Tunable Heteroassembly of a Plant Pseudoenzyme–Enzyme Complex

Irina V. Novikova,

Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99354, United States

Mowei Zhou,

Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99354, United States

James E. Evans,

Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99354, United States; School of Biological Sciences, Washington State University, Pullman, Washington 99164, United States

Chen Du,

Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States; Resource for Native Mass Spectrometry Guided Structural Biology, The Ohio State University, Columbus, Ohio 43210, United States

Marcelina Parra,

School of Biological Sciences, Washington State University, Pullman, Washington 99164, United States

Doo Nam Kim,

Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99354, United States

Zachary L. VanAernum,

Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States; Resource for Native Mass Spectrometry Guided Structural Biology, The Ohio State University, Columbus, Ohio 43210, United States

Jared B. Shaw,

Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99354, United States

Hanjo Hellmann,

School of Biological Sciences, Washington State University, Pullman, Washington 99164, United States

Vicki H. Wysocki

Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States; Resource for Native Mass Spectrometry Guided Structural Biology, The Ohio State University, Columbus, Ohio 43210, United States

Abstract

Pseudoenzymes have emerged as key regulatory elements in all kingdoms of life despite being catalytically nonactive. Yet many factors defining why one protein is active while its homologue is inactive remain uncertain. For pseudoenzyme–enzyme pairs, the similarity of both subunits can often hinder conventional characterization approaches. In plants, a pseudoenzyme, PDX1.2, positively regulates vitamin B_6 production by association with its active catalytic homologues such as PDX1.3 through an unknown assembly mechanism. Here we used an integrative experimental approach to learn that such pseudoenzyme–enzyme pair associations result in heterocomplexes of variable stoichiometry, which are unexpectedly tunable. We also present the atomic structure of the PDX1.2 pseudoenzyme as well as the population averaged PDX1.2–PDX1.3 pseudoenzyme–enzyme pair. Finally, we dissected hetero-dodecamers of each stoichiometry to understand the arrangement of monomers in the heterocomplexes and identified symmetry-imposed preferences in PDX1.2–PDX1.3 interactions. Our results provide a new model of pseudoenzyme–enzyme interactions and their native heterogeneity.

Graphical Abstract
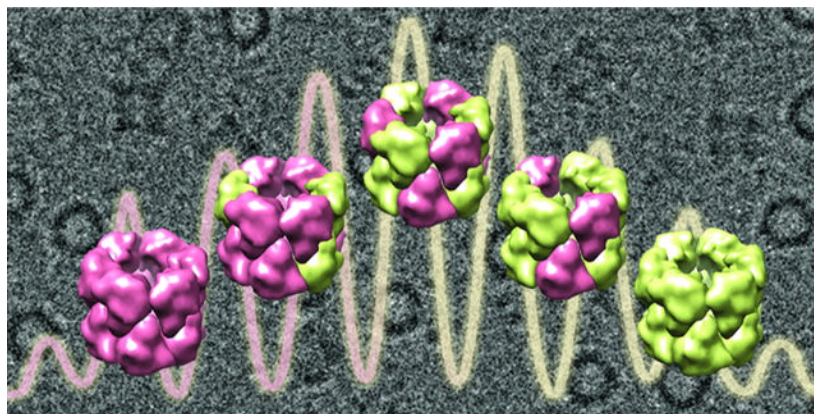
INTRODUCTION

Pseudoenzymes are homologues of enzymes that are catalytically deficient or inactive. Prior research from select protein families of kinases and phosphatases has shown that many pseudoenzymes retain the structural fold of their catalytically active partners and function as protein interaction modules.[1] Structural studies were the key tools in exploring and clarifying the molecular mechanisms of such systems, and they uncovered their roles as allosteric regulators, scaffolds, molecular switches, and substrate competitors.[2–4] These pseudoenzymes have to be diligently targeted or antitargeted (i.e., avoid intervention) for desirable outcomes in drug design or bioengineering applications so as to not be indiscriminately treated as their active homologues. Genomic studies now show that pseudoenzymes can make up as much as 10% of the enzyme families,[5] while very strict homology and rather conservative filtering of biological databases provides a lower-bound estimate of <2%.[6] However, the high structural similarities between pseudoenzymes and canonical enzymes create new experimental challenges in classical molecular and biochemical characterization approaches for unraveling the underlying mechanism. This is the case for the present study, which is focused around pseudoenzyme PDX1.2, a regulator of vitamin B_6 biosynthesis in plants.

Vitamin B_6 functions as a cofactor for over 100 enzymatic processes and also provides an important defense mechanism against oxidative stress and pathogen infection.[7–11] Plants are able to synthesize vitamin B_6 (PLP) *de novo* by two enzymes, PDX1 and PDX2 (pyridoxine biosynthesis 1 and 2).[12] In particular, in *Arabidopsis thaliana*, there are three homologues of PDX1, designated as PDX1.1, PDX1.2, and PDX1.3, and one gene of PDX2. Of these, PDX1.1, PDX1.3, and PDX2 are all catalytically active enzymes, while PDX1.2 has been shown to be an inactive variant by *in vivo* yeast complementation tests.[12] Homology analysis of PDX1.2 showed that key amino acids in the catalytic sites P1 and P2 are mutated.[13] Subsequent *in vivo* experiments in plants showed that PDX1.2 expression is upregulated by heat and oxidative stress and coincides with a boost in vitamin B_6 synthesis, suggesting a positive regulatory impact.[13–15] PDX1.2 was first proven to interact with its catalytic PDX1.1 and PDX1.3 proteins *in vivo* and to form complexes of high molecular weight similar to its catalytic homologues.[16] Recombinant coexpression of PDX1.2 with either PDX1.1 or PDX1.3 showed dodecameric heterocomplexation.[13]

PDX1.2 protein resists crystallization protocols, while PDX1.3 homocomplexes and PDX1.2/PDX1.3 heterocomplexes are compatible with macromolecular crystallography.[17] However, unlike PDX1.3 homocomplexes, the crystals of PDX1.2/PDX1.3 heterocomplexes suffered from a statistical disorder, where the positions of individual proteins were not distinguished despite exhaustive approaches used. The ambiguity in the data prohibited the determination of the exact assembly mechanism. An interhexamer assembly mechanism was proposed as the most likely arrangement where one hexameric ring of active PDX1.1 (or PDX1.3) stacks on top of a hexamer ring of inactive PDX1.2 (Figure 1, panel a) to create a final hetero-dodecameric state, but alternate arrangements were also considered.[13,17] Herein, using an integrated biochemical and structural approach combining cell-free protein synthesis technology, native mass spectrometry (native MS), and cryo-electron microscopy (cryo-EM), we show that PDX1.2 and PDX1.3 form hetero-dodecamers through

intrahexamer association of two symmetric or similar hexamers with varying but tunable stoichiometry.

RESULTS AND DISCUSSION

Cell-Free Expression of PDX Homo- and Heterocomplexes.

We employed a cell-free protein expression pipeline using wheat germ extract[18–20] to provide the closest translational environment for the production of these plant proteins in terms of folding and post-translational modifications. The "open-format" of a cell-free platform provided us with an opportunity for precise stoichiometric control of protein coexpression (Figure 1, panel b) by varying the amounts of corresponding DNA templates (Supplementary Figure 1), not easily achieved by cell-based protein expression methods. We prepared two DNA plasmids constructs (one for PDX1.2 and one for PDX1.3), and combined them in specific molar ratios to achieve the desired protein output in the coexpression conditions.

In the initial cell-free translation experiments, we supplemented all the translation reactions with fluorophore-labeled lysine-charged tRNA as recently described.[20] This procedure allowed detection of newly synthesized proteins in the crude mixture with no need for purification. For coexpression, different DNA template ratios (9:1, 3:1, 1:1, 1:3, and 1:9) were tested to establish that the molar ratio of expressed proteins can be controlled in a precise fashion. For instance, by in-gel quantification, the coexpression conditions 9:1, 3:1, and 1:1 correspond to protein product compositions of 3:1, 1:1, and 1:3 (Figure 1, panel c). The same samples were also analyzed under native gel conditions where all coexpressed samples traveled faster than individual PDX1.2 homocomplexes but slower than PDX1.3 homocomplexes. In addition, the bands for coexpression complexes appeared significantly wider in size suggesting the presence of multiple stoichiometric dodecamer species. The mobilities of these samples also appeared to be dependent on the protein composition where a higher proportion of PDX1.3 in the coexpressed sample caused faster migration of the heterocomplex, and the opposite was true for higher PDX1.2 content.

We also verified that coexpression was essential for the PDX1.2 and PDX1.3 heterocomplex formation, supporting previous discoveries.[13,17] Mixing the two separately expressed and purified proteins post-translationally did not produce detectable amounts of heterocomplexes on native PAGE (Supplementary Figure 2, panel a). The same coexpression experiments were conducted using active PDX1.1 enzyme (instead of PDX1.3) and PDX1.2, and identical pseudoenzyme–enzyme assembly behavior was observed (Supplementary Figure 2, panel b). Because PDX1.3 is the dominant homologue in *A. thaliana*[21,22] and was previously found to be more impacted enzymatically by the presence of PDX1.2,[13] we have limited our further structural survey to PDX1.2 and PDX1.3. We have also verified that these observations are not impacted by the purification tags positioned on their N-terminus by expressing tag-free PDX1.2 and PDX1.3 proteins and analyzing their mobility and their assembly behavior in the crude mixture. No differences were detected between tagged and untagged proteins (compare Figure 1, panel c, with Supplementary Figure 2, panel c).

Tunable Stoichiometry and Activity of PDX Heterocomplexes.

For further structural characterization, we scaleup the synthesis and then purified select samples: PDX1.2 and PDX1.3 homocomplexes and heterocomplexes with protein ratios of 3:1, 1:1, and 1:3 (Supplementary Figure 3). Protein sequences were verified by peptide mapping (Supplementary Figure 4). Liquid chromatography mass spectrometry (LCMS) of intact (denatured) proteins showed masses of 37339.7 and 36405.5 Da for PDX1.2 and PDX1.3, respectively (Supplementary Figure 5), which are consistent with their monoisotopic theoretical masses of 37339.6 and 36405.3 Da. Although approximately 50% of the protein N-termini were found to be acetylated, no preferential acetylation profile was detected suggesting that these post-translational modifications are unlikely to play a role in PLP biosynthesis regulation (Supplementary Figures 5 and 6).

While the native PAGE results indicated the presence of mixed co-complexes in each coexpression condition, the technique did not have the resolution to define the stoichiometry. We therefore employed native MS, where we electrosprayed the proteins under nondenaturing conditions in 200 mM ammonium acetate solution. The intact 12mers were detected around m/z of 9000, carrying ~50 positive charges (Supplementary Figure 7). All species in the coexpressed samples were resolved, and the stoichiometry could be accurately determined from the measured mass, as shown by the native MS spectrum of heterocomplex 3:1 (Figure 2, panel a). From the spectra, mass distributions of all species in the sample can be deconvoluted by combining the intensities across charge states (Figure 2, panel b). As expected, the PDX1.2 and PDX1.3 samples showed predominantly homo-12mers with molecular weights of 449 kDa and 437 kDa, respectively. Each coexpressed sample generated 6–8 heterocomplexes of variable stoichiometry that were all dodecameric and influenced by the initial plasmid DNA ratio input to the cell-free reaction. For example, heterocomplex 3:1, with PDX1.2 representing the major protein constituent, spans stoichiometry values in the 6:6 to 12:0 range (6:6, 7:5, 8:4, 9:3,10:2, 11:1, and 12:0). For the heterocomplex 1:1, there is a shift to the 3:9–11:1 stoichiometry range, while heterocomplex 1:3 was limited to the 0:12–8:4 stoichiometry range (Figure 2, panel b). The observed stoichiometry values were consistent with the amounts of proteins produced in each sample (see Figure 1, panel c, and Supplementary Figure 3). By varying the input DNA ratio (and thus protein ratios), the full range of co-complexes from 0:12 to 12:0 stoichiometry can be sampled, demonstrating the tunability of these protein–protein associations.

Interestingly, a significant amount of free monomer and dimer species were observed for PDX1.3 containing complexes in the low m/z region <6000 (Supplementary Figure 7). In the coexpressed samples, the relative abundance of PDX1.3 monomers and dimers was also found to be positively correlated with the ratio of PDX1.3 in the heterocomplexes. In contrast, PDX1.2 appeared to be more stable under the same experimental conditions. We confirmed that 200 mM ammonium acetate did not cause noticeable protein unfolding or disassembly in solution prior to MS (Supplementary Figure 8). Previous native MS studies have shown that positively charged proteins may interact with silanol groups on borosilicate glass surface at sub-micrometer sized emitter tips, resulting in supercharging and unfolding during electrospray.[23] In addition, supercharging typically makes complexes

easier to dissociate *in vacuo*.[24,25] Therefore, the dissociation of PDX1.3 containing 12mers might be due to unique charging properties of PDX1.3 that are distinct from PDX1.2.

To understand how the PDX1.2/PDX1.3 ratio in heteroassembly affects their enzymatic performance, we employed a diagnostic assay where PLP is used as a cofactor molecule for apoenzyme chemical transformations, the products of which were detected by a colorimetric approach (Figure 2, panel c). Under these assay conditions with saturating substrate amounts, the specific activity of active enzyme PDX1.3 was determined to be 793 (μmol/min)/mg protein. As expected, pseudoenzyme PDX1.2 was found to be inactive and displays a slightly inhibitory effect. In heterocomplexes 1:1 and 1:3, the presence of PDX1.2 positively impacts PLP synthesis with an increase of up to 120% in relative enzymatic activity. However, heterocomplex 3:1, where 75% of the total protein is PDX1.2, displays a decline in relative activity to 67% of that observed for PDX1.3 homo-12mer. Yet, the measured activities represent a sum of PDX1.2 and PDX1.3. If these relative enzymatic activities are normalized based on the amount of catalytically active PDX1.3 in each coexpressed sample, the positive regulatory impact of PDX1.2 is observed for every condition (Figure 2, panel d). The higher ratio of PDX1.2 incorporation showed higher impact on PDX1.3 catalytic activity with the corresponding outputs: 3-fold increase for heterocomplex 3:1, 2.4-fold increase for heterocomplex 1:1, and 1.5-fold increase for heterocomplex 1:3.

Lateral Symmetry within the Seemingly Stochastic Assembly.

To resolve the structures of the purified complexes, we employed single particle cryo-EM. The PDX1.2 sample was first vitrified on EM grids and imaged in super-resolution mode using a Titan Krios cryo-EM with K2 direct electron detector (Supplementary Figure 9). A homogeneous distribution of particles in various orientations are clearly visible in the thin layer of ice, and several rounds of 2D classification produced well-defined 2D classes where a stacked 2-ring architecture was easily distinguished and consistent with the expected dodecameric fold known for catalytically active PDX1.1 and PDX1.3. The final 3D volume was reconstructed at 3.2 Å resolution, and a PDX1.2 homology model was then docked and real space refined (Supplementary Figure 10 and Supplementary Table 1). Residues 29–288 were clearly resolved while N- (residues 1–28) and C-terminal (residues 289–313) regions appear rather flexible due to the lack of associated densities in the volume. We also obtained a high-resolution cryo-EM data set for the heterocomplex 3:1 (Supplementary Figure 11). A well-defined two-ring fold seemed visually identical to the PDX1.2 homo-12mer. Various heterogeneous, nonuniform, and local refinements did not sort out the potential subclasses, which is not surprising due to variability of the stoichiometry of the coexpressed species, the similarity of PDX1.2 and PDX1.3 folds, and the small difference (<1 kDa) between their monomer molecular weights. As a result of the lack of distinct subclasses, the heterocomplex 3:1 particle data set was processed as a single 3D class, and nonuniform refinement yielded the final map. Both the PDX1.2 and PDX1.3 homomers can be fitted independently with high validation scores (Supplementary Figure 12 and Supplementary Table 1) to any of the monomer positions - similar to the result reported recently for X-ray structure.[17] In other words, the two different subunits cannot be easily distinguished in the heterocomplex, suggesting a stochastic heteroassembly.

To better understand the distribution of the subunits within the heterocomplexes, we mass isolated individual heteromers of each stoichiometry and subjected them to surface induced dissociation (SID) in a modified mass spectrometer.[26] The charge states of the proteins were reduced by adding triethylammonium acetate in solution, which suppressed unfolding upon activation and allowed more informative dissociation products to be obtained in SID[27] (Supplementary Figure 13). A representative SID spectrum for the 6:6 heterocomplex is shown (Figure 3, panel a). Collision of the precursor ions into the surface under controlled laboratory frame collision energy (5.6 keV) in SID caused a rapid increase in an internal energy for the 12mer, leading to the dissociation into subcomplexes. Our previous studies have shown that subcomplexes generated in SID are often from dissection of the weakest interfaces and are reflective of the native quaternary structure of protein complexes.[28,29] In contrast, collision-induced dissociation (commercially available as higher-energy collisional dissociation, HCD, in the instrument used here) results in activation of the 12mer via multiple low-energy collisions and, at most, only resulted in stripping of unfolded monomers at the maximum collision energy (10.5 keV) (Supplementary Figure 14). Among the released species in SID, the most abundant was the 6mer (Figure 3, panel b). Other species such as 4mers and 8mers were also detected but at lower abundance. The major products, 6mers, can be generated either exclusively through horizontal cleavage along the interface of two 6mer rings or through vertical cleavages across the 6mer rings while maintaining the lateral interactions (Figure 3, panel c). In case of vertical cleavages, (2 + 10)mer, (4 + 8)mer, and (6 + 6)mer species should all have identical likelihood due the symmetry and thus the abundances of (2 + 10)mers and (4 + 8)mers should be similar to the abundance of (6 + 6)mers. However, we find that (6 + 6)mers were the dominant products, suggesting that horizontal cleavage is the major contributor (Supplementary Figure 15). The vertical cleavages are the minor pathway for cleavage and are responsible for the lower abundance 4mer + 8mer pair and contributed partially to the signal of 6mers. We suspect the 5mers in both spectra may partially originate from secondary dissociation of 6mers and are not from direct cleavage of the 12mers because of the low abundance of the complementary 7mers. Herein, we focus on the major pathway for structural elucidation of the heteroassemblies.

The deconvoluted SID spectra show a relatively monodisperse stoichiometric distributions in the released 6mers. For example (Figure 3, panel b), ~70% of the 6mers from the 6:6 heteromer had the stoichiometry of 3:3, while the remainder were found to be 2:4 and 4:2. This suggests that the 6:6 12mers are formed in either the [3:3 and 3:3] or [2:4 and 4:2] combination of PDX1.2/PDX1.3. For 7:5 heteromer with odd numbers of each protein in the 12mer, 6mers with stoichiometry of both 3:3 and 4:2 were released at similar intensities. The most abundant 6mers released by SID from other 12mers all have half of the stoichiometry of the 12mer precursors (Supplementary Figure 15). We estimated the abundances of the 6mers generated only from the major pathway (see Methods for details) and plotted the relative abundances of the different stoichiometries as a function of the 12mer precursor stoichiometry (Figure 3, panel d). The narrow distribution of the stoichiometry of the 6mer rings strongly favors a lateral symmetry in the 12mers. If PDX1.2 and PDX1.3 were distributed stochastically, a wide variety of stoichiometries should be observed in the released 6mers (Supplementary Figure 16).

Similar Protein Folds but Inverted Electrostatic Surface Potentials.

Because of the challenges in resolving individual proteins in heterocomplexes, we instead focused on the cryo-EM structure of the PDX1.2 homo-12mer for clues on its positive regulatory function by comparing it with several crystallographic structures of PDX1.3 available at different stages of enzymatic action.[30] Precursor binding stage PDX1.3-R5P (PDB 5lns), intermediate stage PDX1.3-I$_{320}$ (PDB 5lnu), postintermediate PDX1.3-I$_{320}$-G3P (PDB 5lnw), and final product stage where PLP is still covalently bound to Lys166 (PDB 5lnr) were all aligned with our structure of PDX1.2 and overlaid for comparison (Figure 4, panel a). The PDX1.3 enzyme employs a lysine relay mechanism where all catalytic action happens within a single domain using two catalytic sites, P1 and P2.[30,31] Two key residues Lys98 and Lys166 are at the heart of this process where they trap the substrates, covalently tether the intermediates and then shuttle the product out. We found that the overall monomer fold of PDX1.2 is very similar to PDX1.3, and the best fit for PDX1.2 is to the X-ray structure of PDX1.3-PLP, which has the P1 site unoccupied.

Importantly, the overall monomer fold of PDX1.2 is identical to that of PDX1.3 at the outer shell where proper higher order protein interactions with neighboring subunits need to be preserved (Figure 4, panel a). The conservation pressure on those amino acids, which are involved in quaternary interactions, appears not to be specific for *A. thaliana* and is imposed on many plant organisms (Supplementary Figure 17). Thus, the maintenance of PDX1.2 quaternary contacts with PDX1.3 is very important for pseudoenzyme functional performance. Major differences between PDX1.2 and PDX1.3 are observed in the area surrounding the catalytic site P1. Specifically, the $\alpha 2'$ helix and $\beta 6-\alpha 6$ loop are significantly altered in PDX1.2. The $\alpha 2'$ helix, present in PDX1.3, is absent in PDX1.2 due to the amino acid truncation in that region and is substituted by a loop that points slightly outward. The amino acid changes in the $\beta 6-\alpha 6$ loop, including an amino acid insertion, create a slight kink in PDX1.2 architecture toward the center of the 12mer, resulting in the potential weakening or aborting of the phosphate binding at this site. Key lysine residues Lys98 and Lys166, which form covalent adducts with the intermediates of PLP formation in PDX1.3, are mutated to Arg100 and Gln168, respectively, in PDX1.2. While Arg100 is positioned similarly to Lys98, Gln168 in PDX1.2 (in place of Lys166 in PDX1.3) points toward the P1 site similar to the intermediate state of the PLP synthesis instead of pointing to P2 as observed in the priming PDX1.3-R5P state. The same positional occupancy of Gln168 was assigned to the composite assignment of Gln168 and Lys166 in the crystallized PDX1.2–PDX1.3 co-complex where individual protein positions were not distinguished.[17] Such structural alterations in the P1 site in PDX1.2 are in agreement with the experimental data, showing the loss of catalytic potential. Unlike P1, the catalytic site P2 in PDX1.2 appears to be undisturbed except for the mutation of the key residue Lys166 to Gln168.

Based on the structural analysis of overall shape and the comparison of areas around the catalytic sites between PDX1.2 and PDX1.3, there is no obvious benefit in keeping such a pseudoenzyme and there is no clear mechanism for how PDX1.2 can positively regulate PDX1.3. We did not find any significant differences in the interfaces between subunits (Supplementary Figure 18). Hydrophobicity plots of PDX1.2 and PDX1.3 did not show any difference between the positions of polar and hydrophobic residues (Supplementary

Figure 19). Molecular modeling using *Rosetta* showed no obvious impact of PDX1.2 on the stability of the complex. Surprisingly, the *Rosetta* calculation estimated that PDX1.3 packs more tightly than PDX1.2 and has more favorable sum of interactions and interface stabilization (Supplementary Table 2). Therefore, other regulatory contributions are likely present (e.g., long-range surface electrostatics changes).

Interestingly, the most stunning differences were found when comparing the electrostatic surface potentials between the two (Figure 4, panel b). The surface potential of outer and inner regions of the PDX1.2 12mer, which spans the P1 and P2 sites, appears inverted relative to PDX1.3, where negative charges replace the positively charged regions. The difference in charge based on the structures was also consistent with the apparent lower stability of PDX1.3 than PDX1.2 in native MS as mentioned earlier. The positive electrostatic surface potential of PDX1.3 could induce charging and unfolding in electrospray, giving rise to significant dissociation in the native MS spectra.

CONCLUSION

Pseudoenzymes represent a largely uncharted territory with much to learn about the additional layer of regulatory control they impose in nature. They are now known for allosteric regulation and scaffolding through conserved binding interfaces or domains like their active enzyme homologues.[5] Yet, technical challenges exist in probing such pesudoenzyme−enzyme interactions due to extremely high structural similarity, as exemplified by the PDX1.2−PDX1.3 pseudoenzyme−enzyme pair detailed in this work.

A single particle cryo-EM approach was used to solve the structure of PDX1.2, a protein proven recalcitrant to crystallization. We found that the PDX1.2 fold closely mimics the fold of PDX1.3, a common trend observed for pseudoenzymes involved in allosteric regulation of their catalytic counterparts. PDX1.2 also displays an altered catalytic site P1, which appears perturbed in concordance with its lack of activity, while the P2 site remains largely unchanged. In the studies described herein, we were able to accurately control the stoichiometry of heterocomplexes using cell-free expression and characterize them using native MS. To our surprise, the catalysis can be tuned by shifting pseudoenzyme−enzyme stoichiometry, a discovery that was not reported previously for any known pseudoenzyme−enzyme pair today.

The tunable nature of PDX1.2 and PDX1.3 heterocomplexes produced using our experimental *in vitro* conditions strongly disfavors a hetero-dodecamer model where PDX1.2 homohexamer stacks on PDX1.3 homo-hexamer (Figure 1, panel a). However, we want to outline that we do not exclude the possibility of higher-order control of the PDX1.2/1.3 heteroassociation by other assembly factors *in vivo*. Our native MS and SID data demonstrated that PDX1.2 coassembly with its catalytic partner PDX1.3 is largely based on stochastic subunit incorporation at different locations but with some degree of symmetry (Figure 5, panel a). Driven by the molar ratios of individual components during coexpression conditions, the entire range of all possible stoichiometry combinations (from 12:0 to 0:12) has been recorded.

Another intriguing finding is the inverted surface charge of PDX1.2 compared to PDX1.3. It has been shown that surface charge–charge interactions can be redesigned to increase protein stability.[32] As *in vivo* studies have demonstrated a stabilizing effect of PDX1.2 on PDX1.3 catalysis under heat stress,[33] the switch in the electrostatic surface potential might serve as an evolutionary measure to heat adaption. Additionally, our proposed structural models (Figure 5, panels b–f) could help to explain the positive regulatory impact of PDX1.2. The catalytic site P1 is buried inside the center of the individual PDX1.3 subunit, and it is therefore unlikely to be impacted by PDX1.2. Instead, in concordance with the observed mechanism of assembly and amino acid conservation, we hypothesize that the most profound effect would be on the catalytic site P2. Site P2 is located at the edge of the PDX1.3 subunit and shares two subunit–subunit interfaces: side-by-side I1′ within the hexamer ring and bottom–bottom I3 at the ring–ring interface (Figure 5, panel b). We constructed several PDX1.2–PDX1.3 assembly models around the catalytic site P2 of PDX1.3 by swapping subunits in the aligned homo-12mer structures (Figure 5, panels c–f). Lateral PDX1.2 incorporation can create a local supercharging effect on the $\beta4-\alpha4$ loop, which could potentially act on the nitrogen atom of its pyridoxine ring (Figure 5, panel d). This supercharging effect could promote imine bond breaking between PLP and Lys166, help in the translocation, and assist with the exit of the final product and thus increase the rate of reaction turnover. The supercharging of the same loop could also be possible when PDX1.2 shares a bottom–bottom interface (Figure 5, panel e). The positioning of PDX1.2 on the bottom also slightly weakens the charge state around the phosphate binding site in $\alpha4$. Both side and bottom consequent positioning of PDX1.2 (Figure 5, panel f) could explain enhanced PDX1.3 activity in the heterocomplex and the benefit of having a pseudoenzyme that heteroassociates with active PDX1.3 in the 12mer.

While our current data cannot determine precisely which interfaces in the heterointeraction are favored, the constructed models strongly suggest the benefit of such interactions for enzyme activity. Given the inverted electrostatic surface potentials of the two proteins, attractive forces between opposite charges may favor the symmetry and drive preferential arrangement in such heterointeractions without additional assembly factors. We thus propose that PDX1.2 and PDX1.3 monomers are placed in an alternating manner within the heterocomplex so that the heterointeractions can be maximized as driven by the opposite charges. This is supported by our observation that homo-subcomplexes were not released in SID for most heterocomplexes (Supplementary Figure 15), meaning the same protomers do not prefer to cluster within the 12mer. Such positioning strategy empowers a simple, tunable, and gentle approach to accelerate vitamin B_6 production by active enzymes such as PDX1.3 and PDX1.1.

Although the monomer positions within individual heterocomplexes cannot be structurally resolved by either cryo-EM or X-ray crystallography, native MS provided accurate determination of stoichiometry and SID revealed the lateral symmetry in the subunit arrangement. Ongoing development in native MS instrumentation and the gas-phase protein behaviors would allow more structural information to be extracted from such experiments in the near future. For instance, the experimental data from native MS and SID could be integrated with computational techniques to better define the subclasses from cryo-EM data. Our study shows that complementing high resolution structural biology

techniques such as cryo-EM with native MS offers new ways to understand heterogeneous pseudoenzyme–enzyme interactions that cannot be elucidated by single techniques alone.

METHODS

Plasmid Construction.

Gene sequences for PDX proteins were sourced from the genomic DNA of *Arabidopsis thaliana*. The clones were amplified by PCR to include the purification tag and then subcloned using Gibson Assembly in the designated vector pEU for the cell-free protein synthesis.[20,34] All reagents for PCR and Gibson reactions such as the Q5 Hot Start High-Fidelity 2X Master Mix and Gibson Assembly Master Mix were acquired from New England Biolabs. DNA primers were from ThermoFisher Scientific. For expression of PDX proteins, the following vectors were constructed: pEU_3XFLAG_PDX1.2, pEU_3XFLAG_PDX1.3, pEU_PDX1.1, pEU_PDX1.2 and pEU_PDX1.3 (the sequences of synthesized proteins are in Supplementary Figure 1). The sequences of all plasmids used in this study were verified by Sanger sequencing (MCLAB).

Cell-Free Expression and Coexpression and Purification of PDX Proteins.

Protein synthesis was carried out using Wheat Germ Protein Research Kits WEPRO7240 from CellFree Sciences. For coexpression, stock plasmids were prepared containing various amounts of pEU_3XFLAG_PDX1.2 and pEU_3XFLAG_PDX1.3. For example, in the case of coexpression at 9:1, 9 μL of 1 μg/μL pEU_3XFLAG_PDX1.2 was mixed with 1 μL of 1 μg/μL pEU_3XFLAG_PDX1.3 before the transcription step. The MINI-scale protein synthesis in the presence of fluorophore-charged lysine tRNA was conducted with the following conditions: (i) for transcription, 1 μL of 1 μg/μL stock plasmid was mixed with 9 μL of Transcription premix solution (1× Transcription buffer, 2.5 mM NTP, 8 U/μL SP6 polymerase, 8 U/μL RNase inhibitor) and left to incubate at 37 °C for 4 h, (ii) for translation, 2 μL of the resulting mRNA was mixed with 1 μL of FluoroTect Green$_{Lys}$ (Promega) and 3 μL of the wheat germ extract and then transferred under 50 μL of the translation buffer (1× SUB-AMIX SGC) within a 96-well half-area plate. The translations were conducted overnight at 15 °C and away from light in the vibration-free setting on an Eppendorf ThermoMixer C.

The MAXI-scale protein synthesis was performed using the robotic system Protemist DT II from CellFree Sciences. All steps were performed at 4 °C and at 140 rpm. The buffer exchange step of the crude extract (in order to remove dithiothreitol, DTT) was skipped. A total of 600 μL of ANTI-M2 affinity resin (50% suspension, pre-equilibrated in 1× TBS buffer, Sigma) was used per each 6 mL translation reaction. Protein binding to the resin was carried out for 1 h. The resin was subsequently washed three times with 2.5 mL of 1× TBS buffer for 5 min each time. The protein products were eluted twice with 1 mL of Elution Buffer (1× TBS buffer, 100 μg/mL 3XFLAG peptide, 4 mM DTT) each time for 30 min. Proteins were further concentrated and buffer-exchanged to the Storage Buffer (1× TBS with 4 mM DTT) using 30 kDa Amicon Ultra-2 centrifugal filter units. The concentrated samples were flash-frozen in liquid nitrogen in 10 μL aliquots and then transferred to −80 °C for storage.

General PAGE electrophoresis conditions employed in this work followed standard procedures. The fluorescent bands on the gels were detected by Typhoon FLA 9500 from GE Healthcare, set for Alexa488 dye excitation and emission settings. As both PDX1.2 and PDX1.3 proteins possess the same number of lysines and were coexpressed in the same vial and at the same experimental conditions, the incorporation rate of fluorophore-labeled lysine was expected to be identical so no adjustments for in-gel quantification measurements were performed or needed. Bioanalyzer runs were performed with the Agilent Bioanalyzer 2100 system using the P80 kit according to the manufacturer's guidelines. Bioanalyzer built-in integration algorithms were used to estimate protein size and concentration and sample purity.

Peptide Mapping.

Four micrograms of protein was diluted in 50% trifluoroethanol and further incubated at 60 °C for 2 h at 800 rpm. Then 2 mM DTT was added, and the mixture was incubated at 37 °C for 1 h with shaking. The samples were diluted in 200 μL of 25 mM ammonium bicarbonate and digested with 0.1 μg of trypsin at 37 °C for 3 h with shaking. Samples were dried down and resuspended at 0.05 μg/μL in water for LCMS analysis. Peptide solution (5 μL) was loaded on a Waters M-Class NanoAcquity LC equipped with C18 reversed phase column (packed in-house, length 70 cm, 75 μm ID, 3 μm Jupiter particle from Phenomenex). Peptides were separated using water/acetonitrile mobile phases with 0.1% formic acid over a gradient of 2 h (ramping acetonitrile gradient 1–40%) at a flow rate of 300 nL/min. Mass spectra of eluting peptides were collected on a Thermo Q-Exactive Orbitrap. Resolution of 35000 and 17500 (at m/z 200) were used for MS^1 and MS^2, respectively. The top 12 precursors were selected for higher energy collision (HCD) at 30% normalized collision energy. A dynamic exclusion of 30 s was used. Singly charged or charge unassigned precursors were excluded. LCMS data were analyzed with Byonic with a semitryptic search of the *Arabidopsis thaliana* protein database (source: *Arabidopsis* Information Resource), including common contaminants. The 3XFLAG purification tag sequence was manually appended to the original PDX1.2 and 1.3 sequences (AT3G16050.1 and AT5G01410.1). Precursor and fragment mass tolerances were set to 7 and 10 ppm. Protein N-terminal acetylation, pyroglutamic acid (Q), and oxidation (H, M, W) were included as dynamic modifications. Raw data and processing results are deposited in MassIVE (massive.ucsd.edu) with accession MSV000085233.

Intact Mass Determination by LCMS.

Protein (0.2–0.3 μg) diluted in water at 0.1 μg/μL was loaded on a Waters NanoAcquity LC equipped with a C2 reversed phase column (packed in-house, length 70 cm, 75 μm ID, 3 μm C2 particle). Intact proteins were separated using water/acetonitrile mobile phases with 0.1% formic acid over a gradient of 1 h (ramping acetonitrile gradient 10–50%) at a flow rate of 300 nL/min. Mass spectra of eluting proteins were collected on a Thermo VelosOrbitrap Elite. MS^1 spectra were acquired at 120000 resolution (at m/z 200) and were averaged over 8 microscans. ProMex[35] was used to deconvolute the data and obtain the intact mass of the proteins. LCMS data were aggregated across retention time and binned into unit mass (sum intensities within 1 Da) to generate the intact mass plot (Supplementary Figure 5). Raw

data and processing results are deposited in MassIVE (massive.ucsd.edu) with accession MSV000085233.

Native Mass Spectrometry (Native MS).

Purified PDX complexes at 0.2–0.4 mg mL^{-1} concentrations (10 μL each in 1× TBS with 4 mM DTT) were buffer exchanged into 200 mM ammonium acetate (pH 6.7) using Zeba 75 μL microspin columns from Thermofisher. Proteins were then loaded into pulled glass capillaries (GlassTip, part number BG12692N20, Scientific Instrument Services, Inc., Ringoes, NJ) with gel loading pipet tips. A platinum wire connected to the electrospray voltage was inserted into the capillary and was in contact with the protein solution. One kilovolt was applied at the capillary to sustain the nanoelectrospray on a Thermo Q-Exactive UHMR mass spectrometer. The source was kept at 200 °C, with the S-lens set to 100%. In-source fragmentation of 100 V and in-source trapping of 50 V were used for optimum desolvation to measure the mass of the complexes without extensive gas-phase dissociation. Even with in-source fragmentation and trapping voltages off, a significant amount of PDX1.3 monomers were still observed, implying that the PDX1.3 monomers were not exclusively generated by the desolvation conditions. HCD gas flow was set to 2. For each final spectrum, 500 microscans were averaged. Mass calibration was performed using cesium iodide clusters. To obtain the mass distribution of the PDX 12mer complexes, mass spectra were processed using UniDec v2.7.[36] Original data within *m/z* 8000–11000 were extracted with curved baseline subtraction, Gaussian smoothing 2.0, and binned every 3 *m/z*. Deconvolution was restricted to sampling mass every 10 Da with peak fwhm of 8.

Activity Assay.

For quantitative PLP detection, we used Enzymatic Vitamin B6 assay from A/C diagnostics (http://www.vitaminb6assay.com/). From an initial set of experiments using active enzyme PDX1.3, we determined that 5–10 μg/mL was the optimal protein concentration for conducting this assay (in place of the original design of using blood plasma) as this range results in the OD675 signal falling within the linear PLP detection range of 0–200 nmol/L (data not shown).

To compare relative enzymatic activities, all PDX samples (individual or coexpressed) were first diluted to 12.5 μg/mL concentration in TBS buffer. Then 5 μL of 12.5 μg/mL protein was combined with 5 μL of 2× Assay buffer (50 mM Tris, pH 8.0, 1 mM ribose 5-phosphate (R5P), 2 mM D-glyceraldehyde 3-phosphate (G3P), 20 mM ammonium sulfate) and incubated away from light for 30 min at 30 °C in the Eppendorf thermocycler. Note that K_m values for R5P (0.057 mM) and G3P (0.22) of PDX1.3 were previously determined;[13] thus the assays were conducted at the saturating amounts of substrates. For the calibration curve, PLP (Sigma, cat. no. 82870) was diluted in a serial manner to 400 nM, 200 nM, 100 nM, 50 nM, and 25 nM concentration in TBS buffer, further combined with an equal volume of 2× Assay buffer, and incubated away from light for 30 min at 30 °C along with PDX samples for consistency.

To quantify the amount of PLP in the mixture, 2.5 μL of each sample was further added to the 37.5 μL of Working Binding Assay buffer (with apoenzyme), provided by the assay

kit from A/C diagnostics and placed in the Corning NBS 384-well microplate. The mixture was incubated at 37 °C for 30 min at 750 rpm on the Eppendorf Thermomixer C away from light. Then 20 μL of Working Assay buffer (from the kit) was added to the mixture, and incubation was continued for additional 20 min. This was followed by the addition of 6.25 μL of Chromogen RI and 3.75 μL of Chromogen RII and 10 min incubation. The color development was measured at 675 nm using the Infinite M200 PRO microplate reader by Tecan. For accuracy, all samples and calibration standards were assayed in triplicate.

Surface Induced Dissocation (SID).

Purified PDX complexes were buffer exchanged into 200 mM ammonium acetate (pH 6.7) using the same protocol described for native MS. Then the protein solution was mixed 1:1 with 100 mM triethylammonium acetate (TEAA), and loaded into in-house pulled glass capillaries (Sutter Instruments P-97 micropipet puller, with glass capillary catalog BF100–78-10). The same instrument conditions were used as for native MS but on a modified Thermo Q-Exactive UHMR mass spectrometer as described in a previous report.[26] Briefly, a custom-made SID cell was inserted after the quadrupole and before the C-trap, replacing the original hexapole in the commercial configuration. PDX complexes were mass-isolated with a 30–35 *m/z* window for the 35+ charge state, and steered toward the surface under a controlled acceleration voltage of 160 V (5.6 keV). Product ions after the surface impact were collected in the C-trap and mass analyzed in the Orbitrap. The electrospray voltage was set to 1.2–1.3 kV, S-lens set at 100%, in-source trapping set at −100 V, and HCD gas flow set to 4. Spectra were collected at a resolution of 6250 (at *m/z* of 400) with an injection time of 1 s and averaged over 200–500 scans. Mass calibration was performed using cesium iodide clusters. SID tuning voltages are listed in Supplementary Table 3. SID spectra were deconvoluted using UniDec v4.2.1.[36] Mass was sampled every 10 Da with the peak fwhm set to 6, charge state smooth set at 2 and point smooth set to 1. Experimental peaks were fit to all possible combinations of PDX1.2/PDX1.3 oligomers using the mass list option in UniDec. A DScore of 30% and a peak threshold of 0.01 were used to filter the deconvolution results.

In the deconvoluted results (Supplementary Figure 15), peak areas were used for calculation of relative abundances of different 6mer species. The peak areas of the 6mers from both the major pathway (horizontal cleavage in Figure 3, panel c) and the minor pathways (vertical cleavage in Figure 3, panel c or secondary dissociation) cannot be directly distinguished by mass. We assume that the minor pathways generate different types of products with similar abundances. Thus, the peak areas of the 6mers from the minor pathways can be estimated by the median peak area of the detected 4mers, 5mers, 7mers, and 8mers in the same SID spectrum. For precursor 9:3, 10:2, and 11:1, only 6mers were consistently detected and the medians were set to 0. The peak areas of the 6mers from the major pathway were calculated by subtracting the contributions of minor pathways from the experimental peak areas. If the values are negative, they are reset to 0. The "trimmed" abundances were normalized and plotted (Figure 3, panel d, and Supplementary Figure 15, panel c). The relative abundances calculated without "trimming" are also presented (Supplementary Figure 15, panel d).

Cryo-EM Data Acquisition and Processing.

Quantifoil grids (658–300-AU, Ted Pella) were first glow-discharged at 15 mA for 1 min using PELCO easiGlow (Ted Pella). The grids were then transferred to a Leica EM GP plunge-freezer and brought to 85% humidity at 25 °C, and 3 μL of protein sample (0.2 mg mL^{-1} in TBS buffer with 4 mM DTT) was pipetted onto the carbon side. The grid with sample was then blotted for 3 s, plunge-frozen in the liquid ethane, and transferred to liquid nitrogen. The frozen samples were imaged on a 300 keV FEI Titan Krios at Pacific Northwest Center for Cryo-EM (https://pncc.labworks.org) equipped with a Gatan K2 summit direct electron detector. The data was collected using automated acquisition software SerialEM.[37] For PDX1.2 sample, 8106 resolution movies with 60 frames each at a dosage of 1.5 electrons per Å^2 per frame and pixel size of 0.253 Å were collected. For heterocomplex 3:1, 6904 super-resolution movies with 50 frames each at a dosage of 2 electrons per Å^2 per frame and pixel size of 0.253 Å were acquired. All data were processed using cryoSPARC v2 software[38] and were corrected for full frame motion using MotionCor2[39] and CTF estimated using CTFFIND4.[40] For PDX1.2 sample, a total of 787 982 particles were identified from template-based autopicking algorithm. Three rounds of reference-free 2D classification narrowed the pool to 265 224 particles. These were used to generate the ab initio 3D map as a reference. Homogenous refinement in cryoSPARC v2 with D_6 symmetry produced a final 3.2 Å reconstruction. For heterocomplex 3:1, a total of 510 660 particles were identified from template-based autopicking algorithm. Three rounds of reference-free 2D classification narrowed the pool to 286 642 particles. These were used to generate the ab initio 3D map as a reference. Homogenous refinement in cryosparc v2 with D_6 symmetry produced 3.2 Å reconstruction while nonuniform refinement (BETA) yielded the 3.16 Å map. Local resolution calculations on the final maps were performed in Relion 3.[41]

Atomic Modeling.

An initial homology model of PDX1.2 monomer was generated using HHPRED[42] and MODELLER[43] using comparative modeling against known structures for PDX from other organisms (2NV1, 5LNR, 4JDY, 2YZR, and 3O07). The Cryo-EM map and the aligned initial model for a single subunit were imported to PHENIX[44] and docked using *Phenix.DockInMap*, and the density region surrounding the monomer unit was mapped out via *Phenix.MapBox*. Several cycles of Phenix.RealspaceRefine were carried out to refine the atomic model for the individual monomer. Map symmetry parameters were then applied on the real space refined monomer model using *Phenix.ApplyNCSoperators* to generate the 12mer, which underwent additional rounds of *Phenix.RealspaceRefine*. Final refinement statistics and validation scores are presented in Supplementary Table 1. Multiscale models, structural comparisons, structure-based alignments, and computing of electrostatic potentials were conducted in Chimera.[45]

Rosetta Calculations.

Rosetta3[46] was used to characterize the features of PDX proteins (Supplementary Table 2). *FastRelax*[47] with the latest score function (*ref 2015_cart*) was perfomed before analyzing full atom energy values. To better pack protein structure, we used *dualspace*[48] when we

relaxed structures. Among 100 decoys ("nstruct") per each relaxation, we used the top 10% decoys in terms of lower ("better") total score values. For interface analysis and ligand binding energy calculation, we used the lowest scored decoy. For interface analysis and packing status, we used *InterfaceAnalyzer*[49] and *RosettaHoles*[50] respectively.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

REFERENCES

(1). Murphy JM; Mace PD; Eyers PA Live and let die: insights into pseudoenzyme mechanisms from structure. Curr. Opin. Struct. Biol 2017, 47, 95–104. [PubMed: 28787627]

(2). Yu JW; Jeffrey PD; Shi YG Mechanism of procaspase-8 activation by c-FLIPL. Proc. Natl. Acad. Sci. U. S. A 2009, 106, 8169–8174. [PubMed: 19416807]

(3). Murphy JM; Czabotar PE; Hildebrand JM; Lucet IS; Zhang JG; Alvarez-Diaz S; Lewis R; Lalaoui N; Metcalf D; Webb AI; Young SN; Varghese LN; Tannahill GM; Hatchell EC; Majewski IJ; Okamoto T; Dobson RCJ; Hilton DJ; Babon JJ; Nicola NA; Strasser A; Silke J; Alexander WS The Pseudokinase MLKL Mediates Necroptosis via a Molecular Switch Mechanism. Immunity 2013, 39, 443–453. [PubMed: 24012422]

(4). Schimpl M; Rush CL; Betou M; Eggleston IM; Recklies AD; van Aalten DMF Human YKL-39 is a pseudo-chitinase with retained chitooligosaccharide-binding properties. Biochem. J 2012, 446, 149–157. [PubMed: 22742450]

(5). Ribeiro AJM; Das S; Dawson N; Zaru R; Orchard S; Thornton JM; Orengo C; Zeqiraj E; Murphy JM; Eyers PA Emerging concepts in pseudoenzyme classification, evolution, and signaling. Sci. Signal 2019, 12, No. eaat9797.

(6). Ribeiro AJM; Tyzack JD; Borkakoti N; Thornton JM Identifying pseudoenzymes using functional annotation. How loss of function correlates with mutations in the catalytic site. FEBS J. 2020, 287, 4128–4140. [PubMed: 31733177]

(7). Parra M; Stahl S; Hellmann H. Vitamin B-6 and Its Role in Cell Metabolism and Physiology. Cells-Basel 2018, 7, 84.

(8). Mooney S; Leuendorf JE; Hendrickson C; Hellmann H. Vitamin B6: A Long Known Compound of Surprising Complexity. Molecules 2009, 14, 329–351. [PubMed: 19145213]

(9). Hsu CC; Cheng CH; Hsu CL; Lee WJ; Huang SC; Huang YC Role of vitamin B6 status on antioxidant defenses, glutathione, and related enzyme activities in mice with homocysteine-induced oxidative stress. Food Nutr Res. 2015, 59, No. 25702.

(10). Zhang YF; Liu B; Li XH; Ouyang ZG; Huang L; Hong YB; Zhang HJ; Li DY; Song FM The de novo Biosynthesis of Vitamin B6 Is Required for Disease Resistance Against Botrytis cinerea in Tomato. Mol. Plant-Microbe Interact 2014, 27, 688–699. [PubMed: 24678833]

(11). Chandrasekaran M; Paramasivan M; Chun SC Bacillus subtilis CBR05 induces Vitamin B6 biosynthesis in tomato through the de novo pathway in contributing disease resistance against Xanthomonas campestris pv. vesicatoria. Sci. Rep 2019, 9, No. 6495. [PubMed: 31019197]

(12). Tambasco-Studart M; Titiz O; Raschle T; Forster G; Amrhein N; Fitzpatrick TB Vitamin B6 biosynthesis in higher plants. Proc. Natl. Acad. Sci. U. S. A 2005, 102, 13687–13692. [PubMed: 16157873]

(13). Moccand C; Boycheva S; Surriabre P; Tambasco-Studart M; Raschke M; Kaufmann M; Fitzpatrick TB The pseudoenzyme PDX1.2 boosts vitamin B6 biosynthesis under heat and oxidative stress in Arabidopsis. J. Biol. Chem 2014, 289, 8203–8216. [PubMed: 24505140]

(14). Leuendorf JE; Mooney SL; Chen LY; Hellmann HA Arabidopsis thaliana PDX1.2 is critical for embryo development and heat shock tolerance. Planta 2014, 240, 137–146. [PubMed: 24748553]

(15). Denslow SA; Rueschhoff EE; Daub ME Regulation of the Arabidopsis thaliana vitamin B6 biosynthesis genes by abiotic stress. Plant Physiol. Biochem 2007, 45, 152–161. [PubMed: 17344055]

(16). Leuendorf JE; Osorio S; Szewczyk A; Fernie AR; Hellmann H. Complex assembly and metabolic profiling of Arabidopsis thaliana plants overexpressing vitamin B(6) biosynthesis proteins. Mol. Plant 2010, 3, 890–903. [PubMed: 20675613]

(17). Robinson GC; Kaufmann M; Roux C; Martinez-Font J; Hothorn M; Thore S; Fitzpatrick TB Crystal structure of the pseudoenzyme PDX1.2 in complex with its cognate enzyme PDX1.3: a total eclipse. Acta Crystallogr. D Struct Biol 2019, 75, 400–415. [PubMed: 30988257]

(18). Madin K; Sawasaki T; Ogasawara T; Endo Y. A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes. Proc. Natl. Acad. Sci. U. S. A 2000, 97, 559–564. [PubMed: 10639118]

(19). Sawasaki T; Hasegawa Y; Tsuchimochi M; Kamura N; Ogasawara T; Kuroita T; Endo Y. A bilayer cell-free protein synthesis system for high-throughput screening of gene products. FEBS Lett. 2002, 514, 102–105. [PubMed: 11904190]

(20). Novikova IV; Sharma N; Moser T; Sontag R; Liu Y; Collazo MJ; Cascio D; Shokuhfar T; Hellmann H; Knoblauch M; Evans JE Protein structural biology using cell-free platform from wheat germ. Adv. Struct Chem. Imaging 2018, 4, 13. [PubMed: 30524935]

(21). Titiz O; Tambasco-Studart M; Warzych E; Apel K; Amrhein N; Laloi C; Fitzpatrick TB PDX1 is essential for vitamin B6 biosynthesis development and stress tolerance in Arabidopsis. Plant J. 2006, 48, 933–946. [PubMed: 17227548]

(22). Wagner S; Bernhardt A; Leuendorf JE; Drewke C; Lytovchenko A; Mujahed N; Gurgui C; Frommer WB; Leistner E; Fernie AR; Hellmann H. Analysis of the Arabidopsis rsr4–1/pdx1–3 mutant reveals the critical function of the PDX1 protein family in metabolism development, and vitamin B6 biosynthesis. Plant Cell 2006, 18, 1722–1735. [PubMed: 16766694]

(23). Mortensen DN; Williams ER Electrothermal supercharging of proteins in native MS: effects of protein isoelectric point buffer, and nanoESI-emitter tip size. Analyst 2016, 141, 5598–5606. [PubMed: 27441318]

(24). Leney AC Subunit pI Can Influence Protein Complex Dissociation Characteristics. J. Am. Soc. Mass Spectrom 2019, 30, 1389–1395. [PubMed: 31077092]

(25). Hall Z; Hernandez H; Marsh JA; Teichmann SA; Robinson CV The Role of Salt Bridges. Structure 2013, 21, 1325–1337. [PubMed: 23850452]

(26). VanAernum ZL; Gilbert JD; Belov ME; Makarov AA; Horning SR; Wysocki VH Surface-Induced Dissociation of Noncovalent Protein Complexes in an Extended Mass Range Orbitrap Mass Spectrometer. Anal. Chem 2019, 91, 3611–3618. [PubMed: 30688442]

(27). Zhou MW; Dagan S; Wysocki VH Impact of charge state on gas-phase behaviors of noncovalent protein complexes in collision induced dissociation and surface induced dissociation. Analyst 2013, 138, 1353–1362. [PubMed: 23324896]

(28). Stiving AQ; VanAernum ZL; Busch F; Harvey SR; Sarni SH; Wysocki VH Surface-Induced Dissociation: An Effective Method for Characterization of Protein Quaternary Structure. Anal. Chem 2019, 91, 190–209. [PubMed: 30412666]

(29). Harvey SR; Seffernick JT; Quintyn RS; Song Y; Ju Y; Yan J; Sahasrabuddhe AN; Norris A; Zhou MW; Behrman EJ; Lindert S; Wysocki VH Relative interfacial cleavage energetics of protein complexes revealed by surface collisions. Proc. Natl. Acad. Sci. U. S. A 2019, 116, 8143–8148. [PubMed: 30944216]

(30). Rodrigues MJ; Windeisen V; Zhang Y; Guedez G; Weber S; Strohmeier M; Hanes JW; Royant A; Evans G; Sinning I; Ealick SE; Begley TP; Tews I. Lysine relay mechanism coordinates intermediate transfer in vitamin B6 biosynthesis. Nat. Chem. Biol 2017, 13, 290. [PubMed: 28092359]

(31). Robinson GC; Kaufmann M; Roux C; Fitzpatrick TB Structural definition of the lysine swing in Arabidopsis thaliana PDX1: Intermediate channeling facilitating vitamin B-6 biosynthesis. Proc. Natl. Acad. Sci. U. S. A 2016, 113, E5821–E5829. [PubMed: 27647886]

(32). Strickler SS; Gribenko AV; Gribenko AV; Keiffer TR; Tomlinson J; Reihle T; Loladze VV; Makhatadze GI Protein stability and surface electrostatics: A charged relationship. Biochemistry 2006, 45, 2761–2766. [PubMed: 16503630]

(33). Dell'Aglio E; Boycheva S; Fitzpatrick TB The Pseudoenzyme PDX1.2 Sustains Vitamin B6 Biosynthesis as a Function of Heat Stress. Plant Physiol. 2017, 174, 2098–2112. [PubMed: 28550206]

(34). Sawasaki T; Hasegawa Y; Tsuchimochi M; Kasahara Y; Endo Y. Construction of an efficient expression vector for coupled transcription/translation in a wheat germ cell-free system. Nucleic Acids Symp. Ser 2000, 44, 9–10.

(35). Park J; Piehowski PD; Wilkins C; Zhou M; Mendoza J; Fujimoto GM; Gibbons BC; Shaw JB; Shen Y; Shukla AK; Moore RJ; Liu T; Petyuk VA; Tolic N; Pasa-Tolic L; Smith RD; Payne SH; Kim S. Informed-Proteomics: open-source software package for top-down proteomics. Nat. Methods 2017, 14, 909–914. [PubMed: 28783154]

(36). Marty MT; Baldwin AJ; Marklund EG; Hochberg GK; Benesch JL; Robinson CV Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. Anal. Chem 2015, 87, 4370–4376. [PubMed: 25799115]

(37). Mastronarde DN Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol 2005, 152, 36–51. [PubMed: 16182563]

(38). Punjani A; Rubinstein JL; Fleet DJ; Brubaker MA cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 2017, 14, 290–296. [PubMed: 28165473]

(39). Zheng SQ; Palovcak E; Armache JP; Verba KA; Cheng YF; Agard DA MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 2017, 14, 331–332. [PubMed: 28250466]

(40). Rohou A; Grigorieff N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol 2015, 192, 216–221. [PubMed: 26278980]

(41). Zivanov J; Nakane T; Forsberg BO; Kimanius D; Hagen WJH; Lindahl E; Scheres SHW New tools for automated highresolution cryo-EM structure determination in RELION-3. Elife 2018, 7, No. e42166.

(42). Zimmermann L; Stephens A; Nam SZ; Rau D; Kubler J; Lozajic M; Gabler F; Soding J; Lupas AN; Alva V. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J. Mol. Biol 2018, 430, 2237–2243. [PubMed: 29258817]

(43). Webb B; Sali A. Protein Structure Modeling with MODELLER. Methods Mol. Biol 2017, 1654, 39–54. [PubMed: 28986782]

(44). Adams PD; Afonine PV; Bunkoczi G; Chen VB; Echols N; Headd JJ; Hung LW; Jain S; Kapral GJ; Grosse Kunstleve RW; McCoy AJ; Moriarty NW; Oeffner RD; Read RJ; Richardson DC; Richardson JS; Terwilliger TC; Zwart PH The Phenix software for automated determination of macromolecular structures. Methods 2011, 55, 94–106. [PubMed: 21821126]

(45). Pettersen EF; Goddard TD; Huang CC; Couch GS; Greenblatt DM; Meng EC; Ferrin TE UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem 2004, 25, 1605–1612. [PubMed: 15264254]

(46). Leaver-Fay A; Tyka M; Lewis SM; Lange OF; Thompson J; Jacak R; Kaufman K; Renfrew PD; Smith CA; Sheffler W; Davis IW; Cooper S; Treuille A; Mandell DJ; Richter F; Ban YE; Fleishman SJ; Corn JE; Kim DE; Lyskov S; Berrondo M; Mentzer S; Popovic Z; Havranek JJ; Karanicolas J; Das R; Meiler J; Kortemme T; Gray JJ; Kuhlman B; Baker D; Bradley P. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. Methods Enzymol. 2011, 487, 545–574. [PubMed: 21187238]

(47). Tyka MD; Keedy DA; Andre I; Dimaio F; Song Y; Richardson DC; Richardson JS; Baker D. Alternate states of proteins revealed by detailed energy landscape mapping. J. Mol. Biol 2011, 405, 607–618. [PubMed: 21073878]

(48). Conway P; Tyka MD; DiMaio F; Konerding DE; Baker D. Relaxation of backbone bond geometry improves protein energy landscape modeling. Protein Sci. 2014, 23, 47–55. [PubMed: 24265211]

(49). Stranges PB; Kuhlman B. A comparison of successful and failed protein interface designs highlights the challenges of designing buried hydrogen bonds. Protein Sci. 2013, 22, 74–82. [PubMed: 23139141]

(50). Sheffler W; Baker D. RosettaHoles: rapid assessment of protein core packing for structure prediction refinement, design, and validation. Protein Sci. 2009, 18, 229–239. [PubMed: 19177366]
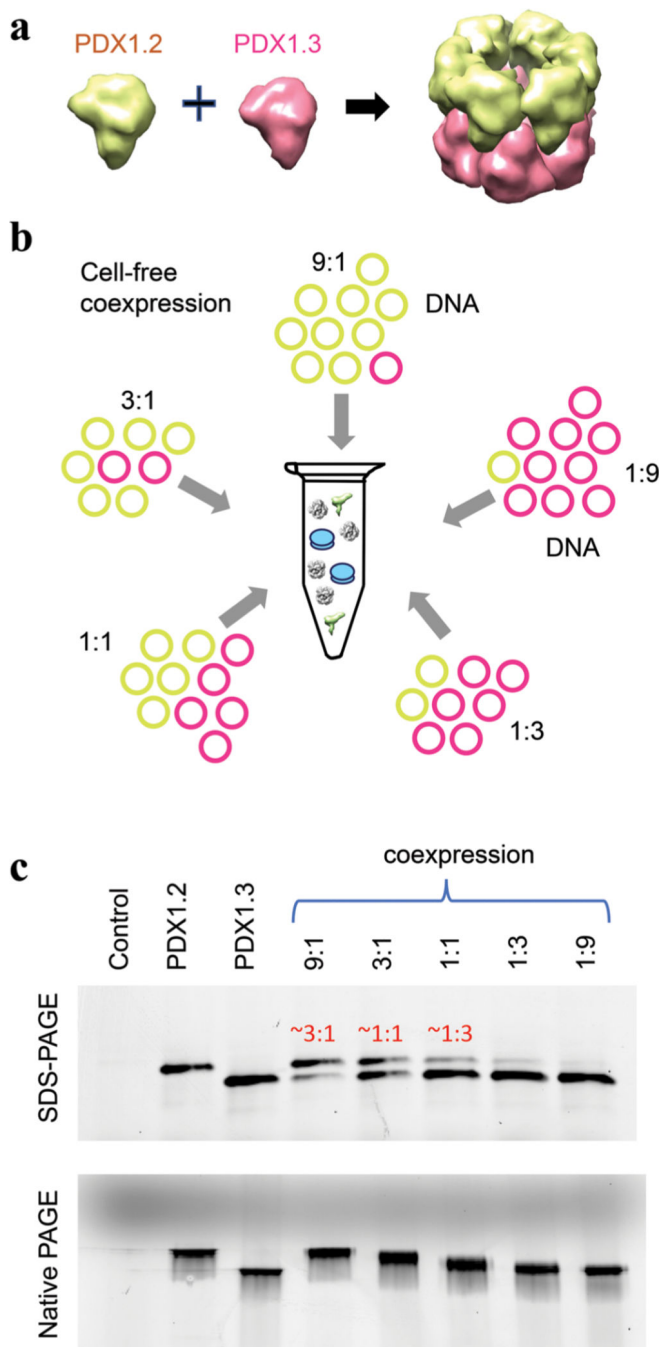
Figure 1.
Coexpression of PDX1.2 and PDX1.3 proteins. (a) Schematic showing the prior proposed heteroassembly of PDX1.2–PDX1.3 association. (b) Illustration of experimental design where variable amounts of DNA plasmids are fed into the cell-free protein translation system to control and study the coexpression complexes. (c) PAGE analysis of newly synthesized PDX proteins, detected by fluorescence in a crude mixture. The control sample is a translation reaction where DNA plasmid is not supplemented. PAGE data include a denaturing SDS-PAGE gel on the top and native PAGE on the bottom. In the coexpressed

samples, protein molar ratios, determined by in-gel quantifications, are shown in red. Coexpressed conditions are denoted as 9:1, 3:1, 1:1, 1:3, and 1:9, where these numbers correspond to DNA template ratios of PDX1.2 to PDX1.3.
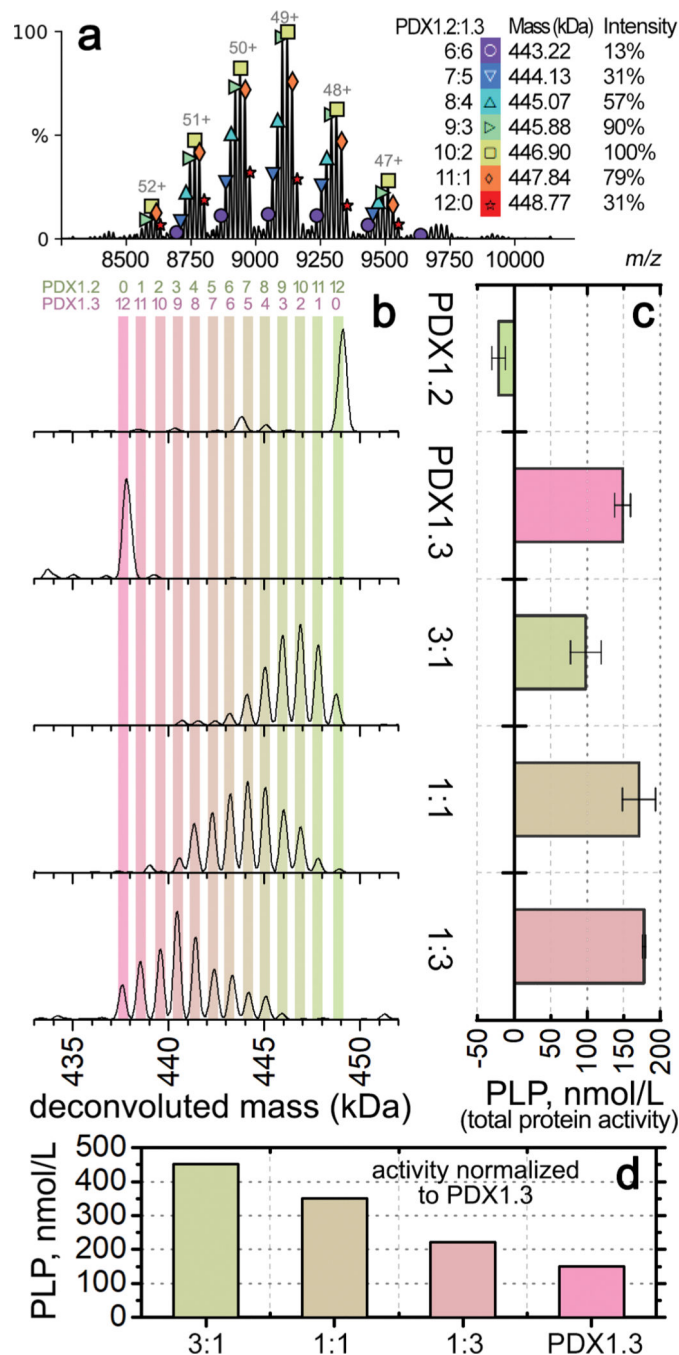
Figure 2.
Oligomeric and enzymatic characterization of individual PDX homo- and heterocomplexes.
(a) Representative native MS spectrum for heterocomplex 3:1, zoomed into the 12mer
region. Each symbol above the peak indicates one 12mer species, with their assignment,
mass, and relative intensity shown on the right. Each peak with the same symbol is the same
species carrying a different number of charges. Annotation was performed using UniDec.[36]
(b) Deconvoluted mass distributions of PDX 12mer complexes from different samples as
labeled on the right. Color bars are overlaid to show the numbers of PDX1.2 and PDX1.3

protomers in the PDX 12mers as noted by the numbers on the top. (c) Histograms showing the enzymatic activities based on total PDX protein concentration. Error bars represent the standard deviations from triplicate experiments. The sample names match to the mass distributions shown horizontally in panel b. (d) Histograms showing the enzymatic activities normalized to the amount of PDX1.3 in each sample.
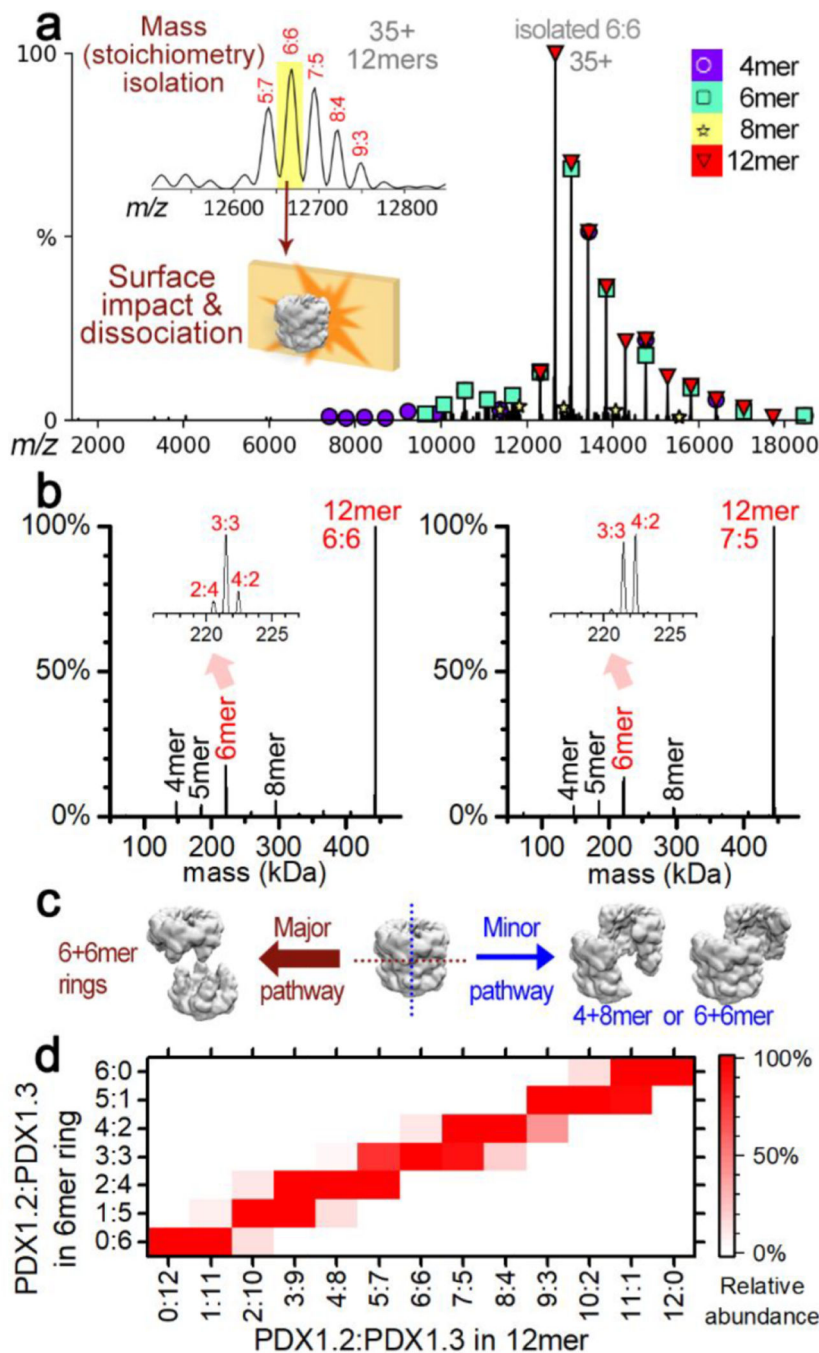
Figure 3.
Dissection of PDX 12mers in the gas phase by surface induced dissociation (SID). (a) Representative SID spectrum for a heteromer with a single stoichiometry of 6:6 (PDX1.2/PDX1.3). Mass isolated 6:6 was collided with a surface to produce 6mers and other subcomplexes detected in the mass spectrum, which are labeled by the different symbols. The key to the symbols is on the top right. (b) Deconvoluted mass distribution of the SID spectra for 6:6 (left) and 7:5 (right). Zoom-in views of the 6mer region are shown in the insets. (c) Proposed dissociation pathways for PDX by SID. The cleavage of the 12mer

in the horizontal direction is a major pathway (left). A minor pathway for cleaving in the vertical direction likely produced several other species at low abundance. (d) Heat map showing the relative abundance (normalized to max) of the released 6mer rings (vertical axis) from each stoichiometry of 12mers (horizontal axis) following the major pathway. The contribution of 6mers from minor pathways were estimated by the abundance of 4-, 5-, 7-, and 8mers and was subtracted from the raw experimental data as described in the Methods.
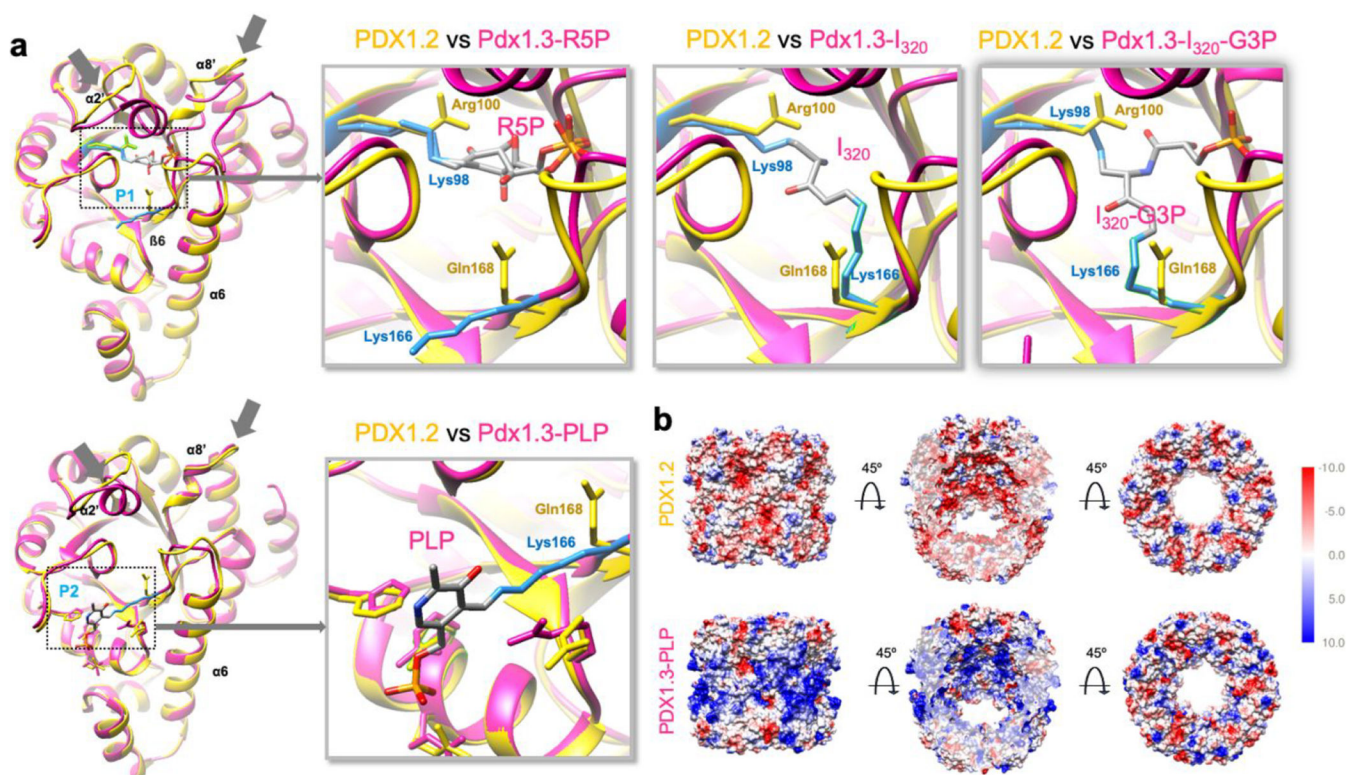
Figure 4.
Cryo-EM structure of PDX1.2 versus the crystal structures of PDX1.3. (a) Comparison of PDX1.2 individual subunit versus four different crystal structures, corresponding to the different stages of PLP synthesis at active sites P1 and P2. The PDX1.2 monomer subunit is in yellow, PDX1.3 subunit is in magenta. Key residues in PDX1.3 catalysis, Lys98 and Lys166, are shown in light blue. Corresponding residues in PDX1.2 include Arg100 and Gln168 (in yellow). Substrate, intermediate products, and PLP are colored by heteroatom. Structure-based sequence alignments and annotations can be found in the Supporting Information (Supplementary Figure 20). (b) Coulombic electrostatic potential maps. The middle images highlight the sliced view to show the interior potential of the complex.
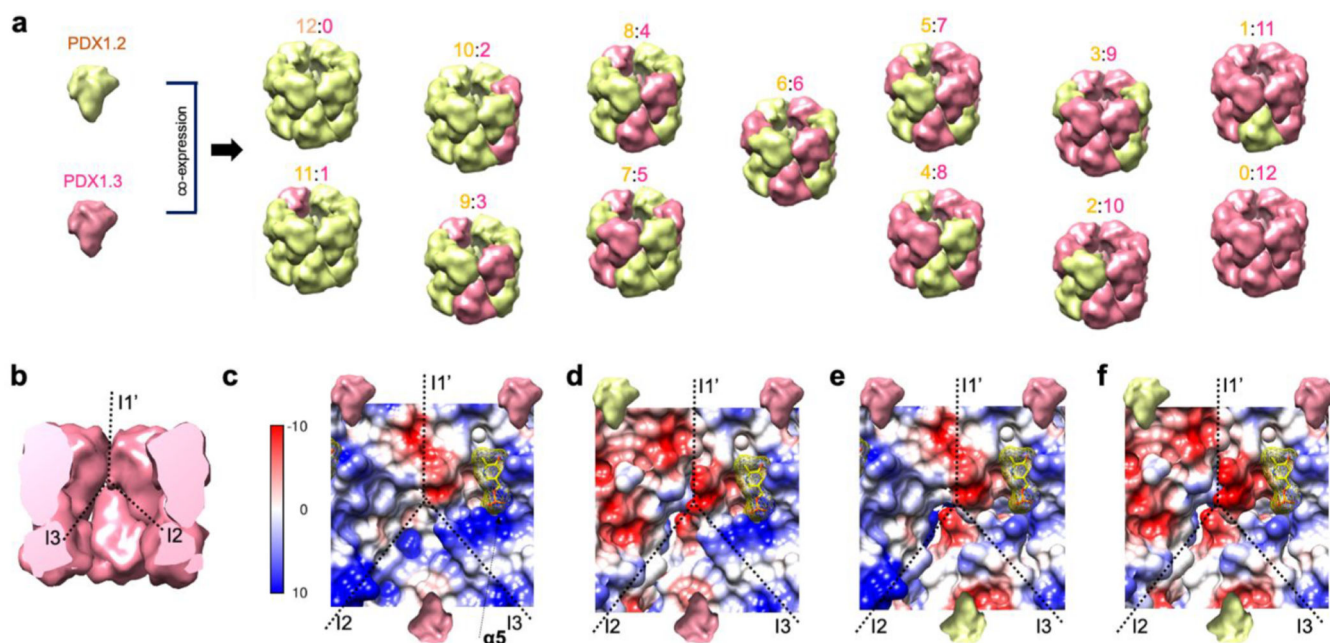
Figure 5.
Heteromeric assembly mechanism and the proposed electrostatic reorganization around the P2 catalytic site. (a) Revised model of PDX1.2 and PDX1.3 coassembly. Each subclass is annotated as 12:0, 11:1, 10:2, 9:3, etc., where the integers correspond to the number of PDX1.2 subunits relative to number of PDX1.3 subunits. Only one representative combination per class is shown while the total number of possible combinations is variable based on stoichiometry, symmetry operations, and lateral favorability in the rings. (b) The potential impact of PDX1.2 on PDX1.3 activity can be through I1′ and I3 interfaces, surrounding the catalytic site P2 in each monomer. (c) Electrostatic surface potential map of the area surrounding the P2 site in PDX1.3 homocomplex. The yellow molecule represents PLP. The models (d–f) were constructed by replacing PDX1.3 monomers with the cryo-EM structure of PDX1.2 monomers. Note the significantly changed electrostatic potential on the upper left of the P2 site in model d.