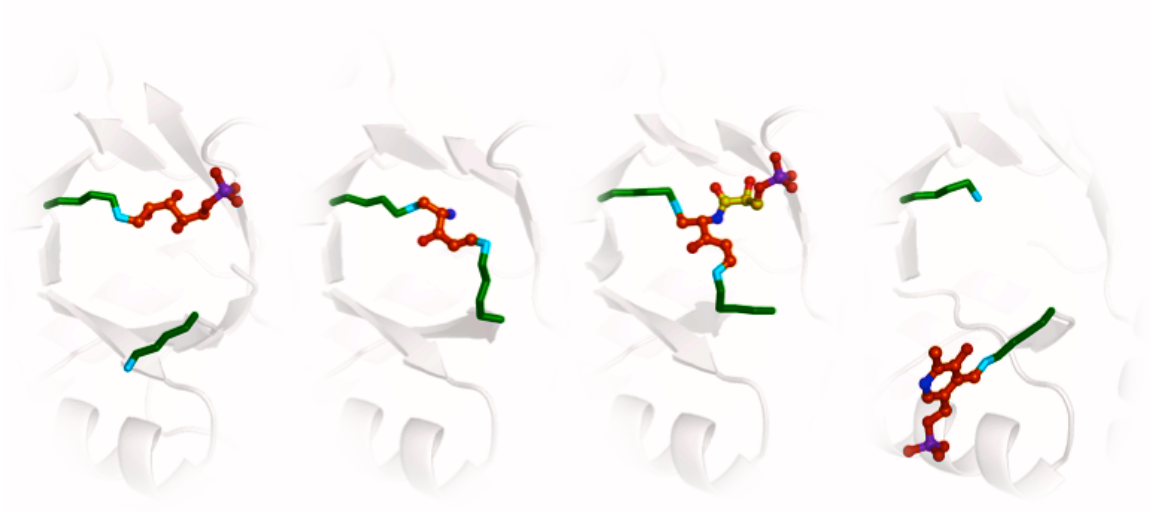
Lysine relay mechanism coordinates intermediate transfer in vitamin B6 biosynthesis

Matthew J. Rodrigues[1,2], Volker Windeisen[1,3], Yang Zhang[4], Gabriela Guédez[3], Stefan Weber[3], Marco Strohmeier[3], Jeremiah W. Hanes[4,5], Antoine Royant[6,7], Gwyndaf Evans[2], Irmgard Sinning[3], Steven E. Ealick[4], Tadhg P. Begley[8], Ivo Tews[1,3,*]

EDITORIAL SUMMARY

AOP: Crystallographic snapshots of its catalytic cycle illuminate the mechanism by which the enzyme Pdx1 shuttles intermediates between lysine residues in its two active sites during the biosynthesis of pyridoxal 5′-phosphate.



Lysine relay mechanism coordinates intermediate transfer in vitamin B6 biosynthesis

Matthew J. Rodrigues[1,2], Volker Windeisen[1,3], Yang Zhang[4], Gabriela Guédez[3], Stefan Weber[3], Marco Strohmeier[3], Jeremiah W. Hanes[4,5], Antoine Royant[6,7], Gwyndaf Evans[2], Irmgard Sinning[3], Steven E. Ealick[4], Tadhg P. Begley[8], Ivo Tews[1,3,*]

[1]Centre for Biological Sciences, University of Southampton, Southampton, SO17 1BJ, UK

[2]Diamond Light Source, Harwell Science and Innovation Campus, Didcot OX11 0DE, UK

[3]Heidelberg University Biochemistry Center (BZH), Im Neuenheimer Feld 328, 69120 Heidelberg, Germany

[4]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853, USA

[5]Pacific Biosciences, 1380 Willow Rd., Menlo Park, CA 94025, USA[6]Institut de Biologie Structurale, Université Grenoble Alpes, CNRS, CEA, 38044 Grenoble, France

[7]European Synchrotron Radiation Facility, 38043 Grenoble, France

[8]Department of Chemistry, Texas A&M University, College Station, TX 77843, USA

*Corresponding author

Substrate channeling has emerged as a common mechanism for enzymatic intermediate transfer. A conspicuous gap in knowledge concerns the use of covalent lysine imines in transfer of carbonyl group-containing intermediates, despite their wide use in enzymatic catalysis. Here we show how imine chemistry operates in transfer of covalent intermediates in pyridoxal 5'-phosphate biosynthesis by the Arabidopsis enzyme Pdx1. An initial ribose 5-phosphate lysine imine converts to the chromophoric I_{320} intermediate, simultaneously bound to two lysine residues and partially vacating the active site, creating space for glyceraldehyde 3-phosphate to bind. Crystal structures show how substrate binding, catalysis and shuttling are coupled to conformational changes around strand β6 of the Pdx1 $(\beta\alpha)_8$-barrel. The dual specificity active site and imine relay mechanism for migration of carbonyl intermediates provide elegant solutions to the challenge of coordinating a complex sequence of reactions, following a path of over 20 Å between substrate and product binding sites.

Pyridoxal 5'-phosphate (PLP) is an active form of vitamin B_6, which functions as an enzyme cofactor and antioxidant[1,2]. More than 140 different PLP-dependent enzymes catalyze reactions in prokaryotic amino acid and carbohydrate metabolism, and an estimated 1.5% of the genes in microbial genomes encode PLP-dependent enzymes[3]. The PLP synthase complex, which consists of the enzymes Pdx1 and Pdx2, is conserved in all domains of life[4,5]. In striking contrast, the related *Escherichia coli* pathway, used only by a small number of bacteria, requires six enzymes to achieve the same biosynthesis[4,6].

The Pdx1 enzyme requires three substrates to catalyze the formation of PLP[7-9]. Ribose 5-phosphate (R5P) reacts with ammonia, which is produced by a separate Pdx2 glutaminase domain to produce the chromophoric I_{320} intermediate with an absorbance maximum at 320 nm (ref. 10). I_{320} formation is followed by addition of the third substrate, glyceraldehyde 3-phosphate (G3P) (Fig. 1a)[9]. Pdx1 has a $(\beta\alpha)_8$-barrel fold in a dodecameric assembly; each subunit contains two phosphate binding sites, named P1 and P2, defining the active site architecture[11,12] (Fig. 1b). P1 and P2 are separated by 21 Å (phosphorus to phosphorus) and are known to bind the phosphate groups of R5P and PLP respectively, but the mechanism by which Pdx1 transfers intermediates between the sites is unknown.

The crystallographic studies described here reveal five Pdx1 intermediate structures and have been used to propose the mechanism of intermediate transfer chemistry not evident from previous characterization of the I_{320} structure by MS analysis, phosphine trapping, and NMR analysis[9,13]. The I_{320} intermediate is simultaneously bound to two lysine residues, allowing the transfer between catalytic centres, and we describe structural changes to support intermediate formation and transfer.

RESULTS

The I_{320} intermediate is central to PLP biosynthesis

Various bacterial and eukaryotic Pdx1 proteins were screened to identify the *Arabidopsis* Pdx1 (ref. 14) crystal system that provided the Pdx1-intermediate complexes reported here (Fig. 2, Supplementary Results, Supplementary Fig. 1 and Table 1). Although other crystal systems such as *Thermus thermophilus* yielded higher resolution diffraction (1.6 Å), they were not robust enough to survive crystal soaking experiments.

PLP formation is initiated with formation of the Pdx1-R5P complex, observed here at 1.9 Å resolution (Fig. 2a) and also previously observed for *Plasmodium berghei* and *Geobacillus stearothermophilus* Pdx1 (ref. 15,16). A covalent imine forms between the ε-NH_2 of Lys98 and the R5P C1, with the R5P phosphate bound in P1 (ref. 10,15-17) (Fig. 2a). The proximity of the ε-NH_2 group of Lys98 to the side chains of Asp119 and Ser121 (Supplementary Fig. 2a and 2b) of the totally conserved DESE sequence motif in PLP synthases suggests that Asp119 or Ser121 may play a role in Schiff base formation, but as yet we were unable to clearly assign the catalytic base catalyzing this step.

Transient association with Pdx1 activates Pdx2 to catalyze glutamine hydrolysis, releasing ammonia as a nitrogen source for PLP biosynthesis[8,18,19]. The ammonia produced by Pdx2 passes through a hydrophobic tunnel in the center of the Pdx1 $(\beta\alpha)_8$-barrel to the Pdx1 P1 active site; channeling of intermediates prevents their diffusion into the bulk solvent and is common in glutamine amidotransferase enzymes such as PLP synthase [11,15,20]. Ammonia reacts with Pdx1-R5P to form the chromophoric I_{320} intermediate[10,21]. Since the product, PLP, binds in the P2 site[22], we hypothesized that

formation of the I_{320} intermediate was central to the migration of reaction intermediates from Lys98 (P1) to Lys166 (P2).

Soaking Pdx1-R5P crystals with ammonium chloride generated Pdx1-I_{320} (Fig. 2c). While the addition of ammonium salts decouples Pdx1-I_{320} formation from Pdx2 dependent glutamine hydrolysis[10], the change in osmotic pressure caused many crystals to become disordered. More than 1000 crystals were tested to optimize the protocol used to provide both high resolution diffraction and homogeneous accumulation of Pdx1 in the I_{320} state. Online UV-Vis micro-spectrophotometry[23] was performed to ensure that crystals used in diffraction experiments contained Pdx1 predominantly in the I_{320} state (Fig. 3a).

Pdx1-I_{320} crystals typically diffracted to 2.3 – 2.0 Å resolution and showed continuous electron density between Lys98 and Lys166, and loss of the phosphate group (Fig 2c, 3b). Interpretation of the highest resolution dataset at 1.7 Å resolution was aided by the earlier MS/NMR data that show that I_{320} retains all five carbons of the R5P substrate[10], with C1 covalently bound to a nitrogen atom[13], observed here to be the ε-NH_2 of Lys98 (Fig. 3b).

Comparison of the ^{13}C NMR spectra of I_{320} reconstituted *in vitro* using ^{14}N and ^{15}N ammonium chloride showed that the nitrogen incorporation occurred at the C2 atom[13], which allowed assignment of nitrogen and oxygen atoms bound to C2 and C3 of the intermediate, respectively (Fig. 2c and 3b). The C3 bound oxygen atom of I_{320} is the sole oxygen that remains from the R5P substrate[13]. The C5 attachment of I_{320} to a nitrogen atom was previously interpreted as being caused by I_{320} decomposition under the

denaturing and acidic conditions in the NMR experiment[13]; however, the crystallographic analysis now resolves the nature of the C5-N interaction, observed here to be a covalent bond between C5 and the ε-NH_2 of Lys166. Thus, the intermediate is not released from Lys98 prior to I_{320} formation as previously thought[13], but is simultaneously bound to two lysine residues.

A network of polar interactions between side chains of Asp119 and Glu122 of the DESE sequence motif and Arg164 (all residues totally conserved in PLP synthases) is in close proximity to the intermediates in the P1 site. Previous studies demonstrate that replacing Arg164 in *Geobacillus stearothermophilus* Pdx1 with alanine causes a 95% reduction in the rate of I_{320} formation[16], and Arg164 is suitably positioned to activate the ε-NH_2 group of Lys166 for Schiff base formation in the Pdx1-I_{320} structure (Supplementary Fig. 2c). This conserved network of polar interactions is reminiscent to acetoacetate decarboxylase, a well-studied example of the electrostatic perturbation model for enzyme activation[24].

Several examples of crystallographic structures with imine bound cofactors are known to suffer from site-specific radiation damage[25,26]. Radiation damage may lead to artifacts in electron density maps not representative of the protein structure prior to X-ray exposure[27]. UV-Vis spectra were collected of Pdx1-I_{320} *in crystallo* during X-ray irradiation. The spectra of the crystals changed in response to irradiation, with 20% of the changes occurring in the first 245 kGy. While their occupancy is below 20%, the species generated by X-ray irradiation are not expected to have a significant effect on the interpretation of electron density maps[28]. We constructed a 2.2 Å resolution multi-crystal dataset, using a protocol similar to that first described for investigation of site-specific

radiation damage in horseradish peroxidase[29]. The X-ray dose absorbed during data collection was calculated using RADDOSE-3D[30]; BLEND was used to merge the diffraction data collected from each crystal below the dose threshold of 245 kGy into a single dataset[31]. The resulting structure confirmed the nature of the bridging structure of I_{320} (Supplementary Fig. 3).

Conformational changes supporting intermediate transfer

Formation of Pdx1-I_{320} requires conformational changes in strand β6 of the Pdx1 $(\beta\alpha)_8$-barrel resulting in reorientation of Lys166 towards P1; this residue points towards the P2 site in the Pdx1-R5P structure (Fig. 3b, Supplementary Fig. 4a). We investigated the Pdx1 K166R variant[10,16] to understand better whether the structural transition is linked to ammonia incorporation. Sequential exposure of Pdx1 K166R to R5P and ammonium chloride results in a complex termed Pdx1-preI_{320} (Fig. 2b). Like Lys166 in the Pdx1-I_{320} complex, Arg166 reorients towards P1, and a peptide flip between Arg166 and Gly167 causes reorientation of the Thr165 and Arg166 side chains as Pdx1 transitions from the Pdx1-R5P state to the Pdx1-preI_{320} and Pdx1-I_{320} complexes (Fig. 3b, Supplementary Fig. 4b and 4c).

An earlier analysis predicted such changes by showing that the β6 strand lacks hydrogen bond stabilization from adjacent β-strands allowing for additional flexibility in this region[12]. The recent structural analysis of *G. stearothermophilus* Pdx1 also identified that Lys166 adopts these two conformations[16].

These conformational changes are likely linked to the passage of ammonia through the transient hydrophobic tunnel in the $(\beta\alpha)_8$-barrel[11,15]. Notably, the absolutely conserved

methionine residue, Met162, in this tunnel is located on strand β6 together with the catalytic residue Lys166[15]. In the previous study, the equivalent residue to Met162 in *Plasmodium*, Met148, was exchanged for leucine, causing an increase in the catalytic rate of Pdx2-dependent PLP biosynthesis with no effect on the rate of NH_4Cl-dependent catalysis, suggesting a coupling of Pdx1 synthase and Pdx2 glutaminase activities involving Met148. We propose that the disruption of Met162 on the β6 strand by the passage of ammonia couples glutamine hydrolysis to the conformational changes required for I_{320} formation (Supplementary Fig 4d).

Elimination of the R5P phosphate group bound in the P1 site immediately precedes the formation of the I_{320} intermediate[9]. This phosphate binding site is located in the typical position for the $(\beta\alpha)_8$-fold at the C-terminal face of the β-barrel and makes additional contacts with the β6–α6 loop that carries Lys166[11,12,32]. The Pdx1-preI_{320} structure highlights the role of this phosphate group as a rigid anchor that, together with the Lys98 imine, ensures correct positioning of the intermediate for ammonia incorporation and covalent attachment to Lys166 in the wild type protein (Fig. 2b). Although incorporation of the ammonia nitrogen atom occurs in the Pdx1 K166R variant, the different functionality of the Arg166 guanidino group prevents the catalysis of phosphate elimination[10,16]. The retention of the phosphate group explains the different UV-Vis absorption spectrum of the pre-I_{320} complex that has a lower absorbance compared to Pdx1-I_{320}[10].

Dual specificity binding site for pentose and triose

The structural rearrangements associated with I_{320} formation result in the displacement of carbon atoms C3, C4 and C5 of the intermediate away from the P1 phosphate binding site (Fig. 3b), and the elimination reaction allows the R5P phosphate group to diffuse out of the P1 site. These actions create space for binding of G3P in the P1 site. In the Pdx1-I_{320}-G3P complex, G3P is covalently bound to the C2 nitrogen atom of I_{320}, with the G3P phosphate in the P1 binding site (Fig. 2d, Fig 3c). Surprisingly, the intermediate is covalently attached to both Lys98 and Lys166, forming a horseshoe-like structure.

In solution, addition of G3P to Pdx1-I_{320} allows the reaction to proceed to the PLP product state; however, this did not occur *in crystallo*. The total loss of diffraction upon prolonged G3P exposure provides strong evidence for extensive structural changes accompanying the final steps of PLP formation. The C-terminus of Pdx1 is disordered in crystal structures and is known to participate in the final stages of PLP biosynthesis[32-34]. The packing of Pdx1 in the crystalline state may prevent the C-terminus from adopting the conformation required for Pdx1 to complete catalysis of PLP biosynthesis.

The Pdx1-PLP state can be observed by directly soaking Pdx1 crystals with PLP. The 1.6 Å resolution electron density reveals an imine attachment of PLP to Lys166 by the carbon atom that initially was C5 of R5P (Fig. 2e, Fig 3d). A noncovalent yeast Pdx1-PLP complex was previously reported[22]; however, examination of the deposited structure (PDB ID 3O05) shows conclusively that the PLP is covalently bound. The covalent nature of this complex is supported by UV-Vis spectroscopy of a Pdx1-PLP crystal that

shows the characteristic absorption maximum at 408 nm rather than 388 nm as is typical for free PLP (Supplementary Fig. 5)[13].

The structural data presented here have been combined with the published biochemical biophysical data from Pdx1 into a mechanistic proposal for catalysis and intermediate transfer (Fig. 4, Supplementary Fig. 6). The schematic shows the central role of the I_{320} intermediate, linking between side-chains Lys98 and Lys166, for transfer of the reaction between substrate and product binding sites.

DISCUSSION

Pdx1 is the first enzyme identified to use a dual transamine intermediate to transfer a reaction across an extended active site within a $(\beta\alpha)_8$-barrel enzyme[35], in addition to coupling delivery of ammonia from Pdx2 to Pdx1 to I_{320} formation. Pdx1 enzymes studied thus far span bacterial species[11,12,16,17], unicellular eukaryotes[15,19,22,36,37] and plants[14]. The mechanism of I_{320} formation and intermediate transfer is assumed to be conserved given the high overall conservation of the Pdx1 enzyme[5], the consistency of findings across all kingdoms of life, and in particular the conservation of Lys98 and Lys166. Central to the function of Pdx1 is the bridging structure of the I_{320} intermediate between these conserved lysine residues (Fig 2c, Fig. 3b).

The structures reveal a mechanism for how Pdx1 uses covalent tethers to prevent loss of intermediates to surrounding solvent, maintains a high local concentration of substrate,

and protects the reactive I_{320} species. The use of covalent tethers is reminiscent of the protein-bound phosphopantothenoyl thioesters that are used to transfer intermediates between active sites in the assembly line enzymology of fatty acid synthases[38], polyketide synthases[39] and non-ribosomal polypeptide synthases[40]. Pdx1 uses a similar transfer strategy to those found in the glycine cleavage system[41] and the classical pyruvate dehydrogenase complex[42], where lipoic acid is used to chaperone intermediates between active sites. In common with Pdx1, these enzymes use covalent tethers to prevent the loss of substrates or intermediates to surrounding solvent and maintain the local concentration of substrate. In contrast to the previous examples, Pdx1 transfers covalent intermediates within a single catalytic domain. The intricate relay mechanism displayed by the Pdx1 subunit of PLP synthase allows the enzyme to maintain precise control of the complex reaction performed across multiple active sites.

A conspicuous absence in Nature's repertoire is the use of lysine imines to channel carbonyl group-containing intermediates; this is surprising because lysine imines are used extensively in enzymatic catalysis and trans-imination reactions are facile. In this study we described PLP synthase as the first example of an enzyme that uses two lysine residues to shuttle intermediates between two active sites. Lys98 anchors intermediates in the P1 site for the first phase of PLP assembly. Lys166 then adds and facilitates intermediate transfer to the P2 site where completion of PLP formation occurs. This transfer strategy enables a single enzyme to catalyze the complex assembly process required for the formation of the pyridine ring of PLP and stands in contrast to the use of six enzymes in *E. coli* to achieve the same biosynthetic goal.

Accession codes. Coordinates and structure factors have been deposited in the Protein Data Bank under the accession codes 5LNS (Pdx1-R5P), 5LNT (Pdx1 K166R-preI_{320}), 5LNU (Pdx1-I_{320}-G3P), 5LNV (Pdx1-I_{320} multi-crystal), 5LNW (Pdx1-I_{320}-G3P), 5LNR (Pdx1-PLP).

Online data: Methods, along with any additional Supplementary Results are available in the online version of the paper; references unique to these sections appear only in the online paper.

Acknowledgements. We thank W.J. Anderson, H. Clarke, C. Phippen, Southampton, A. Wessling, N. Kwak, Heidelberg, for their experimental contributions, T. Fitzpatrick (Geneva, Switzerland) for the generous gift of the expression plasmid for Arabidopsis Pdx1.3; S. Findlow and C. Holes at the Macromolecular Crystallisation Facility, Centre for Biological Sciences, and P. Horton and S. Coles at the Southampton Diffraction Centre, Chemistry, both University of Southampton; J. Kopp and C. Siegmann from the crystallisation platform of the Cluster of Excellence CellNetworks, Heidelberg; staff at the Diamond Light Source and the European Synchrotron Radiation Facility for access and excellent user support without which this project would not have been possible; P. Carpentier, M. Weik, G. Gotthard and D. von Stetten at ESRF for support during data collection and online spectroscopy and D. Flot and G. Leonard for flexible access to the ESRF beamlines; A. Douangamath, P. Aller, R. Owen and M. Walsh for support at Diamond beamlines; B. Kappes, Erlangen, and P. Macheroux, Graz, for critical discussions and L. Kinsland for assistance in preparation of the manuscript. This work was supported in parts by grants by the European Commission (VITBIOMAL-012158) and by the DFG (TE368) to I.T., by NIH (DK44083) and by the Robert A. Welch

Foundation (A-0034) to T.P.B., by ESRF Mx1461, Mx1732, RADDAM and Diamond Light Source Mx8891 to I.T.. M.J.R. was supported by a joint studentship between Diamond Light Source and the University of Southampton.

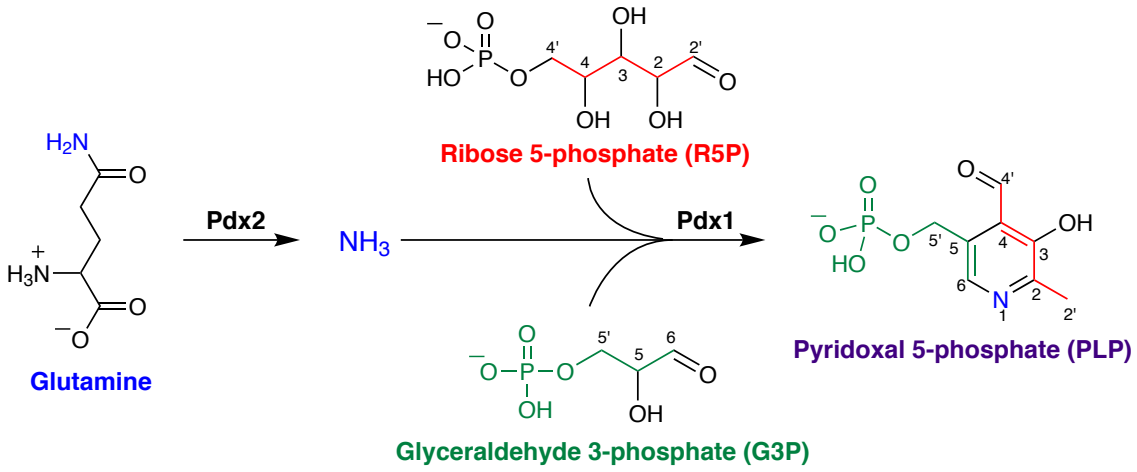
Author contributions. M.J.R., V.W., G.G., S.W., M.S. and J.W.H. performed protein expression, purification, and enzymatic essays. M.J.R., V.W., G.G., S.W., M.S. and I.T. performed crystallization and X-ray diffraction experiments. M.J.R, V.W., A.R. and I.T. performed online spectroscopy experiments. M.J.R., V.W., Y.Z., I.S., S.E.E., T.P.B. and I.T. performed crystallographic analysis and data deposition. M.J.R., A.R., G.E., T.P.B. and I.T. performed spectroscopic data analysis. M.J.R., G.E., I.S., S.E.E., T.P.B. and I.T. wrote the paper. M.J.R. and V.W. contributed equally to this work.

Competing financial interests. The authors declare no competing financial interests.

Additional information. Correspondence should be addressed to I.T. (ivo.tews@soton.ac.uk).

FIGURE CAPTIONS

a



b

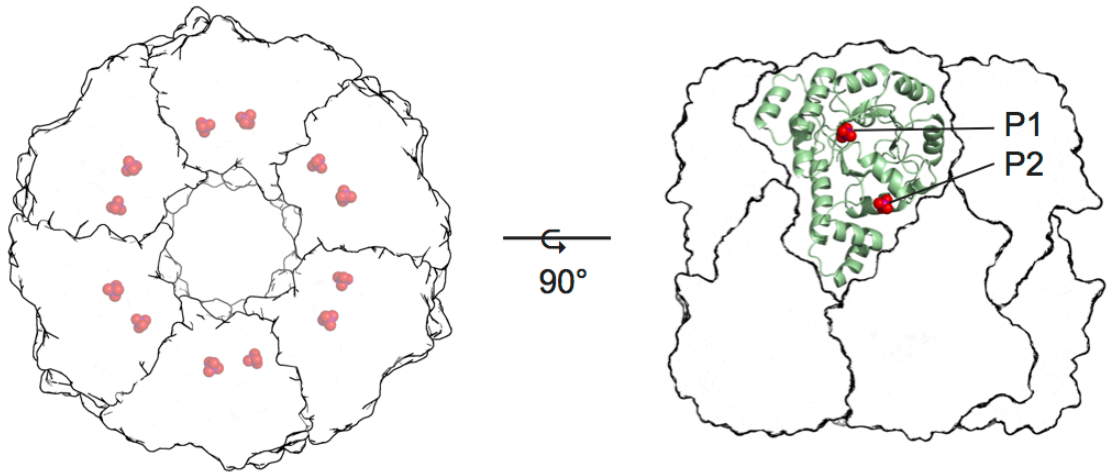
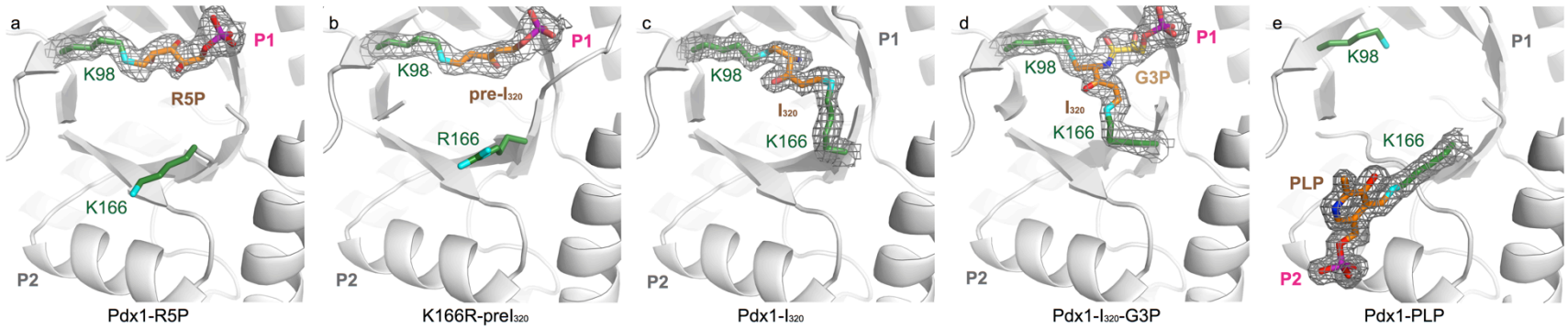


Figure 1 | The reaction catalyzed by the Pdx1 and Pdx2 subunits of the PLP synthase complex and the structure of Pdx1. a, PLP synthase catalyzes the complex condensation reaction between ribose 5-phosphate, glyceraldehyde 3-phosphate and ammonia. **b,** Overall structure of the Pdx1 core. Pdx1 forms a dodecamer from two

interlocked hexameric rings. This core complex is shown in two orientations, with subunit boundaries and the two phosphate-binding sites for each protein monomer indicated. The phosphate binding sites P1 and P2 are required for substrate and product binding, respectively, and are separated by 21 Å (phosphorus to phosphorus). A cartoon representation for one monomer illustrates the $(\beta\alpha)_8$-barrel fold of Pdx1.

Figure 2 | Crystallographic structures of five covalent intermediates in PLP biosynthesis. Pdx1 is shown in cartoon representation, catalytic lysine side chains and intermediate atoms are shown in stick representation with carbon atoms in green (for Lys98 and Lys166) and orange (for intermediates), nitrogen atoms in blue, oxygen atoms in red, and phosphorus atoms in purple. $2F_o$-F_c electron density maps are shown at 1σ for the complexes. **a,** Binding of R5P (first substrate) uses P1 and occurs by covalent attachment through Schiff base formation with Lys98. **b, c,** Addition of ammonia (second substrate) leads to the formation of an intermediate in P1 in the Lys166Arg mutant (**b**), which converts to the I_{320} species in wild type enzyme through formation of a second Schiff base with Lys166 (**c**). **d,** Incorporation of G3P (third substrate) leads to a covalent complex with I_{320}, with the G3P phosphate bound in the P1 site. **e,** The PLP complex shows PLP covalently bound to the enzyme through Schiff base formation with Lys166, with its phosphate bound in the P2 site.


a
P1
K98
R5P
K166
P2
Pdx1-R5P
b
P1
K98
pre-I320
R166
P2
K166R-preI320
c
P1
K98
I320
K166
P2
Pdx1-I320
d
P1
K98
G3P
I320
K166
P2
Pdx1-I320-G3P
e
K98
P1
K166
PLP
P2
Pdx1-PLP

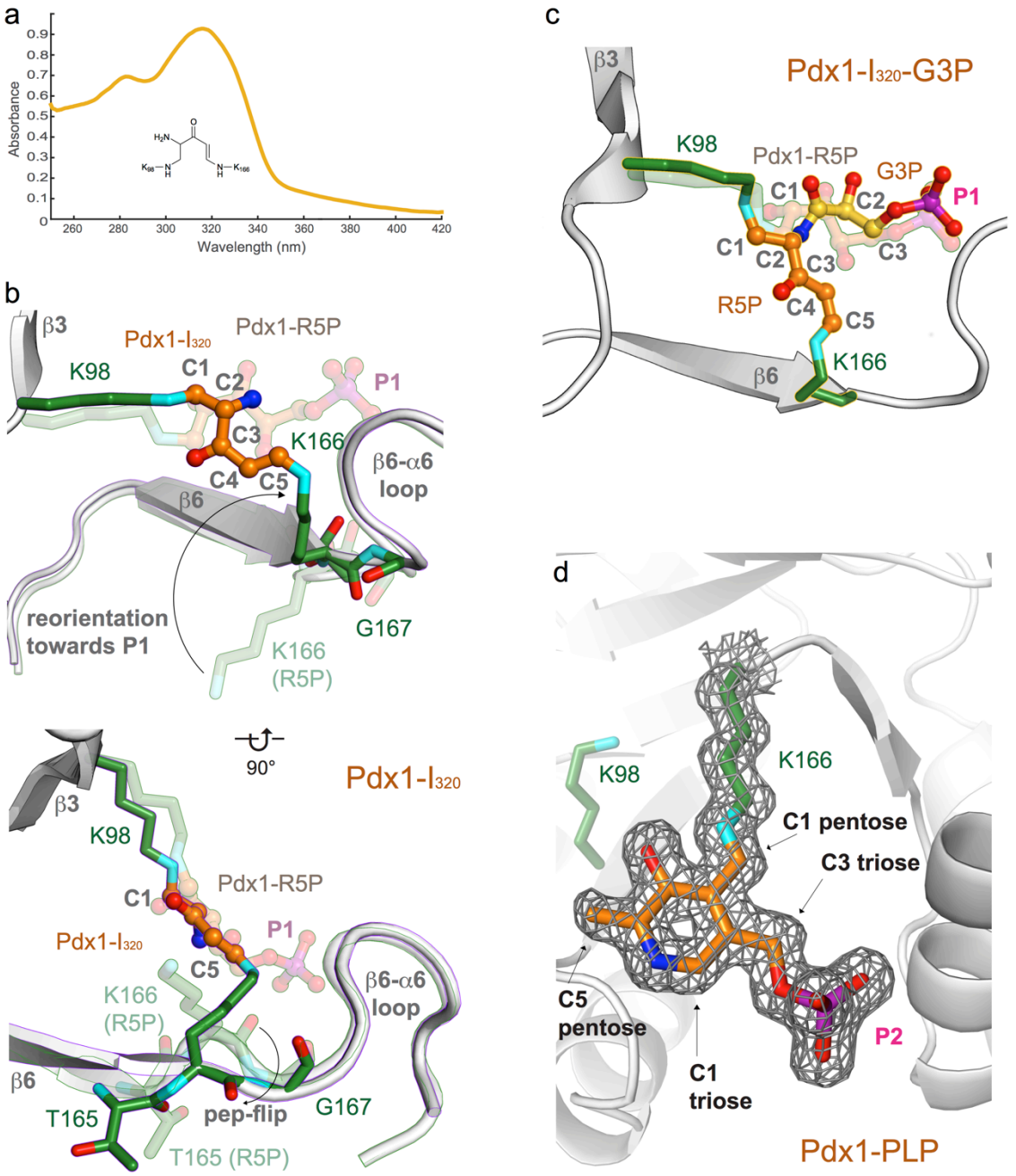


Figure 3 | The dual specificity binding site P1 and the product binding site P2. a, The UV-Vis spectrum of a Pdx1-I_{320} crystal shows an absorption maximum at 280 nm, corresponding to protein, and near 320 nm for the I_{320} intermediate. **b**, The intermediate I_{320} is covalently bound to both Lys98 and Lys166, shown in two views rotated by 90°.

The overlay with the R5P complex (faded) shows the different positioning of Lys166 and adjacent residues Thr165 and Gly167. Selected amino acid atoms and the I_{320} atoms are shown in stick representation with carbon atoms in green (for amino acids) and orange (for intermediates), nitrogen atoms in blue, oxygen atoms in red, and phosphorus atoms in purple. Arrows indicate the different conformations of the Lys166 side chains and the Lys166-Gly167 peptide between Pdx1-I_{320} and Pdx1-R5P complexes. **c**, The I_{320}-G3P complex horseshoe-like intermediate. The intermediate is covalently attached to both Lys98 and Lys166. Formation of the G3P binding site requires prior formation of the I_{320} adduct to dissociate the C3-C5 atoms from the binding site, which allows the G3P phosphate to bind in the P1 site. The overlay with the R5P complex (faded) shows where the ribose and triose portions match. Atoms colored as described for **Figure 3b**. **d,** The PLP adduct is covalently bound by a Schiff base to Lys166. $2F_o$-F_c electron density map of the refined complex shown at 1 σ. Atoms colored as described for **Figure 3b**.

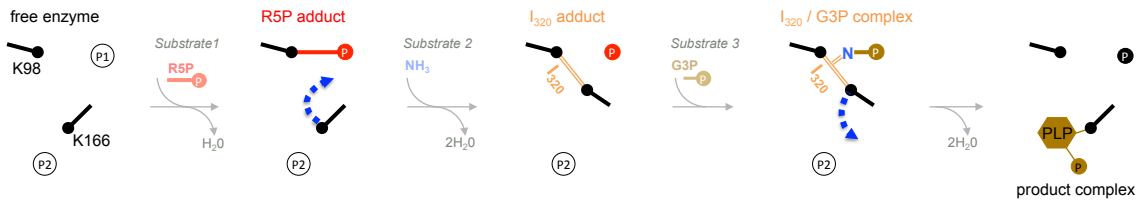


Figure 4 | The central role of I_{320} intermediate transfer in vitamin B6 biosynthesis. a, The proposed Pdx1 reaction mechanism is shown as a schematic. Lys166 points towards P2 in the free enzyme. Structural transitions support reorientation of the side chain of Lys166 towards the P1 site at the time of ammonia incorporation, leading to formation of I_{320}, creation of the G3P binding site, and the subsequent formation of the covalent I_{320}-G3P complex. The reversion of the structural transitions around Lys166 lead to release of

the I_{320}-G3P intermediate from covalent attachment to Lys98 and observation of covalently bound PLP in the P2 site.

REFERENCES

1 Ehrenshaft, M., Bilski, P., Li, M. Y., Chignell, C. F. & Daub, M. E. A Highly Conserved Sequence is a Novel Gene Involved in De Novo Vitamin B6 Synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 9374-9378 (1999).

2 Mooney, S. & Hellmann, H. Vitamin B6: Killing two birds with one stone? *Phytochemistry* **71**, 495-501 (2010).

3 Percudani, R. & Peracchi, A. The B6 database: a tool for the description and classification of vitamin B6-dependent enzymatic activities and of the corresponding protein families. *BMC bioinformatics* **10**, 1-8 (2009).

4 Mittenhuber, G. Phylogenetic analyses and comparative genomics of vitamin B6 (pyridoxine) and pyridoxal phosphate biosynthesis pathways. *Journal of molecular microbiology and biotechnology* **3**, 1-21 (2001).

5 Das, S. *et al.* Biology's new Rosetta stone. *Nature* **385**, 29-30 (1997).

6 Fitzpatrick, T. B. *et al.* Two independent routes of de novo vitamin B6 biosynthesis: not that different after all. *The Biochemical journal* **407**, 1-13 (2007).

7 Raschle, T., Amrhein, N. & Fitzpatrick, T. B. On the two components of pyridoxal 5'-phosphate synthase from Bacillus subtilis. *The Journal of biological chemistry* **280**, 32291-32300 (2005).

8 Burns, K. E., Xiang, Y., Kinsland, C. L., Mc Lafferty, F. W. & Begley, P. T. Reconstitution and Biochemical Characterisation of Pyridoxal-5'-Phosphate Biosynthetic Pathway. *Journal of the American Chemical Society* **127**, 3682-3683 (2005).

9 Hanes, J. W. *et al.* Mechanistic studies on pyridoxal phosphate synthase: The reaction pathway leading to a chromophoric intermediate. *Journal of the American Chemical Society* **130**, 3043-3052 (2008).

10 Raschle, T. *et al.* Reaction Mechanism of Pyridoxal 5'-Phosphate Synthase: Detection of an Enzyme-Bound Chromophoric Intermediate. *The Journal of Biochemistry* **282**, 6098-6105 (2007).

11 Strohmeier, M. *et al.* Structure of a Bacterial Pyridoxal 5'-Phosphate Synthase Complex. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 19824-19829 (2006).

12 Zhu, J., Beurgner, J. W., Harmst, E., Belitsky, B. R. & Smith, J. L. A New Arrangement of $(\beta/\alpha)_8$ Barrels in the Synthase Subunit of PLP Synthase. *The Journal of Biochemistry* **280**, 27914-27923 (2005).

13 Hanes, J. W., Keresztes, I. & Begley, T. P. 13C NMR snapshots of the complex reaction coordinate of pyridoxal phosphate synthase. *Nature chemical biology* **4**, 425-430 (2008).

14 Tambasco-Studart, M. *et al.* Vitamin B6 biosynthesis in higher plants. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 13687-13692 (2005).
15 Guedez, G. *et al.* Assembly of the Eukaryotic PLP-Synthase Complex from Plasmodium and Activation of the Pdx1 Enzyme. *Structure* **20**, 172-184 (2012).
16 Smith, A. M., Brown, W. C., Harms, E. & Smith, J. L. Crystal Structures Capture Three States in the Catalytic Cycle of a Pyridoxal Phosphate (PLP) Synthase. *The Journal of biological chemistry* **290**, 5226-5239 (2015).
17 Zein, F. *et al.* Structure of the PLP Synthase Holoenzyme from Thermatoga Maritima. *Biochemistry* **45**, 14609–14620 (2006).
18 Wallner, S., Neuwirth, M., Flicker, K., Tews, I. & Macheroux, P. Dissection of Contributions from Invariant Amino Acids to Complex Formation and Catalysis in the Heteromeric Pyridoxal 5-Phosphate Synthase Complex from Bacillus subtilis. *Biochemistry* **48**, 1928-1935 (2009).
19 Gengenbacher, M. *et al.* Vitamin B6 biosynthesis by the malaria parasite Plasmodium falciparum: biochemical and structural insights. *The Journal of biological chemistry* **281**, 3633-3641 (2006).
20 Raushel, F. M., Thoden, J. B. & Holden, H. M. Enzymes with Molecular Tunnels. *Accounts of Chemical Research* **36**, 539-548 (2003).
21 Hanes, J. W., Keresztes, I. & Begley, T. P. Trapping of a chromophoric intermediate in the Pdx1-catalyzed biosynthesis of pyridoxal 5'-phosphate. *Angewandte Chemie* **47**, 2102-2105 (2008).
22 Zhang, X. *et al.* Structural Insights into the Catalytic Mechanism of the Yeast Pyridoxal 5-Phosphate Synthase Snz1. *The Biochemical journal* **432**, 445-450 (2010).
23 von Stetten, D. *et al.* In crystallo optical spectroscopy (icOS) as a complementary tool on the macromolecular crystallography beamlines of the ESRF. *Acta crystallographica. Section D, Biological crystallography* **71**, 15-26 (2015).
24 Ho, M.-C., Menetret, J.-F., Tsuruta, H. & Allen, K. N. The origin of the electrostatic perturbation in acetoacetate decarboxylase. *Nature* **459**, 393-397 (2009).
25 Borshchevskiy, V. *et al.* Low-dose X-ray radiation induces structural alterations in proteins. *Acta crystallographica. Section D, Biological crystallography* **70**, 2675-2685 (2014).
26 Dubnovitsky, A. P., Ravelli, R. B., Popov, A. N. & Papageorgiou, A. C. Strain relief at the active site of phosphoserine aminotransferase induced by radiation damage. *Protein science : a publication of the Protein Society* **14**, 1498-1507 (2005).
27 Weik, M. *et al.* Specific Chemical and Structural Damage to Proteins Produced by Synchrotron Radiation *Proceedings of the National Academy of Sciences of the United States of America* **97**, 623-628 (2000).
28 Pearson, A. R. & Owen, R. L. Combining X-ray crystallography and Single-Crystal Spectroscopy to Probe Enzyme Mechanisms. *Biochemical Society Transactions* **37**, 378-381 (2009).

29 Berglund, G. I. *et al.* The Catalytic Pathway of Horseradish Peroxidase at High Resolution. *Nature* **417**, 463-468 (2002).
30 Zeldin, O. B., Gerstel, M. & Garman, E. F. Optimizing the Spatial Distribution of Dose in X-Ray Macromolecular Crystallography. *Journal of Synchrotron Radiation* **20**, 49-57 (2013).
31 Foadi, J. *et al.* Clustering Procedures for the Optimal Selection of Data Sets from Multiple Crystals in Macromolecular Crystallography. *Acta Crystallographica Section D* **69**, 1617-1632 (2013).
32 Derrer, B. *et al.* Defining the structural requirements for ribose 5-phosphate-binding and intersubunit cross-talk of the malarial pyridoxal 5-phosphate synthase. *FEBS Letters* **584**, 4169-4174 (2010).
33 Raschle, T. *et al.* Intersubunit Cross-Talk in Pyridoxal 5'-Phosphate Synthase, Coordinated by the C Terminus of the Synthase Subunit. *The Journal of biological chemistry* **284**, 7706-7718 (2009).
34 Moccand, C., Kaufmann, M. & Fitzpatrick, T. B. It takes two to tango: defining an essential second active site in pyridoxal 5'-phosphate synthase. *PloS one* **6**, e16042 (2011).
35 Nagano, N., Orengo, C. A. & Thornton, J. M. One Fold with Many Functions: The Evolutionary Relationships between TIM Barrel Families Based on their Sequences, Structures and Functions. *Journal of Molecular Biology* **321**, 741-765 (2002).
36 Reeksting, S. B. *et al.* Exploring inhibition of Pdx1, a component of the PLP synthase complex of the human malaria parasite Plasmodium falciparum. *The Biochemical journal* **449**, 175-187 (2013).
37 Neuwirth, M. *et al.* X-Ray Crystal Strucuture of Saccharomyces cerevisiae Pdx1 Provides Insights into the Oligomeric Nature of PLP Synthases. *FEBS Letters* **583**, 2179-2186 (2009).
38 Jenni, S. *et al.* Structure of fungal fatty acid synthase and implications for iterative substrate shuttling. *Science* **316**, 254-261 (2007).
39 Wu, N., Tsuji, S. Y., Cane, D. E. & Kholsa, C. Assessing the Balance between Protein-Protein Interactions and Enzyme-Substrate Interactions in the Channeling of Intermediates between Polyketide Synthase Modules. *Journal of the American Chemical Society* **123**, 6466-6475 (2001).
40 Tanovic, A., Samel, S. A., Essen, L. & Marahiel, M. A. Crystal Structure of the Termination Module of a Nonribosomal Peptide Synthetase. *Science* **321**, 659-663 (2008).
41 Pares, S., Cohen-Addad, C., Sieker, L., Neuburger, M. & Douce, R. X-ray crystal structure determination at 2.6-A resolution of a lipoate-containing protein: The H-protein of the glycine decarboxylase complex from pea leaves. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 4850-4853 (1994).
42 Reed, L. J. Multienzyme Complexes. *Accounts of Chemical Research* **7**, 40-46 (1974).

ONLINE METHODS

Molecular Biology and Protein Production. Pdx1.1 and Pdx1.3, the two alleles of active Pdx1 from *Arabidopsis thaliana*, were cloned as C-terminal his-tagged proteins in *E. coli* BL21 (DE3) cells as previously described, using the NdeI/XhoI restriction sites of the pET21a expression plasmid (Novagen)[14]. The expressed proteins (including tag) have molecular weights of 33,532 Da (Pdx1.1) and 33,886 Da (Pdx1.3). Cells were grown in 1 l cultures at 37 °C to an OD of 0.6. Protein expression was induced using 60 ml of 25 % (w/v) lactose and then grown for a further 16 h at 30 °C.

Protein Purification. Cells were lysed by sonication, and centrifuged for 1 h at 140,000 x g. Proteins were purified from the supernatant using 1 ml Immobilised Metal Affinity Chromatography HiTrap columns (GE Healthcare Life Sciences) loaded with nickel and equilibrated with lysis buffer (50 mM Tris-Cl, pH 7.5, 500 mM sodium chloride, 10 mM imidazole, 2% glycerol). The column was washed with wash buffer (50 mM Tris-Cl, pH 7.5, 500 mM sodium chloride, 50 mM imidazole, 2% glycerol) and the protein was eluted in elution buffer (50 mM Tris-Cl pH 7.5, 500 mM sodium chloride, 500 mM imidazole, 5% glycerol). For subsequent size exclusion chromatography a Superdex 26/60 column (GE Healthcare Life Sciences) was equilibrated with gel filtration buffer (20 mM Tris-Cl pH 8.0, 200 mM KCl).

Complex preparation. Pdx1-R5P using crystals of Pdx1.3 grown in 0.5 M sodium citrate buffered with 0.1 M HEPES, pH 7.5. The prepared crystals were soaked in 2.5 mM R5P. K166R-preI_{320} was prepared using preformed crystals of K166R Pdx1.1 grown in 100 mM sodium cacodylate pH 6.5 and 15 % (w/v) PEG8000. 0.5 µl of 1:1 100 mM

R5P and mother liquor were then added to 1 µl drops containing the crystals, leading to a final concentration of about 15 mM R5P in the drop. The well was closed for 5 min, allowing for equilibration of the sample. 0.5 µl of 1:1 1 M NH_4Cl and mother liquor were added to the same drop resulting in a final drop size of 1.5 µl and final NH_4Cl concentration of 125 mM. The well was sealed and allowed to equilibrate for 2 h. Pdx1-I_{320} was prepared using preformed crystals of Pdx1.3 grown in 100 mM Tris-Cl pH 8.5 and 12.25 % (w/v). 0.5 µl of 50 mM R5P in mother liquor was added to 2 µl drops containing Pdx1.3 crystals resulting in a final concentration of 10 mM R5P in the drop. The well was sealed and equilibrated for 30 min. 0.5 µl of 500 mM NH_4Cl in mother liquor were then added to the same well with Pdx1.3 crystals, resulting in a final NH_4Cl concentration of 100 mM in the drop. The well was sealed and equilibrated for 4 days. Pdx1-I_{320}-G3P was prepared starting with crystals of Pdx1-I_{320} described above. After 4 days, the Pdx1-I_{320} crystals were transferred to a solution containing mother liquor containing 10 mM G3P and allowed to equilibrate for 150 s. Cryoprotectant buffer containing 10 mM G3P was then added to stabilise the crystals prior to flash cooling in liquid nitrogen. The Pdx1-PLP complex was prepared by adding preformed Pdx1.3 crystals grown in 0.1 M Tris-Cl, pH 8.25, containing 8% PEG8000 to 10 mM PLP.

In all cases except for the I_{320}-G3P complex, cryobuffer containing the well solution and an additional 20% glycerol was added, and crystals were either flash cooled in liquid nitrogen (R5P, I_{320}, PLP complexes) or directly transferred into the cryostream (pre-I_{320} complex).

Data collection and structure determination. Offline and online UV-Vis absorption spectra were recorded at the European Synchrotron Radiation Facility (ESRF), France, on

beamline ID14-4 and at the ID29S-Cryobench laboratory[23,43]. Diffraction data for the Pdx1-R5P dataset were collected on beamline ID14-1 of the ESRF at a wavelength of 0.9760 Å. Diffraction data for the Pdx1 K166R-preI_{320} dataset were collected on beamline ID29 of the ESRF, at a wavelength of 0.933 Å. Diffraction data for the Pdx1-I_{320} crystal single crystal dataset were collected on beamline I04-1 of the Diamond Light Source, UK, at a wavelength of 0.9173 Å. Diffraction data for the Pdx1-I_{320} multi-crystal dataset were collected on beamline ID23-1 of the ESRF, at a wavelength of 0.9763 Å. Diffraction data for the Pdx1-I_{320}-G3P crystal were collected on beamline ID23-1 of the ESRF, at a wavelength of 0.9763 Å. Diffraction data for the Pdx1-PLP dataset were collected on beamline ID14-1 of the ESRF, at a wavelength of 0.9334 Å. All datasets were collected at a temperature of 100 K. Data integration and scaling was carried out with XDS, DIALS and AIMLESS[44-46]. Structure determination by molecular replacement used bacterial Pdx1 from *Bacillus subtilis*[4] as the search model (PDB ID 2NV1) and MOLREP[47]. Iterative model building and refinement were carried out using COOT[48], REFMAC5[49] and PHENIX[50]. Pdx1 crystallized in a R3 unit cell with four monomers in the asymmetric unit. All monomers were first refined individually, i.e. without NCS constrains, and inspected. The crystals selected in this study represent structures that were uniform with respect to accumulation of intermediates over all four monomers. For I_{320} multi-crystal and I_{320}-G3P, the final round of refinement used the Cartesian protocol in PHENIX with Chain A as reference and NCS constraints over Lys98, Lys166 and the intermediate. Dictionary files defining ligand geometry restraints for refinement were created using JLigand[51]. All further data manipulations were carried out by the CCP4 suite of programs[52]. The refined crystal structures include Ramachandran favored residues (outliers) as Pdx1-R5P 98.8% (0.0%), Pdx1 K166R-preI_{320} 98.7% (0.0%), single crystal

Pdx1-I_{320} 98.8% (0.0%), multi-crystal Pdx1-I_{320} 97.6% (0.0%), Pdx1-I_{320}-G3P 98.7% (0.0%), Pdx1-PLP 98.3% (0.0%). The illustrations in the main text were prepared using PyMOL (Schrödinger, LLC). Data collection and refinement statistics are shown in Supplementary Table 1. Coordinates and structure factors have been deposited in the Protein Data Bank under the accession codes 5LNS (Pdx1-R5P), 5LNT (Pdx1 K166R-preI_{320}), 5LNU (Pdx1-I_{320}-G3P), 5LNW (Pdx1-I_{320}-G3P), 5LNR (Pdx1-PLP).

Multi-crystal analysis to assess beam damage. About 50 complete datasets for I_{320} complexes were collected at ID23-1, ESRF. Data collected before the 245 kGy threshold were integrated using XDS[44]; all X-ray doses were calculated using RADDOSE-3D[53]. BLEND was run in analysis mode to group datasets with similar unit cell dimensions[31]. Grouped wedges of data were input to synthesis mode to identify groups that merged well. A core group of five wedges with good merging statistics (R_{MEAS} 16.5, R_{PIM} 11.0, 62.3% complete, 2.0 Å resolution) was identified. BLEND was then run in combination mode to add wedges to the core group individually and assess whether any improvements in completeness could be made without significantly worsening merging statistics. This resulted in a 19-wedge dataset from 9 crystals (R_{MEAS} 20.1, 99.5% complete, 2.2 Å resolution) leading to a low dose structure of the I_{320} complex. Coordinates and structure factors of the multi-crystal Pdx1-I_{320} complex have been deposited in the Protein Data Bank under the accession code 5LNV.

METHODS REFERENCES

43 McGeehan, J. *et al.* Colouring cryo-cooled crystals: online microspectrophotometry. *Journal of Synchrotron Radiation* **16**, 163-172 (2009).

44 Kabsch, W. XDS. *Acta crystallographica. Section D, Biological crystallography* **66**, 125-132 (2010).

45 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta crystallographica. Section D, Biological crystallography* **69**, 1204-1214 (2013).

46 Gildea, R. J. *et al.* New methods for indexing multi-lattice diffraction data. *Acta crystallographica. Section D, Biological crystallography* **70**, 2652-2666 (2014).

47 Vagin, A. & Teplyakov, A. MOLREP: an Automated Program for Molecular Replacement. *Journal of Applied Crystallography* **30**, 1022-1025 (1997).

48 Emsley, P. & Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallographica Section D* **60**, 2126-2132 (2004).

49 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D* **53**, 240-254 (1997).

50 Adams, P. D. *et al.* The Phenix software for automated determination of macromolecular structures. *Methods* **55**, 94-106 (2011).

51 Lebedev, A. A. *et al.* JLigand: a graphical tool for the CCP4 template-restraint library. *Acta crystallographica. Section D, Biological crystallography* **68**, 431-440 (2012).

52 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* **67**, 235-242 (2011).

53 Zeldin, O. B., Gerstel, M. & Garman, E. F. RADDOSE-3D: time- and space-resolved modelling of dose in macromolecular crystallography. *Journal of Applied Crystallography* **46**, 1225-1230 (2013).

54 Brunger, A. T. *et al.* Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallographica Section D* **54**, 905-921 (1998).

55 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501 (2010).

56 Wallace, A. C., Laskowski, R. A. & Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering* **8**, 127-134 (1995).

Supplementary Information to

Lysine relay mechanism coordinates intermediate transfer in vitamin B6 biosynthesis

Matthew J. Rodrigues[1,2], Volker Windeisen[1,3], Yang Zhang[4], Gabriela Guédez[3], Stefan Weber[3], Marco Strohmeier[3], Jeremiah W. Hanes[4,5], Antoine Royant[6,7], Gwyndaf Evans[2], Irmgard Sinning[3], Steven E. Ealick[4], Tadhg P. Begley[8], Ivo Tews[1,3,*]

[1]Centre for Biological Sciences, University of Southampton, Southampton, SO17 1BJ, UK

[2]Diamond Light Source, Harwell Science and Innovation Campus, Didcot OX11 0DE, UK

[3]Heidelberg University Biochemistry Center (BZH), Im Neuenheimer Feld 328, 69120 Heidelberg, Germany

[4]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853, USA

[5]Pacific Biosciences, 1380 Willow Rd., Menlo Park, CA 94025, USA[6]Institut de Biologie Structurale, Université Grenoble Alpes, CNRS, CEA, 38044 Grenoble, France

[7]European Synchrotron Radiation Facility, 38043 Grenoble, France

[8]Department of Chemistry, Texas A&M University, College Station, TX 77843, USA

*Corresponding author

Supplementary Results

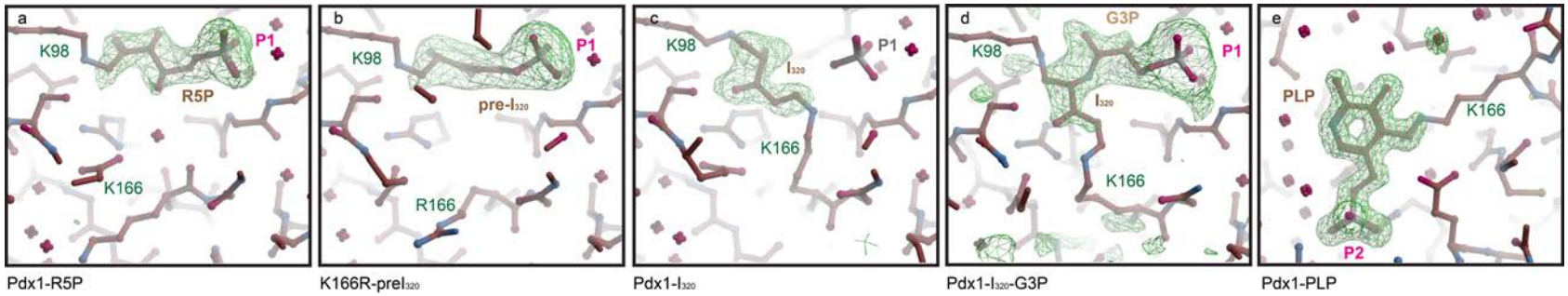
Supplementary Table 1 | Crystallographic Analysis

	R5P	K166R prel$_{320}$	I$_{320}$ single crystal	I$_{320}$ multi crystal[1]	I$_{320}$-G3P	PLP
Data collection						
Space group	R3	R3	R3	R3	R3	R3
Cell Dimensions						
a, b, c (Å)	176.73, 176.73, 114.65	177.90, 177.90, 115.02	178.12, 178.12, 115.96	178.58, 178.58, 117.38	178.26, 178.26, 115.31	178.56, 178.56, 116.88
α, β, γ (°)	90,90,120	90,90,120	90,90,120	90,90,120	90,90,120	90,90,120
Resolution (Å)	24.79-1.91 (1.94-1.91)	24.74-2.32 (2.35-2.32)	52.20-1.73 (1.76-1.73)	46.75-2.24 (2.29-2.24)	44.57-1.90 (1.93-1.90)	40.26-1.61 (1.64-1.61)
R_{merge}	0.056 (0.308)	0.095 (0.336)	0.054 (0.445)	0.216 (1.361)	0.123 (0.861)	0.049 (0.475)
CC(1/2)	0.996 (0.812)	— —	0.997 (0.841)	0.978 (0.426)	0.977 (0.447)	0.998 (0.799)
I / σI	11.4 (2.0)	9.5 (2.8)	9.3 (1.9)	30.5 (2.7)	6.1 (1.3)	9.0 (1.5)
Completeness (%)	99.9 (100.0)	99.9 (100.0)	98.1 (99.7)	99.3 (99.2)	99.4 (97.8)	100.0 (100.0)
Redundancy	3.8 (3.7)	5.7 (5.6)	2.9 (2.9)	4.3 (4.2)	3.2 (3.0)	2.8 (2.8)
Wilson B	21.3	33.7	18.6	26.4	26.6	19.3
Refinement	**5LNS**	**5LNT**	**5LNU**	**5LNV**	**5LNW**	**5LNR**
Resolution (Å)	24.79-1.91	24.74-2.32	46.35-1.73	46.75-2.24	44.57-1.90	40.26-1.61
No. Reflections	103359	58635	141161	65127	107047	179959
R_{work} / R_{free}	0.158/0.189	0.147/0.196	0.162/0.191	0.177/0.220	0.181/0.220	0.123/0.165
No. Atoms						
Protein	8405	7959	8357	8300	8166	8834
Ligand/ion	124	104	68	68	180	144
Water	859	412	1037	573	516	1070
B-factors						
Protein	29.4	36.3	22.9	29.9	34.4	28.4
Ligand/ion	23.8	41.6	30.3	36.8	42.9	27.8
Water	32.5	36.4	33.8	31.4	35.3	43.5
R.m.s. deviations						
Bond lengths (Å)	0.006	0.007	0.006	0.011	0.006	0.005
Bond angles (°)	0.800	0.838	0.813	1.355	0.934	0.709

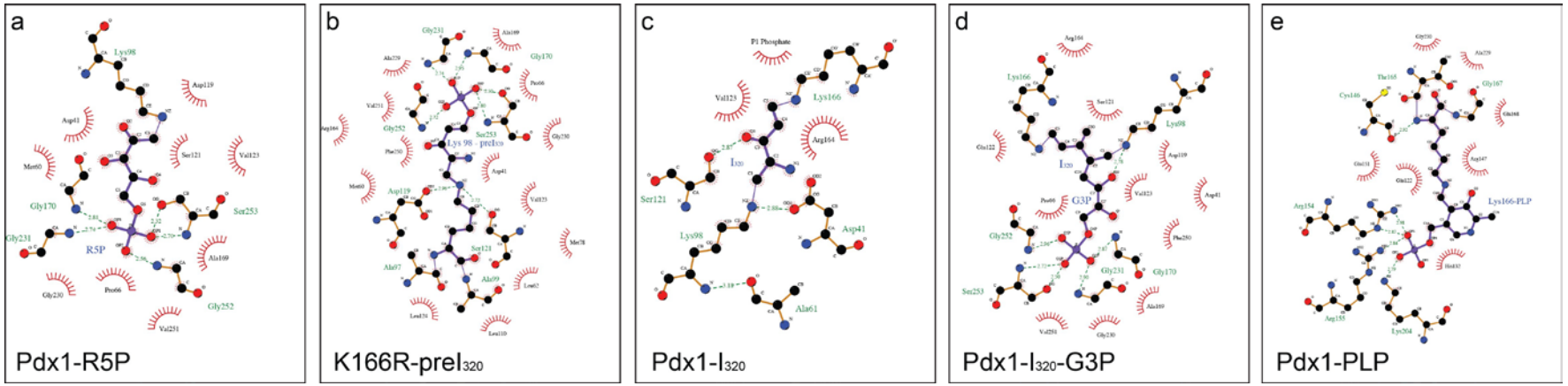
Data were collected from one crystal for each structure.

[1] Data merged from 19-wedges coming from 9 crystals, see Online Methods.

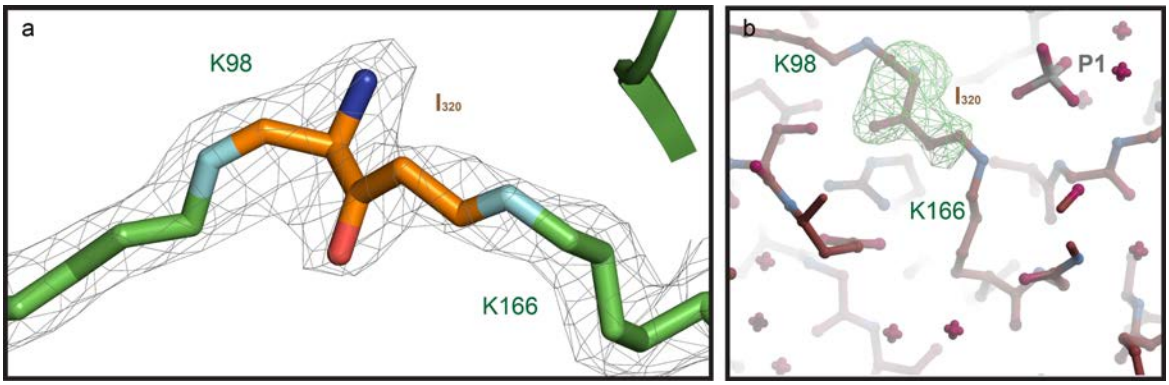
*Values in parentheses are for highest-resolution shell.



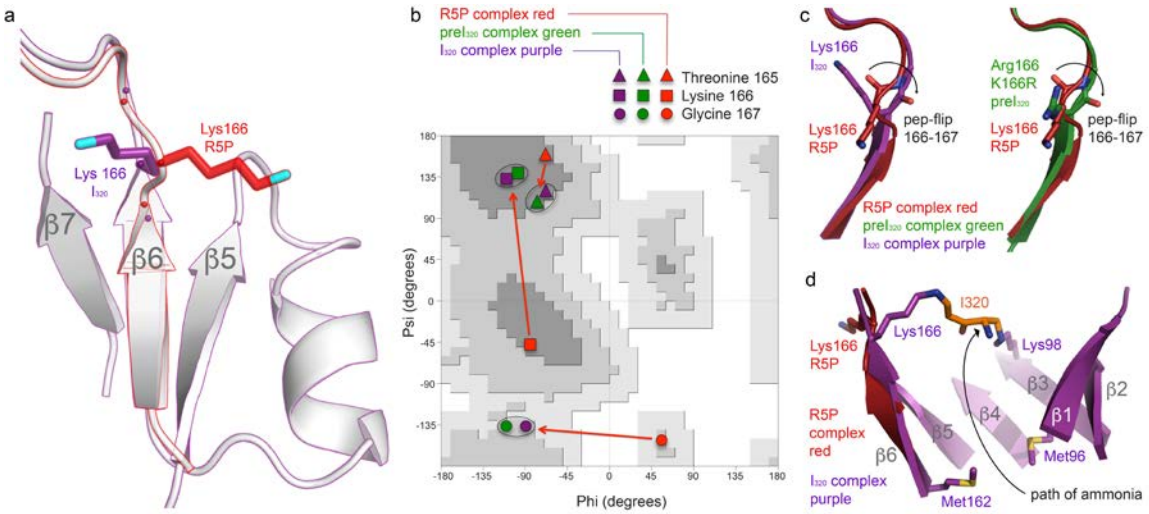
Supplementary Figure 1 | Simulated annealing omit maps for the crystallographic complexes reported in this study. a, Pdx1-R5P. **b**, K166R-preI$_{320}$. **c**, Pdx1-I$_{320}$. **d**, Pdx1-I$_{320}$-G3P. **e**, Pdx1-PLP. In each case, the intermediate ligand atoms were omitted from the map calculation using simulated annealing refinement in CNS (SA-omit-map protocol with default parameters)[54]. The F_o-F_c electron density map is shown as green mesh at 10σ for Pdx1-R5P, K166R-preI$_{320}$, Pdx1-I$_{320}$ and Pdx1-PLP and at 5σ for Pdx1-I$_{320}$-G3P. The model shown is that of the refined complexes, as reported in the main paper. Carbon atoms are red, oxygen atoms purple, phosphorous atoms grey, nitrogen atoms blue. Figure prepared with Coot[55].



Supplementary Figure 2 | Interactions between Pdx1 and intermediates of PLP synthesis. The figure shows interactions for the crystallographic complexes reported in this study: **a**, Pdx1-R5P. **b**, K166R-preI$_{320}$. **c**, Pdx1-I$_{320}$. **d**, Pdx1-I$_{320}$-G3P. **e**, Pdx1-PLP. In each case, the carbon atoms are black with bonds shown for protein amino acids in black and for intermediates in purple, oxygen atoms red, phosphorous atoms purple, nitrogen atoms blue. Hydrophobic interactions are shown as van-der-Waals spheres. Figure prepared with Ligplot[56].

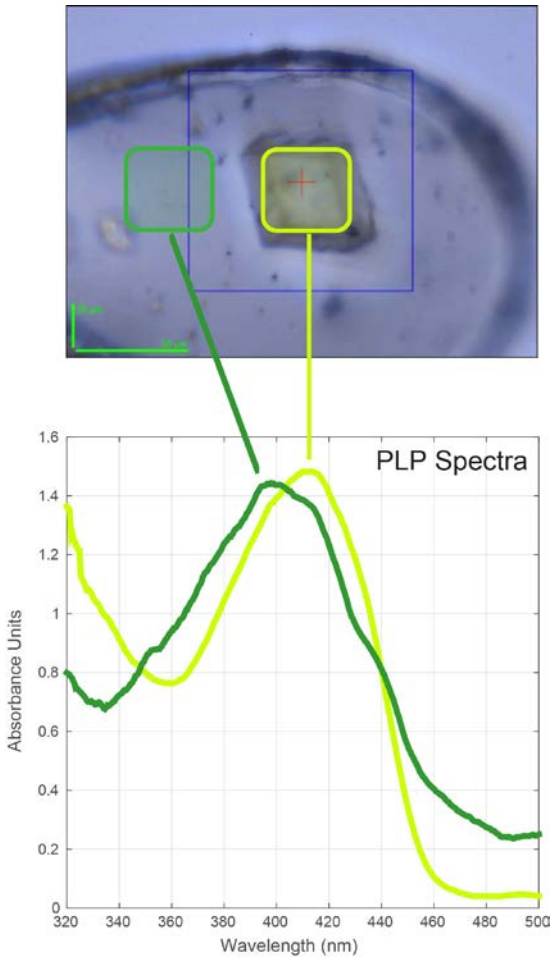


Supplementary Figure 3 | Low dose Pdx1-I_{320} multi-crystal experiment. The diffraction data were collected from nine crystals at limiting dose of 245 kGy and merged in BLEND to 2.2 Å resolution[31]. **a,** $2F_o$-F_c map after refinement contoured at 1σ. Lysine N-ε is cyan, I_{320} carbon atoms orange, nitrogen atoms blue, and oxygen atoms red. **b,** Simulated annealing F_o-F_c electron density map shown as green mesh at 10σ for Pdx1-I_{320}; carbon atoms are red, oxygen atoms purple, phosphorous atoms grey, nitrogen atoms blue; see legend of Supplementary Fig. 1 for details.

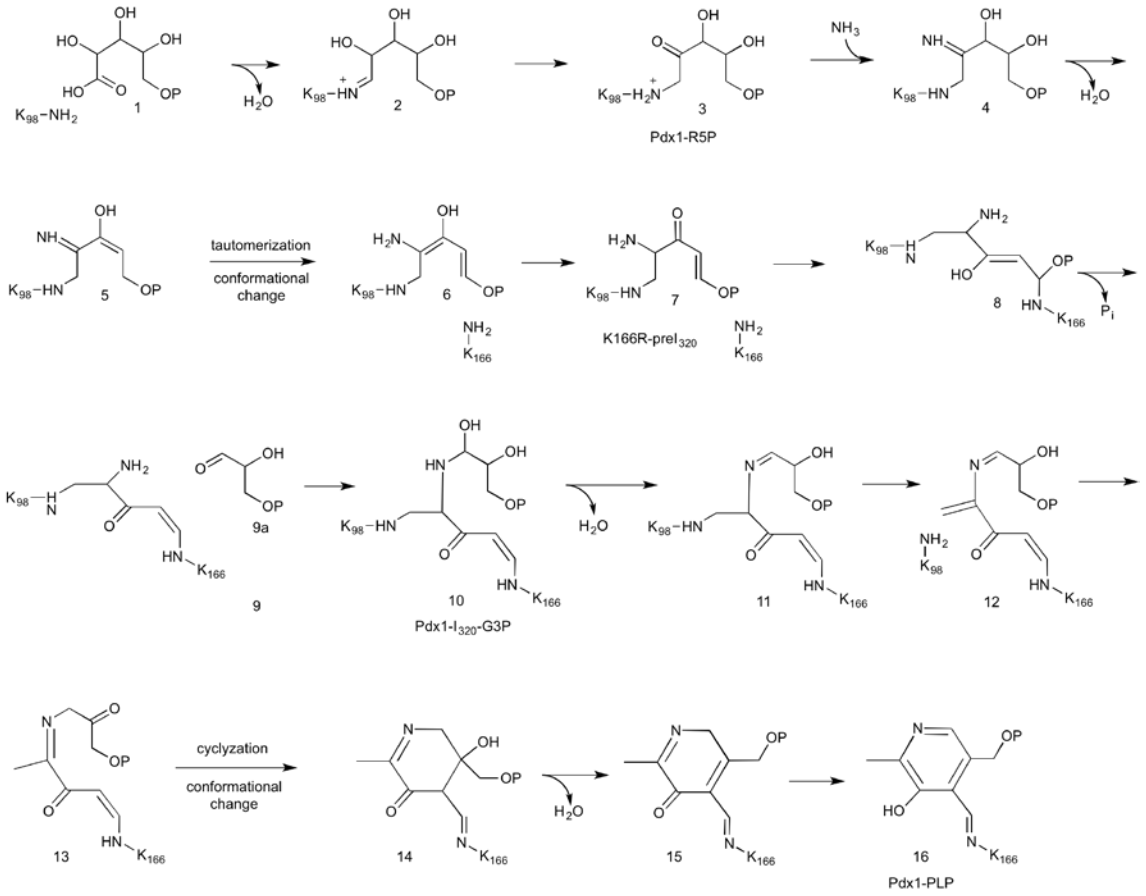


Supplementary Figure 4 | Structural changes of Pdx1 support I_{320} intermediate formation. a, Selected structure elements of Pdx1 are shown in cartoon representation. The side chain of the catalytic Lys166 has a different orientation in the Pdx1-R5P and in the Pdx1-I_{320} complexes and is shown as red and purple sticks, respectively, nitrogen atoms blue. Coloured dots indicate the Cα-positions for Thr165 and Gly167.**b**. UV-Vis spectrum of a Pdx1-I_{320} and a Pdx1-preI_{320} crystal measured at ESRF beamlines ID14-1 and ID29 respectively[23,43]. The spectrum of the Pdx1-I_{320} crystal (yellow) shows an absorption maximum at 280 nm, corresponding to protein, and near 320 nm for the I_{320} intermediate. The K166R variant (purple) shows reduced absorbance at 320 nm with the absorption maximum shifted to ~333 nm. **c,** The Ramachandran plot illustrates the changes in backbone conformation between the Pdx1-R5P and the K166R-preI_{320} / Pdx1-I_{320} complexes for amino acids Thr165, Lys166/Arg166 and Gly167. **d,** Cartoon representation of strand β6 and part of the β6-α6-loop on which the catalytic Lys166 is located. A peptide flip occurs at the Lys166-Gly167 peptide bond, supporting the different side chain orientation of Lys166 between the Pdx1-R5P (red) and the Pdx1-I_{320} complexes (purple, left), and similarly between Arg166-Gly167 in the Pdx1-R5P (red) and the K166R-preI_{320} complexes (green, right). **e,** Ammonia is produced by the Pdx1 associated glutaminase Pdx2 and allows conversion of the R5P adduct to I_{320} (Pdx1-I_{320} complex purple, Pdx1-R5P complex red). Ammonia is transferred from the Pdx2 active site through the central β-barrel of the Pdx1 enzyme, passing a constriction formed by amino acids Met162 and Met96. Mutagenesis data support a coupling of Pdx1 and Pdx2

enzymatic activities through these residues[15], suggesting that the conformational changes around Lys166 are coupled with ammonia transfer.



Supplementary Figure 5 | UV-Vis Spectra of Pdx1-PLP Crystals, UV-Vis spectra recorded at ESRF beamline ID14-1 from a Pdx1-PLP crystal[43]. The spectrum of the crystal shows an absorption maximum at λ = 414 nm for covalently bound PLP (light green), while free PLP in the surrounding buffer has an absorption maximum of λ = 388 nm (dark green). The shift in absorbance maximum from 388 to 414 nm is characteristic for a covalent PLP complex[6,16].



Supplementary Figure 6 | Mechanistic proposal for PLP biosynthesis by Pdx1, The mechanism begins with formation of an R5P imine **2**, which tautomerizes to **3,** and ends with a PLP imine **16**. In this mechanism, Pdx1-preI$_{320}$ **7** reacts with Lys166 to form Pdx1-I$_{320}$ **9**. This then reacts with G3P to form a further covalent intermediate **10**. Release of Lys98 and transfer of this intermediate to the P2 site leads to the observed covalent PLP complex **16**. Structures of intermediates **3**, **7**, **9**, **10**, and **16** were determined.

REFERENCES

6 Fitzpatrick, T. B. *et al.* Two independent routes of de novo vitamin B6 biosynthesis: not that different after all. *The Biochemical journal* **407**, 1-13 (2007).

15 Guedez, G. *et al.* Assembly of the Eukaryotic PLP-Synthase Complex from Plasmodium and Activation of the Pdx1 Enzyme. *Structure* **20**, 172-184 (2012).

16 Smith, A. M., Brown, W. C., Harms, E. & Smith, J. L. Crystal Structures Capture Three States in the Catalytic Cycle of a Pyridoxal Phosphate (PLP) Synthase. *The Journal of biological chemistry* **290**, 5226-5239 (2015).

23 von Stetten, D. *et al.* In crystallo optical spectroscopy (icOS) as a complementary tool on the macromolecular crystallography beamlines of the ESRF. *Acta crystallographica. Section D, Biological crystallography* **71**, 15-26 (2015).

31 Foadi, J. *et al.* Clustering Procedures for the Optimal Selection of Data Sets from Multiple Crystals in Macromolecular Crystallography. *Acta Crystallographica Section D* **69**, 1617-1632 (2013).

43 McGeehan, J. *et al.* Colouring cryo-cooled crystals: online microspectrophotometry. *Journal of Synchrotron Radiation* **16**, 163-172 (2009).

54 Brunger, A. T. *et al.* Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallographica Section D* **54**, 905-921 (1998).

55 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501 (2010).

56 Wallace, A. C., Laskowski, R. A. & Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering* **8**, 127-134 (1995).