J. Mol. Microbiol. Biotechnol. (2001) 3(1): 1-20.

Phylogenetic Analyses and Comparative Genomics of Vitamin B_6 (Pyridoxine) and Pyridoxal Phosphate Biosynthesis Pathways

Gerhard Mittenhuber*

Institut für Mikrobiologie und Molekularbiologie, Ernst-Moritz-Arndt-Universität, F.-L. Jahnstr. 15, D-17487 Greifswald, Germany

Abstract

Vitamin B_6 in its active form pyridoxal phosphate is an essential coenzyme of many diverse enzymes. Biochemistry, enzymology and genetics of *de novo* vitamin B_6 biosynthesis have been primarily investigated in *Escherichia coli*. Database searches revealed that the key enzymes involved in ring closure of the aromatic pyridoxin ring (PdxA; PdxJ) are present mainly in genomes of bacteria constituting the γ subdivision of proteobacteria. The distribution of DXS, a transketolase-like enzyme involved in vitamin B_6 biosynthesis as well as in thiamine and isoprenoid biosynthesis and the distribution of vitamin B_6 modifying enzymes (PdxH: oxidase; PdxK: kinase) was also analyzed. These enzymes are also present in the genomes of animals. Two recent papers (Ehrenshaft *et al.*, 1999, Proc. Natl. Acad. Sci. USA. 96 : 9374-9378; Osmani *et al.*, 1999, J. Biol. Chem. 274 : 23565-23569) show the involvement of an extremely conserved protein (a member of the UPF0019 or SNZ family) found in all three domains of life (bacteria, archaea, eukarya) in an alternative vitamin B_6 biosynthesis pathway. Members of this family were previously identified as a stationary phase inducible protein in yeast, as an ethylene responsible protein in plants and in a marine sponge, as a singlet oxygen resistance protein in *Cercospora nicotianae* and as a cumene hydroperoxide and H_2O_2 inducible protein in *Bacillus subtilis.* In yeast, the SNZ protein interacts with another protein called SNO which also represents a member of a highly conserved protein family (called UPF0030 or SNO family). Phylogenetic trees for the DXS, PdxA, PdxJ, PdxH, PdxK, SNZ and SNO protein families are presented and possible implications of the two different vitamin B_6 biosynthesis pathways in cellular metabolism are discussed. A radically different view of bacterial evolution (Gupta, 2000, Crit. Rev. Microbiol. 26: 111-131) which proposes a linear rather than a treelike evolutionary relationship between procaryotic species indicates that the γ subdivision of proteobacteria represents the most recently evolved bacterial lineage. This proposal might help to explain why the PdxA/PdxJ pathway is largely restricted to this subdivision.

Received August 2, 2000; accepted September 4, 2000.
*For correspondence. Email mittenhu@microbio4.biologie.uni-greifswald.de; Tel. +49-3834-864218; Fax. +49-3834-864202.

Introduction

Pyridoxal 5'-phosphate (PLP), the active form of vitamin B_6, is an essential cofactor of many enzymes involved in amino acid metabolism in all cells (for reviews see Dolphin *et al.*, 1986; Dakshinamurti, 1990). The most common reactions involve (1) transaminases catalyzing the conversion of α-ketoacids to amino acids, and (2) amino acid racemases producing D-amino acids originating from L-amino acids (Grogan, 1988). But PLP is also a coenzyme of (3) sulfinate desulfinase catalyzing the removal of elemental sulfur and selenium atoms from L-cysteine, L-cystine, L-selenocysteine, and L-selenocystine to produce L-alanine (Mihara *et al.*, 1997), (4) of glycogen phosphorylases (Helmreich, 1992), (5) of diaminopelargonic acid synthase, an enzyme involved in biotin synthesis (Kack *et al.*, 1999) and (6) of 1-aminocyclopropane-1-carboxylate synthase (Rottmann *et al.*, 1991) synthesizing 1-aminocyclopropane-1-carboxylic acid (ACC) - the precursor of the plant hormone ethylene - from S-adenosylmethionine. Despite these vast differences in the nature of the reactions catalyzed by PLP-dependent enzymes, their three-dimensional structures share common structural elements (Capitani *et al.*, 1999, Denessiouk *et al.*, 1999, Kack *et al.*, 1999, Schneider *et al.*, 2000).

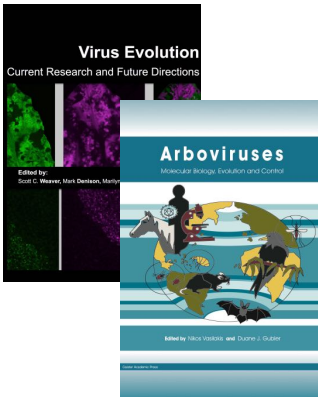
The biochemistry of *de novo* PLP biosynthesis has been studied in the gramnegative model organism *Escherichia coli* (Hill and Spenser, 1996). Also, molecular cloning and characterization of genes coding for enzymes involved in PLP biosynthesis was performed using this organism, mainly by Malcolm E. Winkler and coworkers: The PLP precursor pyridoxine 5'-phosphate (PNP) is synthesized by the PdxA and PdxJ enzymes using 4-phosphohydroxy-L-threonine (4PHT; synonym: 3-hydroxyhomoserine) and 1-deoxy-D-xylulose 5-phosphate (DXP) as substrates (Lam *et al.*, 1992; Zhao and Winkler, 1996; Cane *et al.*, 1998; Laber *et al*, 1999). PNP is oxidized to the active coenzyme PLP by the action of PdxH oxidase, a flavoprotein (Lam and Winkler, 1992; Zhao and Winkler, 1995; Notheis *et al.*, 1995).

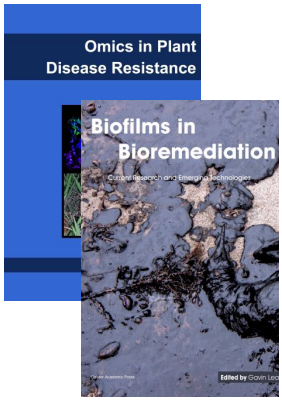
The two substrates DXP and 4PHT are supplied by two independent pathways (Figure 1), which are both linked to carbohydrate metabolism, *i.e.* glycolysis and the pentose phosphate cyle: DXP, which is also a precursor in isoprenoid and thiamine (vitamin B_1) biosynthesis (Begley *et al.*, 1999) is synthesized by the transketolase-like enzyme DXP-synthase (DXS) (Sprenger *et al.*, 1997; Lois *et al.*, 1998) using pyruvate and D-glyceraldehyde-3-

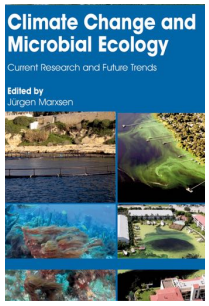
Further Reading

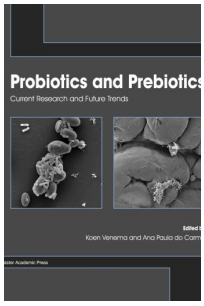
Caister Academic Press is a leading academic publisher of advanced texts in microbiology, molecular biology and medical research. Full details of all our publications at **caister.com**

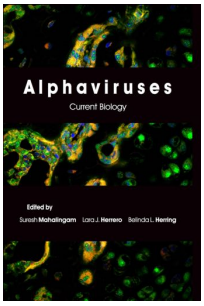
- **MALDI-TOF Mass Spectrometry in Microbiology**
  Edited by: M Kostrzewa, S Schubert (2016)
  www.caister.com/malditof
- ***Aspergillus* and *Penicillium* in the Post-genomic Era**
  Edited by: RP Vries, IB Gelber, MR Andersen (2016)
  www.caister.com/aspergillus2
- **The Bacteriocins: Current Knowledge and Future Prospects**
  Edited by: RL Dorit, SM Roy, MA Riley (2016)
  www.caister.com/bacteriocins
- **Omics in Plant Disease Resistance**
  Edited by: V Bhadauria (2016)
  www.caister.com/opdr
- **Acidophiles: Life in Extremely Acidic Environments**
  Edited by: R Quatrini, DB Johnson (2016)
  www.caister.com/acidophiles
- **Climate Change and Microbial Ecology: Current Research and Future Trends**
  Edited by: J Marxsen (2016)
  www.caister.com/climate
- **Biofilms in Bioremediation: Current Research and Emerging Technologies**
  Edited by: G Lear (2016)
  www.caister.com/biorem
- **Microalgae: Current Research and Applications**
  Edited by: MN Tsaloglou (2016)
  www.caister.com/microalgae
- **Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives**
  Edited by: H Shintani, A Sakudo (2016)
  www.caister.com/gasplasma
- **Virus Evolution: Current Research and Future Directions**
  Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)
  www.caister.com/virusevol
- **Arboviruses: Molecular Biology, Evolution and Control**
  Edited by: N Vasilakis, DJ Gubler (2016)
  www.caister.com/arbo
- ***Shigella*: Molecular and Cellular Biology**
  Edited by: WD Picking, WL Picking (2016)
  www.caister.com/shigella
- **Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment**
  Edited by: AM Romaní, H Guasch, MD Balaguer (2016)
  www.caister.com/aquaticbiofilms
- **Alphaviruses: Current Biology**
  Edited by: S Mahalingam, L Herrero, B Herring (2016)
  www.caister.com/alpha
- **Thermophilic Microorganisms**
  Edited by: F Li (2015)
  www.caister.com/thermophile

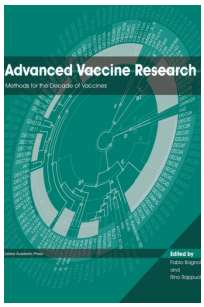


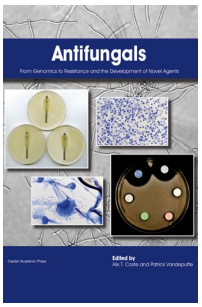












- **Flow Cytometry in Microbiology: Technology and Applications**
  Edited by: MG Wilkinson (2015)
  www.caister.com/flow
- **Probiotics and Prebiotics: Current Research and Future Trends**
  Edited by: K Venema, AP Carmo (2015)
  www.caister.com/probiotics
- **Epigenetics: Current Research and Emerging Trends**
  Edited by: BP Chadwick (2015)
  www.caister.com/epigenetics2015
- ***Corynebacterium glutamicum*: From Systems Biology to Biotechnological Applications**
  Edited by: A Burkovski (2015)
  www.caister.com/cory2
- **Advanced Vaccine Research Methods for the Decade of Vaccines**
  Edited by: F Bagnoli, R Rappuoli (2015)
  www.caister.com/vaccines
- **Antifungals: From Genomics to Resistance and the Development of Novel Agents**
  Edited by: AT Coste, P Vandeputte (2015)
  www.caister.com/antifungals
- **Bacteria-Plant Interactions: Advanced Research and Future Trends**
  Edited by: J Murillo, BA Vinatzer, RW Jackson, et al. (2015)
  www.caister.com/bacteria-plant
- ***Aeromonas***
  Edited by: J Graf (2015)
  www.caister.com/aeromonas
- **Antibiotics: Current Innovations and Future Trends**
  Edited by: S Sánchez, AL Demain (2015)
  www.caister.com/antibiotics
- ***Leishmania*: Current Biology and Control**
  Edited by: S Adak, R Datta (2015)
  www.caister.com/leish2
- ***Acanthamoeba*: Biology and Pathogenesis (2nd edition)**
  Author: NA Khan (2015)
  www.caister.com/acanthamoeba2
- **Microarrays: Current Technology, Innovations and Applications**
  Edited by: Z He (2014)
  www.caister.com/microarrays2
- **Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications**
  Edited by: D Marco (2014)
  www.caister.com/n2

Order from caister.com/order

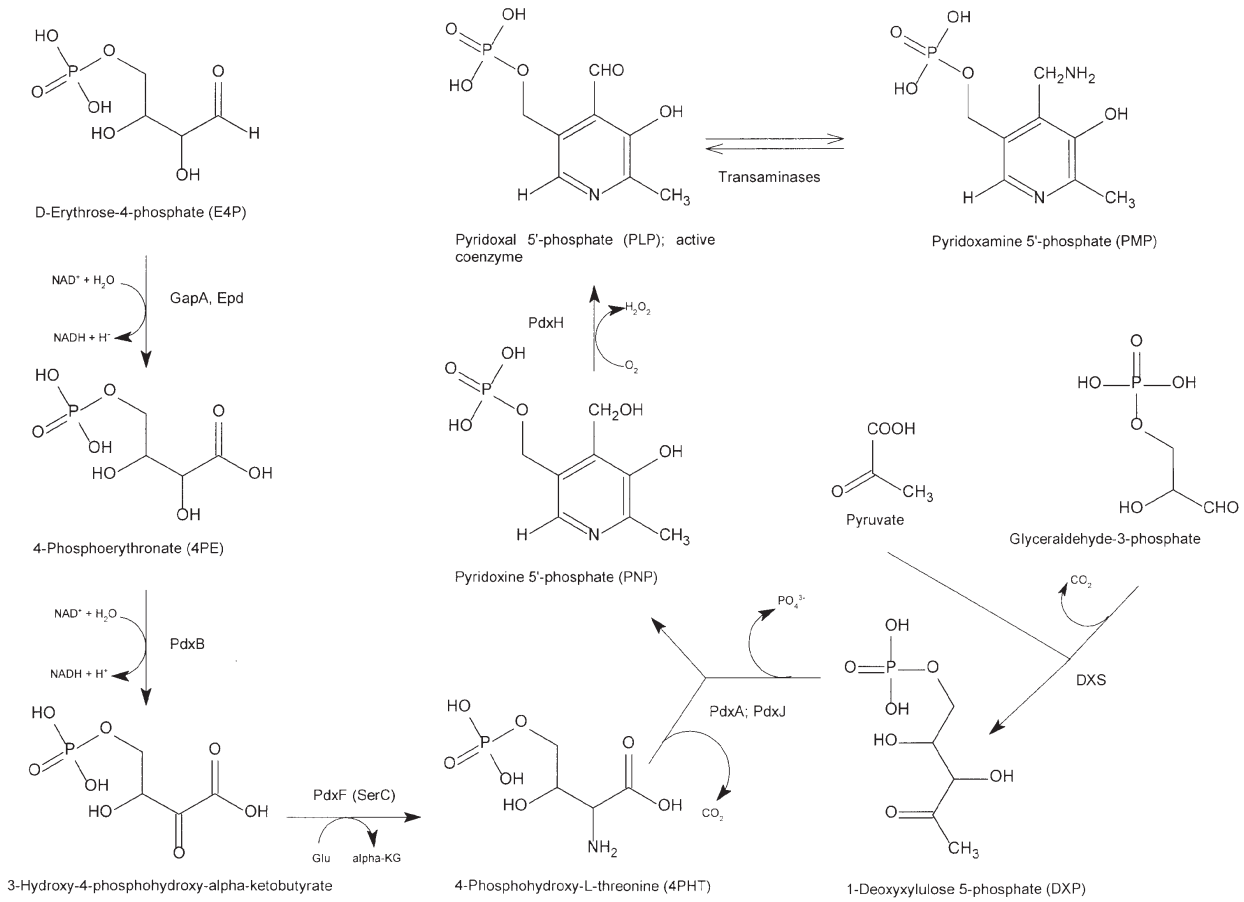


Figure 1. The *de novo* PLP biosynthesis pathway used by *E. coli*. The key enzymes PdxA and PdxJ catalyze synthesis of the pyridine ring. The precursors 4PHT and DXP can be readily produced from intermediates of the pentose phosphate cycle and of glycolysis, respectively.

Table 1. A comparison of the genes encoding pyridoxal 5'-phosphate biosynthesis enzymes of *E. coli* and orthologous open reading frames found in the *B. subtilis* genome.

E. coli genes coding for enzymes involved in pyridoxine biosynthesis	Enzymatic function of gene product	References	Similarities in *B. subtilis*	Predicted enzymatic function of gene product	Expectation value
gapA	Glyceraldehyde 3-P dehydrogenase A	Yang *et al.*, 1998b	*gap*	Glyceraldehyde-3-phosphate dehydrogenase	4.5e-101
			gapB	Glyceraldehyde-3-phosphate dehydrogenase	5.0e-86
epd (*gapB*)	Erythrose-4-P dehydrogenase	Yang *et al.*, 1998b	*gapB*	Glyceraldehyde-3-phosphate dehydrogenase	1.6e-80
			gap	Glyceraldehyde-3-phosphate dehydrogenase	1.3e-73
pdxB	Erythronate-4-phosphate dehydrogenase?	Schoenlein *et al.*, 1989	*serA*	Phosphoglycerate dehydrogenase	3e-22
			yoaD	Unknown; similar to phosphoglycerate dehydrogenase	1e-12
			yvcT	Unknown; similar to glycerate dehydrogenase	2e-09
pdxF (*serC*)	Phosphoserine aminotransferase	Drewke *et al.*, 1996	*serC*	Phosphoserine aminotransferase	8e-78
dxs	DXP (Deoxy-xylulose-P) synthase	Sprenger *et al.*, 1997; Lois *et al.*, 1998	*yqiE*	Unknown; similar to unknown proteins	3.1e-141
			pdhB	pyruvate dehydrogenase (E1 beta subunit).	1.5e-11
pdxA	Pyridoxine biosynthesis	Roa *et al.*, 1989; Cane *et al.*, 1998	*spoIIID*	Regulation of genes controlled by mother cell-specific sigma factors σ^E and σ^K	0.37
pdxJ	Pyridoxine biosynthesis	Lam *et al.*, 1992 Takiff *et al.*, 1992	*spoIIB*	Spore development	0.29
			ytxH	General stress protein	0.71
pdxH	Pyridoxine-phosphate oxidase	Notheis *et al.*, 1995	*ydaG*	General stress protein	0.0091
pdxK	Vitamin B_6 kinase	Yang *et al.*, 1996	*thiD*	Phosphomethylpyrimidine kinase	6.6e-13
			yjbV	unknown; similar to phosphomethylpyrimidine kinase	1.1e-08
pdxY	Alternative vitamin B_6 kinase	Yang *et al.*, 1998a	*yjbV*	Unknown; similar to phosphomethylpyrimidine kinase	0.0008
			thiD	Phosphomethylpyrimidine kinase	0.0014

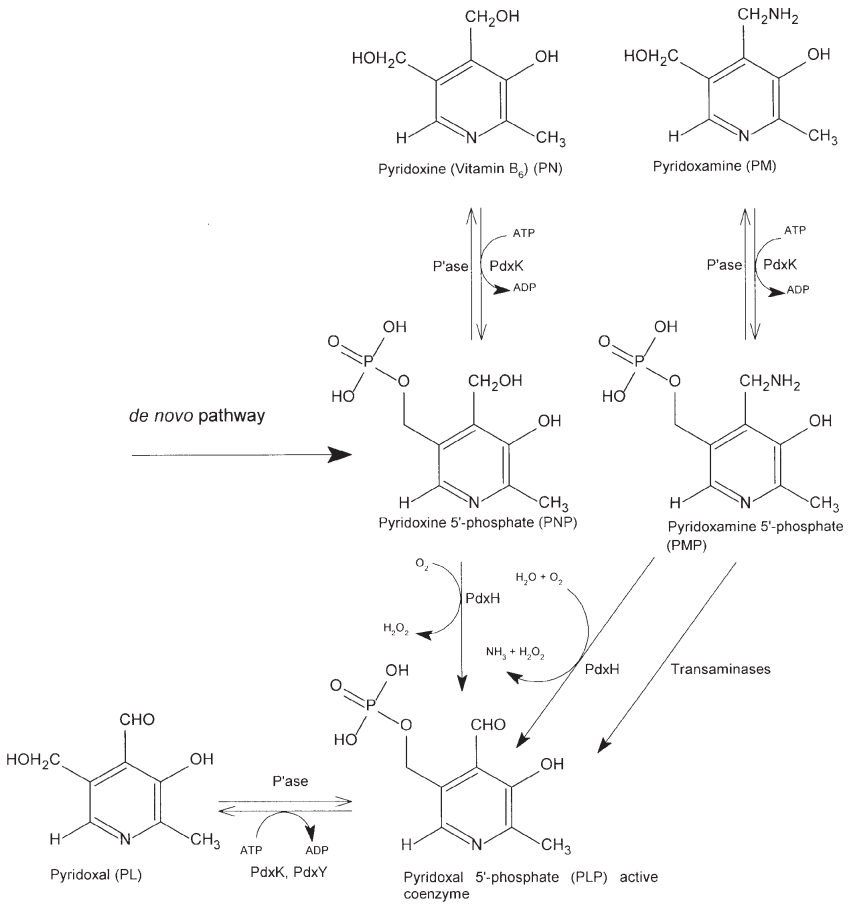


Figure 2. The salvage pathway PLP biosynthesis pathway found in *E. coli*. Similar reactions are performed also by higher organisms, including mammals. Note that *E. coli* possesses two pyridoxal kinases (PdxK, PdxY). PdxK exhibits broader substrate specifity. PLP homoeostasis is achieved by relatively unspecific phosphatases (P'ases).

phosphate as substrates. The other intermediate 4PHT (Zhao and Winkler, 1996) is formed in a series of reactions involving two oxidation steps and one transamination step in a pathway similar to serine biosynthesis (Lam and Winkler, 1990) starting from erythrose-4-phosphate (E4P), a central metabolite of the pentose phosphate pathway. E4P is also a precursor of aromatic amino acids (L-tryptophan, L-phenylalanine and L-tyrosine) and aromatic vitamins (p-aminobenzoate, p-hydroxybenzoate, 2,3-dihydroxybezoate) and it is produced directly by the action of transketolases TktA and TktB (Zhao and Winkler, 1994). These enzymes use D-glycerinaldehyde-3-phosphate and D-fructose-6-phosphate as substrates to produce D-xylulose-5-phosphate and E4P. In the first oxidation step, E4P is converted to 4-phosphoerythronate (4PE) by the action of dehydrogenases GapA or Epd (GapB) (Yang *et al*., 1998b). 4PE is further oxidized by the PdxB dehydrogenase (Lam and Winkler, 1990) to 3-hydroxy-4-phosphohydroxy-α-ketobutyrate. By a transamination reaction using glutamate as donor, this compound is finally transformed into 4PHT by the action of the PdxF (SerC) transaminase, a pyridoxal 5'-phosphate containing enzyme (Drewke *et al*., 1996).

Additionally, PLP can be synthesized by a salvage pathway (Figure 2) that uses B_6-vitamers pyridoxal (PL), pyridoxine (PN) and pyridoxamine (PM) present in the growth medium (Hill and Spenser, 1996; Yang *et al*., 1996, Yang *et al*. 1998a). In this pathway, the substrates PL, PN and PM are phosphorylated by kinases to form PLP, PNP and pyridoxamine 5'-phosphate (PMP). Two different kinases exhibiting a different substrate specifity have been identified in *E. coli*: The PN/PL/PM kinase PdxK (Yang *et al*., 1996) and the PL kinase PdxY (Yang *et al*., 1998a) PNP and PMP are oxidized to PLP by the PdxH oxidase (Zhao and Winkler, 1995) which functions in both pathways. Alternatively, PMP can be converted to PLP by the action of transaminases. Similar salvage pathways involving oxidases and kinases exist in mammalian cells (Choi *et al*., 1987; Hanna *et al*. 1997; McCormick and Chen, 1999) which do - of course - not possess the *de novo* pathway. PLP Homoeostasis is maintaned by relatively unspecific PLP phosphatases.

In contrast to the detailed knowledge concerning PLP biosynthesis in *E. coli*, only one early report (Pflug and Lingens, 1978) deals with vitamin B_6 biosynthesis in the grampositive model organism *Bacillus subtilis*. In order to identify genes whose products might participate in PLP biosynthesis in this organism, BLAST searches were perfomed at the "SubtiList" database (Moszer *et al*., 1995; Moszer, 1998) using the sequences of *E. coli* PLP biosynthesis enzymes as queries. As shown in Table 1, the PdxA/J sequences are not present in the *B. subtilis* genome. It was concluded that *B. subtilis* might use another pathway.

Table 2. Alphabetical list of species analyzed for enzymes of vitamin B_6 metabolism and information about the phylogenetic status. Abbrevations of names are used in the phylogenetic trees and the Tables.

Organism	Phylogenetic position	Abbreviation
Actinobacillus actinomycetemcomitans	Proteobacteria; γ subdivision; Pasteurellaceae	Aac
Aeropyrum pernix	Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae	Ape
Alcaligenes eutrophus (*Ralstonia eutropha*)	Proteobacteria; β subdivision; Ralstonia group	Aeu
Arabidopsis thaliana	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales; Brassicaceae	Ath
Archaeoglobus fulgidus	Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae	Afu
Aquifex aeolicus	Aquificales; Aquificaceae	Aae
Aquifex pyrophilus	Aquificales; Aquificaceae	Apy
Artemisia annua	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae; Anthemideae	*Aan*
Bacillus anthracis	Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus cereus group	Ban
Bacillus halodurans	Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group	Bha
Bacillus stearothermophilus	Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group	Bst
Bacillus subtilis	Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group	Bsu
Bordetella bronchiseptica	Proteobacteria; β subdivision; Alcaligenaceae	Bbr
Bordetella pertussis	Proteobacteria; β subdivision; Alcaligenaceae	Bpe
Bradyrhizobium japonicum	Proteobacteria; α subdivision; Bradyrhizobium group	Bja
Burkholderia cepacia	Proteobacteria; β subdivision; Burkholderia group	Bce
Caenorhabditis elegans	Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae	Cel
Campylobacter jejuni	Proteobacteria; ε subdivision; Campylobacter group	Cje
Candida albicans	Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales; Anamorphic Saccharomycetales	Cal
Capsicum annuum (bell pepper)	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales; Solanaceae; Chloroplast	Can
Catharanthus roseus (rosy periwinkle) (Madagascar periwinkle)	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Gentianales; Apocynaceae	Cro
Caulobacter crescentus	Proteobacteria; α subdivision; Caulobacter group	Ccr
Cercospora nicotianae	Eukaryota; Fungi; Ascomycota; Pleosporales; Leptosphaeriaceae; anamorphic Leptosphaeriaceae	Cni
Chlamydia muridarum	Bacteria; Chlamydiales; Chlamydiaceae	
Chlamydia pneumoniae (*Chlamydophila pneumoniae*)	Bacteria; Chlamydiales; Chlamydiaceae	Cpn
Chlamydia trachomatis	Bacteria; Chlamydiales; Chlamydiaceae	Ctr
Chlorobium tepidum	Green sulfur bacteria	Cte
Clostridium acetobutylicum	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae	Cac
Clostridium difficile	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae	Cdif
Corynebacterium diphteriae	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae	Cdi
Dehalococcoides ethenogenes	Green non-sulfur bacteria; Dehalococcoides group	Det
Deinococcus radiodurans	Thermus/Deinococcus group; Deinococcales	Dra
Desulfovibrio vulgaris	Proteobacteria; δ subdivision	Dvu
Drosophila melanogaster	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae	Dme
Emericella nidulans (*Aspergillus nidulans*)	Eukaryota; Fungi; Ascomycota; Eurotiales; Trichocomaceae; anamorphic Trichocomaceae	Eni
Erwinia herbicola	Proteobacteria; γ subdivision; Enterobacteriaceae	Ehe
Escherichia coli	Proteobacteria; γ subdivision; Enterobacteriaceae	Eco
Euphorbia pulcherrima	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Malpighiales; Euphorbiaceae	Epu
Francisella tularensis	Proteobacteria; γ subdivision; Thiomicrospira group; Francisella group	Ftu
Geobacter sulfurreducens	Proteobacteria; δ subdivision; Geobacteriaceae	Gsu
Haemophilus ducreyi	Proteobacteria; γ subdivision; Pasteurellaceae	Hdu
Haemophilus influenzae	Proteobacteria; γ subdivision; Pasteurellaceae	Hin
Helicobacter pylori	Proteobacteria; ε subdivision; Helicobacter group	Hpy
Hevea brasiliensis (rubber tree)	Eukaryota; Viridiplantae; Streptophyta Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Malpighiales; Euphorbiaceae	Hbr
Homo sapiens (man)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae	Hsa
Klebsiella pneumoniae	Proteobacteria; γ subdivision; Enterobacteriaceae	Kpn
Legionella pneumophila	Proteobacteria; γ subdivision; Legionellaceae	Lpn
Lycopersicon esculentum (tomato)	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales; Solanaceae; Solanum	Les
Mentha piperita (peppermint)	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Lamiales; Lamiaceae;	Mpi
Methanobacterium thermoautotrophicum	Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae	Mth
Methanococcus jannaschii	Archaea; Euryarchaeota; Methanococcales; Methanococcaceae	Mja
Methanococcus vannielii	Archaea; Euryarchaeota; Methanococcales; Methanococcaceae	Mva
Mycobacterium avium	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium avium complex (MAC)	Mav
Mycobacterium bovis	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae Mycobacterium tuberculosis complex	Mbo
Mycobacterium leprae	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae	Mle
Mycobacterium tuberculosis	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium tuberculosis complex	Mtu
Mycoplasma capricolum	Firmicutes; Bacillus/Clostridium group; Mollicutes	Mca
Mycoplasma genitalium	Firmicutes; Bacillus/Clostridium group; Mollicutes	Mge
Myxococcus xanthus	Proteobacteria; δ subdivision; Myxobacteria; Myxcococcales; Cystobacterinae	Mxa
Neisseria gonorrhoeae	Proteobacteria; β subdivision; Neisseriaceae	Ngo
Neisseria meningitidis	Proteobacteria; β subdivision; Neisseriaceae	Nme
Pasteurella multocida	Proteobacteria; γ subdivision; Pasteurellaceae	Pmu
Pichia angusta (*Hansenula polymorpha*)	Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales; Saccharomycetaceae	Pan
Plasmodium falciparum	Eukaryota; Alveolata; Apicomplexa; Haemosporida;	Pfa
Porphyromonas gingivalis	CFB group; Bacteriodaceae	Pgi
Pseudomonas aeruginosa	Proteobacteria; γ subdivision; Pseudomonas group	Pae
Pseudomonas putida	Proteobacteria; γ subdivision; Pseudomonas group	Ppu
Pyrococcus abyssi	Archaea; Euryarchaeota; Thermococcales; Thermococcaceae	Pab
Pyrococcus furiosus	Archaea; Euryarchaeota; Thermococcales; Thermococcaceae	Pfu
Pyrococcus horikoshii	Archaea; Euryarchaeota; Thermococcales; Thermococcaceae	Pho
Rattus norvegicus (rat)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae	Rno
Rhizobium sp. NGR234	Proteobacteria; α subdivision; Rhizobiaceae	Rsp

Rhodobacter capsulatus (*Rhodopseudomonas capsulata*)	Bacteria; Proteobacteria; α subdivision; Rhodobacter group	Rca
Rhodobacter sphaeroides (*Rhodopseudomonas sphaeroides*)	Bacteria; Proteobacteria; α subdivision; Rhodobacter group	Rsph
Rhodovulum sulfidophilum (*Rhodobacter sulfidophilus*)	Bacteria; Proteobacteria; α subdivision; Rhodobacter group	Rsu
Roseobacter denitrificans (*Erythrobacter* sp. OCh 114)	Bacteria; Proteobacteria; α subdivision; Rhodobacter group	Rde
Oryza sativa (rice)	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae	Osa
Ovis aries (sheep)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Caprinae	Oar
Saccharomyces cerevisiae	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales; Saccharomycetaceae	Sce
Salmonella paratyphi	Proteobacteria; γ subdivision; Enterobacteriaceae	Spa
Salmonella typhi	Proteobacteria; γ subdivision; Enterobacteriaceae	Sty
Salmonella typhimurium	Proteobacteria; γ subdivision; Enterobacteriaceae	Stym
Schizophyllum commune	Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Stereales; Schizophyllaceae	Sco
Schizosaccharomyces pombe	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales; Schizomycetaceae	Spo
Shewanella putrefaciens	Proteobacteria; γ subdivision; Alteromonadaceae	Spu
Sinorhizobium meliloti	Proteobacteria; α subdivision; Rhizobiaceae	Sme
Sphingomonas aromaticivorans	Proteobacteria; α subdivision; Zymomonas group	Sar
Solanum tuberosum (potato)	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Potato	Stu
Staphylococcus aureus	Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group	Sau
Stellaria longipes	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Caryophyllidae; Caryophyllales; Caryophyllaceae	Slo
Stellaria media	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Caryophyllidae; Caryophyllales; Caryophyllaceae	Smed
Streptococcus pneumoniae	Firmicutes; Bacillus/Clostridium group; Streptococcaceae	Spn
Streptomyces coelicolor	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae	Sco
Streptomyces sp CL190	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae	SspCL190
Suberites domuncula	Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha; Hadromerida Suberitidae	Sdo
Suberites ficus	Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha; Hadromerida; Suberitidae	Sfi
Sulfolobus solfataricus	Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae	Sso
Sus scrofa (pig)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae;	Ssc
Synechocystis sp.	Cyanobacteria; Chroococcales	Ssp
Synechococcus leopoliensis	Cyanobacteria; Chroococcales	Sle
Thermotoga maritima	Bacteria; Thermotogales	Tma
Thiobacillus ferrooxidans	Proteobacteria; γ subdivision	Tfe
Treponema denticola	Bacteria; Spirochaetales; Spirochaetaceae	Tde
Treponema pallidum	Bacteria; Spirochaetales; Spirochaetaceae	Tpa
Trypanosoma brucei	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae	Tbr
Vibrio cholerae	Proteobacteria; γ subdivision; Vibrionaceae	Vch
Yersinia pestis	Proteobacteria; γ subdivision; Enterobacteriaceae	Ype
Zymomonas mobilis	Proteobacteria; α subdivision; Sphingomonas group	Zmo

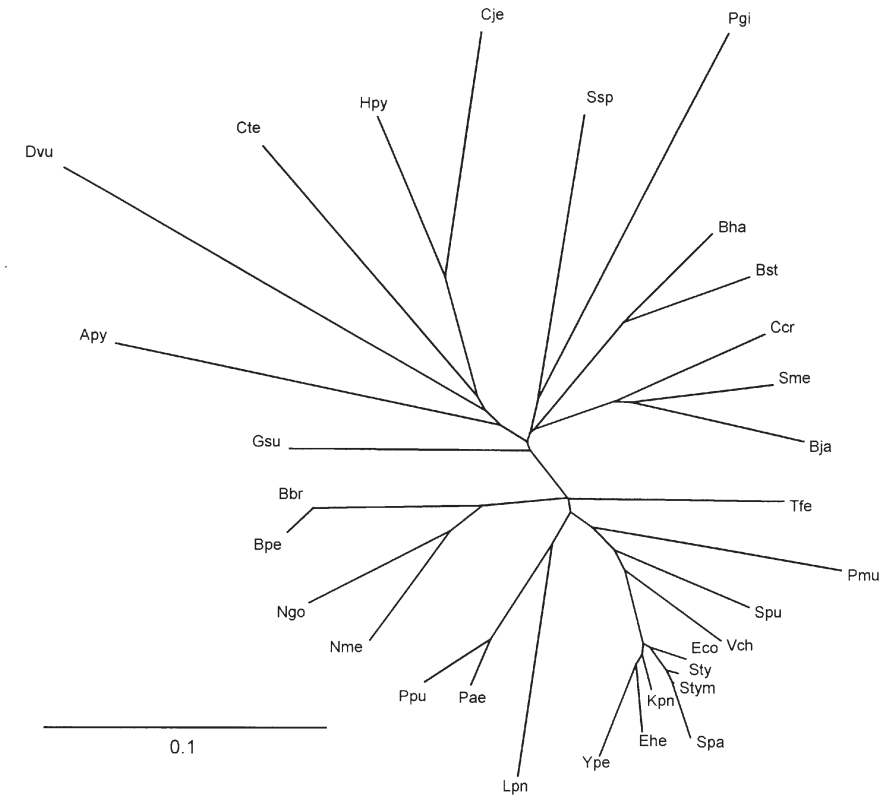


Figure 3. Unrooted phylogenetic tree based on 16S rRNA sequences. Sequences from all organisms in which a *pdxA* (and/or a *pdxJ* gene) was detected (see Tables 5, 6; Figures 6, 7) , were included in the dataset used for construction of the tree.

Table 3. Information on small subunit rRNA sequences was not available for some organisms in which gene sequences coding for enzymes involved in vitamin B_6 metabolism were identified. A closely related organism was chosen instead for construction of the rRNA derived tree.

Organism without information about small subunit rRNA	Related organism included in the rRNA derived phylogenetic tree
Hevea brasiliensis	*Euphorbia pulcherrima*
Stellaria longipes	*Stellaria media*
Suberites domuncula	*Suberites ficus*

Two recent reports (Ehrenshaft *et al*., 1999; Osmani *et al*., 1999) which investigate pyridoxine biosynthesis in the fungus *Emericella* (*Aspergillus*) *nidulans* and singlet oxygen resistance in the fungus *Cercospora nicotianae* show that these organisms do not possess PdxA/J-like enzymes but instead an alternative pathway is involved in PLP biosynthesis. This conclusion was based on the observation that mutations in an extremely conserved gene result in pyridoxine autotrophy in both species. These mutations also result in methylene blue sensitivity. Methylene blue, a photosensitizer, is involved in the production of deleterious singlet oxgen molecules in the presence of light.

No biochemical information about this pathway is yet available but this pathway appears to be conserved in bacteria, archaea and eukarya, the three domains of life according to the classification of Woese *et al*. (1990). In this paper, an attempt is made to characterize the phylogenetic relationships of the different PLP biosynthesis pathways. Phylogenetic trees for the DXS, PdxA, PdxJ, PdxH and PdxK proteins as well as for proteins representative of the newly described pathway are presented. In selected cases, these trees are compared to rRNA derived phylogenetic trees. Furthermore, possible physiological and evolutionary implications of the different pathways are discussed.

Two 16S/18S Ribosomal RNA Derived Phylogenetic Trees

In order to provide a basis for the phylogenetic analyses of *de novo* PLP biosynthesis pathways, two phylogenetic trees based on the sequences of the small subunit RNA (16S or 18S rRNA) were constructed. The same parameters were used in the CLUSTALW alignments in every case. Every organism in whose genome an enzyme related to *de novo* PLP biosynthesis was detected was included in one of the trees. These organisms are listed in alphabetical order in Table 2. Furthermore, information of the phylogenetic status (based on the analysis of rRNA sequences) and a three letter abbreviation for the identification of the organism in the tables and trees are given. However, for a few of the organisms harbouring genes whose products are related to PLP biosynthesis, no information about the rRNA sequences was available in the databases. In all but one cases, rRNA sequences from closely related organisms were available and included in the trees (see Table 3).

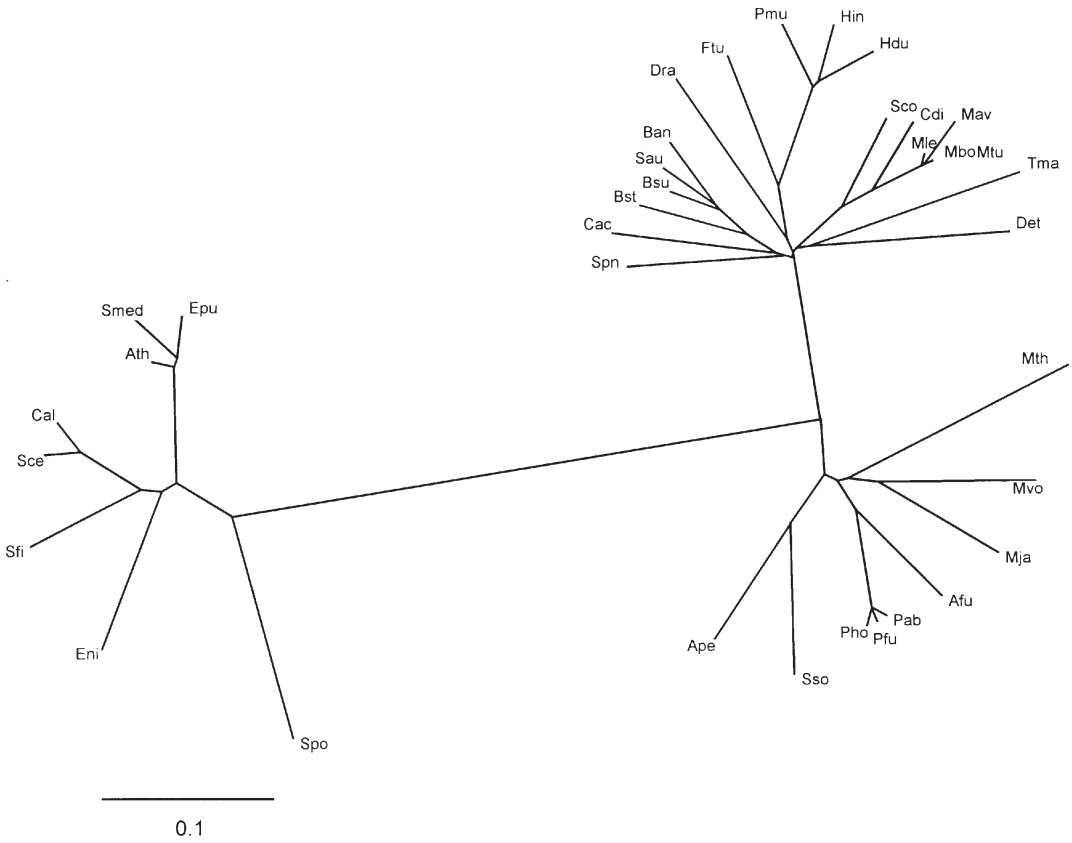


Figure 4. Unrooted phylogenetic tree based on 18S/16S rRNA sequences. Sequences from all organisms in which a *yaaD*-like gene (see Table 9; Figure 10) was detected, were included in the dataset used for construction of the tree.

The first tree (Figure 3) comprises organisms (belonging mostly to the γ subdivision of proteobacteria) in which the PdxA/PdxJ heterodimer is responsible for vitamin B_6 biosynthesis whereas the second contains organisms in which the alternative pathway is operating. In the second tree (Figure 4), a clear classification of the organisms into the three domains of life bacteria, archaea and eukarya can be observed (Woese *et al*., 1990).

The DXS Derived Tree

DXS orthologs (COG number COG0743; see Tatusov *et al*., 2000) identified by database searches are listed in Table 4A. Figure 5 shows an unrooted phylogenetic tree of DXS orthologs and other enzymes. Sprenger *et al*. (1997) also included transketolases and dehydrogenase subunits as well as other related proteins in their phylogenetic tree. For normalization, these and other newly identified protein sequences (see Table 4B) were also included in the alignment used for construction of this tree. Clustering of these proteins which are more or less related to DXS in four distinct groups and the clear separation of these groups from the remaining proteins included in the tree supports the idea that these proteins indeed represent DXP synthases.

Organisms harbouring genes encoding DXP synthases also include organisms which do not possess vitamin B_6 biosynthesis genes (see below) (*i.e. Chlamydia* and *Treponema* species, *Plasmodium falciparum*; Note that these organisms were not included in the rRNA derived trees). This might reflect the fact that DXP is also an intermediate in mevalonate independent biosynthesis of isopentenyl diphosphate which in turn is universal precursor of isoprenoid synthesis (for a phylogenetic analysis of both pathways see Boucher and Doolittle, 2000). This pathway is present in bacteria and plants, but absent in animals. Therefore, this pathway might constitute targets for novel antibacterials (Kuzuyama *et al*., 2000a,b). In plants, mutations in genes encoding DXP synthases affect a wide variety of physiological functions due to the inability of isoprenoid biosynthesis (Bouvier *et al*., 1998; Lange *et al*., 1998). The plant enzymes are most closely related to the DXS found in the genomes of the α subdivision of proteobacteria (*C. crescentus*, *S. meliloti*, photosynthetic bacteria). In photosynthetic bacteria, genes encoding DXS are located in the vicinity of genes encoding components of the photosynthetic apparatus (Youvan *et al*., 1984; Kortlüke *et al*., 1997; Masuda *et al*., 1999). Interestingly, DXS from cyanobacteria (*Synechococcus*) (Miller *et al*., 1999) are closely related to those originating from grampositive bacteria.

The PdxA Derived Tree

PdxA orthologs (COG1995) identified by database searches are listed in Table 5. Figure 6 shows an unrooted phylogenetic tree of PdxA orthologs. Most of the PdxA orthologs were identified in genomes of proteobacteria (Stackebrandt *et al*., 1988). Many of these can be attributed to the γ subdivision of proteobacteria which form a tight cluster around the PdxA protein of *E. coli*. Remerkable exceptions are *Erwinia herbicola*, whose PdxA ortholog is related to two *Bordetella* species (β subdivision) and *Pasteurella multocida* whose PdxA ortholog is related to *Geobacter sulfurreducens* (δ subdivision). The PdxA orthologs of *Helicobacter pylori* and *Campylobacter jejuni* (ε subdivision) as well as those of *Sinorhizobium meliloti* and *Caulobacter crescentus* (α subdivision) are also clustered. Two of the PdxA orthologs are from grampositive bacteria, namely *Bacillus halodurans* and *Bacillus stearothermophilus* which appear to be closely related. Two of the PdxA paralogs are located on catabolic plasmids: pNL1 of *Sphingomonas aromaticivorans* (Romine *et al*., 1999) (α subdivision) and plasmid pMOP of *Burkholderia* (*Pseudomonas*) *cepacia* (Saint and Romas, 1996) (β subdivision). Both plasmids encode degradation pathways for aromatic substances and the plasmid encoded PdxA orthologs are closely related. Because the plasmid encoded PdxA sequences might not represent the PdxA sequences of the hosts, these paralogs were called "ORF" in Table 5 and in Figure 6.

The PdxJ Derived Tree

PdxJ (COG0854) orthologs identified by database searches are listed in Table 6. Figure 7 shows an unrooted phylogenetic tree of PdxJ orthologs. As demonstrated for the PdxA orthologs, most of the PdxJ orthologs were identified in the γ subdivision of proteobacteria. However, the cluster formed by this group is not as tight as seen in the PdxA derived tree but nevertheless exhibiting an overall similar pattern of relatedness. No PdxJ paralogs were observed in the unfinished genomes of the grampositive species *B. halodurans* and *B. stearothermophilus* and in the complete sequence of the catabolic plasmid pNL1 from *S. aromaticivorans*. Also, no PdxJ paralogs were detected in the unfinished genomes of *P. multocida*, *S. meliloti*, and *E. herbicola*. However, PdxJ orthologs were observed in the unfinished genomes of *Aquifex pyrophilus* (the PdxJ ortholog is closely related to that one of *A. aeolicus*), *L. pneumophila*, *B. japonicum* (this PdxJ ortholog is related to the ortholog of *C. crescentus* which both belong to the α subdivision). The PdxJ orthologs of *C. jejuni* and *H. pylori* (ε subdivision) are closely related whereas all orthologs of all other species (*P. gingivalis*, *D. vulgaris*, *S. spec.*, *G. sulfurreducens*) appear at different positions compared to the PdxA derived tree.

The PdxH Derived Tree

PdxH orthologs (COG0259) identified by database searches are listed in Table 7. Figure 8 shows an unrooted phylogenetic tree of PdxH orthologs. Since PdxH is an enzyme participating in the salvage pathway of PLP biosynthesis, this tree also includes PdxH orthologs found in animals. PdxH orthologs of the γ subdivision of proteobacteria again form a tight cluster. Exceptions are the PdxH orthologs of *B. bronchiseptica* and *B. pertussis*, of *P. aeruginosa*, of *N. meningitidis* and *N. gonorhoeae*, of *H. influenzae* and *H. ducreyi*, of *T. ferroxidans* and the ortholog of *L. pneumophila* which curiously appears to form a cluster with the PdxH orthologs of animals (*C. elegans*, *R. norvegicus*, *H. sapiens*). The PdxH ortholog from *D. melanogaster* is related to those found in mycobacteria.

Table 4A. Dxs orthologs. The sequence of the Dxs protein of *E. coli* was used as query sequence. Members of this family are listed in alphabetical order. In several cases, the combination "DXS species name" was replaced by a number in the DXS derived tree (Figure 5). This number is listed in the rightmost column.

Abbreviation	Number of residues	Accession or contig number	Number in Figure 5
Dxs Aae	628	O67036	
DXS1 Aan	713 (fragment)	AF182286	
Dxs Aac	616	gnl|OUACGT_714|A.actin_Contig381	1
CLA1 Ath	739	O49738	
Dxs Ban	633 (fragment)	gnl|TIGR_1392|banth_1573	
Dxs Bbr	402 (fragment)	gnl|Sanger_518|bbronchi_Contig1028	
Dxs Bpe	600 (fragment)	gnl|Sanger_520|B.pertussis_Contig273	
Dxs Bst	626 (fragment)	gnl|UOKNOR_1422|bstear_Contig1135	
Dxs Bsu	633	P54523	
TKT2 Can	719	O78328	2
Dxs Cac	619	gnl|GTC|C.aceto_gnl	
Dxs Ccr	611 (fragment)	gnl|TIGR|C.crescentus_4671	
Dxs Cdi	635	gnl|Sanger_1717|cdiph_Contig146	
Dxs Cdif	621	gnl|Sanger_1496|cdifficile_Contig995.1	
Dxs Cje	615	CAB72788	
TC0608 Cmu	632	AAF39439	
TktB2 Cpn	644	Q9Z6J9	
DXS Cre	735	O81954	3
DXS Cro	716	O82676	
Dxs Cte	644	gnl|TIGR|C.tepidum_3419	
Dxs Ctr	640	O84335	4
Dxs Det	633	gnl|TIGR_61435|deth_1549	
Dxs Dra	629	Q9RUB5	
Dxs Eco	620	AF035440	5
Dxs Hin	625	P45205	6
Dxs Hpy	618	Q9ZM94	
Dxs Hdu	617	gnl|HTSC_730|ducreyi	7
Dxs Kpn	201	gnl|WUGSC_573|kpneumo_B_KPN.Contig683	
DXS Les	719	Q9XH50	8
Dxs Mav	641	gnl|TIGR|M.avium_5	
Dxs Mbo	638	gnl|Sanger_1765|mbovis_Contig727	9
TktB Mle	736	U15181	
DXS Mpi	724	O64904	
Dxs Mtu	638	O07184	10
Rv3379c Mtu	536	O50408	
Dxs Ngo	637	gnl|OUACGT_485|Ngon_Contig10	11
Dxs Nme	637	CAB83880	12
CLA1 Osa	594	O22567	13
ESTS AU078063 Osa	628	Q9SNQ1	
Dxs Pae	627	gnl|PAGP_287|Paeruginosa_Contig1	14
Dxs Pgi	633	gnl|TIGR|P.gingivalis_GPG.con	
DXS Pfa	1205	O96694	
Dxs Pmu	614	gnl|CBCUMN_747|Pmultocida	15
Dxs Ppu	631	gnl|TIGR|pputida_all_432	16
Dxs Rca	641	P26242	
Orf1 Rde	354 (fragment)	O69774	
Orf641 Rsph	105 (fragment)	Q9RFB7	
Orf641 Rsu	107 (fragment)	Q9WXE4	
SC6A5.17 Sco	656	Q9X7W3	
SC7B7.10 Sco	353 (fragment)	O50507	
SC1C3.01 Sco	341 (fragment)	O69843	
Dxs Ssp	640	P73067	
Dxs SspCL190	631	Q9RBN6	
Dxs Sle	636	Q9R6S7	
Dxs Sme	210 (fragment)	gnl|Stanford_382|smelil_423035E12.x1	
Dxs Spu	622	gnl|TIGR_24|sputre_6412	
Dxs Sty	620	gnl|Sanger_601|S.typhi_Contig23	17
Dxs Tde	304 (fragment)	gnl|TIGR_158|tdent_gtd242	
Dxs Tfe	624	gnl|TIGR|t_ferrooxidans_3343	
TM1770 Tma	608	Q9X291	
TktB Tpa	630	O83796	
Dxs Vch	626	gnl|TIGR|V.cholerae_666_1760_Bert	
Dxs Ype	619	gnl|Sanger_632|Y.pesits_Contig763	

The PdxH orthologs of yeasts are closely related; the PdxH ortholog of S. coelicolor is also located in this cluster. The expression of the PdxH ortholog of *M. xanthus* is developmentally regulated (the gene was previously named *fprA* by Shimkets, 1990); this protein forms a cluster with those of other species of the α subdivision.

The PdxK Derived Tree

PdxK orthologs (COG2240) identified by database searches are listed in Table 8. Figure 9 shows an unrooted phylogenetic tree of PdxK orthologs. As the PdxH tree, this tree also contains orthologs found in animals and

Table 4B. For comparison, several proteins distantly related to DXS were also included in the DXS derived tree (Figure 5). This table lists and names these enzymes.

Description/ Abbreviation	Number of residues	Accession number
2-oxoisovalerate dehydrogenase β subunit		
OdbB Bsu	327	P37941
OdbB Ppu	339	P09061
Pyruvate dehydrogenase E1 β subunit		
OdpB Mca	339	MCU62057
ODPB Sce	366	P32473
PdhB Tfe	355	TFU81808
Pyruvate dehydrogenase E1 α subunit		
ODPA Ath	389	P52901
OdpA Bsu	370	P21881
ODPA Hsa	1410	M24848
OdpA Mca	382	MCU62057
PdhA Tfe	337	TFU81808
Transketolase cluster 1		
TktC Aeu	627	P21725
TktA Bsu	667	P45694
TktA Hin	665	P43757
TktA Mge	643	P47312
DAS Pan	710	P06834 (formaldehyde transketolase)
TKL1 Sce	679	P23254
TktC Stu	694 (fragment)	S58083
Transketolase cluster 2		
APE0583 Ape	322	Q9YEJ5
TKT1 Hsa	623	P29401
TKT2 Hsa	557	P51854
TktC Mja	316	Q58092
Tkt2 Pab	317	Q9V1I1
Y4MN Rsp (on plasmid PNGR234a)	345	P55573

yeasts. The PdxK orthologs found in the γ subdivision again form a cluster. The PdxY paralog of *E. coli* is closely related to the orthologs of *Y. pestis* and *K. pneumoniae*. It might be interesting to investigate whether the less related paralogs of both *Pseudomonas* species and those of *H. ducreyi, P. multocida*, *A. actinomycetemcomitans* share the limited substrate specificity of PdxY of *E. coli* (Yang *et al*., 1998a). The PdxK orthologs of yeast form a cluster; some species (*C. albicans*, *S. pombe*) possess paralogous enzymes. PdxK orthologs found in animals form a cluster which also (in contrast to the PdxH derived tree) includes PdxK of *D. melanogaster*. *O. aries* (sheep) possesses two paralogous enzymes. The PdxK protein of *B. burgdorferi* does not seem related to any other enzyme.

A Novel Pathway for *De Novo* PLP Biosynthesis

Two recent papers (Ehrenshaft *et al*., 1999; Osmani *et al*, 1999) show that an extremely conserved gene found in all three domains of life (Ehrenshaft *et al*., 1998) is involved in *de novo* synthesis of vitamin B_6 in two fungal species (named *pyroA* in *Emericella* (*Aspergillus*) *nidulans* and *SOR1* in *Cercospora nicotianae*. Both papers include database searches in complete and unfinished genomes which demonstrate the (1) absence of the *pdxA*/*J* genes in genomes of organisms in which *pyroA/SOR1* orthologs are

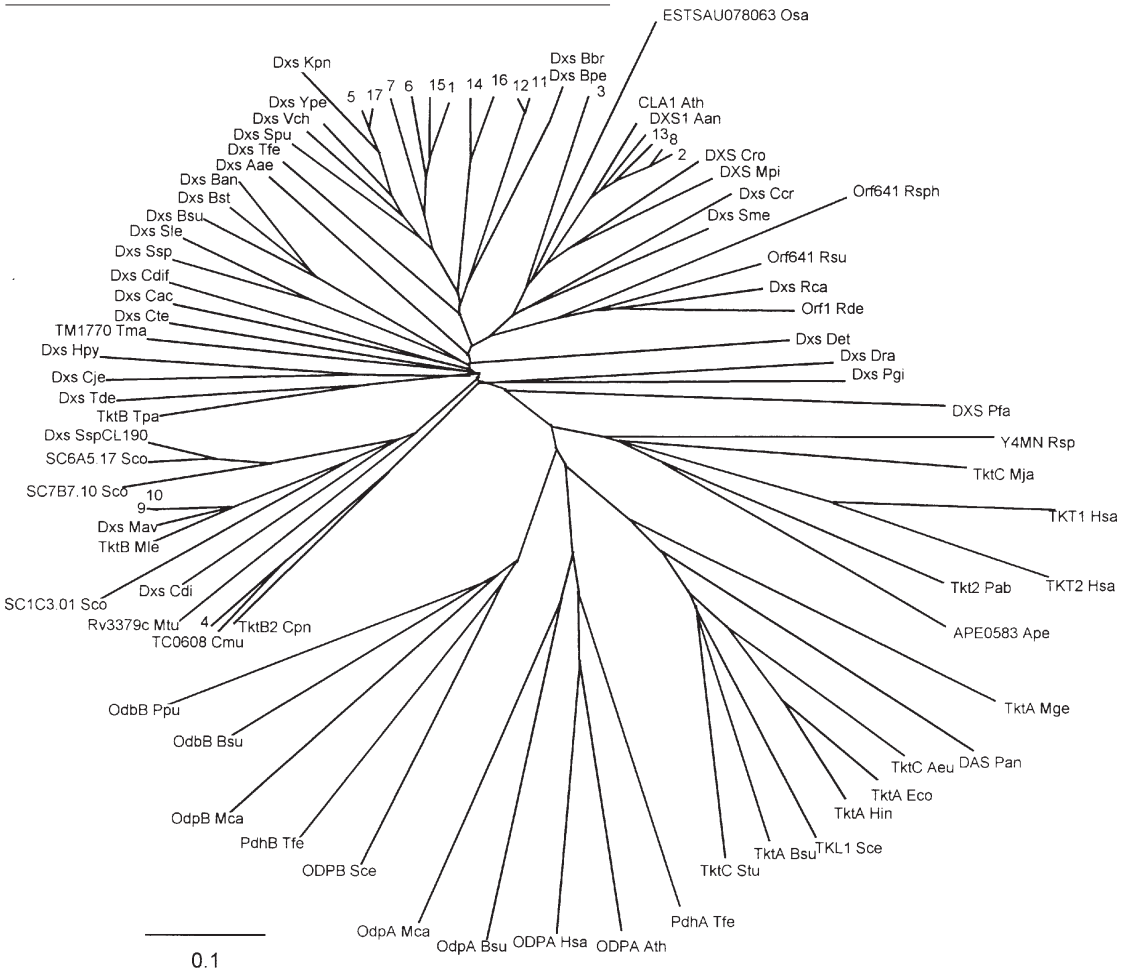


Figure 5. The DXS derived, unrooted phylogenetic tree (see Table 4A). This tree also includes enzymes which are closely related to DXS (see Table 4B).

Table 5. PdxA orthologs. The sequence of the PdxA protein of *E. coli* was used as query sequence. Members of this family are listed in alphabetical order.

Abbreviation	Number of residues	Accession or contig number
PdxA Aae	320	O67019
PdxA Bbr	324	gnl|Sanger_518|bbronchi_Contig1035
Orf Bce (on plasmid pMOP)	258 (fragment)	P94285
PdxA Bha	334	Q9RC88
PdxA Bpe	325	gnl|Sanger_520|B.pertussis_Contig273
PdxA Bst	281 (fragment)	gnl|UOKNOR-1422|bstear_Contig389
PdxA Ccr	378	gnl|TIGR|C.crescentus_4741
PdxA Cje	364	CAB73493
PdxA Cte	334	gnl|TIGR|C.tepidum_3495
PdxA Dvu	218	gnl|TIGR_881|dvulg_gdv165
PdxA Ehe	105 (fragment)	Q47824
PdxA Eco	329	P19624
PdxA Gsu	233 (fragment)	gnl|TIGR_35554|gsulf_GGS_1274
PdxA Hpy	307	Q9ZJ28
PdxA Kpn	330	gnl|WUGSC_573|kpneumo_B_KPN.Contig823
PdxA Lpn	193 (fragment)	gnl|CUCGC_446|lpneumo_WG.008.47-R.0820
PdxA Ngo	330	gnl|OUACGT_485|Ngon_Contig11
PdxA Nme	335	AAF40652
PdxA Pae	325	gnl|PAGP_287|Paeruginosa_Contig1
PdxA Pgi	365	gnl|TIGR|P.gingivalis_GPC.con
PdxA Pmu	337	gnl|CBCUMN_747|Pm70seq.fasta
PdxA Ppu	333	gnl|TIGR|pputida_all_133
Orf Sar (Orf 1158 on plasmid pNL1)	330	O85987
PdxA Spa	327	gnl|WUGSC_32027|spara_B_SPA.0.23420
PdxA Sme	147 (fragment)	gnl|Stanford_382|smelil_423023E01.x1
PdxA Ssp	349	Q55982
PdxA Spu	330	gnl|TIGR_24|sputre_6410
PdxA Sty	329	gnl|Sanger_601|S.typhi_Contig1691
PdxA Stym	151 (fragment)	gnl|WUGSC_99287|stmlt2-E2.Contig220
PdxA Tfe	297 (fragment)	gnl|TIGR|t_ferrooxidans_2902
PdxA Vch	330	gnl|TIGR|V.cholerae_666_1760_Bert
PdxA Ype	334	gnl|Sangerl_632|Y.pesits_Contig766

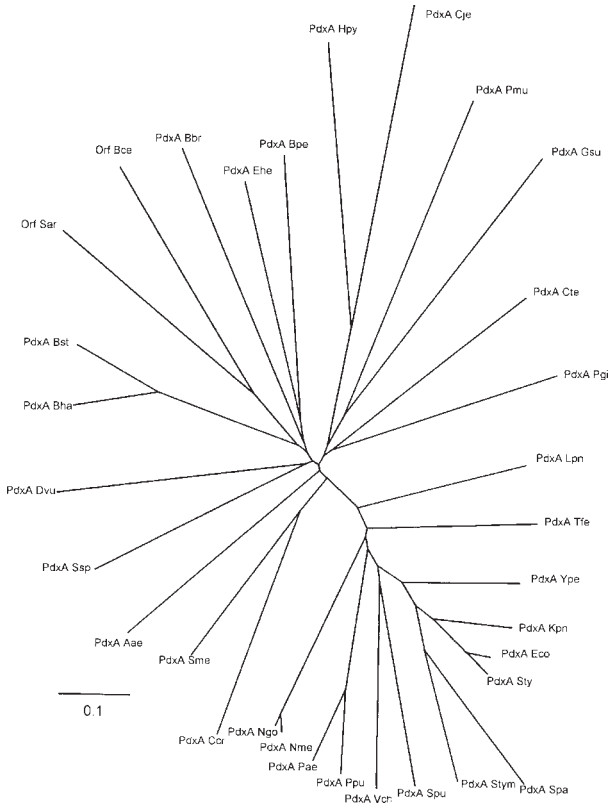


Figure 6. The PdxA derived, unrooted phylogenetic tree (see Table 5).

present (2) the presence of the *pdxA*/*J* genes in genomes of organisms in which *pyroA*/*SOR1* orthologs are absent. In accordance with these studies, the present work also did not detect *pdxA*/*J* orthologs in completely sequenced genomes of Archaea (*M. jannaschii, P. horikoshii*, *M. thermoautotrophicum*, *A. fulgidus*) and in genomes of obligately parasitic bacteria (*Mycoplasma pneumoniae*, *Mycoplasma genitalium*, *Treponema pallidum, Chlamydia trachomatis*, *Rickettsia prowazekii*). Both papers note that many parasites lack both *pdxA*/*J* and *pyroA/SOR1* orthologs. In the databases, members of the latter group are included in the "UPF0019 (SNZ) family". Several different functions (singlet oxygen resistance, ethylene responsive protein, stress inducible protein) have been assigned to this protein (see Table 9). In yeast, it has been shown that paralogs (called *SNO1-3*) of another highly conserved sequence in the databases are located upstream of each of three paralogs (called *SNZ1-3*) (Braun *et al*., 1996; Padilla *et al*., 1998). This family is called "UPF0030 (SNO) family". Expression of adjacent *SNZ* and *SNO* genes is coregulated and the Snz1 and Sno1 proteins interact in a two-hybrid assay (Padilla *et al*., 1998). Interestingly, studies on the origin of the nitrogen atom in pyridoxine synthesized *de novo* by yeast cells have demonstrated that this atom is derived from the amido group of glutamine in this organism, whereas the labelled nitrogen atom was not incorporated into pyridoxine by *E. coli* (Tazuya *et al*. 1995). The authors concluded that different pyridoxine pathways are present in these species.

In an elegant *in silico* analysis of an operon encoding these protein families (called SnzA and SnzB in their paper),

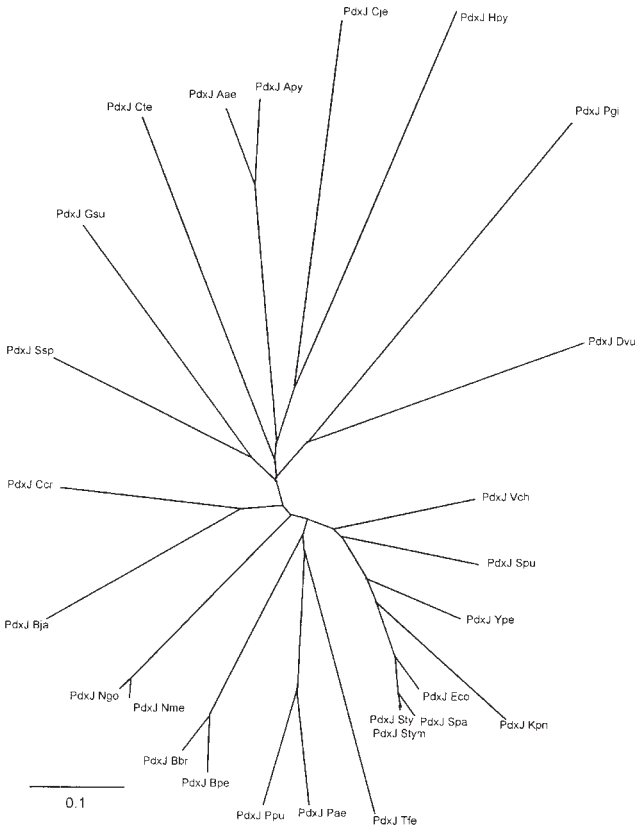


Figure 7. The PdxJ derived, unrooted phylogenetic tree (see Table 6).

Galperin and Koonin (1997) analyze the molecular anatomy of the putative proteins, postulate an interaction between both proteins and propose a novel link between histidine and purine biosynthesis. They also note the absence of the operon in the completely sequenced genomes of *E. coli* and *S. spec*. In *B. subtilis*, the *yaaD* and *yaaE* genes which are organized as a two-gene operon encode these highly conserved proteins.

The YaaD Derived Phylogenetic Tree (the UPF0019 or SNZ Family)

YaaD orthologs (COG0214; the COG database describes these proteins as predicted phosphate utilizing enzymes involved in pyridoxine/purine/histidine biosynthesis) identified by database searches are listed in Table 9. Figure 10 shows an unrooted phylogenetic tree of YaaD orthologs. The YaaD of *B. subtilis* ortholog is located within a cluster consisting of orthologs found in other grampositive bacteria and interestingly, also two Orf's originating from *P. multocida* and *H. ducreyi* (γ subdivision of proteobacteria, family *Pasteurellaceae*). The HI1647 protein of *H. influenzae* is closely related to an Orf found in the genome of *S. pneumoniae* (*Bacillus*/*Clostridium* group and less related to IP1 of *F. tulariensis* (γ subdivision), DR1367 of *D. radiodurans* (*Thermus*/*Deinococcus* group) and to an Orf identified in the genome of *C. acetobutylicum* (*Bacillus*/*Clostridium* group). Orthologs found in the Euryarchaeota *M. thermoautotrophicum, A. fulgidus*, *M. jannaschii*, *P. furiosus* and two other *Pyrococcus* species also form a cluster, whereas the ortholog of another species grouped within the Euryarchaeota, *M. vanielli*, is associated with those of two Crenarchaeota species, *A. pernix* and *S. solfataricus*. Mycobacterial orthologs form a cluster with those found in *S. coelicolor* and *C. diphteriae*. This is in good agreement with the phylogenetic position of these bacteria (Actinobacteridae). The two yeast paralogs SNZ1 and SNZ2 form a cluster with an Orf found in the genome of *C. albicans*, whereas the third, YEM4, is more closely

Table 6. PdxJ orthologs. The sequence of the PdxJ protein of *E. coli* was used as query sequence. Members of this family are listed in alphabetical order.

Abbreviation	Number of residues	Accession or contig number
PdxJ Aae	242	O67417
PdxJ Apy	236 (fragment)	P46212
PdxJ Bbr	246	gnl|Sanger_518|bbronchi_Contig2500
PdxJ Bja	249	AAD02936
PdxJ Bpe	242	gnl|Sanger_520|B.pertussis_Contig307
PdxJ Ccr	254	gnl|TIGR|C.crescentus_4777
PdxJ Cje	257	CAB73492
PdxJ Cte	237	gnl|TIGR|C.tepidum_3495
PdxJ Dvu	243	gnl|TIGR|dvulg_gdv190
PdxJ Eco	243	U36841
PdxJ Gsu	169 (fragment)	gnl|TIGR_35554|gsulf_GGS_1618
PdxJ Hpy	262	O26102
PdxJ Kpn	243	gnl|WUGSC_573|kpneumo_B_KPN.Contig1545
PdxJ Ngo	242	gnl|OUACGT_485|Ngon_Contig11
PdxJ Nme	242	Q9RQV9
PdxJ Pae	243	gnl|PAGP_287|Paeruginosa_Contig1
PdxJ Pgi	238	Q51843
PdxJ Ppu	240	gnl|TIGR|pputida_all_749
PdxJ Spa	200 (fragment)	gnl|WUGSC_32027|spara_B_SPA.0.18908
PdxJ Spu	245	gnl|TIGR_24|sputre_6422
PdxJ Ssp	221	P72776
PdxJ Sty	243	gnl|Sanger_601|S.typhi_Contig1689
PdxJ Stym	243	gnl|WUGSC_99287|stmlt2-.Contig1447
PdxJ Tfe	241	gnl|TIGR|t_ferrooxidans_2940
PdxJ Vch	243	gnl|TIGR|V.cholerae_666_1760_Bert
PdxJ Ype	243	Gnl|Sanger_632|Y.pesits_Contig769

Table 7. PdxH orthologs. The sequence of the PdxH protein of *E. coli* was used as query sequence. Members of this family are listed in alphabetical order.

Abbreviation	Number of residues	Accession or contig number
PdxH Aac	210	gnl|OUACGT_714|A.actin_Contig268
PdxH Bbr	210	gnl|Sanger_518|bbronchi_Contig2279
PdxH Bpe	210	gnl|Sanger_520|B.pertussis_Contig299
PDXH Cal	247	gnl|Stanford_5476|C.albicans_Con4-2649
PdxH Ccr	222	gnl|TIGR|C.crescentus_4753
PDXH Cel	253	Q20939
CG2649 Dme	484	Q9VHZ5
PdxH Dra	214	Q9RX20
PdxH Eco	218	AE000259
PdxH Hdu	209	gnl|HTSC_730|ducreyi
PdxH Hin	229	P44909
CDNAFLJ10535FIS Hsa	261	BAA91668
PdxH Kpn	207 (fragment)	gnl|WUGSC_573|kpneumo_B_KPN.Contig410
PdxH Lpn	215	gnl|CUCGC_446|lpneumo_WG.MF.177.110597
PdxH Mav	209	gnl|TIGR|M.avium_27
PdxH Mbo	218	gnl|Sanger_1765|mbovis_Contig637
PdxH Mle	219	O33065
PdxH Mtu	224	O06207
FprA Mxa	270	M29288
PdxH Ngo	210	gnl|OUACGT_485|Ngon_contig_10
PdxH Nme	210	AAF41734
PdxH Pae	215	gnl|PAGP_287|Paeruginosa_Contig1
PdxH Pgi	214	gnl|TIGR|P.gingivalis_GPC.con
PDXH Rno	261	O88794
PDX3 Sce	228	P38075
PdxH Sco	229	O74250
PdxH Spa	218	gnl|WUGSC_32027|spara_B_SPA.0.12565
SPAC1093.02 Spo	231	Q9UTQ1
PdxH Spu	212	gnl|TIGR|sputre_6418
PdxH Ssp.	230	P74211
PdxH Stym	220	gnl|WUGSC_99287|stmlt2-.Contig1471
PdxH Tfe	218	gnl|TIGR|t_ferrooxidans_2934
PdxH Vch	211	gnl|TIGR|V.cholerae_666_1758_Ernie
PdxH Ype	217	gnl|Sanger_632|Y.pesits_Contig743.0
FprA Zmo	192	Q9RNP3

related to the fungal proteins PYROA of *E. nidulans* and PDX1 (SOR1) of *C. nicotianae*. In several plant species, members of this family have been identified by screening for ethylene inducible proteins: The ER1 protein of *H. brasiliensis*, the rubber tree, was identified as a stress (ethylene and salicylic acid responsive) inducible protein (Sivasubramaniam *et al*., 1995). This protein is highly related to two paralogs of *A. thaliana*, whereas a third paralog of this species is distantly related to the ortholog of another plant, *S. longipes*. Interestingly, a member of this family was also identified as an ethylene responsive protein in a primitive animal, the marine sponge *S. domuncula* (Krasko *et al*., 1999). This protein exhibits the deepest branching within the YaaD derived tree.

The YaaE Derived Phylogenetic Tree (the UPF0030 or SNO Family)

YaaE orthologs (COG0311; the COG database describes these proteins as glutamine amidotransferases possibly involved in histidine and purine biosynthesis) identified by database searches are listed in Table 10. Figure 11 shows an unrooted phylogenetic tree of YaaE orthologs. Compared to the YaaE derived tree similar relationships between the members of this family can be observed. Interesting differences include: (1) The absence of a cluster consisting of euryarchaeotal YaaE orthologs (*A. fulgidus*, *M. thermoautotrophicum*, *M. jannaschii*, three *Pyrococcus* species) near the group of gram positive orthologs (including YaaE). The orthologs of these species are grouped within the other archaeal orthologs *(A. pernix*, *S. solfataricus*). (2) YaaE orthologs were not detected in plants and in filamentous fungi. Some proteins have been annotated as HisH-like proteins in the databases. Galperin and Koonin (1997) noted that the annotation as HisH-like proteins might be a mistake: For example, the *H. influenzae* genome contains a complete histidine biosynthesis operon containing a HisH ortholog exhibiting high similarity to other HisH proteins. Nevertheless, the HI1648 protein belonging to this conserved family is described as putative amidotransferase HisH in the databases.

Discussion

A PdxA ortholog has been identified in the incomplete genomes of *B. halodurans* (Takami *et al*., 1999) and *B. stearothermophilus*. However, no PdxJ-like counterpart has yet been identified in the genome of these species. These bacteria might be the only representatives of grampositive bacteria in which the vitamin B_6 biosynthesis pathway catalyzed by PdxA/J enzymes is functional. Also, no PdxJ ortholog is encoded in the complete plasmid sequence of the catabolic plasmid pNL1 of *S. aromaticivorens.* This plasmid contains 79 genes that encode enzymes associated with the transport or degradation of biphenyl, naphthalene, m-xylene, and p-cresol (Romine *et al*., 1999).

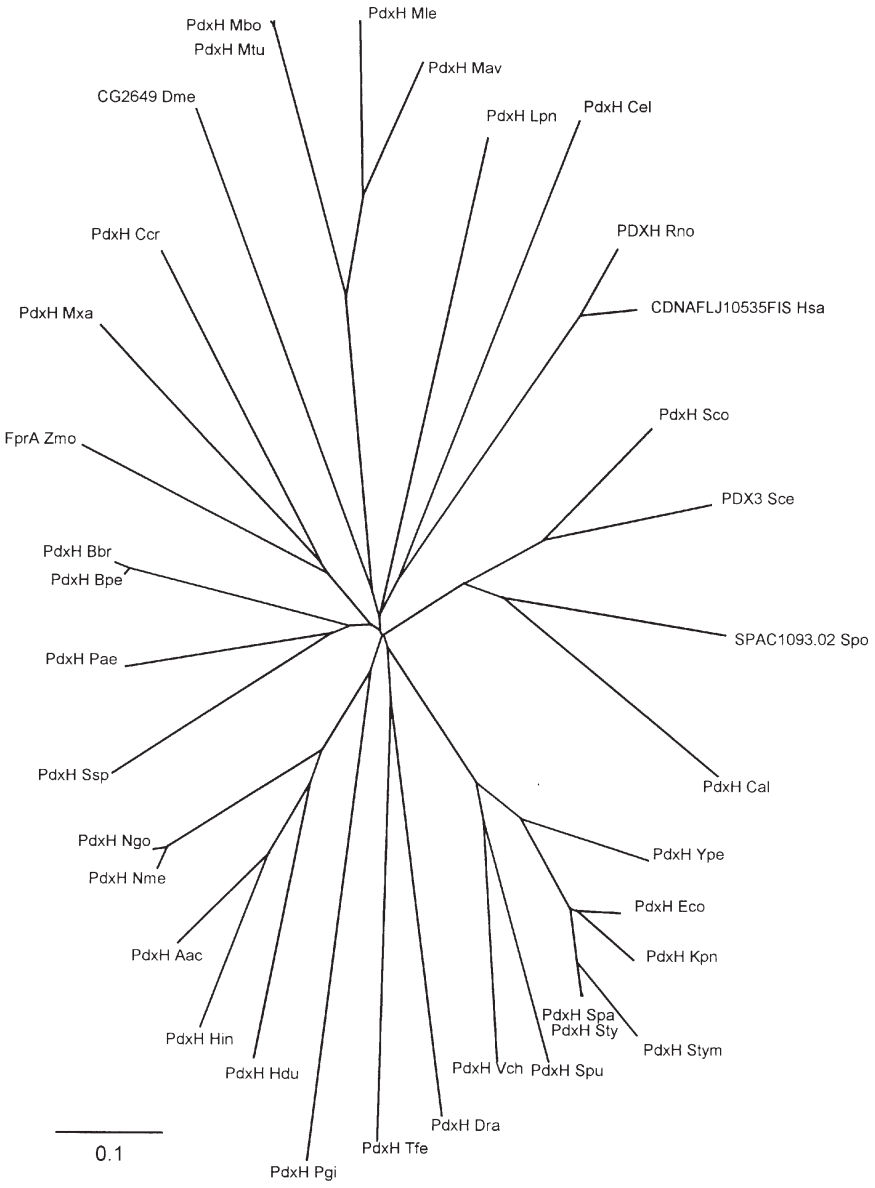


Figure 8. The PdxH derived, unrooted phylogenetic tree (see Table 7).

pNL1 encodes several genes associated with integration and recombination, including two group II intron-associated maturases. These genes were identified in the replication region, suggesting that pNL1 is able to undergo integration and excision events with the chromosome and/or other portions of the plasmid. It might be therefore possible that the PdxA ortholog of pNL1 represents the chromosomally encoded *pdxA* gene of *S. aromaticivorans.* Interestingly, pMOP, a plasmid involved in 4-methylphthalate catabolism in *Burkholderia cepacia* also encodes a PdxA ortholog which is closely related to that encoded on pNL1. These observations might indicate a previously unrecognized participation of vitamin B_6 in the biotechnologically important catalysis of degradation of aromatic compounds.

It remains unclear why the present study failed to identify orthologs of PdxH and PdxK in the genomes of archaea and grampositive bacteria. The only exceptions are the PdxH ortholog of *S. coelicolor* and PdxH enzymes of mycobacteria. In animals, both enzymes exhibit closer relationships to their bacterial paralogs than one might expect as indicated by the rRNA derived tree.

The alternative *de novo* vitamin B_6 biosynthesis pathway encoded by the SNZ/SNO family is clearly more widely distributed and evolutionary well conserved when compared to the rRNA derived tree. Woese (1998) and also Doolittle (1999a) postulated that evolution and history of life cannot be properly represented as tree. The "universal ancestor" is viewed as a "diverse community of cells that survives and evolves as a biological unit" (Woese, 1998). A web- or net-like pattern of horizontal or lateral gene transfers between lineages of organisms – as observed by comparative genomics of twenty completely sequenced microbial genomes – has been proposed to describe accurately relationships between all living species (Doolittle, 1999b).

Table 8. PdxK orthologs. The sequence of the PdxK protein of *E. coli* was used as query sequence. Members of this family are listed in alphabetical order.

Abbreviation	Number of residues	Accession or contig number
PdxK Aae	285	gnl|OUACGT_714|A.actin_Contig384
PdxK Bbr	326 (fragment)	gnl|Sanger_518|bbronchi_Contig2120.1
PdxK Bpe	283	gnl|Sanger_520|B.pertussis_contig345
PDXK Cal	349	gnl|Stanford_5476|C.albicans_Con4-2753
CAC20C1_15 Cal	295	O94003
PdxK Ccr	283	gnl|TIGR|C.crescentus_4669
PdxK Cdi	286	gnl|Sangre_1717|cdiph_Contig 101
PDXK Cel	348	AF003142
CG4446 Dme	494	Q9VSW3
PdxK Dra	329	Q9RYX0
PdxK Eco	283	P40191
PdxY Eco	287	P77150
PdxY Hin	288	P44690
PDXK Hsa	312	O00764
PdxK Kpn	287	gnl|WUGSC_573|kpneumo_B_KPN.Contig1561
PDXK Oar	312	P82197
PDXY Oar	297 (fragment)	Q9XSD8
PdxK Pae	288	gnl|PAGP_287|Paeruginosa_Contig1
PdxK Pmu	286	gnl|CBCUMN-747|Pmultocida
PdxK Ppu	290	gnl|TIGR|pputida_all_105
PDXK Rno	312	O35331
YEC9 Sce	312	P39988
PdxK Spa	269 (fragment)	gnl|WUGSC_32027|spara_B_SPA.0.263
SPCC18.10 Spo	340	O74860
SPAC6F6.11c Spo	309	O14242
PDXK Ssc	322	O46560
PdxK Sty	278 (fragment)	gnl|Sanger_601|S.typhi_Contig7
PdxK Stym	287	P40192
PDXK Tbr	300	O15927
PdxK Ype	286	gnl|Sanger_632|Y.pesits_Contig743.0

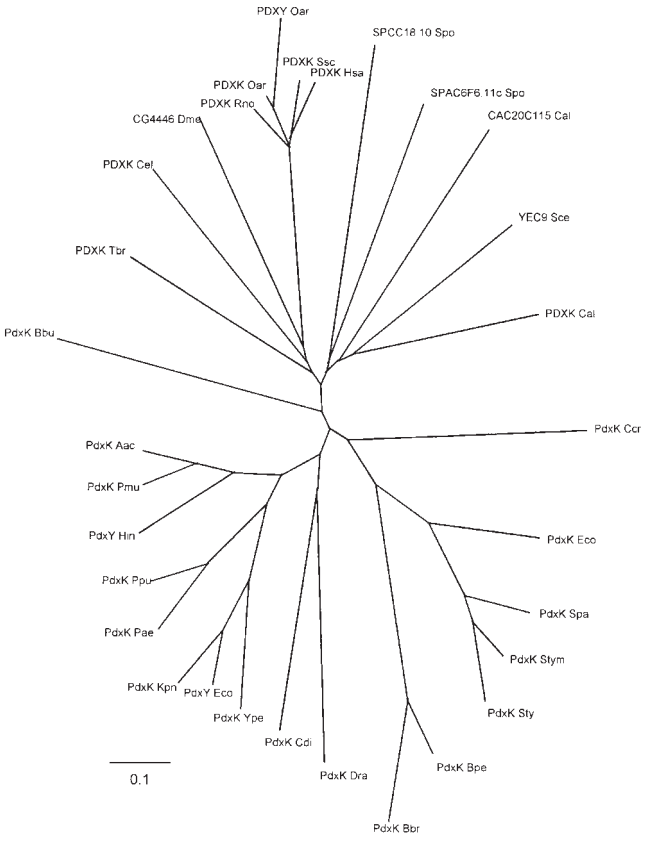


Figure 9. The PdxK derived, unrooted phylogenetic tree (see Table 8).

Evaluating the available dataset, a few organisms might possess both pathways: These are *B. stearothermophilus* and *P. multocida*. In the unfinished genomes of both species, PdxA (but not yet PdxJ) orthologs and both SNZ/SNO orthologs were identified. Three species of the family Pasteurellaceae, γ subdivision of proteobacteria (*P. multocida*, *H. influenzae*, *H. ducreyi*) possess the SNZ/SNO pathway, whereas all other proteobacteria use the PdxA/J pathway for vitamin B_6 biosynthesis. All these observations might indicate that the common precursor of the γ subdivision of proteobacteria used the SNZ/SNO pathway in the first instance. During evolution of the proteobacteria the *snz*/*sno* like genes were most likely deleted for some unknown reason and the PdxA/J dependent pathway evolved in this proteobacterial lineage.

Very recently, the widely accepted three domain proposal by Woese *et al.* (1990) has been questioned and critized by Gupta (Gupta, 2000). Briefly, his approach includes the inclusion of ultrastructural characteristics of cells (*i.e.* organisms surrounded by a single membrane (monoderms) are phylogenetically older and distinct from cells bounded by both an inner and an outer membrane (diderms) whereas a periplasmatic space is created) and the examination of “signature sequences” (Gupta, 1998a, 1998b) (*i.e.* conserved insertions or deletions restricted to specific taxa; “indels”) in dif ferent proteins for the analysis of phylogenetic relationships. These comparisons resulted in the observation of a close relatedness between grampositive bacteria and archaea (Gupta, 1998a) and in the proposal of a linear evolutionary relationship among bacterial taxa (Gupta, 2000) postulating that each major

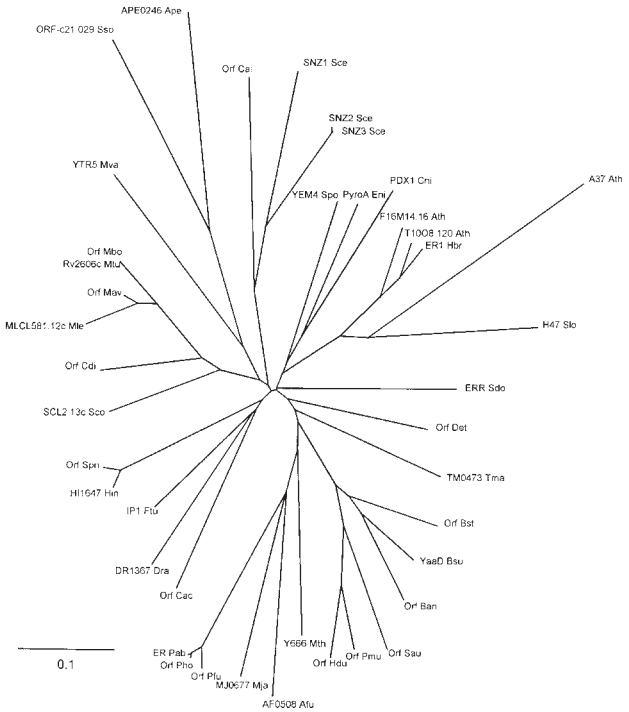


Figure 10. The YaaD derived, unrooted phylogenetic tree (see Table 9).

eubacterial phylum evolved from the preceding one instead randomly from any of the previously existing taxa. According to this linear scheme (Figure 3 in Gupta, 2000) the γ subdivision of proteobacteria represents the most current lineage of prokaryotic evolution whereas grampositive bacteria and archaea are among the oldest representatives of prokaryotes. These observations might indicate that the PdxA/PdxJ enzyme system represents a relatively novel invention of evolution. The distribution spectrum of the two different vitamin B_6 biosynthesis pathways among the prokaryotes is in agreement with this proposal.

Osmani *et al.* (1999) ask why two different pathways for *de novo* vitamin B_6 biosynthesis have evolved. This question cannot be answered right now but a few thoughts might shed some light on this problem:

(1) In the genome of the "modern" bacterium *E. coli*, the *pdxA* and *pdxJ* genes are located within complex operons (*surA-pdxA-ksgA-apaG-apaH* operon; *rnc-era-recO-pdxJ-acpS* operon) containing a variety of other

Table 9. YaaD orthologs. The sequence of the YaaD protein of *B. subtilis* was used as query sequence. The function assigned to each protein - as present in the databases - is also listed. Members of this family are listed in alphabetical order.

Abbreviation	Function	Number of residues	Accession or contig number
AF0508 Afu	Hypothetical ethylene-inducible protein homolog	336	O29742
APE0246 Ape	hypothetical ethylene-responsive protein	337	Q9YFK2
F16M14.16 Ath	putative ethylene inducible protein	309	O80448
T10O8_120 Ath	pyridoxine biosynthesis protein-like	309	CAB81924
A37 Ath	hypothetical protein	314	Q9ZNR6
Orf Ban	-	295	gnlITIGR_1392lbanth_1786
Orf Bst	-	256 (fragment)	gnlIUOKNOR_1422lbstear_Contig1128
YaaD Bsu	Unknown	294	D26185
Orf Cac	-	291	gnlIGTCIC.aceto_gnl
Orf Can	-	292	gnlIStanford_5476lC.albicans_Con4-2682
Orf Cdi	-	300	gnlISanger_1717lcdiph_Contig131
PDX1 Cni (SOR1)	pyridoxine biosynthesis protein; singlet oxygen resistance gene	343	O59905
Orf Det	-	293	gnlITIGR_61435lDeth 1562
DR1367 Dra	singlet oxygen resistance protein, putative	307	Q9RUL7
PYROA Eni	pyridoxine biosynthesis protein	304	Q9UW83
IP1 Ftu	hypothetical protein	239	O69190
ER1 Hbr	ethylene inducible protein	309	Q39963
Orf Hdu	-	295	gnlIHTSC_730lducreyi
HI1647 Hin	hypothetical protein	291	P45293
Orf Mav	-	292	gnlITIGRIM.avium_155
Orf Mbo	-	299	gnlISanger_1765lmbovis_Contig637
MJ0677 Mja	ethylene-inducible protein homolog	330	Q58090
MLCL581.12c Mle	hypothetical protein	333	O07145
Y666 Mth	ethylene-inducible protein	293	O26762
Rv2606c Mtu	hypothetical protein	299	O06208
YTR5 Mva	hypothetical protein in tRNA/5S rRNA cluster	237 (fragment)	Q50841
Er Pab	ethylene-responsive protein	335	Q9V0J7
Orf Pfu	-	332	gnlIUCHGR_2261IMM11-MM11 00540
Orf Pho	-	326 (fragment)	bp 1222808-1221831 of chromosome
Orf Pmu	-	295	gnlICBCUMN_747lPm70seq.fasta
Orf Sau	-	295	gnlISanger_1280_3lsaureusmr_Contig766
SNZ1 Sce	stress induced protein	297	Q03148
SNZ2 Sce	stress induced protein	298	P53824
SNZ3 Sce	member of the stationary phase-induced gene family	298	NP_011127
SCL2.13c Sco	hypothetical protein	303	CAB70925
ERR Sdo	ethylene responsive receptor	306	Q9U5K
H47 Slo	hypothetical protein	235 (fragment)	Q41348
Orf Spn	-	296	gnlITIGRIS.pneumoniae_3478
YEM4 Spo	putative stress-induced protein	296	O14027
ORF-c21_029 Sso	hypothetical protein	338	Q9UWX3
TM0473 Tma	conserved hypothetical protein	293	Q9WYU4

Table 10. YaaE orthologs. The sequence of the YaaE protein of *B. subtilis* was used as query sequence. The function assigned to each protein - as present in the databases - is also listed. Members of this family are listed in alphabetical order.

Abbreviation	Function	Number of residues	Accession or contig number
HisH Afu	imidazoleglycerolphosphate synthase, subunit H, putative	198	O29741
APE0244 Ape	hypothetical protein	186	Q9YFK4
Orf Ban	-	196	gnlITIGR_1392Ibanth_1786
Orf Bst	-	196	gnlIUOKNOR_1422Ibstear_Contig1128
YaaE Bsu	hypothetical protein	196	P37528
Orf Cac	-	186	gnlIGTCIC.aceto_gnl
Orf Cal	-	249	gnl Stanford 5476 C. albicans Con4-2696
Orf Cdi	-	185	gnlISanger_1717Icdiph_Contig131
Orf Det	-	195	gnlITIGR_61435Ideth_1513
DR1366 Dra	probable amidotransferase HisH	196	Q9RUL8
Orf Hdu	-	189	gnlIHTSC_730Iducreyi
HI1648 Hin	conserved hypothetical protein	175	P45294
Orf Mav	-	193	gnlITIGRIM.avium_155
Orf Mbo	-	225 (fragment)	gnlISanger_1765Imbovis_Contig637
MJ1661 Mja	hypothetical protein	186	Q59055
HisH Mle	Amidotransferase HisH homolog	219	Q49637
MTH190 Mth	conserved protein	192	O26292
Rv2604c Mtu	hypothetical protein	198	O06210
HisH-like Pab	Imidazoleglycerolphosphate synthase, subunit H, putative	196	Q9V0J6
Orf Pfu	-	197	gnlIUCHGR_2261IMM11-MM11
PH1354 Pho	hypothetical protein	196	O59079
Orf Pmu	-	193	gnlICBCUMN_747IPm70seq.fasta
Orf Sau	-	186	gnlIOUACGT_1280Is.aureus_Contig628
SNO1 Sce	hypothetical protein	224	Q03144
SNO2 Sce	hypothetical protein	222	P53823
SNO3 Sce	member of a stationary phase induced protein family	222	NP_011126.1
SCL2.12c Sco	hypothetical protein	202	CAB70924
Orf Spn	-	193	gnlITIGRIS.pneumoniae_3478
SPAC222.08c Spo	hypothetical protein	234	Q9UTE4
ORF-c21_028 Sso	hypothetical protein	200	Q9UWX4
TM0472 Tma	putative amidotransferase	188	Q9WYU3

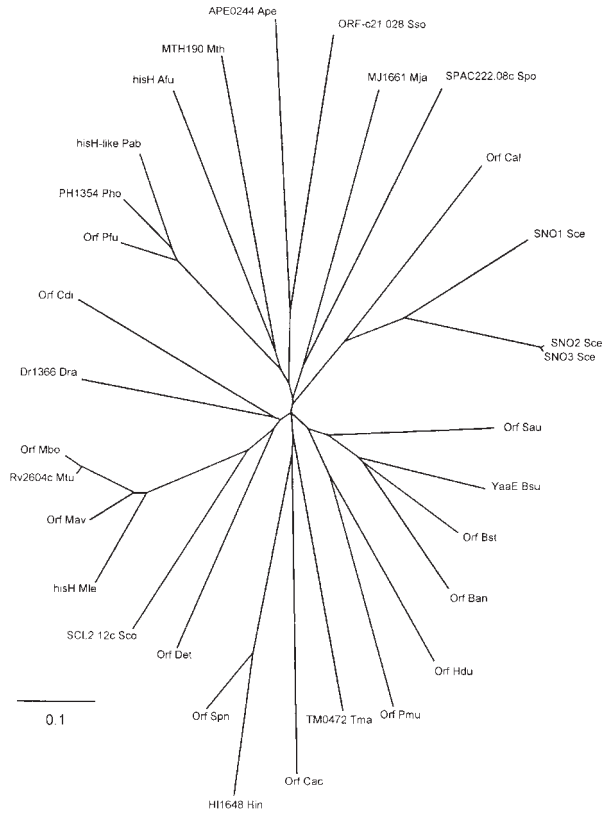


Figure 11. The YaaE derived, unrooted phylogenetic tree (see Table 10).

genes involved in important cellular functions (Table 11; Table 11 gives also an overview about possible orthologs of these genes identified in *B. subtilis*). Both operons are subjected to complex transcriptional regulation mechanisms (Roa *et al.*, 1989; Lam *et al.*, 1992; Matsunaga *et al.*, 1996). Many of these proteins are somehow associated with the ribosomal RNA: The KsgA protein is a 16S rRNA methyltransferase (van Gemen *et al.*, 1987), the *rnc* gene product RNAse III is involved in maturation of ribosomal RNA (Matsunaga *et al.*, 1996), the essential Era protein binds to 16S rRNA (Meier *et al.*, 2000). Interestingly, suppressor mutations of a cold sensitive *era* mutation have been mapped in the *ksgA* gene (Lu and Inouye, 1998), indicating that of the gene products of both operons not only PdxA and PdxJ but also KsgA and Era perform protein-protein interactions. Era is a small G-protein widely conserved in eubacteria and eukaryotes. Although essential for bacterial growth and implicated in diverse cellular processes (Britton *et al.*, 1998, Johnstone *et al.*, 1999), its actual function remains unclear. The ApaH protein also performs an important function in bacterial growth: It inactivates AppppA (Ap4A), an alarmone synthesized by aminoacyl-tRNA-synthetases (Brevet *et al.*, 1989) in response to oxidative stress (Bochner *et al.*, 1982; Lee *et al.*, 1983). Interestingly, mutations in *apaH* result in an increased frequency of cell division (Lévêque *et al.*, 1990, Nishimura *et al.*, 1997). At least the *rnc-era-recO-pdxJ-acpS* operon seems to be conserved in *B. japonicum* (Bairl and Müller, 1998) and *P. aeruginosa* (Powell *et al.*, 1999). In *E. coli* (and presumably its close relatives), transcription of the *pdxA* and *pdxJ* genes seems to be coupled fundamental cellular processes.

Table 11. The complex *pdxA* and *pdxJ* operons of *E. coli* contain genes involved in diverse cellular processes. The orthologous genes of *B. subtilis* which do not exhibit this type of organization are also listed.

E. coli genes in the operon containing *pdxA*	Enzymatic function of gene product; comments	References	Similarities in *B. subtilis*	Enzymatic function of gene product; comments	Expectation value	References
surA	periplasmic peptidyl-prolyl isomerase required for stationary phase survival	Tormo *et al.*, 1990; Lazar *et al.* 1998	*prsA* *yacD*	protein secretion (post-translocation chaperonin) unknown; similar to protein secretion PrsA homolog.	9.3e-10 0.00014	Kontinen *et al.*, 1999 Kontinen and Sarvas, 1993; Jacobs *et al.*, 1993
pdxA	a heterodimer consisting of PdxA and PdxJ catalyzes final step of pyridoxine 5'-phosphate biosynthesis using 4-phospho-hydroxy-L-threonine and 1-Deoxy-D-xylulose 5-phosphate	Laber *et al.*, 1999	none			
ksgA	S-Adenosylmethionine-6-N',N'-adenosyl (rRNA) dimethyltransferase	van Gemen *et al.*, 1987	*ksgA*	dimethyladenosine transferase; second gene in a three gene operon containing also *yabF* and *yabG*	1.1e-33	
apaG	Unknown	Lévêque *et al.*, 1990	*trxB*	thioredoxin reductase	0.045	
apaH	Diadenosine tetraphosphatase	Lévêque *et al.*, 1990	*yjbP*	unknown; similar to diadenosine tetraphosphatase; monocistronic	1.2e-10	
E. coli genes in the operon containing *pdxJ*	Enzymatic function of gene product; comments	References	Similarities in *B. subtilis*	Enzymatic function of gene product; comments	Expectation value	References
rnc	RNase III; cleaves double-stranded RNA	Matsunaga *et al.*, 1996	*rncS*	Ribonuclease III; monocistronic	2.0e-34	Wang and Bechhofer, 1997
era	Essential GTPase; ras-likeG-protein; binds to 16S rRNA	Britton *et al.*, 1998 Johnstone *et al.*, 1999 Meier *et al.*, 2000	*bex*	GTP-binding protein; complements *era* mutants; upstream of *recO*	5.1e-61	
recO	Conjugational recombination and repair; DNA-binding protein; RecA-like strand assimilation	Morrison *et al.*, 1989 Luisi-De Luca and Kolodner, 1994	*recO*	involved in DNA repair and homologous recombination null mutant leads to high sensitivity to DNA-damaging agents and toreduced efficiency of plasmid transformation (intramolecular recombination) and chromosomal transformation (intermolecular recombination)	2.2e-08	Fernandez *et al.*, 1999
pdxJ	see *pdxA*		see Table 2			
acpS (*dpj*)	Acyl-carrier protein	Lambalot and Walsh 1995	*ydcB*	unknown; similar to holo-acyl-carrier protein synthase; operon with *ydcC*	2.1e-16.	

(2) The *de novo* PLP biosynthesis pathway of *E. coli* contains a PLP-dependent transamination reaction catalyzed by the PdxF (SerC) enzyme, thus representing a "hen and egg"-paradox. It seems very likely that the PdxA/PdxJ dependent *de novo* pathway evolved after vitamin B_6 vitamers (which can be modified by the salvage pathway) were already present in the environment of *E. coli* due to the "alternative" pathway. This pathway most probably avoids PLP-dependent transamination steps and instead uses vitamin B_6 independent (Massière and Badet-Denisot, 1998; Huang and Raushel, 1999) HisH-like amidotransferases for transfer of nitrogen atoms. Interestingly, PdxF (SerC) is also involved in regulation of the frequency of septation (Mouslim *et al.*, 2000).

(3) The discovery of the "alternative" pathway has been connected with the realization that vitamin B_6 is implicated in stress responses and particularly in cellular antioxidant defense (Ehrenshaft *et al.*, 1999, Bilski *et al.*, 2000), namely in singlet oxygen resistance. Also, the YaaD protein of *B. subtilis* has been identified as a cumene hydroperoxide and H_2O_2 inducible protein (Antelmann *et al.*, 1997). In cells grown in sporulation medium, the YaaD protein is subjected to an unusual posttranslational modification, guanylation (Mitchell *et al.*, 1992). Two reports indicate that vitamin B_6 might also have some protective function against these oxidants. Manganese dipyridoxyl diphosphate, a magnetic resonance imaging contrast agent possesses antioxidative and cardioprotective properties (Brurok *et al.*, 1999). Vitamin B_6 provides a similar protective effect against cumene hydroperdoxide as does this compound (Meyer *et al.*, 1992). Assuming the correctness of Gupta's theory of linear procaryotic evolution, these observations might indicate that vitamin B_6 biosynthesis represents an ancient response to oxidative stress.

In conclusion, the well studied PdxA/PdxJ *de novo* vitamin B_6 biosynthesis pathway is restricted primarily to *E. coli* and its relatives where it seems to be coupled to cellular growth; the other, widely distributed pathway is stress inducible. *B. subtilis*, which has been used in biotechnology for the overproduction of another vitamin, riboflavin (Mack *et al.*, 1998; Sauer and Bailey, 1999) might be also the organism of choice to elucidate the biochemical and enzymological aspects of this conserved pathway. Given the magnitude of genes of unknown function in completely sequenced microbial genomes, it will be also interesting to learn whether comparative analysis of the presence and/or absence of certain genes in individual genomes might also be used to prove (or disprove) Gupta's hypothesis of linear evolutionary relationships in prokaryotes.

Experimental Procedures

For the initial comparison of *E. coli* and *B. subtilis* genes, protein sequences of *E. coli* were obtained at the "ColiBri" database (http://genolist.pasteur .fr/Colibri/) and a BLAST Search was performed at the "SubtiList" website (http://genolist.pasteur.fr/SubtiList) (Moszer *et al.*, 1995, Moszer, 1998). The sequences of small subunit rRNA's (16S or 18S rRNA) were retrieved from databases (SRS6 system (http://srs6.ebi.ac.uk/) or Genbank (http://www.ncbi.nlm.nih.gov:80/entrez/)). Using the deduced protein sequences of the *pdxA*, *pdxJ*, *pdxH*, *pdxK* genes of *E. coli* and of the *yaaD* and *yaaE* genes of *B. subtilis* as query sequences, BLAST searches (Altschul *et al*, 1997) were performed at http://dove.embl-heidelberg.de/Blast2/. BLAST searches of unfinished microbial genomes were performed at http://www.ncbi.nlm.nih.gov/Microb_blast/unfinishedgenome.html. Raw DNA sequences were translated using the program "Translate tool" at http://www.expasy.ch/tools/dna.html. The COG database can be assessed at http://www.ncbi.nlm.nih.gov/COG. Alignments were generated using CLUSTALW (Thompson *et al.*, 1994) at http://www.ebi.ac.uk/clustalw/. Phylogenetic trees were constructed using the data from the alignments

with the help of the program TreeView (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html; Page, 1996) and edited using the program Metafile Companion (http://www.companionsoftware.com/).

Acknowledgements

This work was performed in the lab of Michael Hecker, whose support is gratefully acknowledged. I thank Michael Hecker and Ulrich Zuber for comments on the manuscript. Sequences of unfinished microbial genomes were obtained at the website of "The Institute of Genomic Research" (TIGR) at http://www.tigr.org. Sequencing of unfinished genomes is in progress at several institutions funded by a variety of organizations: This listing includes "name of organism (institution(s); source of funding)" and can be also viewed at http://www.tigr.org/tdb/mdb/mdbinprogress.html. *A. actinomycetemcomitans* (University of Oklahoma; National Institute of Dental Research (NIDR)); *B. anthracis* (TIGR; Office of Naval Research (ONR)/Department of Energy (DOE)/National Institute of Allergy and Infectious Diseases (NIAD)); *B. halodurans* (Japan Marine Science and Technology Center); *B. stearothermophilus* (University of Oklahoma; National Science Foundation (NSF); *B. bronchiseptica*, *B. pertussis*, *C. difficile*, *N. meningitidis*, *Y. pestis* (Sanger Centre; Beowulf Genomics); *C. albicans* (Sanger Centre; Beowulf Genomics and also at Stanford University; NIDR/National Institutes of Health (NIH)/Burroughs Wellcome Fund); *C. crescentus*, *C. tepidum*, *D. ethenogenes*, *D. vulgaris*, *S. putrefaciens*, *T. ferrooxidans* (TIGR/DOE); *C. acetobutylicum* (Genome Therapeutics; DOE); *C. diphteriae* (Sanger Centre/World Health Organization (W.H.O.)/Public Health Laboratory; Beowulf Genomics); *C. trachomatis, C. pneumoniae* (Genset); *F. tularensis* (European and North American Consortium); *G. sulfurreducens* (TIGR/University of Massachusetts, Amherst/DOE); *H. ducreyi* (NIAID); *K. pneumoniae*, *S. paratyphi*, *S. typhimurium* (Washington University Consortium); *L. pneumophila* (Columbia Genome Center; NIAID); *M. avium*, *V. cholerae* (TIGR/NIAID); *M. bovis* (Sanger Centre/Institut Pasteur/VLA Weybridge; MAFF/Beowulf Genomics); *M. leprae* (Sanger Centre; The New York Community Trust); *N. gonorrhoeae* (University of Oklahoma; NIAID); *P. multocida* (University of Minnesota; US Department of Agriculture - National Reserach Initiative (USDA-NRI)/Minnesota Turkey Growers Association); *P. gingivalis* (TIGR/Forsyth Dental Center; NIDR); *P. aeruginosa* (University of Washington/PathoGenesis; Cystic Fibrosis Foundation/PathoGenesis); *P. putida* (TIGR/German Consortium; DOE/Bundesministerium für Bildung und Forschung (BMBF); *P. furiosus* (Center of Marine Biotechnology/University of Utah; DOE); *S. typhi*, *S. pombe* (Sanger Centre; Wellcome Trust); *S. meliloti* (European and Canadian Consortium/Stanford University; European Union); *S. aureus* (TIGR; NIAID/Merck Genome Research Institute (MGRI)); *S. pneumoniae* (TIGR; TIGR/NIAID/MGRI); *S. coelicolor* (Sanger Centre/John Innes Centre; Biotechnology and Biological Sciences Research Council (BBSRC)/Beowulf Genomics); *S. solfataricus* (Canadian and European Consortium); *T. denticola* (TIGR/University of Texas).

This paper is dedicated to the memory of Gerald D. Shockman.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389-3402.

Antelmann, H., Bernhardt, J., Schmid, R., Mach, H., Völker U., and Hecker, M. 1997. First steps from a two-dimensional protein index towards a response-regulation map for *Bacillus subtilis*. Electrophoresis 18: 1451-1463.

Bairl, A., and Müller, P. 1998. A second gene for type I signal peptidase in *Bradyrhizobium japonicum*, *sipF*, is located near genes involved in RNA processing and cell division. Mol. Gen. Genet. 260: 346-356.

Begley, T.P., Downs, D.M., Ealick, S.E., McLafferty, F.W., van Loon, A.P., Taylor, S., Campobasso, N., Chiu, H.J., Kinsland, C., Reddick, J.J., and Xi, J. 1999. Thiamin biosynthesis in prokaryotes. Arch. Microbiol. 171: 293-300.

Bilski, P., Li, M.Y., Ehrenshaft, M., Daub, M.E., and Chignell, C.F. 2000. Vitamin B_6 (pyridoxine) and its derivatives are efficient singlet oxygen quenchers and potential fungal antioxidants. Photochem. Photobiol. 71: 129-134.

Bochner, B.R., Lee, P.C., Wilson, S.W., Cutler, C.W., and Ames, B.N. 1984. AppppA and related adenylylated nucleotides are synthesized as a consequence of oxidation stress. Cell 37: 225-232.

Boucher, Y., and Doolittle, W.F. 2000. The role of lateral gene transfer in the evolution of isoprenoid biosynthesis pathways. Mol. Microbiol. 37: 703-716.

Bouvier, F., D'Harlingue, A., Suire, C., Backhaus, R.A., and Camara, B. 1998. Dedicated roles of plastid transketolases during the early onset of isoprenoid biogenesis in pepper fruits. Plant Physiol. 117: 1423-1431.

Braun, E.L., Fuge, E.K., Padilla, P.A., and Werner-Washburne, M. 1996. A stationary-phase gene in *Saccharomyces cerevisiae* is a member of a novel, highly conserved gene family. J. Bacteriol. 178: 6865-6872.

Brevet, A., Chen, J., Lévêque, F., Plateau, P., and Blanquet, S. 1989. *In vivo* synthesis of adenylylated bis(5'-nucleosidyl) tetraphosphates (Ap4N) by *Escherichia coli* aminoacyl-tRNA synthetases. Proc. Natl. Acad. Sci. USA. 86: 8275-8279.

Britton, R.A., Powell, B.S., Dasgupta, S., Sun, Q., Margolin, W., Lupski, J.R., and Court, D.L. 1998. Cell cycle arrest in Era GTPase mutants: a potential growth rate-regulated checkpoint in *Escherichia coli*. Mol. Microbiol. 27: 739-750. (Erratum: Mol. Microbiol. 1998. 28: 1391-1393).

Brurok, H., Ardenkjaer-Larsen, J.H., Hansson, G., Skarra, S., Berg, K., Karlsson, J.O., Laursen. I., and Jynge, P. 1999. Manganese dipyridoxyl diphosphate: MRI contrast agent with antioxidative and cardioprotective properties? Biochem. Biophys. Res. Commun. 254: 768-772.

Cane, D.E., Hsiung, Y., Cornish, J.A., Robinson, K., and Spenser, I.D. 1998. Biosynthesis of vitamin B_6: the oxidation of 4-(phosphohydroxy)-L-threonine by PdxA. J. Am. Chem. Soc. 120: 1936-1937.

Capitani, G., Hohenester, E., Feng, L., Storici, P., Kirsch, J.F., and Jansonius, J.N. 1999. Structure of 1-aminocyclopropane-1-carboxylate synthase, a key enzyme in the biosynthesis of the plant hormone ethylene. J. Mol. Biol. 294: 745-756.

Choi, S.Y., Churchich, J.E., Zaiden, E., and Kwok, F. 1987. Brain pyridoxine-5-phosphate oxidase. Modulation of its catalytic activity by reaction with pyridoxal 5-phosphate and analogs. J. Biol. Chem. 262: 12013-12017.

Dakshinamurti, K. (ed.). 1990. Vitamin B_6. Ann. N.Y. Acad. Sci. Vol. 585.

Denessiouk, K.A., Denesyuk, A.I., Lehtonen, J.V., Korpela, T., and Johnson, M.S. 1999. Common structural elements in the architecture of the cofactor-binding domains in unrelated families of pyridoxal phosphate-dependent enzymes. Proteins 35: 250-261.

Dolphin, D., Poulson, R., and Avramovic, O. 1986. Vitamin B_6 pyridoxal phosphate: chemical, biochemical and medical aspects. Wiley Interscience, New York, New York.

Doolittle, W.F. 1999a. Phylogenetic classification and the universal tree. Science 284: 2124-2129.

Doolittle, W.F. 1999b. Lateral genomics. Trends Cell Biol. 9: M5-8.

Drewke, C., Klein, M., Clade, D., Arenz, A., Müller, R., and Leistner, E. 1996. 4-O-phosphoryl-L-threonine, a substrate of the *pdxF*(*serC*) gene product involved in vitamin B_6 biosynthesis. FEBS Lett. 390: 179-182.

Ehrenshaft, M., Jenns, A.E., Chung, K.-R., and Daub, M.E. 1998. SOR1, a gene required for photosensitizer and singlet oxygen resistance in *Cercospora* fungi, is highly conserved in divergent organisms. Mol. Cell 1: 603-609.

Ehrenshaft, M., Bilski, P., Li, M.Y., Chignell, C.F., and Daub, M.E. 1999. A highly conserved sequence is a novel gene involved in de novo vitamin B_6 biosynthesis. Proc. Natl. Acad. Sci. USA. 96: 9374-9378.

Fernandez, S., Kobayashi, Y., Ogasawara, N., and Alonso, J.C. 1999. Analysis of the *Bacillus subtilis recO* gene: RecO forms part of the RecFLOR function. Mol. Gen. Genet. 261: 567-573.

Galperin, M.Y., and Koonin, E.V. 1997. Sequence analysis of an exceptionally conserved operon suggests enzymes for a new link between histidine and purine biosynthesis. Mol Microbiol. 24: 443-445.

Grogan, D.W. 1988. Temperature-sensitive murein synthesis in an *Escherichia coli pdx* mutant and the role of alanine racemase. Arch. Microbiol. 150: 363-367.

Gupta, R.S. 1998a. What are archaebacteria: life's third domain or monoderm prokaryotes related to gram-positive bacteria? A new proposal for the classification of prokaryotic organisms. Mol. Microbiol. 29: 695-707.

Gupta, R.S. 1998b. Protein phylogenies and signature sequences: A reappraisal of evolutionary relationships among archaebacteria, eubacteria, and eukaryotes. Microbiol. Mol. Biol. Rev. 62: 1435-1491.

Gupta, R.S. 2000. The natural evolutionary relationships among prokaryotes. Crit. Rev. Microbiol. 26: 111-131.

Hanna, M.C., Turner, A.J., and Kirkness, E.F. 1997. Human pyridoxal kinase. cDNA cloning, expression, and modulation by ligands of the benzodiazepine receptor. J. Biol. Chem. 272: 10756-10760.

Helmreich, E.J. 1992. How pyridoxal 5'-phosphate could function in glycogen phosphorylase catalysis. Biofactors 3: 159-172.

Hill, R.E., and Spenser, I.D. 1996. Biosynthesis of vitamin B_6. In: Escherichia coli and Salmonella: cellular and molecular biology, 2nd ed., F.C. Neidhardt, R. Curtiss III, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S Reznikoff, M. Riley. M. Schaechter, and H.E. Umbarger, eds. ASM Press, Washington, D.C. p. 695-703.

Huang, X., and Raushel, F.M. 1999. Deconstruction of the catalytic array within the amidotransferase subunit of carbamoyl phosphate synthetase. Biochemistry 38: 15909-15914.

Jacobs, M., Andersen, J.B., Kontinen, V., and Sarvas M. 1993. *Bacillus subtilis* PrsA is required *in vivo* as an extracytoplasmic chaperone for

secretion of active enzymes synthesized either with or without pro-sequences. Mol. Microbiol. 8: 957-966.
Johnstone, B.H., Handler, A.A., Chao, D.K., Nguyen, V., Smith, M., Ryu, S.Y., Simons, E.L., Anderson, P.E., and Simons, R.W. 1999. The widely conserved Era G-protein contains an RNA-binding domain required for Era function *in vivo*. Mol. Microbiol. 33: 1118-1131.
Kack, H., Sandmark, J., Gibson, K., Schneider, G., and Lindqvist, Y. 1999. Crystal structure of diaminopelargonic acid synthase: evolutionary relationships between pyridoxal-5'-phosphate-dependent enzymes. J. Mol. Biol. 291: 857-876.
Kontinen, V.P., Saris, P., and Sarvas, M. 1991. A gene (*prsA*) of *Bacillus subtilis* involved in a novel, late stage of protein export. Mol. Microbiol. 5: 1273-1283.
Kontinen, V.P., and Sarvas, M. 1993. The PrsA lipoprotein is essential for protein secretion in *Bacillus subtilis* and sets a limit for high-level secretion. Mol. Microbiol. 8: 727-737.
Kortlüke, C., Breese, V., Gad'on, N., Labahn, A., and Drews, G. 1997. Structure of the *puf* operon of the obligately aerobic, bacteriochlorophyll alpha-containing bacterium *Roseobacter denitrificans* OCh114 and its expression in a *Rhodobacter capsulatus puf puc* deletion mutant. J. Bacteriol. 179: 5247-5258.
Krasko, A., Schroeder, H.C., Perovic, S., Steffen, R., Kruse, M., Reichert, W., Mueller, I.M., and Mueller, W.E.G. 1999. Ethylene modulates gene expression in cells of the marine sponge *Suberites domuncula* and reduces the degree of apoptosis. J. Biol. Chem. 274: 31524-31530.
Kuzuyama, T., Takagi, M., Takahashi, S., and Seto, H. 2000a. Cloning and characterization of 1-deoxy-D-xylulose 5-phosphate synthase from *Streptomyces sp.* strain CL190, which uses both the mevalonate and nonmevalonate pathways for isopentenyl diphosphate biosynthesis. J. Bacteriol. 182: 891-897.
Kuzuyama, T., Takahashi, S., Takagi, M., and Seto, H. 2000b. Characterization of 1-deoxy-D-xylulose 5-phosphate reductoisomerase, an enzyme involved in isopentenyl diphosphate biosynthesis, and identification of its catalytic amino acid residues. J. Biol. Chem. 275: 19928-19932.
Laber, B., Maurer, W., Scharf, S., Stepusin, K., and Schmidt, F.S. 1999. Vitamin B_6 biosynthesis: formation of pyridoxine 5'-phosphate from 4-(phosphohydroxy)-L-threonine and 1-deoxy-D-xylulose-5-phosphate by PdxA and PdxJ protein. FEBS Lett. 449: 45-48.
Lam, H.M., and Winkler, M.E. 1990. Metabolic relationships between pyridoxine (vitamin B_6) and serine biosynthesis in *Escherichia coli* K-12. J. Bacteriol. 172: 6518-6528.
Lam, H.M., Tancula, E., Dempsey, W.B., and Winkler, M.E. 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, H.M., and Winkler, M.E. 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.
Lambalot, R.H., and Walsh, C.T. 1995. Cloning, overproduction, and characterization of the Escherichia coli holo-acyl carrier protein synthase. J. Biol. Chem. 270: 24658-24661.
Lange, B.M., Wildung, M.R., McCaskill, D., and Croteau, R. 1998. A family of transketolases that directs isoprenoid biosynthesis via a mevalonate-independent pathway. Proc. Natl. Acad. Sci. USA. 95: 2100-2104.
Lazar, S.W., Almiron, M., Tormo, A., and Kolter, R. 1998. Role of the *Escherichia coli* SurA protein in stationary-phase survival. J. Bacteriol. 180: 5704-5711.
Lee, P.C., Bochner, B.R,, and Ames, B.N. 1983. AppppA, heat-shock stress, and cell oxidation. Proc. Natl. Acad. Sci. USA. 80: 7496-7500.
Lévêque, F., Blanchin-Roland, S., Fayat, G., Plateau, P., and Blanquet, S. 1990. Design and characterization of *Escherichia coli* mutants devoid of Ap4N-hydrolase activity. J. Mol. Biol. 212: 319-329.
Lois, L.M., Campos, N., Putra, S.R., Danielsen, K., Rohmer, M., and 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. USA. 95: 2105-2110.
Lu, Q., and Inouye, M. 1998. The gene for 16S rRNA methyltransferase (*ksgA*) functions as a multicopy suppressor for a cold-sensitive mutant of *era*, an essential RAS-like GTP-binding protein in *Escherichia coli*. J. Bacteriol. 180: 5243-5246.
Luisi-DeLuca, C., and Kolodner, R. 1994. Purification and characterization of the *Escherichia coli* RecO protein. Renaturation of complementary single-stranded DNA molecules catalyzed by the RecO protein. J. Mol. Biol. 236: 124-138.
Mack, M., van Loon, A.P., and Hohmann, H.P. 1998. Regulation of riboflavin biosynthesis in *Bacillus subtilis* is affected by the activity of the flavokinase/flavin adenine dinucleotide synthetase encoded by *ribC*. J. Bacteriol. 180: 950-955.
Massière, F., and Badet-Denisot, M.A. 1998. The mechanism of glutamine-dependent amidotransferases. Cell. Mol. Life Sci. 54: 205-222.
Masuda, S., Yoshida, M., Nagashima, K.V.P, Shimada, K., and Matsuura, K. 1999. A new cytochrome subunit bound to the photosynthetic reaction center in the purple bacterium, *Rhodovulum sulfidophilum*. J. Biol. Chem. 274: 10795-10801.
Matsunaga, J., Dyer, M., Simons, E.L., and Simons, R.W. 1996. Expression and regulation of the *rnc* and *pdxJ* operons of *Escherichia coli*. Mol. Microbiol. 22: 977-989.
McCormick, D.B., and Chen, H. 1999. Update on interconversions of vitamin B_6 with its coenzyme. J. Nutr. 129: 325-327.
Meier, T.I., Peery, R.B., McAllister, K.A., and Zhao, G. 2000. Era GTPase of *Escherichia coli*: binding to 16S rRNA and modulation of GTPase activity by RNA and carbohydrates. Microbiology 146: 1071-1083.
Meyer, B.U., Schneider, W., and Elstner, E.F. 1992. Peroxide-dependent amino acid oxidation and chemiluminescence catalysed by magnesium-pyridoxal phosphate-glutamate complex. Biochem. Pharmacol. 44: 505-508.
Mihara, H., Kurihara, T., Yoshimura, T., Soda, K., and Esaki, N. 1997. Cysteine sulfinate desulfinase, a NIFS-like protein of *Escherichia coli* with selenocysteine lyase and cysteine desulfurase activities. Gene cloning, purification, and characterization of a novel pyridoxal enzyme. J. Biol. Chem. 272: 22417-22424.
Miller, M., Heuser, T., and Zimmer, W. 1999. A *Synechococcus leopoliensis* SAUG 1402-1 operon harbouring the 1-deoxyxylulose 5-phosphate synthase gene and two additional openreading frames is functionally involved in the dimethylallyl diphosphate synthesis. FEBS Lett. 460: 485-490.
Mitchell, C., Morris, P.W, and Vary, J.C. 1992. Amino acid sequences of several *Bacillus subtilis* proteins modified by apparent guanylylation. Mol. Microbiol. 6:1579-1581.
Morrison, P.T, Lovett, S.T., Gilson, L.E., and Kolodner, R. 1989. Molecular analysis of the *Escherichia coli recO* gene. J. Bacteriol. 171: 3641-3649.
Moszer, I., Glaser, P., and Danchin, A. 1995. SubtiList: a relational database for the *Bacillus subtilis* genome. Microbiology 141: 261-268.
Moszer, I. 1998. The complete genome of *Bacillus subtilis*: From sequence annotation to data management and analysis. FEBS Lett. 430: 28-36.
Mouslim, C., Cano, D.A., Flores, A., and Casadesús, J. 2000. Regulation of septation: a novel role for SerC/PdxF in *Salmonella*? Mol. Gen. Genet. 264: 184-192.
Nishimura, A., Moriya, S., Ukai, H., Nagai, K., Wachi, M., and Yamada, Y. 1997. Diadenosine 5',5'''-P1,P4-tetraphosphate (Ap4A) controls the timing of cell division in *Escherichia coli*. Genes to Cells 2: 401-413.
Notheis, C., Drewke, C., and Leistner, E. 1995. Purification and characterization of the pyridoxol-5'-phosphate:oxygen oxidoreductase (deaminating) from *Escherichia coli*. Biochim. Biophys. Acta 1247: 265-271.
Osmani, A.H., May G.S., and Osmani, S.A. 1999. The extremely conserved *pyroA* gene of *Aspergillus nidulans* is required for pyridoxine synthesis and is required indirectly for resistance to photosensitizers. J. Biol. Chem. 274: 23565-23569.
Padilla, P.A., Fuge, E.K., Crawford, M.E., Errett, A., and Werner-Washburne, M. 1998. The highly conserved, coregulated SNO and SNZ gene families in *Saccharomyces cerevisiae* respond to nutrient limitation. J. Bacteriol. 180: 5718-5726.
Page, R.D.M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. CABIOS. 12: 357-358.
Pflug, W., and Lingens, F. 1978. Vitamin B_6 biosynthesis in *Bacillus subtilis*. Hoppe Seylers Z. Physiol. Chem. 359: 559-570. (Article in German).
Powell, B., Peters, H.K. III, Nakamura, Y., and Court, D. 1999. Cloning and analysis of the *rnc-era-recO* operon from *Pseudomonas aeruginosa*. J. Bacteriol. 181: 5111-5113.
Roa, B.B., Connolly, D.M., and Winkler, M.E. 1989. Overlap between *pdxA* and *ksgA* in the complex *pdxA-ksgA-apaG-apaH* operon of *Escherichia coli* K-12. J. Bacteriol. 171: 4767-4777.
Romine, M.F., Stillwell, L.C., Wong, K.K., Thurston, S.J., Sisk, E.C., Sensen, C., Gaasterland T., Fredrickson, J.K., and Saffer, J.D. 1999. Complete sequence of a 184-kilobase catabolic plasmid from *Sphingomonas aromaticivorans* F199. J. Bacteriol. 181: 1585-1602.
Rottmann, W.H., Peter, G.F., Oeller, P.W., Keller, J.A., Shen. N.F., Nagy, B.P., Taylor, L.P., Campbell, A.D., and Theologis, A. 1991. 1-aminocyclopropane-1-carboxylate synthase in tomato is encoded by a multigene family whose transcription is induced during fruit and floral senescence. J. Mol. Biol. 222: 937-961.
Saint, C.P., and Romas, P. 1996. 4-Methylphthalate catabolism in *Burkholderia* (*Pseudomonas*) *cepacia* Pc701: a gene encoding a phthalate-specific permease forms part of a novel gene cluster. Microbiology 142: 2407-2418.

Sauer, U., and Bailey, J.E. 1999. Estimation of P-to-O ratio in *Bacillus subtilis* and its influence on maximum riboflavin yield. Biotechnol. Bioeng. 64: 750-754.

Schneider, G., Kack, H., and Lindqvist, Y. 2000. The manifold of vitamin B_6 dependent enzymes. Structure Fold Des. 8: R1-6.

Schoenlein, P.V., Roa, B.B., and Winkler, M.E. 1989. Divergent transcription of *pdxB* and homology between the *pdxB* and *serA* gene products in *Escherichia coli* K-12. J. Bacteriol. 171: 6084-6092.

Shimkets, L.J. 1990. The *Myxococcus xanthus* FprA protein causes increased flavin biosynthesis in *Escherichia coli*. J. Bacteriol. 172: 24-30.

Sivasubramaniam, S., Vanniasingham, V.M., Tan, C.T. and Chua, N.H. Characterisation of HEVER, a novel stress-induced gene from *Hevea brasiliensis*. 1995. Plant Mol. Biol. 29: 173-178.

Sprenger, G.A., Schorken, U., Wiegert, T., Grolle, S., de Graaf, A.A., Taylor, S.V., Begley, T.P., Bringer-Meyer, S., and 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. USA. 94: 12857-12862.

Stackebrandt, E., Murray, R.G.E., and Trüper, H.G. 1988. Proteobacteria classis nov., a name for the phylogenetic taxon that includes the "purple bacteria and their relatives." Int. J. Syst. Bacteriol. 38 : 321-325.

Takami, H., Takaki, Y., Nakasone, K., Sakiyama, T., Maeno, G., Sasaki, R., Hirama, C., Fuji, F., and Masui, N. 1999. Genetic analysis of the chromosome of alkaliphilic *Bacillus halodurans* C-125. Extremophiles 3: 227-233.

Takiff, H.E., Baker, T., Copeland, T., Chen, S.M., and Court, D.L. 1992. Locating essential *Escherichia coli* genes by using mini-Tn10 transposons: the *pdxJ* operon. J. Bacteriol. 174: 1544-1553.

Tatusov, R.L., Galperin, M.Y., Natale, D.A., and Koonin, E.V. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28: 33-36.

Tazuya, K., Adachi, Y., Masuda, K., Yamada, K., and Kumaoka, H. 1995. Origin of the nitrogen atom of pyridoxine in *Saccharomyces cerevisiae*. Biochim. Biophys. Acta 1244: 113-116.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673-4680.

Tormo, A., Almiron, M., and Kolter, R. 1990. *surA*, an *Escherichia coli* gene essential for survival in stationary phase. J. Bacteriol. 172: 4339-4347.

van Gemen, B., Koets, H.J., Plooy, C.A., Bodlaender, J., and van Knippenberg, P.H. 1987. Characterization of the *ksgA* gene of *Escherichia coli* determining kasugamycin sensitivity. Biochimie 69: 841-848.

Wang, W., and Bechhofer, D.H. 1997. *Bacillus subtilis* RNase III gene: cloning, function of the gene in *Escherichia coli*, and construction of *Bacillus subtilis* strains with altered *rnc* loci. J. Bacteriol. 179: 7379-7385.

Woese, C.R., Kandler, O., and Wheelis, M.L. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. USA. 87: 4576-4579.

Woese, C. 1998. The universal ancestor. Proc. Natl. Acad. Sci. USA. 95: 6854-6859.

Yang, Y., Zhao, G., and Winkler, M.E. 1996. Identification of the *pdxK* gene that encodes pyridoxine (vitamin B_6) kinase in *Escherichia coli* K-12. FEMS Microbiol. Lett. 141: 89-95.

Yang, Y., Tsui, H.C., Man, T.K., and Winkler, M.E. 1998a. 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. J. Bacteriol. 180:1814-1821.

Yang, Y., Zhao, G., Man, T.K., and Winkler, M.E. 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.

Youvan, D.C., Bylina, E.J., Alberti, M., Begusch, H., and Hearst, J.E. 1984. Nucleotide and deduced polypeptide sequences of the photosynthetic reaction-center, B870 antenna, and flanking polypeptides from *R. capsulata*. Cell 37: 949-957.

Zhao, G., and Winkler, M.E. 1994. An *Escherichia coli* K-12 *tktA tktB* mutant deficient in transketolase activity requires pyridoxine (vitamin B_6) as well as the aromatic amino acids and vitamins for growth. J. Bacteriol. 176: 6134-6138.

Zhao, G., and Winkler, M.E. 1995. Kinetic limitation and cellular amount of pyridoxine (pyridoxamine) 5'-phosphate oxidase of *Escherichia coli* K-12. J. Bacteriol. 177: 883-891.

Zhao, G., and Winkler, M.E. 1996. 4-Phospho-hydroxy-L-threonine is an obligatory intermediate in pyridoxal 5'-phosphate coenzyme biosynthesis in *Escherichia coli* K-12. FEMS Microbiol. Lett. 135: 275-280.