is possible to see in the figure 3.20, the presence of PN and PL stimulates the expression of *pdxY* during the exponential phase with a major effect of PL, maybe due to the specificity of the protein PdxY for this vitamer. The stationary phase is characterized by a general suppression of the gene expression with no effects of the vitamers.

The presence of CAA have an effect totally different relatively to the expression levels of pdxK, in fact it is shown an equivalent expression during the exponential phase but a drastic reduction of the number of mRNA copies during the stationary phase respect to the exponential phase and a less drastic,

but still important reduction effect is visible between the M9 and the M9+CAA condition in the stationary phase.

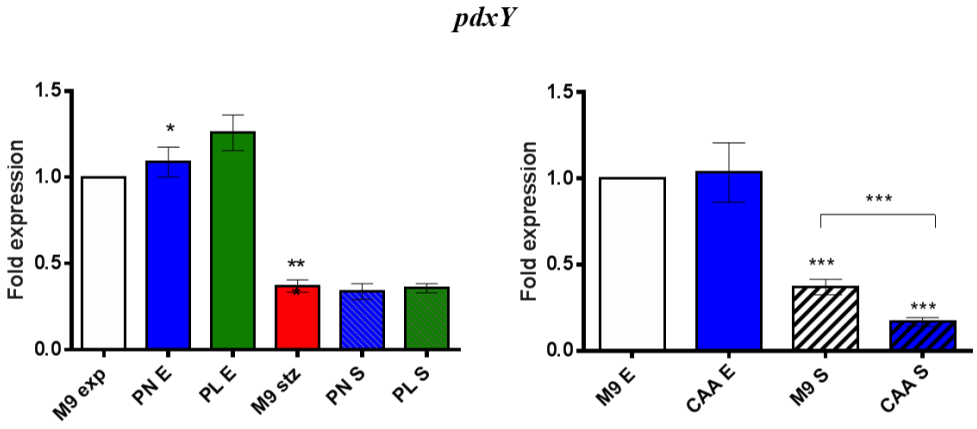


Fig 3.20: Trascriptional level analysis of the gene *pdxY*. Exp or E: indicate the exponential phase; stz or S: indicate the stationary phase. PN: pyridoxine, PL: pyridoxal, CAA: casaminoacids

ybhA

The analysis conducted on the expression of *ybhA* leads to the observation that PN and PL are almost irrelevant on the expression of the gene both in the exponential and in the stationary phase. It is evident that during the stationary phase this gene is almost silenced with a very low expression just slightly stimulated by the PN (Fig. 3.21).

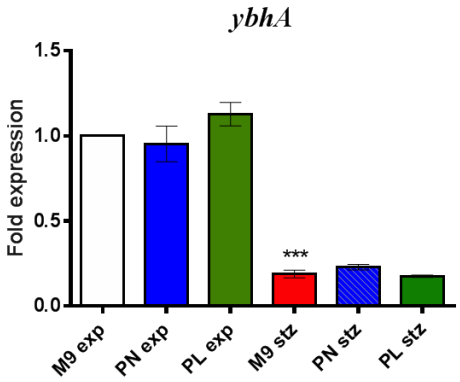


Fig 3.21: Trascriptional level analysis of the gene *ybhA*. Exp or E: indicate the exponential phase; stz or S: indicate the stationary phase. PN: pyridoxine, PL: pyridoxal,

yggS

The expression of this gene, as well as for *ybhA*, is not influenced by the presence of PN and PL (despite the excess of PLP produced when PL is supplemented to the medium), as shown in figure 3.22. The stationary phase is characterised by a lower expression of this gene in all the conditions, maybe because there is enough protein accumulated during the exponential phase or because the production of PLP is slowered. The presence of CAA in the growth medium has an important effect on the expression of this gene, in

fact this condition represses the transcription of *yggS* both in the exponential and in the stationary phase, these results could be given because the supplementation of CAA stimulates the production of PLP-dependent enzymes contributing to the homeostasis of PLP and diminishing the need of YggS for this role.

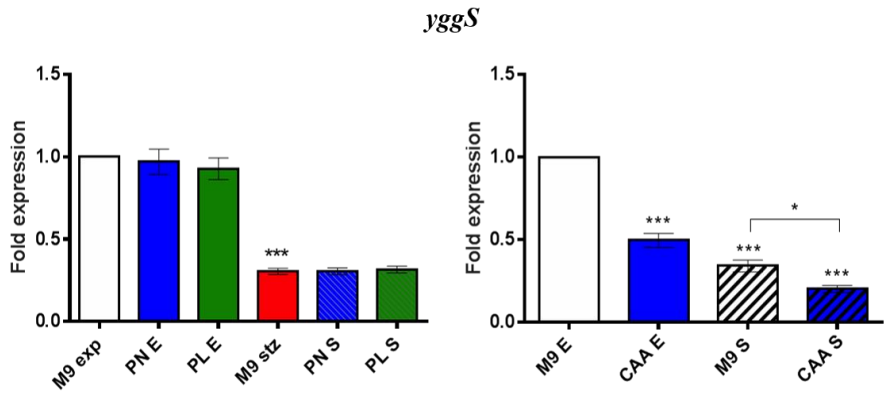


Fig 3.22: Trascriptional level analysis of the gene *yggS*. Exp or E: indicate the exponential phase; stz or S: indicate the stationary phase. PN: pyridoxine, PL: pyridoxal, CAA: casaminoacids

In general, the expression of all these selected genes decreases in stationary phase with respect to the exponential phase and is not affected by the presence of vitamers. This suggested that the transcription of genes involved in PLP biosynthesis and salvage pathways, or at least the selected genes, depends on RNA polymerase containing the the major sigma factor, σ^{70}, which is responsible for transcription of most genes expressed during the exponential cell growth (Helmann and Chamberlin 1988). Only *pdxA*, as part of the polycristronic operons *lptD-surA-pdxA* and *surA-pdxA-rmsA-apaGH*, has been reported to be transcribed from sigma factor σ^{E} – dependent RNA polymerase (see above).

So far, no specific transcriptional regulators of PLP biosynthesis and salvage pathways have been found in *E. coli*. In *Salmonella typhimu*rium, the expression of the *pdxK* gene is repressed from the MocR-like regulator PtsJ, whose gene is divergently transcribed. PLP acts as effector molecule of this regulator, because its binding to PtsJ induces a protein conformational change that increases affinity for DNA and reinforces repression (Tramonti et al. 2017). As stated above, in several bacteria, which synthesize PLP through the so-called DXP-independent pathway, the transcription of the *pdxST* operon encoding PLP synthase is regulated by another MocR-TF (PdxR). A similar transcriptional regulation of PLP biosynthesis has never been identified in *E. coli*, where the function of the two MocR-TFs present in the genome (Table 3.1) has never been identified. We also analysed the promoter region of all the genes and operons involved in PLP biosynthesis and salvage pathways, to check for the presence of common motifs that may correspond to transcriptional regulation binding sites. A 13-bp conserved sequence was identified in all the 18 promoter regions analysed, with *pdxJ*, *pdxK*, *tyrS* and *yjiR* showing the highest score (Fig. 3.23).

In most cases (11 over 18) it encompasses the promoter region (-35 hexamer, -10 hexamer or +1 start site) or is located downstream the transcription start site, suggesting that it may correspond to a repressor-binding site. The actual role of this sequence in transcriptional regulation is far from being recognized; however, the fact that it is present in the promoter region of both *E. coli* MocR-TFs (in the case of YjiR is one of the targets with highest score) encourages an experimental investigation.

Gene	strand	score	Sequence
pdxJ	–	1.40e-7	CGCTTAGGCATAAACTGCCGGAACAGTTCCCTG
yjiR	–	1.55e-7	ATCTGTATTGTGAACTGCCAGAAGGTGGATGAT
pdxK	+	6.43e-7	GATAGCCGATTTGACTGCCAGAATCAGGTAAAG
tyrS	–	7.91e-7	AGACATAGAATAAAGTGCCAGAATCAGGAGTAC
yggS	+	1.38e-5	ATGCTTTACACTAAGGGCCACAATTTCTTCCAT
pdxB	+	1.70e-5	GTTGCCCAGGAGGAGGGCCGGAAATAGGTTGTA
ydcR	+	2.07e-5	ACAACTGATATAGACTGCCGAATCATCTGCACA
pdxY	+	3.68e-5	GCCATCACCATTAACGGTGAAAAACAGTCCGAT
ybhA	–	4.02e-5	TAATTCAGCATTCACCGCCAAAAGCGACTAATT
pdxH	+	4.02e-5	TGGCGCGCGTTACACTGACGGAATCTATGTTTT
epd	–	5.21e-5	CGGGGAAACCTTACCTGTCGCAAACTGCGACTG
serC	–	5.68e-5	AGTAAAACCGTTGAGGGCAAAAATGTGGCTTGC
rnc	–	7.89e-5	CGTTGTCGCCATAACGGCGACAACGTGAACGAA
pdxA	+	9.19e-5	GGCATTTAATCGAACTGCTGGATACCCGTAATG
surA	+	9.89e-5	GAGCGGAAGCTGAACGGTTGGGATAACGATAAA
lptD	–	1.06e-4	TTGCGCGTGACAAACGGTCAGAGACTAACGTAC
xseB	+	1.74e-4	TCCGTTACCATGACGGGGCGAAAAATATTGAGA
pdxI	+	4.07e-4	TCCGTTTTATTACAGGGCAGGGTGCGATG

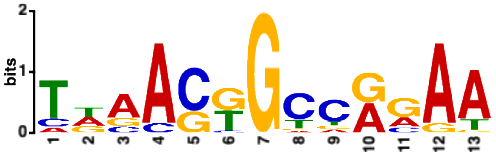


Figure 3.23: Alignment of the promoter regions containing the putative motif involved in the regulation of genes involved in PLP homeostasis. The promoter regions were analyzed using the MEME suite webserver to identify a common motif. The orientation (+ or -) of the promoter region and the score are reported. Logo of the sequence highlights conservation of the identified motif.

3.3.3 Traslational regulation

In bacteria, although transcription represents the main point of gene expression control, regulation at the translational level also occurs (Lu et al. 2007; Picard et al. 2012). In particular, the translational regulation is important especially in the expression of genes forming operons, in which each ORF may be translated at different rate, and as consequence the cellular level of the synthesized proteins may be different. Coherently, several works suggest that translational regulation plays a role in the expression of genes involved in the PLP biosynthetic pathway, such as *epd*, *pdxB*, *pdxA* and *pdxJ* genes, which are members of complex operons. The molecular mechanism through which the translational control of these genes occurs has not been characterized yet. However, some features of mRNAs containing *epd*, *pdxB*, *pdxA* and *pdxJ* ORFs, indicate that their expression is probably modulated at this level. First of all, *epd*, *pdxB* and *pdxA* genes contain infrequent and rare codons that could negatively affect translational rates and lead to low levels of the corresponding protein (Roa et al. 1989; Schoenlein et al. 1989; Bardey et al. 2005). Therefore, for these genes the codon usage represents an intrinsic mechanism at the basis of the translational modulation that ensures low levels of Epd, PdxB and PdxA proteins. Specifically, as well as being (nearly) universal the genetic code is degenerate, since 18 of 20 amino acids are encoded by alternative codons. However, the multiple synonymous codons are not used with equal frequency in genomes; this phenomenon is called codon usage bias (Sharp and Li 1987). Codon usage also non-randomly varies among genes of the same genome and within the same gene. In addition, the frequency of codons in *E. coli* positively correlates with the concentration of tRNAs carrying complementary anticodons (Ikemura 1985). This means that frequent codons are recognised by abundant tRNAs and are

more representative in genes encoding highly expressed proteins, while rare codons are decoded by low-abundant tRNAs and are typical of genes encoding lowly expressed proteins (Kanaya et al. 1999; Karlin et al. 2001). In detail, some experimental evidences suggest that the codon usage specifically affects translation initiation and elongation efficiency. For instance, rare triplets located around start codons or at the centre of the mRNA, may decrease the translation initiation or elongation rates, respectively. Interestingly, it has been shown that the slowing protein synthesis produced by codon usage can also allow a proper protein folding, as well as guarantee low levels of the protein (Chen and Inouye 1994; Quax et al. 2015; Hanson and Coller 2018). Two structural features related to its translation initiation region (TIR) cause the poor expression of the *epd* ORF. Indeed, the ribonucleotides located upstream the ribosome binding site (RBS) can form a stem and loop structure and the distance between the RBS and the start codon AUG is unusually short, and it is evolutionarily conserved among *epd* ORFs (Bardey et al. 2005).

Roa et al. have highlighted another feature of *pdxA* that contributes to low cellular levels of the encoded protein. Indeed, it lacks a Shine-Dalgarno sequence, the typical RBS of bacteria. This may cause a minor affinity of the ribosome for the translational initiation site (Roa et al. 1989).

Finally, the start codons of *pdxB*, *pdxA* and *pdxJ* ORFs are close to a conserved motif, ACGT(G/T)AAAATCC, named “PDX box” (Roa et al. 1989; Schoenlein et al. 1989; Lam et al. 1992). Noteworthy, other PDX boxes were found within the *pdxJ* coding sequence. It has been proposed that this consensus motif might play a role in the regulation of translation, being recognised by a specific regulatory protein or shape particular secondary structures, affecting ribosome activity. Definitely, the translational regulation

mediated by PDX boxes may represent the mechanism through which the expression of *pdxB, pdxA* and *pdxJ* is coordinated in the cell.

To summarize, the properties of *epd, pdxB, pdxA* and *pdxJ* ORFs, discussed so far, do limit their translational levels, allowing a differential expression with respect to other coding sequences contained in the same polycistronic mRNA.

3.3.4 Cellular levels of proteins involved in PLP homeostasis

The combined effect of the regulation at the transcriptional and translational level determines the cellular content of the proteins. Usually, proteins involved in a specific process show similar abundance (Ishihama et al. 2008; Wiśniewski and Rakus 2014). This is not the case for proteins involved in vitamin B_6 metabolism. In fact, we analysed the results obtained from two quantitative proteomic studies, carried out on *E. coli* lysates from cells collected at either the exponential (Ishihama et al. 2008) or stationary (Wiśniewski and Rakus 2014) phase. Although different approaches have been used in combination with LC-MS/MS technology, the copy numbers per cell of proteins involved in PLP homeostasis from the two studies have been compared with respect to two housekeeping proteins, GyrB and RecA (Maheux et al. 2014) (Table 3.4). Among the enzymes of the PLP biosynthesis, the concentrations of SerC and PdxJ are the highest in both proteomic analyses. In particular, in the exponential phase, SerC is 10-fold more abundant than the housekeeping proteins, and 8-fold more represented than PdxJ. This may happen because SerC despite to be involved in PLP biosynthesis, it is essential for the biosynthesis of L-serine, as reported

before. In the exponential phase, the levels of PdxB, PdxH, PdxK and PdxY are within the same order of magnitude. In contrast, PdxA and Epd are the proteins of PLP biosynthesis less abundant. This is probably a direct consequence of the translational control, which negatively affects the relative protein synthesis. In line with gene expression studies (Table 3), the level of proteins decreases in the stationary phase with respect to the exponential phase, with the exception of PdxB, PdxK and PdxH. Concerning these two latter enzymes we can hypothesise that salvage pathway, differently from *de novo* biosynthesis, is essential during the stationary phase when PLP can be recycled from protein turnover.

Table 3.4. Cellular levels of *E. coli* proteins. The copy number of each protein involved in PLP metabolism is reported. The proteins encoded by genes forming all operons (cited in the text) are grouped in the grid and their copy number is shown. At the bottom, levels of the constitutive proteins, RecA and GyrB, are in bold and used as reference values. The proteins copy number was normalized respect to both RecA and GyrB and the mean reported. Exp, exponential phase; st, stationary phase; -, protein level not determined in proteomic analyses.
* Ishihama et al. 2008
** Wiśniewski, Rakus 2014

	protein copy number (exp)*	normalized protein copy number (exp)	protein copy number (st)**	normalized protein copy number (st)
Epd	153	0,26	67	0,12
Pgk	124000	214,20	6137	10,90
FbaA	-	-	3039	5,40
PdxB	353	0,61	330	0,59
Usg	268	0,46	158	0,28

TruA	-	-	66	0,12
DedA	-	-	-	-
SerC	5740	9,91	837	1,49
AroA	673	1,16	83	0,15
LptD	372	0,64	166	0,30
SurA	2200	3,80	283	0,50
PdxA	-	-	30	0,05
RmsA	546	0,90	101	0,18
ApaG	-	-	71	0,13
ApaH	-	-	8	0,01
XseB	368	0,64	-	-
IspA	-	-	53	0,10
Dxs	-	-	74	0,13
YajO	-	-	71	0,12
Rnc	382	0,66	294	0,52
Era	-	-	28	0,05
RecO	-	-	-	-
PdxJ	737	1,27	416	0,74
AcpS	-	-	104	0,18
PdxH	315	0,54	500	0,89
TyrS	1630	2,81	507	0,90
PdxY	546	0,94	232	0,41
PdxK	289	0,50	301	0,53

YbhA	392	0,68	46	0,08
PdxI	-	-	70	0,12
YggS	546	0,94	88	0,16
YggT	-	-	77	0,14
YggU	-	-	35	0,06
RdgB	392	0,68	111	0,20
HemW	-	-	-	-
RecA	**686**		**1153**	
GyrB	**501**		**372**	

4. *In vivo* analysis of genes involved in PLP metabolism

To understand the role *in vivo* of different genes involved in vitamin B_6 metabolism in *E. coli*, it was evaluated the ability of different strains of *E. coli* (Δ*pdxJ*, Δ*pdxH*, Δ*pdxK*, Δ*pdxY* and Δ*yggS* knock out strains from Keio collection, (Baba et al. 2006) to grow on different media, such as M9 minimal medium and on LB growth medium. Both the media were supplemented with PL, PN or the toxic derivative of the pyridoxine: 4-deoxypyridoxine (4dPN).

4.1 Effect of PL and PN

PLP is not imported into *E. coli* cells, while dephosphorylated vitamers are imported (Dempsey and Pachler 1966). Dated evidence has suggested that in *S. typhimurium* and in lactic bacteria PN and PL, but not PM, can enter the cell by facilitated diffusion. Then, phosphorylation of the 5' hydroxyl group by pyridoxal kinase efficiently traps the vitamers inside the cell (Mulligan and Snell 1976; Mulligan and Snell 1977). The identity of the transmembrane protein involved in the transport of B_6 vitamers across the membrane is not known, neither is it known whether a specific protein plays such a role.

High levels of supplemented vitamers are toxic for the *E. coli* cell, as shown in Fig. 4.1A for the *E. coli* strain BW25113. The addition of small concentration of PL, up to 1 mM, in the glucose-based minimal growth medium does not affect bacterial growth. However, even a slightly higher concentration of PL (1.4 mM) has a great effect on cell growth, and a complete arrest in growth is observed when the medium contains 2 mM PL (Fig. 4.1A). In contrast, *E. coli* seems to be more tolerant to the presence of PN in the growth medium. In fact, small amounts are beneficial for the

bacterial growth, which is slowed down only when a high concentration of PN (50 mM) is present (Fig. 4.1D). As expected, *pdxJ* and *pdxH* knock out strains (from Keio collection, (Baba et al. 2006)) cannot growth in minimal medium (Fig. 4.1B and C). Interestingly, while the growth of the Δ*pdxJ* strain fully recovers by adding a small concentration (0.001 mM is enough) of PL in the growth medium, the growth curve of the Δ*pdxH* strain has a big lag, also when 1 mM PL is added to the medium, and does not reach the maximum optical density obtained by both wild type and Δ*pdxJ* strains (Fig. 4.1B and C). Moreover, the addition of PN has no effect on the Δ*pdxH* strain, whereas it fully restores the growth of the Δ*pdxJ* strain (Fig. 4.1E and F). Compared to the wild type strain, in minimal medium and without supplementation, the growth rate of the Δ*pdxK* strain (which is still able to express PdxY) is two times slower, although the OD_{max} reaches the same level (Fig. 4.1H). On the other hand, the growth curve of the Δ*pdxY* strain (which is still able to express PdxK) shows a lag, but it reaches an OD_{max} that is bigger than that of the wild type strain (Fig. 4.1I). The addition of PL (0.001 and 0.01 mM) improves the growth of the Δ*pdxK* strain, but bigger amount of PL (from 0.1 mM) has a negative effect on bacterial growth, until that a failure to grow is observed with 2 mM PL (Fig. 4.1H vs. A). Also the addition of PN has a positive effect on the bacterial growth, whereas 50 mM PN has the same negative effect found in the wild type strain (Fig. 4.1L). Different is the effect of PL on the Δ*pdxY* strain: a low amount (from 0.001 to 0.1 mM) does not change both growth rate and OD_{max}, high concentration of PL (from 1 mM) decreases both parameters, however with 2 mM PL growth is still visible, differently from the wild type strain (Fig. 4.1I vs. A). Adding PN (up to 50 mM) in the medium has no effect on the growth of the Δ*pdxY* strain (Fig. 4.1M). These studies show that PdxK plays an important role in

counteracting the toxicity of excess of vitamers, whereas PdxY is important at the beginning of bacterial growth, but later has a negative effect. The growth curve of *E. coli* Δ*yggS* compared with the wild type in minimal medium and without any supplements does not show any differences. When PL is supplemented in both the strain, the behavior is similar until 1.4mM PL; at this concentration the growth rate is very similar, but the OD_{600} reached change only in the WT strain. Supplementation of 2mM PL or higher doses are equally toxic between the strain. Both WT and Δ*yggS* strains equally tolerate up to 10mM PN, showing long lag phases when 50mM PN is added to the medium. In presence of 50mM PN, Δ*yggS* presents the longest lag phase and a growth rate lower than the WT.

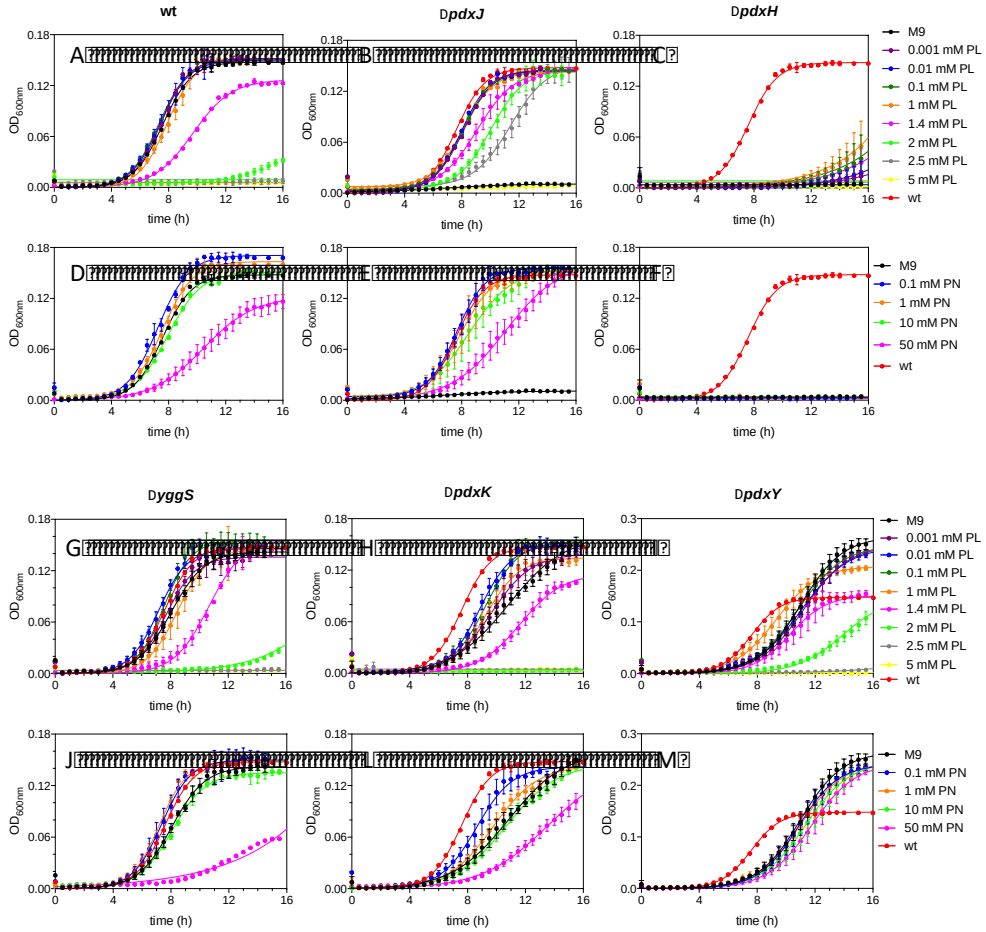


Fig 4.1: Comparison of the *E. coli* BW25113 strains and its derivative knock out strains.

Growth curves of *E. coli* BW25113 (wt), and *pdxJ*, *pdxH*, *yggS*, *pdxK* and *pdxY* knock out strains (from Keio collection, (Baba et al. 2006)) were collected and analysed.
Y axis scale (OD_{600nm}) for the strain ΔpdxY is bigger than the other because of the major optic density reached.

4.2 Effect of 4-deoxypyridoxine

The 4-deoxypyridoxine is an antagonist of the vitamin B_6 (Coburn *et al.*, 1981) that simulates the condition of PLP lack in the cells. This molecule is a toxic derivative of the pyridoxine, in which the alcoholic group on the carbon C4' is substituted by a methyl group and could acts as a competitive inhibitor of those enzymes that binds pyridoxine and other vitamin B_6 vitamers (Fig 4.2).

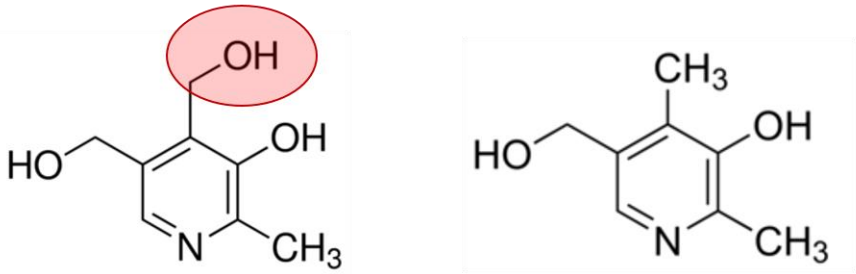


Fig 4.2: Molecular structure of pyridoxine (A) and 4-deoxypyridoxine (B)

In this work, the metabolism and the toxicity mechanism of the 4-dPN have been characterised.

To understand if the 4-dPN could be phosphorylayted by the *E. coli* pyridoxal kinase consuming MgATP, as the PN, it has been performed an assay using the commercial kit "ADP-Glo^{TM} Kinase Assay - Promega". The experiment is based on a luciferase enzyme coupled reaction where it is measured the light emitted from the luciferase consuming the ADP formed during the kinase reaction. The kit permits to stop the kinase reaction and quantify the ATP consumed over time. As suggested by the Fig 4.3, 4-dPN is can be phosphorylated by ePLK in *E.coli*.

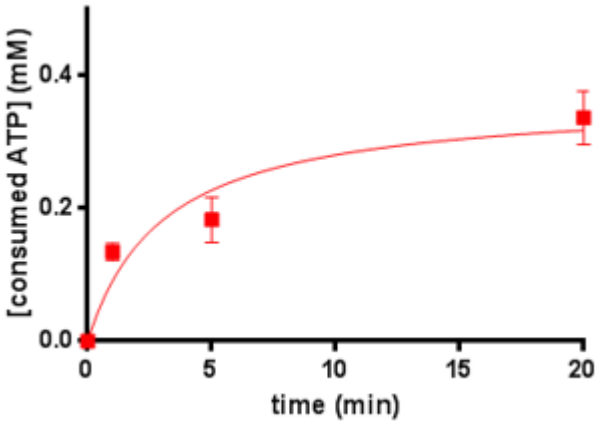


Fig 4.3: Pyridoxal kinase assay on 4-dPN.

The assay has been performed in 20 mM KHEPES pH 7.5 + 0.5 mM $MgCl_2$, at 37°C adding in the following order: 1 mM MgATP, 1 mM 4dPN and 0.25 µM ePLK. The ADP produced will be measured at time interval.

One of the possible toxicity mechanisms of the 4-dPN is due to its capacity to bind PLP dependent enzymes and to inhibit the activity of the fundamental protein PNPOx in *E. coli* as shown by *in vitro* experiments (Fig. 4.4). The experiment clearly shows the inhibition of the PNPO due to the 4-dPNP, demonstrating that also the phosphorylated form of the 4-dPN could act as an inhibitor in the cell.

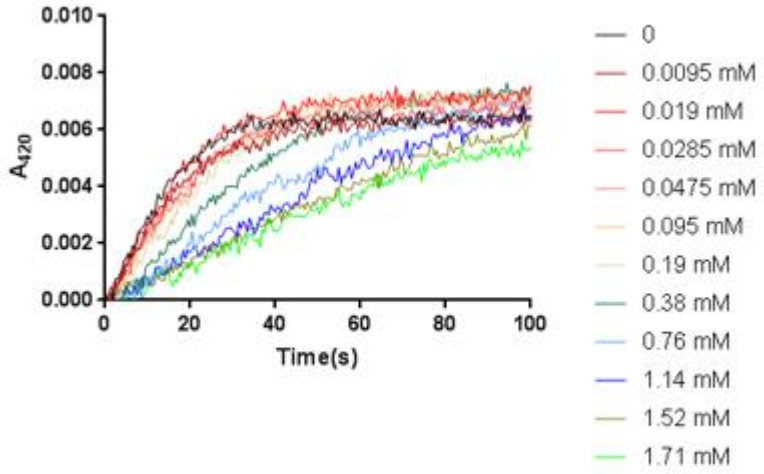


Fig 4.4: Inhibition experiments on ePNPOx using 4-dPNP

The experiment was carried out using 1µM ePNPOx in presence of 1.2µM PNP in 50 mM TRIS-HCl, pH 7.5 at 37 °C increasing the concentration of the inhibitor 4-dPNP from 0 mM to 1.71 mM as shown on the right-hand side of the picture.

The previous experiments have confirmed that 4-dPN can be metabolised and phosphorylated in *E. coli* and still act as an inhibitor as 4-dPNP. This information was precious, in fact 4-dPN was supplemented in M9 minimum medium with 0.4% glucose in order to test the ability of *E. coli* to grow in presence of this compound and determine its toxic concentration. Different concentrations of 4-dPN were used on the strains *E. coli* WT and Δ*yggS*.

Figure 4.5 reported the unexpected results obtained in these experiments; in fact, it seems that 4-dPN in *E. coli* has a double effect being both toxic and non-toxic depending on the concentration and on the strain. The panel A of the figure shows as 4-dPN, up to 2 mM, has no effect on the WT strain, but interestingly this concentration range has a drastic effect on the growth of *E. coli* Δ*yggS* (panel B).In particular the lowest concentrations of 4-dPN are

more toxic than higher concentrations. The panel C and D reports the effect that concentrations from 2 mM to 10 mM of 4-dPN have on the two examined strains: in *E. coli* WT these concentrations have a light toxic effect decreasing the specific growth rate and the number of doublings; in *E. coli* Δ*yggS* the same concentrations of 4-dPN have a stronger effect than in the WT with similar physiological effects and a drastic inhibitory ability of the concentrations 8 mM and 10 mM. From the graphs, *E. coli* Δ*yggS* is more sensitive to the 4-dPN suggesting a protecting role of the protein YggS against this compound.

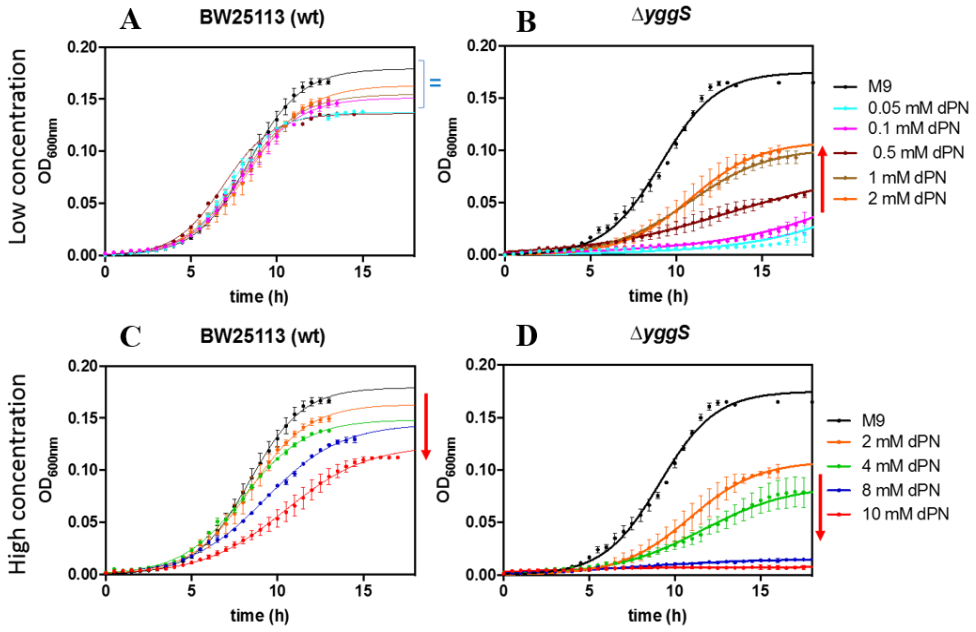


Fig 4.5: Comparison of the *E. coli* WT and *yggS* knock out strains.

Growth curves of *E. coli* BW25113 (wt), and *yggS* knock out strains were collected and analysed.

The same experiments were conducted using concentrations of 4-dPN from 1 mM to 10 mM on the strains *E. coli* Δ*pdxK*, Δ*pdxY*, Δ*pdxJ* and Δ*pdxH* in minimal medium M9 supplemented with 0.4% glucose.

The Δ*pdxJ* and Δ*pdxH E. coli* strains do not grow in minimal medium, and the addition of 4-dPN has no effect on growth of these strains suggesting that 4-dNP cannot act as PLP source.

Figure 4.6 reports the inhibition effects of 4-dPN on *E. coli* Δ*pdxK* and Δ*pdxY* showing how the toxicity of this compound is dose dependent. The results are similar for both strains, but maybe in Δ*pdxY* there is a major toxic effect increasing the lag phase. This similarity could be due to the ability of these two kinases to complement the other one in case of necessity.

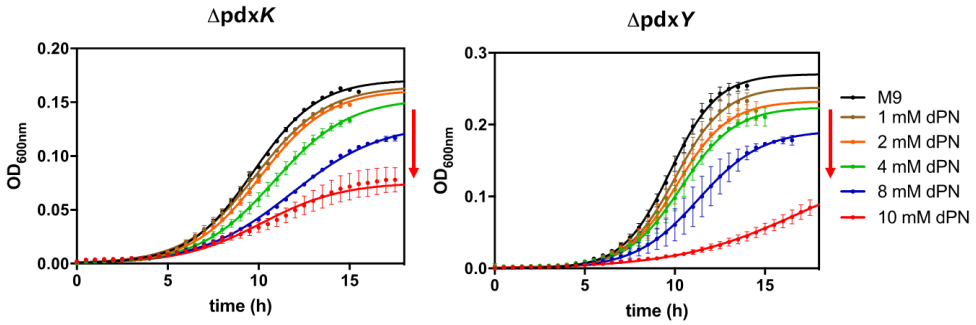


Fig 4.6: Comparison of the *E. coli pdxK* and *pdxY* knock out strains.

Growth curves of *E. coli* BW25113 (wt), and *pdxK* and *pdxY* knock out strains (from Keio collection, (Baba et al. 2006)) were collected and analysed.

The experiments performed in minimum medium M9 were repeated in the complex growth broth LB in order to verify if the abundance of nutrients, such as vitamins, amino acids and other molecules could alleviate the toxic

effect of the 4-dPN. In LB, the different strains were grown in presence of concentrations of 4-dPN from 1 mM to 8 mM obtaining different results.

About *E. coli* WT, as visible in the figure 4.7 panel A, there is a toxic effect less important than in the M9 minimum media with similar 4-dPN concentrations and non dose-dependent. *E. coli* Δ*yggS* (panel B) is most affected by 4-dPN than the WT, showing a bigger gap between the absence of 4-dPN and 1 mM 4-dPN and a major resistance to this compound than in M9, maybe because the rich broth furnishes some vitamers competing with 4-dPN for the PLP dependent enzyme.

Surprisingly Δ*pdxJ* (panel D), that cannot grow in the minimum medium M9, is able to grow up on LB also in the presence of 4-dPN showing a behaviour similar to the strain Δ*yggS* on LB. Another very interesting result is the growth of the mutant Δ*pdxH* in panel C. This strain, that in M9 minimal medium does not grow unless external PL is furnished, in LB is able to grow. As evident in the figure 4.7 panel C, Δ*pdxH* doesn't grow up as the WT in LB, in fact it is characterised by a longer lag-phase and a minor optic density. 4-dPN in this strain is very effective and seems to have a dose dependent inhibition on the bacterium, probably because of a combination of effect due to the fragile system lacking the pyridoxine 5'-phosphate oxidase. The panels E and F, respectively, represent the results for Δ*pdxK* and Δ*pdxY* showing little difference with their behaviour in M9. Respect to the WT, these two strains show a dose dependent inhibition and no significant gap between zero 4-dPN and 1 mM 4-dPN showing a gradual inhibition. A curious result is that *E. coli* Δ*pdxY* in LB has a smaller optic density than in M9 but the values in LB are like all the other strains analysed.

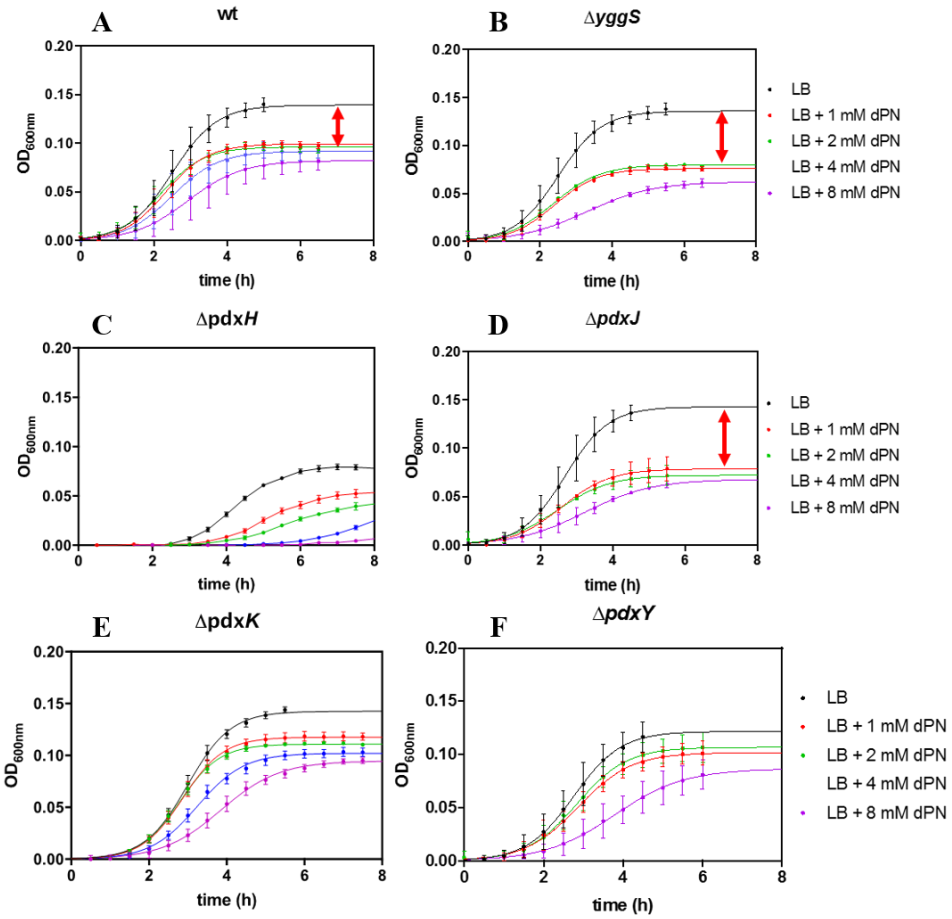


Figure 4.7: Effect of 4-dPN supplementation on the *E. coli* BW25113 WT, yggS, *pdxH, pdxJ, pdxK and pdxY* knock out strains.

Growth curves of *E. coli* BW25113 (wt), and mutants were obtained by measuring the optical density at 600 nm with a Microplate Reader (Thermo Fisher). The *E. coli* strains were grown in LB supplemented with 4-dPN at the indicated concentrations.

In this work we examined if the Δ*yggS* phenotype could be complemented by an *in trans* gene copy of *yggS* on a plasmid.

E. coli BW25113 strain can't metabolise arabinose because of the deletions Δ*araBAD*$_{AH33}$ and Δ*(araB–D)567* in the operon that codes for the genes necessary for the arabinose metabolism. This genetic background avoids expression of the *in trans yggS* gene when not voluntary induced. A second copy of *yggS* has been cloned in the plasmid pBAD24 whose promoter region is positively regulated by arabinose and can be induced only when it is necessary. *E. coli* WT, Δ*yggS* and Δ*yggS+pBAD24yggS* were grown in LB medium in presence of arabinose and their optical density were measured at 600 nm. The results show (Figure 4.8, panel C) that growth curves in the strain carrying the plasmid containing *yggS* perfectly reply the effect of the 4-dPN on the WT. This result is very important because confirms that a *in trans* gene copy of *yggS* can totally recover the loss of the genomic sequence, also confirming the role of *yggS* in the resistance to 4-dPN.

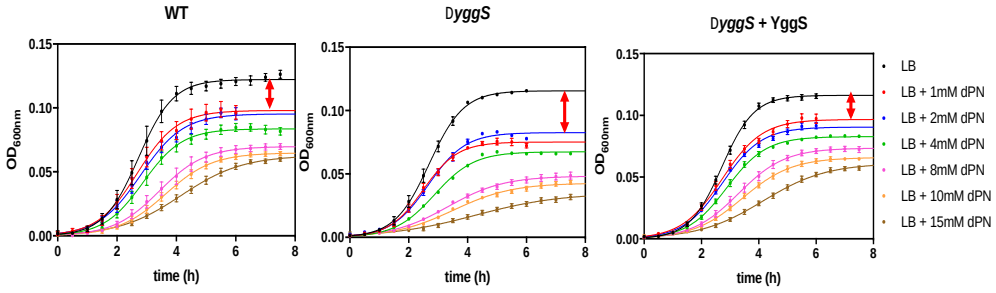


Figure 4.8: Effect of 4-dPN supplementation on the *E. coli* BW25113 WT, *ΔyggS*, and *ΔyggS +yggS* on the plasmid pBAD24.

Growth curves of *E. coli* BW25113 (wt), *ΔyggS*, and *ΔyggS +yggS* on the plasmid pBAD24 (inducibile with arabinose) were obtained by measuring the optical density at 600 nm with a Microplate Reader (Thermo Fisher). The *E. coli* strains were grown in LB supplemented with 4-dPN at the indicated concentrations.

Due to the results obtained in LB and their differences with those obtained in minimum medium M9, in this work has been also analysed the influence of the vitamer PN and PL, that could be present in the LB medium, on the different strains and how they impact on the effect of the 4-dPN.

To analyse the impact of the vitamers PN and PL, the experiments were performed in M9 minimum medium supplemented with 0.1 mM or 1 mM PL that represent a non-toxic and a toxic concentration of PL, and 1 mM PN (non-toxic). Every strain has been treated with 2 mM 4-dPN that is a toxic concentration, permitting to observe the effect on the bacterium without totally inhibit its growth. All the graphs relative to this growth curves are present in the figure 4.9; in this figure the results for each strain are presented relatively to the condition tested.

The first condition examined is the addition of 0.1 mM PL to the medium. In this condition are present a lot of differences between the strains:

In the WT, the presence of 0.1 mM PL doesn't influence the growth of the bacterium and doesn't recover the presence of 2 mM 4-dPN as visible in the panel A of the figure below where 4-dPN causes a slightly decrease in the final OD_{max}.

Results for Δ*pdxH* and Δ*pdxJ* are very similar (panel D and G). In fact, both strains are unable to grow on minimum medium and they recover this ability in presence of 0.1 mM PL. About Δ*pdxH,* the phenotype is stronger showing a lag-phase 7 hours longer than the lag-phase of Δ*pdxJ*.

In Δ*pdxK,* the presence of 0.1 mM PL (panel L) showed a soft effect in minimum media, but interestingly it partially recovers the effect of 2 mM 4-dPN.

In Δ*pdxY* (panel O), the situation was similar to that of Δ*pdxK*; in fact, also in this case the presence of 0.1 mM alone doesn't cause effects, but in combination with 2 mM 4-dPN, it partially recovers the normal growth.
Δ*yggS* has a high sensitivity to 4-dPN; in fact, in presence of 2 mM of this compound the OD_{max} reached in minimum medium is about the half of the OD in its absence. This strain positively responds to the addition of 0.1 mM PL (panel R) that not only partially recovers the effect of the 4-dPN, but seems to slightly stimulate the growth of the bacterium, reducing the lag-phase.

The second condition examined is a higher concentration of PL, precisely 1mM PL; this concentration is generally toxic because PL is phosphorylated in PLP in the cell, unbalancing its homeostasis.
E. coli WT responds to 1 mM PL with a decrease of the OD_{max} at the same level of the treatment with 2 mM 4-dPN (panel B) demonstrating that PL 1 mM is a toxic concentration in normal genetic condition. As it is visible in the figure 4.9, the addition of 1 mM PL and 2 mM 4-dPN combines the negative effects of these two compounds with a stronger growth inhibition.
The strain Δ*pdxH* (panel E), when treated with 1 mM PL, doesn't recover the ability to growth on minimum medium as it does in presence of 0.1 mM PL; this effect is done by the toxicity of the high concentration of PL. Also for this strain the additive toxic effect of 1 mM PL and 2 mM 4-dPN causes a strong inhibition of the growth.

In *E. coli* Δ*pdxJ* (panel H) a concentration 1 mM PL permits a vigorous growth of the bacterium that in absence of this vitamer isn't able to grow. The additive effect of PL and 4-dPN in this strain produces a small toxic effect, reducing the optical density reached from the bacterium treated with only PL.

About *E. coli* Δ*pdxK (panel M)*, the presence of 1 mM PL causes a longer lag phase and a decrease of the final OD to the same level reached in presence of 2 mM 4-dPN alone. In a similar way to the other strains previously analysed, it is possible to notice the additive effect of 1 mM PL and 2 mM 4-dPN that reduces more the optical density reached.

Δ*pdxY* is characterised by a sensitivity to 1 mM PL greater than that for 2 mM 4-dPN (panel P) and in this case is present an additive negative effect on the growth of *E. coli*.

For unknown reasons, Δ*yggS* (panel S) is the only strain that positively responds to the addition of 1 mM PL. In fact, the lack of the PLP homeostasis protein should increase the toxicity of this amount of PL, instead it has only a light negative effect on the OD_{max}, and in presence of 2 mM 4-dPN, 1mM PL helps to recover the severe phenotype. This strange observation could be explained with a stronger competition of the unbound PLP with the PNP and the 4-dPN accumulated for the PLP dependent enzymes.

The third condition tested to study the influence of the vitamers on the different strains is the addition of 1 mM PN in the medium, a concentration that is not toxic for bacteria.

E. coli WT (panel C) does not show any improvement when 1 mM PN is added to the medium alone, but partially recover the effect of 2 mM 4-dPN because these two molecules compete for the same enzyme being analogous.

As expected, in Δ*pdxH* (panel F) the addition of 1 mM PN does not give any amelioration because this strain is unable to use PNP to produce PLP.

The non-growing phenotype of Δ*pdxJ* is totally recovered by the addition of 1 mM PN (panel I). This result was expected, in fact in this case it has been bypassed the lack of PNP synthesis probably thanks to PdxK that

phosphorylates PN in PNP. The presence of 1 mM PN ameliorates the phenotype due to the 4-dPN competing with this toxic compound.

In Δ*pdxK* and in Δ*pdxY* (panels N and Q) very similar growths were observed, in fact both strains do not seem affected from the added pyridoxine being this compound non-toxic in a non-severe genetic background. The effect of 1 mM PN complements the effect of 2 mM 4-dPN restoring the normal phenotype; probably this is due to the competition of PN with 4-dPN for the same enzymes and maybe for a catalytic preference of the enzymes for PN instead of 4-dPN.

The effect of 1 mM PN on the strain Δ*yggS* is very interesting (panel T), in fact this strain accumulates intracellular PNP when PN is furnished as external PN being sensitive to this vitamer as a characteristic phenotype (Prunetti *et al.*, 2016). In minimum medium, not only 1 mM PN does not affect *E. coli* Δ*yggS*, but it also recovers the strong toxic effect of a dose 2 mM 4-dPN. The positive effect of 1 mM PN on *E. coli* Δ*yggS* treated with 2 mM 4-dPN is due, maybe, to its competition with the 4-dPN or to its phosphorylation in PNP and then oxidation to PLP. In this case, the situation should be similar to that in the panel S where 1 mM PL was added and PLP compete with the 4-dPN for the PLP dependent enzymes.

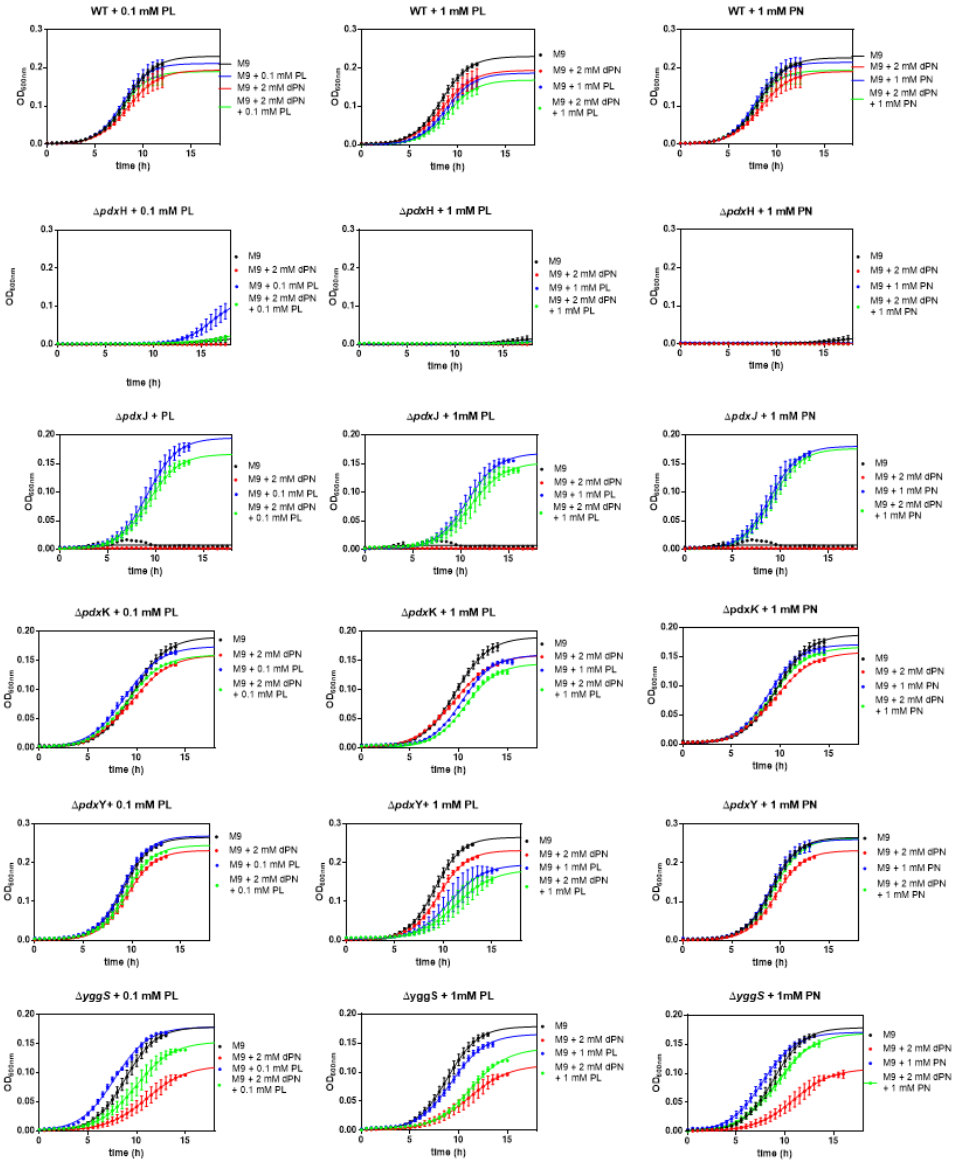


Fig 4.9: Effect of PN and PL supplementation on *E. coli* BW25113 WT, *ΔyggS*, *ΔpdxJ*, *ΔpdxH*, *ΔpdxK* and *ΔpdxY* supplemented with 2mM 4-dPN.

Growth curves of *E. coli* BW25113 WT, and mutants were obtained by measuring the optical density at 600 nm with a Microplate Reader (Thermo Fisher). The *E. coli* strains were grown in M9 supplemented with 2mM 4-dPN and treated with vitamers

The *in vivo* analysis of the genes involved in the PLP metabolism have shown that the gene *yggS* has a central role in the PLP metabolism, in fact when this gene is deleted the smallest concentration analysed of 4-dPN (0.05 mM) is highly toxic for *E. coli*. This result reflects the importance of this protein against the toxic effect of the free PLP and of its agonist, as shown in the strain *ΔyggS* that is extremely sensitive to 4-dPN. The sensitive to PL, similar to that of the WT, is probably due to the fact that PL must be first phosphorylated from the pyridoxal kinase that, in condition of excessive PL, is inhibited and does not produce all the expected PLP.

These data suggest that YggS also binds 4-dPN and it detoxifies the cell from its effect on the PLP-dependent enzymes. The binding between these two molecules could be a good topic for future studies.

5 Characterization of *E. coli* YggS

Escherichia coli YggS is a PLP binding protein whose role has not yet been unveiled. Its role in the cells could be very important, in fact it is involved in the homeostasis of the free PLP whose concentration must be kept low in order to avoid toxic reactions.

It has been hypothesized a carrier role for this protein that is supposed to be able to transport PLP to the apo PLP-dependent enzymes.

To understand the role in the homeostasis and the mechanism of transport of the PLP, in this study YggS has been characterized.

5.1 Molecular cloning of YggS

The molecular cloning of the gene *yggS* started with the digestion using the restriction enzymes *Xho*I and *Nco*I of the plasmid pET-28a(+) (previously described in the section "Materials") and of the gene *yggS* located on the plasmid pBAD24 that was sent to our laboratory from our collaborators in Richmond University, Virginia (USA). After digestion, ligation was performed between the digested gene and the digested plasmid following the procedures described in "Methods". The digestion and ligation processes were verified using a gel electrophoresis analysis in presence of ethidium bromide (as described in "Methods") to check that the plasmid and the gene *yggS* were correctly cut. After the ligation reaction it was expected a band on the agarose gel with the molecular weight of about 5700bp that correspond to the plasmid molecular weight (5000bp) summed to the molecular weight of the gene *yggS* (700bp).

The ligation product has been used to transform some aliquots of competent *E. coli* DH5α cells and after their selection on Kanamycin, the colony forming units were grown in LB overnight. The plasmids were extracted

using the commercial kit “Mackerey Nagel - Nucleospin plasmid” and the correct sequence of the gene, in frame with the T7 promoter on the plasmid were check through DNA Sanger sequencing using the service offered by Microsynth AG, Switzerland.

The same procedures were carried out to clone mutated *yggS* genes produced through site directed mutagenesis. The mutated forms of YggS were produced to characterize the binding sites of YggS and its structure. In particular it has been mutated the lysine 36 involved in the bond with the PLP in the WT YggS, and other four lysine (K38, K137, K233 and K234).

5.2 Protein expression and purification of YggS

The coding region of *E. coli yggS* was cloned into the plasmid pET-28a(+) in order to have a six histidine tag on the N-terminal part of the protein to preserve the structure of the α-helical extension of the the C-terminal β-strand involved in the binding of the phosphate group of PLP (Tremiño *et al.*, 2017; Tremiño, Forcada-Nadal and Rubio, 2018; Labella *et al.*, 2020). This kind of tag allows us to purify the tagged protein using the immobilized metal affinity chromatography (IMAC) system; in fact, the protein YggS was successfully purified on a NiNTA-agarose (**Ni**chel **N**itrilo**T**ri**A**cetate) column using increasing concentration of imidazole as eluent.

The plasmid construct pET-28a/*yggS* was used to transform 100 µL of *E. coli* BL21 (DE3) competent cells (note: only wild type YggS was expressed in the strain BL21(DE3)), all mutant variants of YggS were expressed in the strain *E. coli* HMS174(DE3) in order to avoid undesired homologous recombination; in fact this strain is Δ*recA* and cannot recombine because it lacks the recombinase protein RecA.

The selected transformants colonies were used to express the protein using the procedures described in the Materials and Methods section.

E. coli transformants were grown under constant agitation in LB medium all night long and the day after, the bacteria were harvested. The expressed protein was purified using a first salting out fractioning, collecting the pellet precipitated between 30% and 70% of ammonium sulphate saturation concentration. A NiNTA-agarose chromatographic separation were carried out as a second step to separate the tagged protein from all other proteins, that would interact with the column with less affinity. As shown in figure 5.1, SDS-PAGE analysis confirmed the success of the purification; in fact, in the analysed fraction, a single band with a molecular weight of 26 KDa can be observed on the gel.

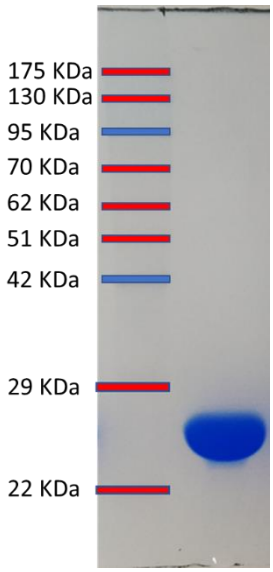


Figura 5.1: SDS-PAGE of purified WT YggS
Protein marker is the prestained protein ladder sold by Abcam PLC

The mutant forms of YggS were purified from *E. coli* HMS174(DE3) using the same method of the WT (protein salting out between 30% and 70% of saturation concentration of ammonium sulphate and then chromatographic separation on NiNTA column) and the resulting purifications were checked

using SDS-PAGE observing that, as for the WT, the mutant forms of YggS were present as single band of 26KDa molecular weight in the gel (Fig. 5.2).

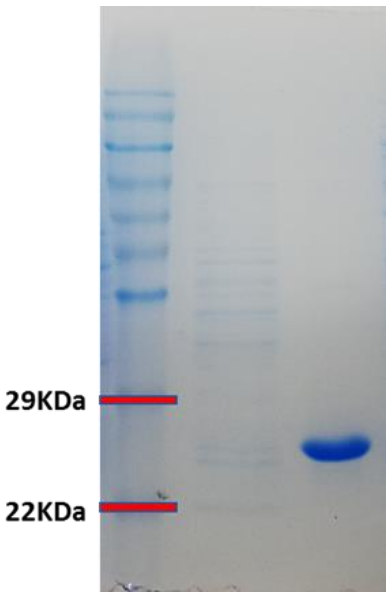


Fig 5.2: Gel SDS-PAGE of purified K36A YggS

The protein marker used is the prestained protein ladder used also in fig 5.1 made by Abcam PLC.

As a first characterization of WT YggS and its mutants the analysis of the absorption spectrum of the purified proteins has been done, which shows a major absorption band at 280 nm with a pronounced shoulder at 290 nm (typical of YggS spectrum) and another band at around 420 nm, which identifies the bond of the PLP through a Shiff base with a lysine residue, suggesting that the protein was purified bound to its ligand PLP (Fig. 5.3).

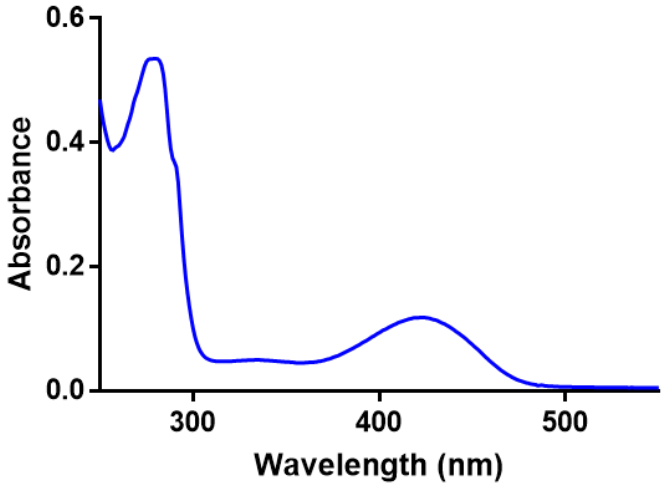


Fig 5.3: Absorbtion spectra of WT YggS

5.3 *E. coli* Serine hydroxy methyltransferase

E. coli serine hydroxymethyltransferase (eSHMT) was purified from a culture of *E. coli* GS1993, a g*lyA* deficient *E. coli* strain engineered to permit the expression of mutant forms of *E. coli* SHMT and also heterologous SHMT without risking an homologous recombination with the gene *glyA* previously removed (Iurescia *et al.*, 1996).

Escherichia coli GS1993 has been transformed with the plasmid pBS containing the coding sequence for g*lyA* WT and *glyA* single and double mutant on the residues Asp89 and Lys354 suggested from the docking experiments performed by Martin Safo and collegues.

The promoter on the plasmid pBS that positively control the gene *glyA* doesn't need to be induced (as described in "Materials and Methods"), so after the inoculum and the bacterial growth overnight, the bacteria were harvested and lysed to extract the protein.

The lysate containing the *E. coli* SHMT was centrifuged in presence of 50% of saturation concentration of ammonium sulphate separating the pellet from the supernatant. The salting out of the *E. coli* SHMT happen between 50% and 75% of saturation concentration and after the solubilization of the protein pellet the purification of the protein has continued through a first chromatographic step on a DEAE-sephadex column and then with a second chromatographic separation on a Phenyl-Sepharose column (purification condition were better explain in "Methods").

The proteins were fractioned in order to keep separated the contaminants from the target protein and then, the collected fractions were checked with the spectrophotometer in order to detect that ones with a band at around 420 nm because eSHMT binds PLP through a Schiff base that have an absorption spectrum with a maximum at 420 nm. The fractions with an absorption at 420 nm were analysed using the SDS-PAGE technique revealing that eSHMT was purified as a pure protein.

5.4 Docking experiments between YggS and SHMT

YggS is a homeostasis protein for PLP and it is supposed to be also a PLP carrier to activate the apo PLP-dependent enzymes. In order to deliver correctly this molecule, it must recognize the PLP-dependent enzymes but the way it does this is still unknown.

To understand which possible mechanism is used to recognize the PLP-dependent enzymes and to identify the interacting surfaces, *in silico* computational studies (including docking and molecular dynamic simulation) between the crystallographic structures of *E. coli* YggS (PDB: 1W8G) and of *E. coli* SHMT (PDB: 1DFO) have been performed by Safo and his collaborators.

These studies have identified two salt bridge interactions: the first one is between the residue Glu134 of the protein YggS and the residue Lys354 of the SHMT and the second one is between the residue Arg229 of the protein YggS and the residue Asp89 of the SHMT (Fig. 5.4 A).

A sequence alignment of a set of phylogenetically distant YggS protein, shows that Glu134 and Arg229 are conserved (Fig. 5.,4B) and, given their charged nature, their location near the PLP binding site and their cross-species conservation, it can be hypothesised that these two residues are important for binding and recognizing PLP-dependent enzymes.

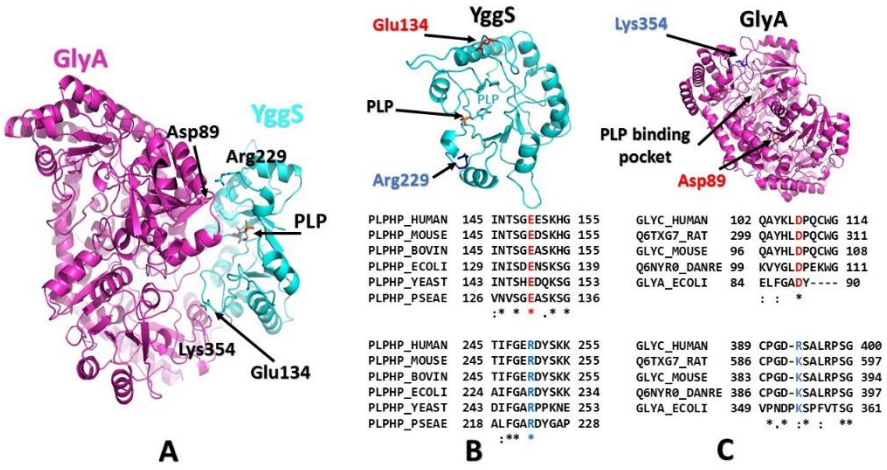


Fig 5.4: Docking experiments

A: Result of the docking between YggS (cyano) and eSHMT (purple) where the salt bridge between the proteins were identified.

B: PLPBP amino acid sequence alignment with the conserved E134 and R229 highlighted in red and cyano respectively.

C: SHMT amino acid sequence alignment with the conserved D89 and K354 highlighted in red and cyano respectively.

A similar analysis has been done with SHMT from different organisms, revealing two conserved exposed residues: Asp89 and Lys354 (Fig. 5.4C; in human and in some other organisms, the positively charged residue Lys354 is substituted with the positively charged residue arginine. This substitution conserves the same charge and a similar dimension of the amino acids conserving the function of the residues) that are close to the active site. These two residues have charges and positions on the structure of the SHMT that it can be hypothesised the possibility that these two residues are complementary to the Glu134 and Arg229 of the protein YggS to form a complex. For these reasons, the genes *yggS* and *glyA* of *E. coli* were punctually mutated to produce the two double mutant proteins YggS E134A/R229A and *e*SHMT D89A/K354A.

5.5 Characterization of YggS

5.5.1 Size Exclusion Chromatography

E. coli YggS have a large structural identity with alanine racemase, the enzyme that interconverts D- and L-alanine. Alanine racemase, as commonly for PLP-dependent enzymes, is a dimeric protein. To characterize YggS, one of the questions that look for an answer was if this protein is monomeric or if could be multimeric. The technique used to understand it, was the Size Exclusion Chromatography (SEC) on a gel filtration column Superdex 200 10/300 GL (GE Healthcare, Little Chalfont, UK). We run different chromatographic analysis with different forms or preparation of YggS protein and it has been observed that, usually, this protein is a monomer, but it might have a percentage of dimeric form in particular condition. A fresh (prepared and used within a couple of days) and an old (prepared and stored in ice for

more than a week) preparation of YggS WT were compared between them in order to control if the hypothesis of dimerization due to possible oxidation of the disulphide bonds could exist also in *E. coli* YggS.

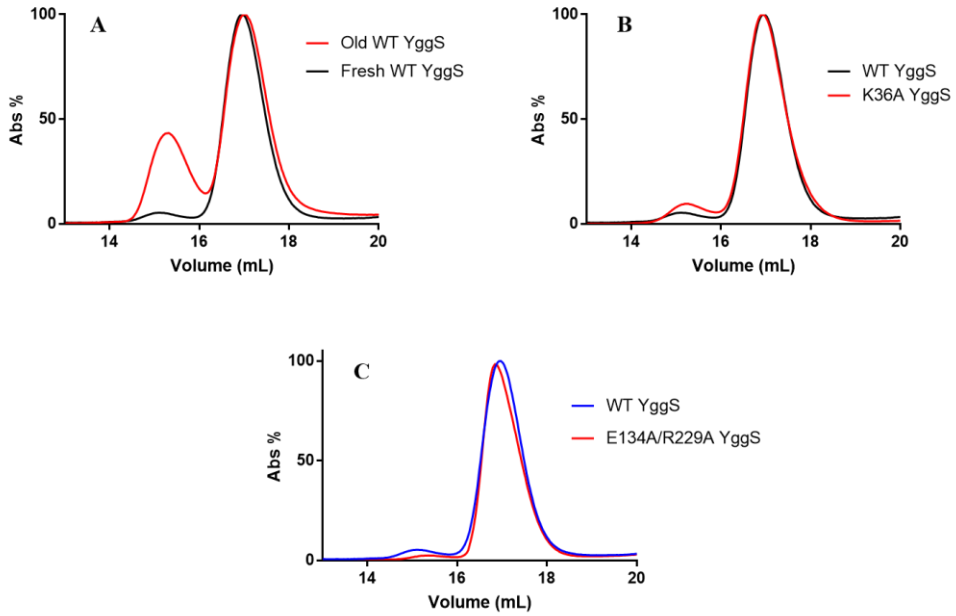


Fig 5.5: Size exclusion chromatography of different forms of YggS.

SEC was performed as described in the "Method" section.

As shown in figure 5.5A, the old preparation of YggS, where the β-mercaptoethanol is not present, has two well defined peaks that are not so marked in the fresh purified YggS. The first pick eluted at around 15 mL is relative to the dimer and is mainly present only in the old protein. This result confirms the hypothesis of Tremino and collegues (Tremiño, Forcada-Nadal and Rubio, 2018) that probably some sulphide bonds are oxidased in protein non exposed to reducing agents. In the panels B and C the fresh purified YggS is compared with K36A and double mutant E134A/R229A preserved

in presence of 5 mM β-mercaptoethanol. These two forms of YggS do not have significative difference with WT YggS about dimer/monomer ratio.

The presented results confirm that preservation conditions of the protein are important to prevent the formation of YggS dimers and that the mutations examined do not influence the formation of dimers or other multimeric structure of YggS.

5.5.2 Second binding site

Based on the literature and on the study of the crystallographic structure of the *Escherichia coli* protein YggS (PDB: 1W8G), the PLP seems to be bound to the Lysine 36 through a Schiff base.

During the experiments of this work, it has been produced and purified the protein YggS where the lysine 36 were mutated in alanine to obtain a protein unable to bind PLP. Contrarily to our expectation, the mutated form of YggS has been purified bound to a PLP molecule, and the spectrophotometric analysis has revealed that the PLP molecule was bond through a Schiff base (absorption band at around 420nm).

Firstly, it has been checked the correctness of the nucleotide sequence and, after this, it was assumed that the mutated form K36A YggS was still able to bind PLP, probably on a second unknown lysine. The absorption spectra of the WT YggS and of K36A YggS were compared and it has been possible to notice that the maximum absorption peaks of the Schiff base were different (Fig 5.6). The wavelength relative to the Schiff base of the WT YggS and its mutant form are pointed on two different values:

WT YggS → 421nm

K36A YggS → 419nm

This difference in the Schiff base absorption pick could be a first signal that PLP is still bound through this kind of bond, but that the area around the bond is different, so it is actually a bond on a different lysine.

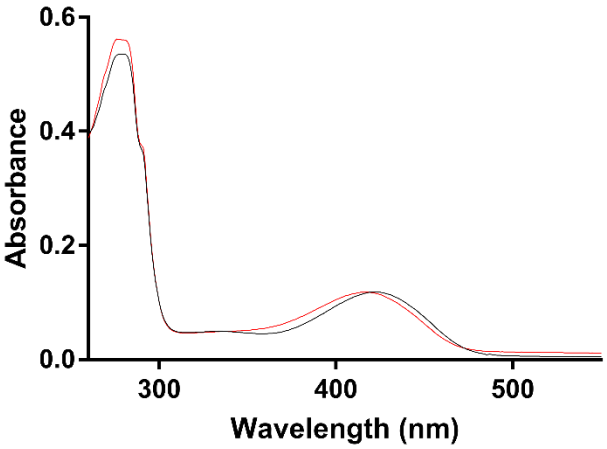


Fig 5.6: Absorbtion spectra of WT YggS (black) and K36A YggS (red)

5.5.3 Dissociation constant

The *E. coli* YggS protein has a high structural homology with Fold-Type III PLP-dependent enzymes; one of the representatives of this class of enzyme is alanine racemase.

As reported from Ito and collaborators (Ito *et al.*, 2013), YggS doesn't seems to have any catalytic activity, so in our work we characterized the PLP binding equilibrium in terms of dissociation constant, a very important parameter for the hypothetic role of PLP carrier and of PLP binding protein.

5.5.4 Fluorimetric assay

In order to characterise YggS and its dissociation constant for PLP, it has been studied the ability of the apo-YggS to bind the molecule of PLP through spectrofluorimetric analysis. To measure the dissociation constant of the PLP

from YggS, it was necessary to verify the binding stoichiometry titrating a solution containing 2µM PLP, adding different concentrations of apo-YggS and monitoring the loss of fluorescence of the PLP. This experiment has shown that the binding stoichiometry is one PLP molecule for one molecule of YggS (Fig 5.7).

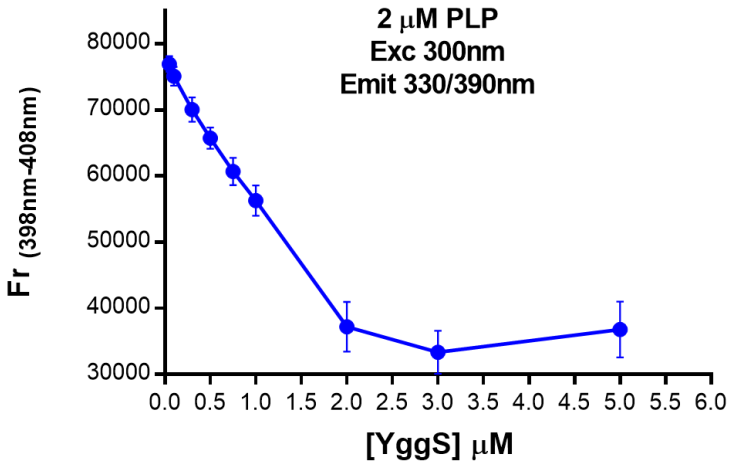
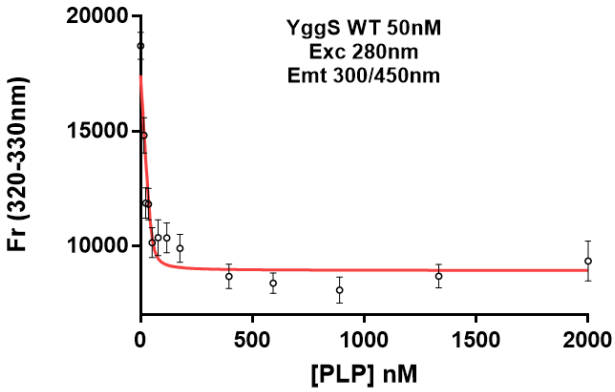


Fig 5.7: Stoichiometry between apo-YggS and PLP.

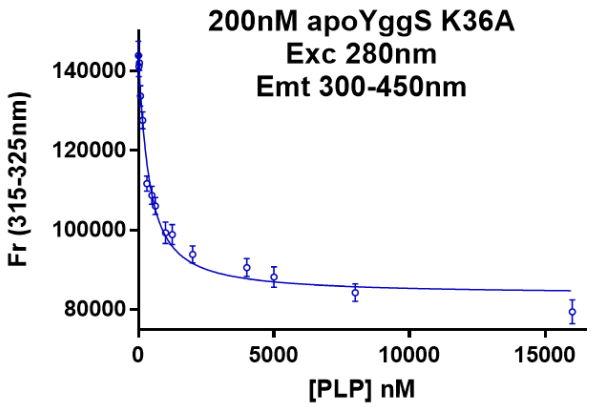
Titration of 2 µM PLP with variable concentrations of apo-YggS

Thanks to the stochiometric ratio, it is possible to set up the experiment to measure the dissociation constant. The variation of fluorescence emission of apo-YggS functionally to the concentration of PLP, was measured using a spectrofluorometer Fluoromax-3 by Horiba Jobin Yvon. After having excited the tryptophan residues using light at 280 nm wavelength, it has been recorded the fluorescence emission between 300 nm and 450 nm. The graphic obtained from the data analysis (analysis methods are reported in

"Method") describe the average fluorescence between 320 nm and 330 nm (pick of the fluorescence emission) over the relative PLP concentration. The saturation curve, fitted with a quadratic equation (reported in the method section), permits to estimate the dissociation constant. The same experiment has been performed for WT YggS and for the mutant forms K36A YggS.



Best-fit values	
Amax	8519
Kd	2.1
APOi	= 50
PLPi	= 0
const	8914
Std. Error	
Amax	280.9
Kd	0.8785
const	137.2



Best-fit values	
Amax	60149
Kd	279.3
APOi	= 200
PLPi	= 0
const	83697
Std. Error	
Amax	761.8
Kd	17.08
const	600.5

Fig 5.8: Fluorescence saturation assay on YggS WT and YggS K36A.

The experiments were performed in 50 mM NaHEPES pH 7,6 and 25 °C. The dissociation constant values are reported in the box on the right for each protein and are: 279,3±17 nM and 2.1±0.87 nM for K36A and WT YggS respectively

The dissociation constant was calculated and the best fit values are those visible on the right in figure 5.8. Due to the higher K_d, in order, to characterise this aspect of K36A YggS it has been necessary to use four time as much protein compared to WT. The values of the dissociation constants are very different. In fact, K36A YggS seems to have a K_d of 279,3±17 nM that is a hundred times less (in terms of affinity for the PLP) than the WT YggS that has a K_d, estimated in 2.1±0.87 nM. These results are consistent with the existence of a second binding site with a minor affinity and a slightly different binding pocket.

5.5.5 Circular Dichroism

During the fluorimetric assays it has been observed that WT YggS and YggS K36A have different dissociation constants. During the experiments to realise this work, it has been purified also a quintuple mutant form of YggS mutated on the lysine K36A, K38A, K137A, K233A and K234A that doesn't bind PLP, sorting the lysine near the PLP binding site and the terminal lysine for thei exposure to the solvent. Given that the quintuple mutant form does not bind PLP, it was necessary to check the structural integrity of the protein in order to verify that the difference of affinity for PLP was due to a different bind and not to a loss of structure of YggS.

In order to check if the structure was conserved between WT YggS and the mutants, the proteins were analysed my means of circular dichroism spectroscopy in the far-UV spectrum (190-260 nm) and from the results, the secondary structure of the WT YggS and its mutants are conserved. Assuming that, if the secondary structures are conserved then the tertiary ones will be maintained. In this scenario PLP binding is not influenced by the

protein structure, because the quintuple mutant form of YggS keep its structure but does not bind PLP.

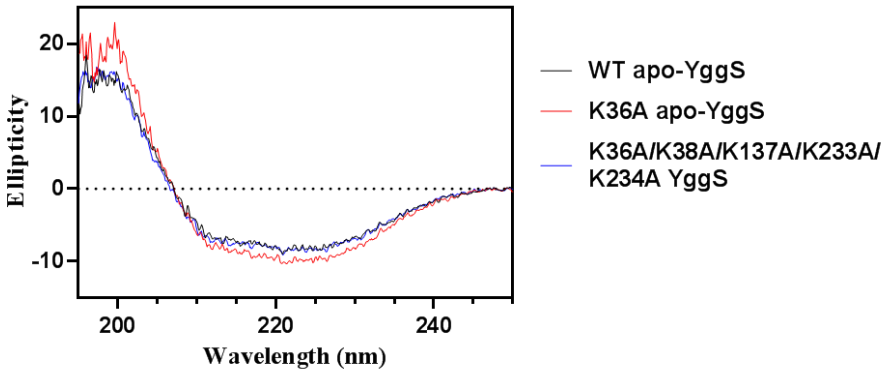


Fig 5.9: CD spectra of apo-YggS WT, K36A and quintuple mutant.

The quintuple mutant is mutated on the lysine 36, 38, 137 and on the two final lysines of the aminoacidic sequence Lys233 and Lys234.

CD spectra were measured in 20mM KPi, pH 7.6

5.6 PLP transfer studies

In this work we want to understand if YggS, in the cell, acts as a PLP carrier to activate the PLP-dependent enzymes, delivering this molecule through the formation of a protein-protein complex or not.

The SEC analysis has excluded the formation of a stable protein complex between YggS and eSHMT (data not shown), so it has been necessary to understand if YggS works through a transient complex or if PLP is transferred to apo PLP-dependent enzymes after dissociation from YggS. In order to better understand how the transfer takes place, it has been used an enzyme coupled assay with a major sensitivity that measures the activity of

the *e*SHMT to convert L-serine to glycine using tetrahydrofolate (THF), as specified in the Method section. Using this assay, it was possible to measure the activity of the reactivated apo-*e*SHMT and, indirectly, the amount of PLP transferred.

As visible in the graph below (Fig 5.10A), following the transfer reaction for a long time it's possible to estimate that, using a ratio of 70 µM apo-*e*SHMT and 20 µM of YggS, about 13% of apo-*e*SHMT is reactivated using almost half of the PLP present in YggS. Another consideration is that, when increasing the human PLP phosphatase (hPLPP) concentration, the amount of PLP transferred diminishes because the phosphatase competes with apo-*e*SHMT for the free PLP.

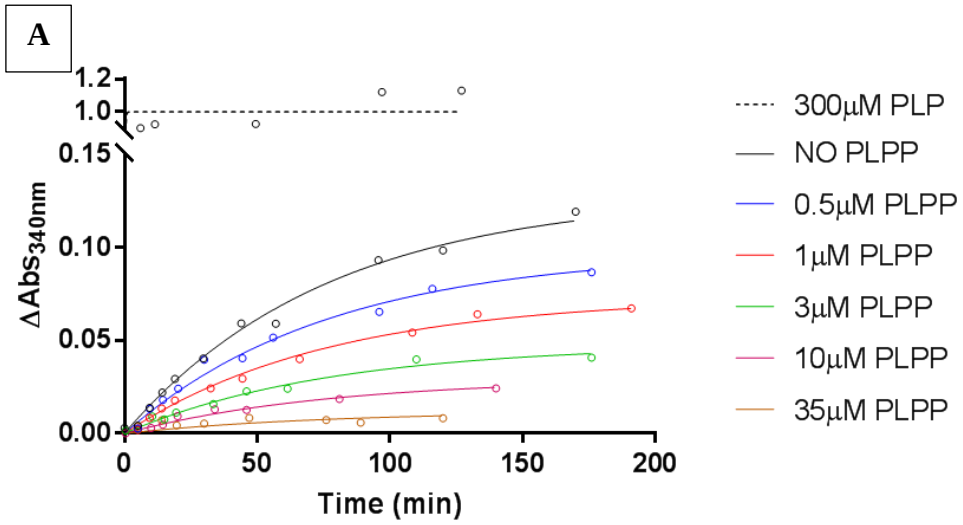


Fig. 5.10A: PLP transfer experiments in presence of increasing concentration of human PLPP.

20 µM holo-YggS were used to transfer PLP to 70 µM apo-eSHMT in 50 mM NaHEPES pH 7.6 at the temperature of 37°C.

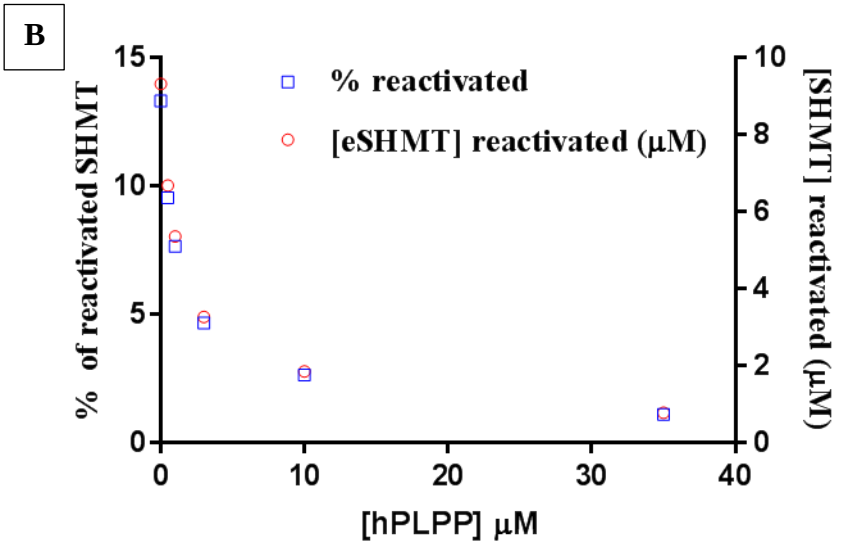


Fig 5.10B: Analysis of the amount of reactivated SHMT in function of the PLPP concentration

From the graph in figure 5.10B, it is evident that even increasing the concentration of hPLPP up to very high values, there is a limit in the PLP hydrolysed. This is an important finding, because it means that a part of the PLP is protected from the action of the phosphatase and it suggests the possibility of interaction between YggS and *e*SHMT.

To verify the hypothesis that YggS and *e*SHMT could interact through the residues resulting from the docking experiments: E134 and R229 on YggS and D89 and K354 on the eSHMT it has been performed the following kinetic experiments.

Firstly, the activity of the mutant forms of *e*SHMT for their ability to convert L-serine in glycine it has been checked, and it resulted that the two mutants *e*SHMT D89A/K354A and *e*SHMT D89K/K354A were almost inactive probably because K354 is necessary to bind THF. Since it was not possible to

use this assay, in this work it has been decided to use a second assay described in "Method". It has been checked if also the threonine aldolase activity of the *e*SHMT was compromised in these two mutants and they resulted only partially affected from the mutations. The threonine aldolase activity was 100% in the WT *e*SHMT, 19,3% in *e*SHMT D89A/K354A and 45% in *e*SHMT D89K/K354A so, the latter was chosen.

The experiments in fig 5.11 have shown that there are no significant differences in the ability of apo-*e*SHMT D89K/K354A to bind PLP when the cofactor is transferred from YggS WT or from the double mutant YggS E134A/R229A. These mutated residues are that ones targeted with docking experiments that should be involved in the protein complex formation and, since there are no differences between the double mutant and the WT YggS, these residues seem, on the contrary, not to be involved in the transfer mechanism or in the complex formation.

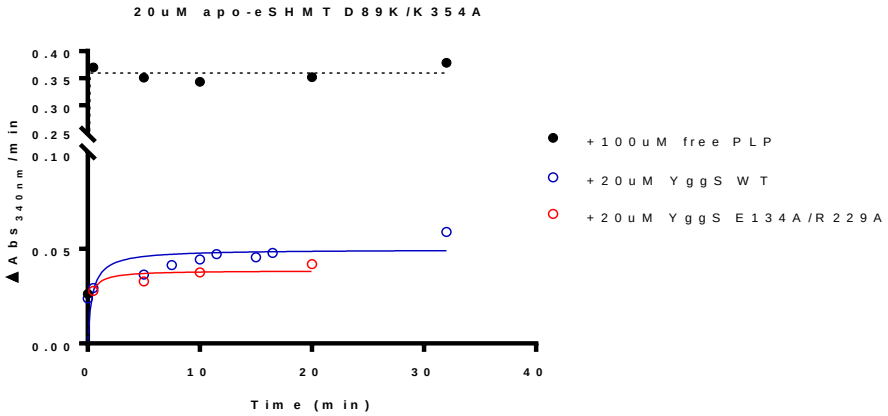


Fig 5.11: PLP transfer kinetic experiments from WT YggS and double mutant YggS E134A/R229A to the double mutant form eSHMT D89K/K354A.

6. *Conclusions*

The physiological role of the protein YggS has not yet been unveiled. Up to now, YggS is supposed to be a PLP-binding protein serving as a carrier and a reservoir of PLP in the cell. In this view, the PLP homeostasis power of YggS would avoid the negative effects of an excess of free PLP, thanks to a sort of "sponge effect".

In this work, the metabolism of the PLP has been deepened in order to understand how it is regulated and which could be the role of the protein YggS in this metabolism. We have found 57 PLP binding proteins in the proteome of *E. coli* MG1655, and 26% of them resulted to be essential for the bacterium survival. The gene expression analysis performed during the experiments for this work showed that the vitamers PN and PL don't influence the expression of the genes involved in the PLP metabolism, except for the expression of the gene pdxJ that is downregulated by the presence of these two vitamers, supposedly to avoid overproduction of PNP. In this study, it was also observed that during the stationary phase the PLP metabolism is generally downregulated.

The transcriptional regulation mechanism is not clear yet. It has been observed a consensus sequence conserved in the promoter region of the genes involved in PLP homeostasis, so a common regulator or a common regulation system could exist, although it has not yet been discovered. This study, and data herein reported, could contribute to a novel point of view in the PLP metabolism regulation.

From the literature, YggS is not an essential protein in physiological condition. In fact, the *E. coli ΔyggS* strain grow without showing differences with the wild type strain (see Fig. 4.1), except for the completely different pool of intracellular amino acids, whose relative concentrations have significant differences, specially isoleucine and valine whose metabolism

seems to be regulated by YggS (Ito *et al.*, 2013) and the accumulation of PNP that may have an effect in perturbing threonine biosynthesis pathway (Prunetti *et al.*, 2016; Ito *et al.*, 2019).

Our *in vivo* study has shown that *E. coli* growth is not enhanced by the presence of PL and PN, and that these vitamers are instead toxic at the concentration of 1,4 mM and 50 mM, respectively.

To understand the PLP metabolism in this bacterium, the effect of 4-deoxypyridoxine and how this compound is metabolized in *E. coli* have been studied. The experiments showed that the 4-dPN inhibits the PNPOx, limiting the PLP production, and it is possible to hypothesise that this molecule binds the other PLP dependent enzymes through a competitive inhibition mechanism for the PLP binding site.

As observed in this study, in *Escherichia coli* YggS plays an important regulatory role in the homeostasis of the PLP acting as a "sponge" for the free form of PLP (keeping its concentration low) and of its agonist like the 4-deoxypyridoxine reducing the toxicity of these compound that is resulted to be highly toxic especially in the *E. coli ΔyggS* strain. In this strain of *E. coli*, it was observed that an *in trans* copy of the gene *yggS* is able to recover the WT phenotype, and that the vitamers PL and PN help the bacterium to recover the effect of the 4-dPN. It is possible to hypothesise that the effect of the toxic compound 4-dPN is counteracted by the protein YggS binding its phosphorylated form 4-dPNP, and limiting the chances to bind PLP-dependent enzymes. A similar hypothesis is possible for the role of PL and PN, that after their phosphorylation in the cell compete with 4-dPNP for the binding to the enzymes.

Given the importance of the protein YggS in *E. coli*, the characterization of this protein was deepened to better understand its role in the bacterium and the mechanisms for the carrier function hypothesised for it.

This work has demonstrated that YggS monomeric form is prevalent and that a dimeric form of this protein can be an artifact depending on the experimental conditions. Our studies have shown that the mutation K36A and the quintuple mutation do not influence the structure of the protein or, in the case of the mutation K36A, the monomer/dimer equilibrium.

YggS, as shown in our experiment, has a second binding site for PLP that binds this molecule through a Schiff base and with a hundred times less affinity than the normal binding site on the lysine 36. The function and the position of this second binding site are still unknown and open a new topic for a future research.

Finally, the PLP carrier role of YggS was confirmed characterizing the ability to transfer PLP to the PLP-dependent enzyme serine hydroxymethyltransferase. The PLP transfer mechanism is still unveiled, but the limit on the hydrolysis of the PLP bound to YggS observed in the transfer experiments in presence of PLPP suggest the possibility of a complex formation that must be studied in future. Our PLP transfer studies have shown that the predicted residues that should have been involved in the complex formation between YggS and eSHMT are not correct, or it is possible that their mutation does not influence the PLP transfer mechanism.

Given these results, in the future more studies will involve YggS and its homologous protein to understand the regulation mechanisms of the PLP homeostasis in the various organisms and the intracellular transport of this cofactor to the apo PLP-dependent enzymes.

7. *Bibliography*

Abe S, Okutsu T, Nakajima H, Kakuda N, Ohtsu I, Aono R. 2003. n-Hexane sensitivity of Escherichia coli due to low expression of imp/ostA encoding an 87 kDa minor protein associated with the outer membrane. Microbiology149: 1265-1273.

Al Mamun AA, Lombardo MJ, Shee C, Lisewski AM, Gonzalez C, Lin D, Nehring RB, Saint-Ruf C, Gibson JL, Frisch RL et al. 2012. Identity and function of a large gene network underlying mutagenic repair of DNA breaks. Science 338: 1344-1348.

Albersen, Monique, Marjolein Bosma, Nine V. V. A. M. Knoers, Berna H. B. de Ruiter, Eugène F. Diekman, Jessica de Ruijter, Wouter F. Visser, Tom J. de Koning, and Nanda M. Verhoeven-Duif. 2013. "The Intestine Plays a Substantial Role in Human Vitamin B6 Metabolism: A Caco-2 Cell Model." PLoS ONE 8(1).

Arps PJ, Winkler ME 1987b. Structural analysis of the Escherichia coli K-12 hisT operon by using a kanamycin resistance cassette. J Bacteriol169: 1061-1070.

Arps PJ, Winkler ME. 1987a. An unusual genetic link between vitamin B6 biosynthesis and tRNA pseudouridine modification in Escherichia coli K-12. J Bacteriol169: 1071-1079.

Atkinson, Mark A., Mark A. Bowman, Lalita Campbell, Bethany L. Darrow, Daniel L. Kaufman, and Noel K. Maclaren. 1994. "Cellular Immunity to a Determinant Common to Glutamate Decarboxylase and Coxsackie Virus in Insulin-Dependent Diabetes." Journal of Clinical Investigation 94(5):2125–29.

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2: 2006.0008

Banks J, Cane DE. 2004. Biosynthesis of vitamin B6: direct identification of the product of the PdxA-catalyzed oxidation of 4-hydroxy-l-threonine-4-phosphate using electrospray ionization mass spectrometry. Bioorg Med Chem Lett14: 1633-1636.

Bardey V, Vallet C, Robas N, Charpentier B, Thouvenot B, Mougin A, Hajnsdorf E, Régnier P, Springer M, Branlant C. 2005. Characterization of

the molecular mechanisms involved in the differential production of erythrose-4-phosphate dehydrogenase, 3-phosphoglycerate kinase and class II fructose-1,6-bisphosphate aldolase in Escherichia coli. Mol Microbiol 57: 1265-1287.

Barile A, Tramonti A, di Salvo ML, Nogués I, Nardella C, Malatesta F, Contestabile R. "Allosteric feedback inhibition of pyridoxine 5'-phosphate oxidase from *Escherichia coli*". J Biol Chem.;294(43):15593-15603. doi: 10.1074/jbc.RA119.009697.

Barile, Anna, Isabel Nogués, Martino L. di Salvo, Victoria Bunik, Roberto Contestabile, and Angela Tramonti. 2020. "Molecular Characterization of Pyridoxine 5′-Phosphate Oxidase and Its Pathogenic Forms Associated with Neonatal Epileptic Encephalopathy." Scientific Reports 10(1):1–15.

Blanchin-Roland S, Blanquet S, Schmitter JM, Fayat G. 1986. The gene for Escherichia coli diadenosine tetraphosphatase is located immediately clockwise to folA and forms an operon with ksgA. Mol Gen Genet 205: 515-522.

Bohney, James P., Margaret L. Fonda, and Richard C. Feldhoff. 1992. "Identification of Lys190 as the Primary Binding Site for Pyridoxal 5′-Phosphate in Human Serum Albumin." FEBS Letters 298(2–3):266–68.

Booth, A. Ashley, Raja G. Khalifah, Parvin Todd, and Billy G. Hudson. 1997. "In Vitro Kinetic Studies of Formation of Antigenic Advanced Glycation End Products (AGEs). Novel Inhibition of Post-Amadori Glycation Pathways." Journal of Biological Chemistry 272(9):5430–37.

Boschi-Muller S, Azza S, Pollastro D, Corbier C, Branlant G. 1997. Comparative enzymatic properties of GapB-encoded erythrose-4-phosphate dehydrogenase of Escherichia coli and phosphorylating glyceraldehyde-3-phosphate dehydrogenase. J Biol Chem 272: 15106-15112.

Bradshaw JS, Kuzminov A. 2003. RdgB acts to avoid chromosome fragmentation in Escherichia coli. Mol Microbiol 48: 1711-1725.

Bramucci E, Milano T, Pascarella S. 2011. Genomic distribution and heterogeneity of MocR-like transcriptional factors containing a domain belonging to the superfamily of the pyridoxal-5'-phosphate dependent enzymes of fold type I. Biochem Biophys Res Commun 415: 88-93.

Braun M, Silhavy TJ. 2002. Imp/OstA is required for cell envelope biogenesis in Escherichia coli. Mol Microbiol 45: 1289-1302.

Calvo JM, Matthews RG. 1994. The leucine-responsive regulatory protein, a global regulator of metabolism in Escherichia coli. Microbiol Rev 58: 466-490.

Cellini, Barbara, Riccardo Montioli, Elisa Oppici, Alessandra Astegno, and Carla Borri Voltattorni. 2014. "The Chaperone Role of the Pyridoxal 5'-Phosphate and Its Implications for Rare Diseases Involving B6-Dependent Enzymes." Clinical Biochemistry 47(3):158–65.

Chen GT, Inouye M. 1994. Role of the AGA/AGG codons, the rarest codons in global gene expression in Escherichia coli. Genes Dev 8: 2641-2652.

Cho BK, Knight EM, Barrett CL, Palsson B. 2008. Genome-wide analysis of Fis binding in Escherichia coli indicates a causative role for A-/AT-tracts. Genome Res 18: 900-910.

Coburn, S. P. et al. (1981) 'Effects of Vitamin B-6 Deficiency and 4′-Deoxypyridoxine on Pyridoxal Phosphate Concentrations, Pyridoxine Kinase and Other Aspects of Metabolism in the Rat', The Journal of Nutrition, 111(2), pp. 391–398. doi: 10.1093/jn/111.2.391.

Contestabile, Roberto, di Salvo Martino Luigi, Bunik Victoria, Tramonti Angela, and Vernì Fiammetta (2020). "The Multifaceted Role of Vitamin B6 in Cancer: Drosophila as a Model System to Investigate DNA Damage." Open Biology 10(3).

Court DL, Gan J, Liang YH, Shaw GX, Tropea JE, Costantino N, Waugh DS, Ji X. 2013. RNase III: Genetics and function; structure and mechanism. Annu Rev Genet 47: 405-431.

Critcher, Matt. 2015. "Vitamin B_6: Low and Very High Concentrations." (September).

Darin, Niklas, Emma Reid, Laurence Prunetti, Lena Samuelsson, Ralf A. Husain, Matthew Wilson, Basma El Yacoubi, Emma Footitt, W. K. Chong, Louise C. Wilson, Helen Prunty, Simon Pope, Simon Heales, Karine Lascelles, Mike Champion, Evangeline Wassmer, Pierangelo Veggiotti, Valérie de Crécy-Lagard, Philippa B. Mills, and Peter T. Clayton. 2016. "Mutations in PROSC Disrupt Cellular Pyridoxal

Phosphate Homeostasis and Cause Vitamin-B6-Dependent Epilepsy." American Journal of Human Genetics 99(6):1325–37.

Dartigalongue C, Missiakas D, Raina S. 2001. Characterization of the Escherichia coli sigma E regulon. J Biol Chem 276: 20866-20875.

Dempsey WB, Pachler PF. 1966. Isolation and characterization of pyridoxine auxotrophs of Escherichia coli. J Bacteriol 91: 642-645.

Deutscher MP. 2009. Maturation and degradation of ribosomal RNA in bacteria. Prog Mol Biol Transl Sci 85: 369-391.

di Salvo, M., Hunt, S. and Schirch, V. (2004) 'Expression, purification, and kinetic constants for human and Escherichia coli pyridoxal kinases', Protein Expr Purif, 36.

di Salvo, Martino L., Safo Martin K., and Contestabile Roberto (2012) "Biomedical Aspects of Pyridoxal 5'-Phosphate Availability." Frontiers in Bioscience (Elite Edition) 4:897–913.8.

di Salvo, Martino Luigi, Roberto Contestabile, and Martin K. Safo. 2011. "Vitamin B 6 Salvage Enzymes: Mechanism, Structure and Regulation." Biochimica et Biophysica Acta - Proteins and Proteomics 1814(11):1597–1608.

Dick, Thomas, Ujjini Manjunatha, Barbara Kappes, and Martin Gengenbacher. 2010. "Vitamin B6 Biosynthesis Is Essential for Survival and Virulence of Mycobacterium Tuberculosis." Molecular Microbiology 78(4):980–88.

Drewke C, Klein M, Clade D, Arenz A, Müller R, Leistner E. 1996. 4-O-phosphoryl-L-threonine, a substrate of the pdxC(serC) gene product involved in vitamin B6 biosynthesis. FEBS Lett390: 179-182.

Duncan K, Coggins JR. 1986. The serC-aro A operon of Escherichia coli. A mixed function operon encoding enzymes from two different amino acid biosynthetic pathways. Biochem J234: 49-57.

Ebrahimi, Neda, Caroline Maltepe, and Adrienne Einarson. 2010. "Optimal Management of Nausea and Vomiting of Pregnancy." International Journal of Women's Health 2(1):241–48.

Ehrenshaft, Marilyn, Piotr Bilski, Ming Y. Li, Colin F. Chignell, and Margaret E. Daub. 1999. “A Highly Conserved Sequence Is a Novel Gene Involved in de Novo Vitamin B6 Biosynthesis.” 96(August):9374–78.

Eswaramoorthy, S., S. Gerchman, V. Graziano, H. Kycia, F. W. Studier, and S. Swaminathan. 2003. “Structure of a Yeast Hypothetical Protein Selected by a Structural Genomics Approach.” Acta Crystallographica - Section D Biological Crystallography 59(1):127–35.

Filenko N, Spiro S, Browning DF, Squire D, Overton TW, Cole J, Constantinidou C. 2007. The NsrR regulon of Escherichia coli K-12 includes genes encoding the hybrid cluster protein and the periplasmic, respiratory nitrite reductase. J Bacteriol 189: 4410-4417.

Fitzpatrick, Teresa B., Cyril Moccand, and Céline Roux. 2010. “Vitamin B6 Biosynthesis: Charting the Mechanistic Landscape.” ChemBioChem 11(9):1185–93.

Fitzpatrick, Teresa B., Nikolaus Amrhein, Barbara Kappes, Peter Macheroux, Ivo Tews, and Thomas Raschle. 2007. “Two Independent Routes of de Novo Vitamin B6 Biosynthesis: Not That Different after All.” Biochemical Journal 407(1):1–13.

Fleming, Angeleen and Andrew J. Copp. 1998. “Embryonic Folate Metabolism and Mouse Neural Tube Defects.” Science 280(5372):2107–9.

Flugel RS, Hwangbo Y, Lambalot RH, Cronan JE, Walsh CT. 2000. Holo-(acyl carrier protein) synthase and phosphopantetheinyl transfer in Escherichia coli. J Biol Chem 275: 959-968.

Foca, F. J. 1985. “Motor and Sensory Neuropathy Secondary to Excessive Pyridoxine Ingestion.” Archives of Physical Medicine and Rehabilitation 66(9):634–36.

Fujisaki S, Nishino T, Katsuki H. 1986. Isoprenoid synthesis in Escherichia coli. Separation and partial purification of four enzymes involved in the synthesis. J Biochem 99: 1327-1337.

Galluzzi, L., E. Vacchelli, J. Michels, P. Garcia, O. Kepp, L. Senovilla, I. Vitale, and G. Kroemer. 2013. “Effects of Vitamin B6 Metabolism on Oncogenesis, Tumor Progression and Therapeutic Responses.” Oncogene 32(42):4995–5004.

Gao, Guo Jian and Margaret L. Fonda. 1994. “Kinetic Analysis and Chemical Modification of Vitamin B6 Phosphatase from Human Erythrocytes.” Journal of Biological Chemistry 269(10):7163–68.

Ghatge, Mohini S., Roberto Contestabile, Martino L. di Salvo, Jigar V. Desai, Amit K. Gandhi, Christina M. Camara, Rita Florio, Isabel N. González, Alessia Parroni, Verne Schirch, and Martin K. Safo. 2012. “Pyridoxal 5′-Phosphate Is a Slow Tight Binding Inhibitor of E. Coli Pyridoxal Kinase.” PLoS ONE 7(7):16–19.

Giardina, Giorgio, Alessio Paone, Angela Tramonti, Roberta Lucchi, Marina Marani, Maria Chiara Magnifico, Amani Bouzidi, Valentino Pontecorvi, Giulia Guiducci, Carlotta Zamparelli, Serena Rinaldo, Alessandro Paiardini, Roberto Contestabile, and Francesca Cutruzzolà. 2018. “The Catalytic Activity of Serine Hydroxymethyltransferase Is Essential for de Novo Nuclear DTMP Synthesis in Lung Cancer Cells.” FEBS Journal 285(17):3238–53.

Grubman, Alexandra, Alexandra Phillips, Marie Thibonnier, Maria Kaparakis-Liaskos, Chad Johnson, Jean Michel Thiberge, Fiona J. Radcliff, Chantal Ecobichon, Agnès Labigne, Hilde de Reuse, George L. Mendz, and Richard L. Ferrero. 2010. “Vitamin B6 Is Required for Full Motility and Virulence in Helicobacter Pylori.” MBio 1(3):1–9.

Gylling, Björn, Robin Myte, Jörn Schneede, Göran Hallmans, Jenny Häggström, Ingegerd Johansson, Arve Ulvik, Per M. Ueland, Bethany Van Guelpen, and Richard Palmqvist. 2017. “Vitamin B-6 and Colorectal Cancer Risk: A Prospective Population-Based Study Using 3 Distinct Plasma Markers of Vitamin B-6 Status.” American Journal of Clinical Nutrition 105(4):897–904.

Hanson G, Coller J. 2018. Codon optimality, bias and usage in translation and mRNA decay. Nat Rev Mol Cell Biol 19: 20-30.

Haskamp V, Karrie S, Mingers T, Barthels S, Alberge F, Magalon A, Müller K, Bill E, Lubitz W, Kleeberg K et al. 2018. The radical SAM protein HemW is a heme chaperone. J Biol Chem 293: 2558-2572.

Hatch, J., D. Coman, P. Clayton, P. Mills, S. Calvert, R. I. Webster, and Kate Riney. 2016. “Normal Neurodevelopmental Outcomes in PNPO Deficiency: A Case Series and Literature Review.” JIMD Reports 26:91–97.

Helmann JD, Chamberlin MJ. 1988. Structure and function of bacterial sigma factors. Annu Rev Biochem 57: 839-872.

Henikoff S, Wallace JC. 1988. Detection of protein similarities using nucleotide sequence databases. Nucleic Acids Res16: 6191-6204.

Huq, M. D. Mostaqul, Nien-Pei Tsai, Ya-Ping Lin, LeeAnn Higgins, and Li-Na Wei. 2007. "Vitamin B6 Conjugation to Nuclear Corepressor RIP140 and Its Role in Gene Regulation." Nature Chemical Biology 3(3):161–65.

Ikemura T. 1985. Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol 2: 13-34.

Ishihama Y, Schmidt T, Rappsilber J, Mann M, Hartl FU, Kerner MJ, Frishman D. 2008. Protein abundance profiling of the Escherichia coli cytosol. BMC Genomics 9: 102.

Ito T, Downs DM. 2020. Pyridoxal reductase, PdxI, is critical for salvage of pyridoxal in. J Bacteriol.

Ito T, Uozumi N, Nakamura T, Takayama S, Matsuda N, Aiba H, Hemmi H, Yoshimura T. 2009. The implication of YggT of Escherichia coli in osmotic regulation. Biosci Biotechnol Biochem 73: 2698-2704.

Ito, T. et al. (2013) 'Conserved pyridoxal protein that regulates ile and val metabolism', Journal of Bacteriology, 195(24), pp. 5439–5449. doi: 10.1128/JB.00593-13.

Ito, T. et al. (2019) 'Conserved Pyridoxal 5'-Phosphate-Binding Protein YggS Impacts Amino Acid Metabolism through Pyridoxine 5'-Phosphate in Escherichia coli', Applied and Environmental Microbiology. American Society for Microbiology, 85(11). doi: 10.1128/aem.00430-19.

Iurescia, S. et al. (1996) 'Site-directed mutagenesis techniques in the study of Escherichia coli serine hydroxymethyltransferase', Protein Expression and Purification, 7(3), pp. 323–328. doi: 10.1006/prep.1996.0046.

Kanaya S, Yamada Y, Kudo Y, Ikemura T. 1999. Studies of codon usage and tRNA genes of 18 unicellular organisms and quantification of Bacillus subtilis tRNAs: gene expression level and species-specific diversity of codon usage based on multivariate analysis. Gene 238: 143-155.

Kannan, Krishnaswamy and Sushil K. Jain. 2004. "Effect of Vitamin B6 on Oxygen Radicals, Mitochondrial Membrane Potential, and Lipid

Peroxidation in H2O2-Treated U937 Monocytes." Free Radical Biology and Medicine 36(4):423–28.

Karlin S, Mrázek J, Campbell A, Kaiser D. 2001. Characterizations of highly expressed genes of four fast-growing bacteria. J Bacteriol 183: 5025-5040.

Kayashima, Tomoko, Kenta Tanaka, Yukako Okazaki, Kiminori Matsubara, Noriyuki Yanaka, and Norihisa Kato. 2011. "Consumption of Vitamin B6 Reduces Colonic Damage and Protein Expression of HSP70 and HO-1, the Anti-Tumor Targets, in Rats Exposed to 1,2-Dimethylhydrazine." Oncology Letters 2(6):1243–46.

Kim J, Flood JJ, Kristofich MR, Gidfar C, Morgenthaler AB, Fuhrer T, Sauer U, Snyder D, Cooper VS, Ebmeier CC et al. 2019. Hidden resources in the. Proc Natl Acad Sci U S A 116: 24164-24173.

Kim J, Kershner JP, Novikov Y, Shoemaker RK, Copley SD. 2010. Three serendipitous pathways in E. coli can bypass a block in pyridoxal-5'-phosphate synthesis. Mol Syst Biol 6: 436.

Kirby J, Nishimoto M, Chow RW, Baidoo EE, Wang G, Martin J, Schackwitz W, Chan R, Fortman JL, Keasling JD. 2015. Enhancing Terpene yield from sugars via novel routes to 1-deoxy-d-xylulose 5-phosphate. Appl Environ Microbiol 81: 130-138.

Kolb A, Busby S, Buc H, Garges S, Adhya S. 1993. Transcriptional regulation by cAMP and its receptor protein. Annu Rev Biochem 62: 749-795.

Kotake, Y., T. Ueda, T. Mori, S. Igaki, and M. Hattori. 1975. "Abnormal Tryptophan Metabolism and Experimental Diabetes by Xanthurenic Acid (XA)." Acta Vitaminologica et Enzymologica 29(1–6):236–39.

Kuznetsova, E. et al. (2006) 'Genome-wide analysis of substrate specificities of the Escherichia coli haloacid dehalogenase-like phosphatase family', Journal of Biological Chemistry, 281(47), pp. 36149–36161. doi: 10.1074/jbc.M605449200.

Labella, J. I. et al. (2020) 'Distinctive features of PipX, a unique signaling protein of cyanobacteria', Life, 10(6). doi: 10.3390/life10060079.

Labella, José I., Raquel Cantos, Javier Espinosa, Alicia Forcada-Nadal, Vicente Rubio, and Asunción Contreras. 2017. "PipY, a Member of the

Conserved COG0325 Family of PLP-Binding Proteins, Expands the Cyanobacterial Nitrogen Regulatory Network." Frontiers in Microbiology 8(JUL):1–16.

Laber, B. et al. (1999) 'Vitamin B6 biosynthesis: Formation of pyridoxine 5'-phosphate from 4-(phosphohydroxy)-L-threonine and 1-deoxy-D-xylulose-5-phosphate by PdxA and PdxJ protein', FEBS Letters, 449(1). doi: 10.1016/S0014-5793(99)00393-2.

Lam HM, Tancula E, Dempsey WB, Winkler ME. 1992. Suppression of insertions in the complex pdxJ operon of Escherichia coli K-12 by lon and other mutations. J Bacteriol 174: 1554-1567.

Lam HM, Winkler ME. 1990. Metabolic relationships between pyridoxine (vitamin B6) and serine biosynthesis in Escherichia coli K-12. J Bacteriol 172: 6518-6528.

Lam HM, Winkler ME. 1992. Characterization of the complex pdxH-tyrS operon of Escherichia coli K-12 and pleiotropic phenotypes caused by pdxH insertion mutations. J Bacteriol 174: 6033-6045.

Lazar SW, Almirón M, Tormo A, Kolter R. 1998. Role of the Escherichia coli SurA protein in stationary-phase survival. J Bacteriol 180: 5704-5711.

Liu, Zhan, Peng Li, Zhi Hong Zhao, Yu Zhang, Zhi Min Ma, and Shuang Xi Wang. 2016. "Vitamin B6 Prevents Endothelial Dysfunction, Insulin Resistance, and Hepatic Lipid Accumulation in Apoe -/- Mice Fed with High-Fat Diet." Journal of Diabetes Research 2016.

Lois LM, Campos N, Putra SR, Danielsen K, Rohmer M, Boronat A. 1998. Cloning and characterization of a gene from Escherichia coli encoding a transketolase-like enzyme that catalyzes the synthesis of D-1-deoxyxylulose 5-phosphate, a common precursor for isoprenoid, thiamin, and pyridoxol biosynthesis. Proc Natl Acad Sci U S A 95: 2105-2110.

Lu P, Vogel C, Wang R, Yao X, Marcotte EM. 2007. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nat Biotechnol 25: 117-124.

Luciano DJ, Levenson-Palmer R, Belasco JG. 2019. Stresses that Raise Np. Mol Cell 75: 957-966.e958.

MacFarlane, Amanda J., Donald D. Anderson, Per Flodby, Cheryll A. Perry, Robert H. Allen, Sally P. Stabler, and Patrick J. Stover. 2011. “Nuclear Localization of de Novo Thymidylate Biosynthesis Pathway Is Required to Prevent Uracil Accumulation in DNA.” Journal of Biological Chemistry 286(51):44015–22.

Maheux AF, Boudreau DK, Bergeron MG, Rodriguez MJ. 2014. Characterization of Escherichia fergusonii and Escherichia albertii isolated from water. J Appl Microbiol 117: 597-609.

Man TK, Pease AJ, Winkler ME. 1997. Maximization of transcription of the serC (pdxF)-aroA multifunctional operon by antagonistic effects of the cyclic AMP (cAMP) receptor protein-cAMP complex and Lrp global regulators of Escherichia coli K-12. J Bacteriol179: 3458-3469.

Matsunaga J, Dyer M, Simons EL, Simons RW. 1996. Expression and regulation of the rnc and pdxJ operons of Escherichia coli. Mol Microbiol 22: 977-989.

McCormick, D. B. 1989. “Two Interconnected B Vitamins: Riboflavin and Pyridoxine.” Physiological Reviews 69(4):1170–98.

Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, Taboada B, Jimenez-Jacinto V, Salgado H, Juárez K, Contreras-Moreira B et al. 2009. Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in E. coli. PLoS One 4: e7526.

Middleton, Henry M. 1990. “Intestinal Hydrolysis of Pyridoxal 5′-Phosphate in Vitro and in Vivo in the Rat - Effect of Amino Acids and Oligopeptides.” Digestive Diseases and Sciences 35(1):113–20.

Mills, Philippa B., Robert A. H. Surtees, Michael P. Champion, Clare E. Beesley, Neil Dalton, Peter J. Scamber, Simon J. R. Heales, Anthony Briddon, Irene Scheimberg, Georg F. Hoffmann, Johannes Zschocke, and Peter T. Clayton. 2005. “Neonatal Epileptic Encephalopathy Caused by Mutations in the PNPO Gene Encoding Pyridox(Am)Ine 5′-Phosphate Oxidase.” Human Molecular Genetics 14(8):1077–86.

Mittenhuber, G. 2001. “Phylogenetic Analyses and Comparative Genomics of Vitamin B6 (Pyridoxine) and Pyridoxal Phosphate Biosynthesis

Pathways.” Journal of Molecular Microbiology and Biotechnology 3(1):1–20.

Moreno-Navarrete, José María, Mariona Jove, Francisco Ortega, Gemma Xifra, Wifredo Ricart, Èlia Obis, Reinald Pamplona, Manuel Portero-Otin, and José Manuel Fernández-Real. 2016. “Metabolomics Uncovers the Role of Adipose Tissue PDXK in Adipogenesis and Systemic Insulin Sensitivity.” Diabetologia 59(4):822–32.

Mulligan JH, Snell EE. 1976. Transport and metabolism of vitamin B6 in Salmonella typhimurium LT2. J Biol Chem 251: 1052-1056.

Mulligan JH, Snell EE. 1977. Transport and metabolism of vitamin B6 in lactic acid bacteria. J Biol Chem 252: 835-839.

Musayev, Faik N., Martino L. Di Salvo, Tzu-Ping Ko, Verne Schirch, and Martin K. Safo. 2003. “Structure and Properties of Recombinant Human Pyridoxine 5′-Phosphate Oxidase.” Protein Science 12(7):1455–63.

Nonaka G, Blankschien M, Herman C, Gross CA, Rhodius VA. 2006. Regulon and promoter analysis of the E. coli heat-shock factor, sigma32, reveals a multifaceted cellular response to heat stress. Genes Dev 20: 1776-1789.

Ohtsu I, Kakuda N, Tsukagoshi N, Dokyu N, Takagi H, Wachi M, Aono R. 2004. Transcriptional analysis of the ostA/imp gene involved in organic solvent sensitivity in Escherichia coli. Biosci Biotechnol Biochem 68: 458-461.

Paone, A., M. Marani, A. Fiascarelli, S. Rinaldo, G. Giardina, R. Contestabile, A. Paiardini, and F. Cutruzzolà. 2014. “SHMT1 Knockdown Induces Apoptosis in Lung Cancer Cells by Causing Uracil Misincorporation.” Cell Death and Disease 5(11):1–11.

Partridge JD, Bodenmiller DM, Humphrys MS, Spiro S. 2009. NsrR targets in the Escherichia coli genome: new insights into DNA sequence requirements for binding and a role for NsrR in the regulation of motility. Mol Microbiol 73: 680-694.

Pease AJ, Roa BR, Luo W, Winkler ME. 2002. Positive growth rate-dependent regulation of the pdxA, ksgA, and pdxB genes of Escherichia coli K-12. J Bacteriol184: 1359-1369.

Percudani, Riccardo and Alessio Peracchi. 2003. “A Genomic Overview of Pyridoxal-Phosphate-Dependent Enzymes.” EMBO Reports 4(9):850–54.

Pfaffl, M W. “A new mathematical model for relative quantification in real-time RT-PCR.” *Nucleic acids research* vol. 29,9 (2001): e45. doi:10.1093/nar/29.9.e45.

Picard F, Milhem H, Loubière P, Laurent B, Cocaign-Bousquet M, Girbal L. 2012. Bacterial translational regulations: high diversity between all mRNAs and major role in gene expression. BMC Genomics 13: 528.

Plecko, Barbara, Markus Zweier, Anaïs Begemann, Deborah Mathis, Bernhard Schmitt, Pasquale Striano, Martina Baethmann, Maria Stella Vari, Francesca Beccaria, Federico Zara, Lisa M. Crowther, Pascal Joset, Heinrich Sticht, Sorina Mihaela Papuc, and Anita Rauch. 2017. “Confirmation of Mutations in PROSC as a Novel Cause of Vitamin B 6-Dependent Epilepsy.” Journal of Medical Genetics 54(12):809–14.

Prunetti, L. et al. (2016) ‘Evidence that COG0325 proteins are involved in PLP homeostasis’, Microbiology (United Kingdom), 162(4), pp. 694–706. doi: 10.1099/mic.0.000255.

Quax TE, Claassens NJ, Söll D, van der Oost J. 2015. Codon Bias as a Means to Fine-Tune Gene Expression. Mol Cell 59: 149-161.

Ramseier TM. 1996. Cra and the control of carbon flux via metabolic pathways. Res Microbiol 147: 489-493.

Raschle, Thomas, Nikolaus Amrhein, and Teresa B. Fitzpatrick. 2005. “On the Two Components of Pyridoxal 5’-Phosphate Synthase from Bacillus Subtilis.” 280(37):32291–300.

Rippa, M., M. Signorini, and S. Pontremoli. 1967. “Purification and Properties of Two Forms of 6-Phosphogluconate Dehydrogenase from Candida Utilis.” European Journal of Biochemistry 1(2):170–78.

Roa BB, Connolly DM, Winkler ME. 1989. Overlap between pdxA and ksgA in the complex pdxA-ksgA-apaG-apaH operon of Escherichia coli K-12. J Bacteriol171: 4767-4777.

Rorsman, Fredrik, Eystein S. Husebye, Ola Winqvist, Elisabeth Björk, F. Anders Karlsson, and Olle Kämpe. 1995. “Aromatic-L-Amino-Acid Decarboxylase, a Pyridoxal Phosphate-Dependent Enzyme, Is a β-Cell

Autoantigen." Proceedings of the National Academy of Sciences of the United States of America 92(19):8626–29.

Rubí, B. 2012. "Pyridoxal 5'-Phosphate (PLP) Deficiency Might Contribute to the Onset of Type I Diabetes." Medical Hypotheses 78(1):179–82.

Ryan-Harshman, Milly and Walid Aldoori. 2007. "Carpal Tunnel Syndrome and Vitamin B6." Canadian Family Physician 53(7):1161–62.

Safo MK, Musayev FN, Hunt S, di Salvo ML, Scarsdale N, Schirch V. 2004. Crystal structure of the PdxY Protein from Escherichia coli. J Bacteriol 186: 8074-8082.

Safo, Martin K., Faik N. Musayev, Martino L. Di Salvo, and Verne Schirch. 2001. "X-Ray Structure of Escherichia Coli Pyridoxine 5′-Phosphate Oxidase Complexed with Pyridoxal 5′-Phosphate at 2.0 Å Resolution." Journal of Molecular Biology 310(4):817–26.

Said, Hamid M. 2004. "Recent Advances in Carrier-Mediated Intestinal Absorption of Water-Soluble Vitamins." Annual Review of Physiology 66:419–46.

Schneider BL, Hernandez VJ, Reitzer L. 2013. Putrescine catabolism is a metabolic response to several stresses in Escherichia coli. Mol Microbiol 88: 537-550.

Schoenlein PV, Roa BB, Winkler ME. 1989. Divergent transcription of pdxB and homology between the pdxB and serA gene products in Escherichia coli K-12. J Bacteriol 171: 6084-6092.

Sévin DC, Fuhrer T, Zamboni N, Sauer U. 2017. Nontargeted in vitro metabolomics for high-throughput identification of novel enzymes in Escherichia coli. Nat Methods 14: 187-194.

Sharp PM, Li WH. 1987. The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res 15: 1281-1295.

Spector, Reynold and Conrad E. Johanson. 2007. "Vitamin Transport and Homeostasis in Mammalian Brain: Focus on Vitamins B and E." Journal of Neurochemistry 103(2):425–38.

Sprenger GA, Schörken U, Wiegert T, Grolle S, de Graaf AA, Taylor SV, Begley TP, Bringer-Meyer S, Sahm H. 1997. Identification of a thiamin-

dependent synthase in *Escherichia coli* required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol. Proc Natl Acad Sci U S A 94: 12857-12862.

Stanulović, M., V. Jeremić, V. Leskovac, and S. Chaykin. 1976. "New Pathway of Conversion of Pyridoxal to 4-Pyridoxic Acid." Enzyme 21:357–69.

Sugimoto R, Saito N, Shimada T, Tanaka K. 2018. Identification of YbhA as the pyridoxal 5'-phosphate (PLP) phosphatase in Escherichia coli: Importance of PLP homeostasis on the bacterial growth. J Gen Appl Microbiol 63: 362-368.

Sugimoto, R. et al. (2017) 'Identification of YbhA as the pyridoxal 5'-phosphate (PLP) phosphatase in escherichia coli: Importance of PLP homeostasis on the bacterial growth', Journal of General and Applied Microbiology, 63(6), pp. 362–368. doi: 10.2323/jgam.2017.02.008.

Suvorova IA, Rodionov DA. 2016. Comparative genomics of pyridoxal 5'-phosphate-dependent transcription factor regulons in. Microb Genom 2: e000047.

Thouvenot B, Charpentier B, Branlant C. 2004. The strong efficiency of the Escherichia coli gapA P1 promoter depends on a complex combination of functional determinants. Biochem J 383: 371-382.

Tramonti A, Milano T, Nardella C, di Salvo ML, Pascarella S, Contestabile R. 2017. Salmonella typhimurium PtsJ is a novel MocR-like transcriptional repressor involved in regulating the vitamin B6 salvage pathway. FEBS J 284: 466-484.

Tramonti, Angela, Caterina Nardella, Martino L. di Salvo, Stefano Pascarella, and Roberto Contestabile. 2018. "The MocR-like Transcription Factors: Pyridoxal 5′-Phosphate-Dependent Regulators of Bacterial Metabolism." FEBS Journal 285(21):3925–44.

Tremiño, L. et al. (2017) 'Studies on cyanobacterial protein PipY shed light on structure, potential functions, and vitamin B6-dependent epilepsy', FEBS Letters, 591(20), pp. 3431–3442. doi: 10.1002/1873-3468.12841.

Tremiño, L., Forcada-Nadal, A. and Rubio, V. (2018) 'Insight into vitamin B 6 -dependent epilepsy due to PLPBP (previously PROSC) missense

mutations', Human Mutation. John Wiley and Sons Inc., 39(7), pp. 1002–1013. doi: 10.1002/humu.23540.

Turgeon, Marc Olivier, Nicholas J. S. Perry, and George Poulogiannis. 2018. "DNA Damage, Repair, and Cancer Metabolism." Frontiers in Oncology 8(FEB).

Urbonavicius J, Qian Q, Durand JM, Hagervall TG, Björk GR. 2001. Improvement of reading frame maintenance is a common function for several tRNA modifications. EMBO J 20: 4863-4873.

van Gemen B, Twisk J, van Knippenberg PH. 1989. Autogenous regulation of the Escherichia coli ksgA gene at the level of translation. J Bacteriol 171: 4002-4008.

Vu, Huong N., Tomokazu Ito, and Diana M. Downs. 2020. " The Role of YggS in Vitamin B 6 Homeostasis in Salmonella Enterica Is Informed by Heterologous Expression of Yeast SNZ3 ." Journal of Bacteriology (September).

Wang, Huei Shyong and Meng Fai Kuo. 2007. "Vitamin B6 Related Epilepsy during Childhood." Chang Gung Medical Journal 30(5):396–401.

Ware, Tyson L., John Earl, Gajja S. Salomons, Eduard A. Struys, Heidi L. Peters, Katherine B. Howell, James J. Pitt, and Jeremy L. Freeman. 2014. "Typical and Atypical Phenotypes of PNPO Deficiency with Elevated CSF and Plasma Pyridoxamine on Treatment." Developmental Medicine and Child Neurology 56(5):498–502.

Wilson, Matthew P., Barbara Plecko, Philippa B. Mills, and Peter T. Clayton. 2019. "Disorders Affecting Vitamin B 6 Metabolism." Journal of Inherited Metabolic Disease.

Windebank, A. J. 1985. "Neurotoxicity of Pyridoxine Analogs Is Related to Coenzyme Structure." Neurochemical Pathology 3(3):159–67.

Wiśniewski JR, Rakus D. 2014. Multi-enzyme digestion FASP and the 'Total Protein Approach'-based absolute quantification of the Escherichia coli proteome. J Proteomics 109: 322-331.

Wu T, McCandlish AC, Gronenberg LS, Chng SS, Silhavy TJ, Kahne D. 2006. Identification of a protein complex that assembles

lipopolysaccharide in the outer membrane of Escherichia coli. Proc Natl Acad Sci U S A103: 11754-11759.

Wyatt, Katrina M., Paul W. Dimmock, Peter W. Jones, and P. M. Shaugh. O'brien. 1999. "Efficacy of Vitamin B-6 in the Treatment of Premenstrual Syndrome: Systematic Review." Bmj 318(7195):1375–81.

Xie, Fang, Gang Li, Yalei Wang, Yanhe Zhang, Long Zhou, Chengcheng Wang, Shuanghong Liu, Siguo Liu, and Chunlai Wang. 2017. "Pyridoxal Phosphate Synthases PdxS/PdxT Are Required for Actinobacillus Pleuropneumoniae Viability, Stress Tolerance and Virulence." PLoS ONE 12(4):1–17.

Yang Y, Zhao G, Man TK, Winkler ME. 1998b. Involvement of the gapA- and epd (gapB)-encoded dehydrogenases in pyridoxal 5'-phosphate coenzyme biosynthesis in Escherichia coli K-12. J Bacteriol 180: 4294-4299.

Yang, Emily S. and Verne Schirch. 2000. "Tight Binding of Pyridoxal 5'-Phosphate to Recombinant Escherichia Coli Pyridoxine 5'-Phosphate Oxidase." Archives of Biochemistry and Biophysics 377(1):109–14.

Yang, Y. et al. (1998) 'Identification and function of the pdxY gene, which encodes a novel pyridoxal kinase involved in the salvage pathway of pyridoxal 5'-phosphate biosynthesis in Escherichia coli K-12', Journal of Bacteriology, 180(7), pp. 1814–1821. doi: 10.1128/jb.180.7.1814-1821.1998.

Yang, Y., Zhao, G. and Winkler, M. E. (1996) 'Identification of the pdxK gene that encodes pyridoxine (vitamin B6) kinase in Escherichia coli K-12', FEMS Microbiology Letters, 141(1), pp. 89–95. doi: 10.1016/0378-1097(96)00203-0.

Yeh, J. I. et al. (2002) 'Multistate binding in pyridoxine 5′-phosphate synthase: 1.96 Å crystal structure in complex with 1-deoxy-D-xylulose phosphate', Biochemistry, 41(39), pp. 11649–11657. doi: 10.1021/bi026292t.

Zaslaver A, Bren A, Ronen M, Itzkovitz S, Kikoin I, Shavit S, Liebermeister W, Surette MG, Alon U. 2006. A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. Nat Methods 3: 623-628.

Zein, Fairuz, Yan Zhang, You-na Kang, Kristin Burns, Tadhg P. Begley, and Steven E. Ealick. 2006. “Structural Insights into the Mechanism of the PLP Synthase Holoenzyme From.” 14609–20.

Zhao G, Pease AJ, Bharani N, Winkler ME. 1995. Biochemical characterization of gapB-encoded erythrose 4-phosphate dehydrogenase of Escherichia coli K-12 and its possible role in pyridoxal 5'-phosphate biosynthesis. J Bacteriol 177: 2804-2812.

Zheng D, Constantinidou C, Hobman JL, Minchin SD. 2004. Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. Nucleic Acids Res 32: 5874-5893.

Zuo, H., P. M. Ueland, Midttun, G. S. Tell, A. Fanidi, W. Zheng, X. Shu, Y. Xiang, J. Wu, R. Prentice, M. Pettinger, C. A. Thomson, G. G. Giles, A. Hodge, Q. Cai, W. J. Blot, M. Johansson, J. Hultdin, K. Grankvist, V. L. Stevens, M. L. McCullough, S. J. Weinstein, D. Albanes, R. G. Ziegler, N. D. Freedman, N. E. Caporaso, A. Langhammer, K. Hveem, M. Næss, J. E. Buring, I. Lee, J. M. Gaziano, G. Severi, X. Zhang, M. J. Stampfer, J. Han, A. Zeleniuch-Jacquotte, L. L. Marchand, J. Yuan, R. Wang, W. Koh, Y. Gao, U. Ericson, K. Visvanathan, M. R. Jones, C. Relton, P. Brennan, M. Johansson, and A. Ulvik. 2019. “Vitamin B6 Catabolism and Lung Cancer Risk: Results from the Lung Cancer Cohort Consortium (LC3).” Annals of Oncology 30(3):478–85.

Collaborations

Prof. Francesco Malatesta
Department of Biochemical Sciences “A. Rossi Fanelli”
Sapienza University of Rome

Prof. Martin Safo
Department of Medicinal Chemistry
Virginia Commonwealth University

Prof. Valerie de Crecy-Lagard
Department of Microbiology and Cell Science
University of Florida

Publications

1. “Manganese is a *Deinococcus radiodurans* growth limiting factor in rich culture medium”.
   Borsetti F, Dal Piaz F, D'Alessio F, Stefan A, Brandimarti R, Sarkar A, Datta A, Montón Silva A, den Blaauwen T, Alberto M, Spisni E, Hochkoeppler A.
   Microbiology 2018, doi: 10.1099/mic.0.000698.

2. “Purification from *Deinococcus radiodurans* of a 66 kDa ABC transporter acting on peptides containing at least 3 amino acids”.
   Stefan A, Gentilucci L, Piaz FD, D'Alessio F, Santino F, Hochkoeppler A.
   Biochemical and Biophysical Research Communications 2020, doi: 10.1016/j.bbrc.2020.06.060.

3. “An Engineered *Escherichia coli* Strain with Synthetic Metabolism for in-Cell Production of Translationally Active Methionine Derivatives”.
   Schipp CJ, Ma Y, Al-Shameri A, D'Alessio F, Neubauer P, Contestabile R, Budisa N, di Salvo M. L.
   ChemBioChem 2020, doi: 10.1002/cbic.202000257.